Available online at [www.sciencedirect.com](http://www.sciencedirect.com/)

[Electronic Notes in Theoretical Computer Science 319 (2015) 137–164](http://dx.doi.org/10.1016/j.entcs.2015.12.010)

[www.elsevier.com/locate/entcs](http://www.elsevier.com/locate/entcs)

Dirichlet is Natural

Vincent Danos[1](#_bookmark0)

*D´epartement d’Informatique Ecole normale sup´erieure Paris, France*

Ilias Garnier[2](#_bookmark0)

*School of Informatics University of Edinburgh Edinburgh, United Kingdom*

**Abstract**

Giry and Lawvere’s categorical treatment of probabilities, based on the probabilistic monad *G*, offer an elegant and hitherto unexploited treatment of higher-order probabilities. The goal of this paper is to follow this formulation to reconstruct a family of higher-order probabilities known as the *Dirichlet process*. This family is widely used in non-parametric Bayesian learning.

Given a Polish space *X*, we build a family of higher-order probabilities in *G*(*G*(*X*)) indexed by *M∗*(*X*) the set of non-zero finite measures over *X*. The construction relies on two ingredients. First, we develop a method to map a zero-dimensional Polish space *X* to a projective system of finite approximations, the limit of which is a zero-dimensional compactification of *X*. Second, we use a functorial version of Bochner’s probability extension theorem adapted to Polish spaces, where consistent systems of probabilities over a projective system give rise to an actual probability on the limit. These ingredients are combined with known combinatorial properties of Dirichlet processes on finite spaces to obtain the Dirichlet family *DX* on *X*. We prove that the family *DX* is a *natural transformation* from the monad *M∗* to *G◦G* over Polish spaces, which in particular is continuous in its parameters. This is an improvement on extant constructions of *DX* [[17](#_bookmark53),[26](#_bookmark62)].

*Keywords:* probability, topology, category theory, monads

# Introduction

It has been argued that exact bisimulations between Markovian systems are better conceptualized using the more general notion of bisimulation metrics [[29](#_bookmark65)]. This is because there are frequent situations where one can only estimate the transition probabilities of a Markov chain (MC). [3](#_bookmark1) Such uncertainties lead one naturally to

1 [vincent.danos@ens.fr](mailto:vincent.danos@ens.fr)

2 [igarnier@inf.ed.ac.uk](mailto:igarnier@inf.ed.ac.uk)

3 Even though the existence of symmetries in physical systems can sometimes lead to exact bisimulations which depend only on structure and not on the actual values of transition probabilities [[28](#_bookmark64)]. There are at- tempts, parallel to bisimulation metrics, at defining robustly the satisfaction of a temporal logic formula [[14](#_bookmark50)]

<http://dx.doi.org/10.1016/j.entcs.2015.12.010>

1571-0661/© 2015 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

using a metric-based notion of approximate equivalence as a more robust way of comparing processes than exact bisimulations. Here, we wish to take a new look at this issue of uncertainty in the model and suggest a novel and richer framework to deal with it. We keep the idea of using a robust means of comparison (typically the Kantorovich or Prohorov metrics lifted to MCs), but we add a second idea: namely to introduce a way of quantifying the uncertainty in the chains being compared.

To quantify uncertainty in the Markov chains, we propose to explore in the longer term concepts of “uncertain Markov chains” as elements of type *X → G*2(*X*), where *X* is an object of **Pol**, the category of Polish spaces (separable and completely metrisable spaces) and *G* is the Giry probability functor. This is to say that the chain takes values in “random probabilities” (ie probabilities of probabilities). [4](#_bookmark2)

This natural treatment of behavioural uncertainty in probabilistic models will al- low one to formulate a notion of (Bayesian) learning and therefore to obtain notions of 1) models which can learn under observations and 2) of behavioural comparisons which can incorporate data and reduce uncertainty. Bisimulation metrics between processes become random variables and learning should decrease their variability.

One needs to set up a sufficiently general framework for learning under ob- servation within the coalgebraic approach. Learning a probability in a Bayesian framework is naturally described as a (stochastic) process of type *G*2(*X*) *→ G*2(*X*) (so *G*2(*X*) *→ G*3(*X*) really!) driven by observations. For finite *X*s this setup poses no difficulty, but for more general spaces, one needs to construct a computational handle on *G*2(*X*) - the space of uncertain or higher-order probabilities. This is what we do in this paper.

To this effect, we build a theory of Dirichlet-like processes in **Pol**. Dirichlet processes [[1](#_bookmark37),[16](#_bookmark52)] form a family of elements in *G*2(*X*) indexed by finite measures over *X* [[1](#_bookmark37), p.17] [5](#_bookmark3) and which is closed under Bayesian learning.

Integral to our construction is a method of “decomposition/recomposition” which allows us to build higher-probabilities via finite approximations of the under- lying space (the limit of which lead to a compactification of the original space). In order to lift finite higher-probabilities we use a bespoke extension theorem of the Kolmogorov-Bochner type in **Pol** (Sec. [2.3](#_bookmark9)). Kolmogorov consistent assignments of probabilities on finite partitions of measurable spaces (or finite joint distributions of stochastic processes) can be seen systematically as points in the image under *G* of projective (countable co-directed) diagrams in **Pol**.

Using the above we show that Dirichlet-like processes in **Pol** can be seen as natural transformations from *M∗* (the monad of non-zero finite measures on **Pol**) to *G*2 built up from finite discrete spaces. The finite version of naturality goes under the name of “aggregation laws” in the statistical literature and can be traced back to the “infinite divisibility” of the one building block, namely the Γ distribution. (This opens up the possibility of an axiomatic version of the construction presented here, see conclusion.)

4 Another possibility is to consider uncertain chains as elements of *G*(*X → G*(*X*)), but, unless *X* is compact, this takes us outside of **Pol**.

5 Eg as for Poisson point processes.

# Notations & basic facts

We provide a primer of general topology as used in the paper in Appendix [A](#_bookmark72). A useful reference on the matter is [[11](#_bookmark42)]. Weak convergence of probability measures is treated in [[7](#_bookmark43),[27](#_bookmark63)].

* 1. *Finite measures on Polish spaces and the Giry monad*

## Weak topology

A measure *P* on a topological space *X* is a positive countably additive set function defined on the Borel *σ*-algebra *B*(*X*) verifying *P* (*∅*) = 0. We will only consider finite measures on Polish spaces, i.e. *P* (*X*) *< ∞*. When *P* (*X*)= 1, *P* is a *probability measure*. We write *G*(*X*) for the space of all probability measures over *X* with the *weak topology* [[7](#_bookmark43),[27](#_bookmark63)], the initial topology for the family of evaluation

∫

maps *EVf* = *P '→ X fdP* where *f* ranges in *Cb*(*X*) and where (*Cb*(*X*)*, ·∞*) is the Banach space of real-valued continuous bounded functions over *X* with the sup norm. A neighbourhood base for a measure *P ∈ G*(*X*) is given by the sets

*NP* (*f*1*,..., fn, ϵ*1*,..., ϵn*)= *Q* ∫ *fidP −* ∫ *fidQ* *< ϵi,* 1 *≤ i ≤ n*

where *fi ∈ Cb*(*X*)*, ϵi >* 0. One can restrict w.l.o.g. to the subset of real-valued bounded uniformly continuous functions, noted *Ub*(*X*). Importantly, if *X* is Polish the weak topology on *G*(*X*) is also Polish (see e.g. Parthasarathy, [[27](#_bookmark63)] Chap. 2.6) and metrisable by the Wasserstein-Monge-Kantorovich distance [[31](#_bookmark67)]. We denote the convergence of a sequence (*Pn ∈ G*(*X*))*n∈*N to *P ∈ G*(*X*) in the weak sense by *Pn - P* . The “Portmanteau” theorem ([[7](#_bookmark43)], Theorem 2.1) asserts that *Pn - P* is equivalent to *Pn*(*B*) *→ P* (*B*) for all *P* -continuity sets *B*, i.e. Borel sets s.t. *P* (*∂B*)= 0. *P* -continuity sets form a Boolean algebra ([[27](#_bookmark63)], Lemma 6.4). The *support* of a probability *P ∈ G*(*X*) is noted *supp*(*P* ) and is defined as the smallest closed set such that *P* (*supp*(*P* )) = 1. For *X, Y* Polish and *P ∈ G*(*X*)*,Q ∈ G*(*Y* ), we write *P ⊗ Q ∈ G*(*X × Y* ) the *product probability*, so that (*P ⊗ Q*)(*BX × BY* ) = *P* (*BX* )*Q*(*BY* ).

## Giry monad

The operation *G* can be extended to a functor *G* : **Pol** *→* **Pol** compatible with the Giry monad structure (*G, δ, μ*) [[19](#_bookmark55)]. For any continuous map *f* : *X → Y* we set *G*(*f* )(*P* )= *B ∈ B*(*Y* ) *'→ P* (*f−*1(*B*)), i.e. *G*(*f* )(*P* ) is the pushforward measure. For a given *X*, *δX* : *X → G*(*X*) is the Dirac delta at *x* while *μX* : *G*2(*X*) *→ G*(*X*) is defined as averaging: *μX* (*P* ) = *B ∈ B*(*X*) *'→ G*(*X*) *EVBdP* where *EVB* = *Q ∈ G*(*X*) *'→ Q*(*B*) evaluates a probability on the Borel set *B*. We have the “change of variables” formula: for all *P ∈ G*(*X*), *f* : *X → Y* and *g* : *Y →* R bounded measurable, *Y gdG*(*f* )(*P* )= *X g◦fdP* . Finally, *G* preserves surjectivity, injectivity and openness:

∫

∫ ∫

**Lemma 2.1** (i) *f* : *X → Y is injective if and only if G*(*f* ) *is injective;*

1. *f is surjective if and only if G*(*f* ) *is surjective.*
2. *If f is an embedding, so is G*(*f* )*.*

**Proof.** We recall that elements of *G*(*X*) for *X* Polish verify the *Radon property* : for all Borel set *B ∈ B*(*X*) and all *P ∈ G*(*X*), *P* (*B*)= sup *{P* (*K*) *| K ⊆ B, K* compact*}* (see [[9](#_bookmark44)], Chap. 7). (i) Let *f* be an injective continuous map. Let *P, Q ∈ G*(*X*) be such that *P* (*B*) */*= *Q*(*B*) for some Borel set *B*. Then, there must exist a compact *K ⊆ B* such that *P* (*B*) */*= *Q*(*B*). The set *f* (*K*) is compact, hence Borel; by in- jectivity *G*(*f* )(*P* )(*f* (*K*)) = *P* (*K*) and similarly for *Q*, therefore *G*(*f* )(*P* )(*f* (*K*)) */*= *G*(*f* )(*Q*)(*f* (*K*)). Conversely, if *G*(*f* ) is injective then it is in particular injective on the set *{δx | x ∈ X} ⊆ G*(*X*), therefore *f* is injective. (ii) Let *f* be surjective continuous. Let *Q ∈ G*(*Y* ) be some probability. By the measurable selection theo- rem [[32](#_bookmark68)], there exists a measurable function *g* : *Y → X* such that *g*(*y*) *∈ f—*1(*y*), which implies *f ◦ g* = *idY* . Let *P* be the pushforward measure of *Q* through *g*,

i.e. *P* (*B*) , *Q ◦ g—*1. By surjectivity of *f* , *P* (*X*) = 1, therefore *P ∈ G*(*X*). The

identity *f ◦ g* = *idY* entails *G*(*f* )(*P* ) = *Q*. Conversely, assume *G*(*f* ) is surjective. Since *{δy | y ∈ Y } ⊆ G*(*Y* ), there must exist for each *y* a *Py ∈ G*(*X*) such that *δy*(*y*)= (*P ◦ f—*1)(*y*) *>* 0, therefore *f* is surjective. (iii) Assume *f* is an embedding. Let *NP* (*g*1*,..., gn, ϵ*1*,..., ϵn*) be some basic neighbourhood of some *P ∈ G*(*X*), and let *Pj ∈ NP* be in the neighbourhood of *P* , i.e. *X gidP − X gidPj < ϵi* for 1 *≤ i ≤ n*. Note that since *f* is an embedding, for each *gi ∈ Cb*(*X*) there exists a

∫ ∫

*gj ∈ Cb*(*f* (*X*)) verifying *gj*(*f* (*x*)) = *gi*(*x*). Therefore:

*i* *i*

∫ *gjdG*(*f* )(*P* ) *−* ∫ *gjdG*(*f* )(*P j*) = ∫ *gj ◦ fdP −* ∫ *gj ◦ fdPj*

*i i* *i*

*Y Y*  ∫*X*

*i*

∫ *X*

*gidP −*

=

*X*

*gidPj < ϵi*

*X* *2*

## Finite measures

The set of all finite non-negative Borel measures on a Polish space, noted *M* (*X*), is a Polish space when endowed with the weak topology ([[9](#_bookmark44)] Theorem 8.9.4). *M* : **Pol** *→* **Pol** is a functor extending *G*, mapping continuous functions to the corresponding pushforward morphism. The monad multiplication *μX* can be conservatively extended to a morphism from *M* 2(*X*) to *M* (*X*) by defining *μX* (*P* ) = *B ∈ B*(*X*) *'→ M* (*X*) *EVBdP* . The everywhere zero measure, noted **0**, is an element of *M* (*X*) that we might want to exclude: *M* (*X*) being Hausdorff im- plies that the set of nonzero measures *M∗* (*X*) , *M* (*X*) *\ {***0***}* is open, hence *Gδ*, hence Polish as a subspace of *M* (*X*). A measure *Q ∈ M* (*X*) is *strictly positive* if for all nonempty open sets *U ⊆ X*, *Q*(*U* ) *>* 0. Equivalently, *Q* is strictly positive if and only if *supp*(*Q*)= *X*.

∫

**Lemma 2.2** *Strictly positive ﬁnite measures on a Polish space X form (when they exist) a Polish subspace of M* (*X*)*. We denote this subspace by M* +(*X*)*.*

**Proof.** It is sufficient to show that *M* +(*X*) is a *Gδ* set in *M* (*X*). Let *{On}n∈*N

be a countable base of *X*. Strict positivity of a measure *Q* is equivalent to having

*Q*(*On*) *>* 0 for all nonempty *On*, therefore *M* +(*X*)= *n {Q ∈ M* (*X*) *| Q*(*On*) *>* 0*}*. Clearly *{Q | Q*(*On*)= 0*}* is closed in the weak topology, therefore *M* +(*X*) is a *Gδ*, and forms a Polish subspace of *M* (*X*). *2*

Summing up, we have for *X* Polish the following inclusions of Polish spaces of finite measures:

*M* +(*X*) *⊆ M∗* (*X*) *⊆ M* (*X*)

Note also that *M* and *M∗* are endofunctors on **Pol** but *M* + is *not*, unless one restricts to the subcategory of epimorphisms.

## Normalisation of measures

We note *νX* : *M∗* (*X*) *→ G*(*X*) the continuous map taking any measure *Q ∈ M∗* (*X*) to its normalisation *νX* (*Q*) , *B ∈ B*(*X*) *'→ Q*(*B*)*/ |Q|*, where *|Q|* , *Q*(*X*) is the total mass of the measure. *νX* verifies an useful property:

**Lemma 2.3** *ν* : *M∗ ⇒ G is natural.*

**Proof.** Let *f* : *X → Y* be a continuous map. We have:

(*G*(*f* ) *◦ νX* )(*Q*)= *νX* (*Q*) *◦ f—*1 =

*Q ◦ f—*1

*Q*(*X*)

*Q ◦ f—*1

= *Q*(*f—*1(*Y* ))

= *νY ◦ M∗* (*f* )(*Q*)

*2*

## Densities and convolution

The Radon-Nikodym theorem asserts that measures in *G*(R) absolutely contin- uous with respect to the Lebesgue measure admit integral representations such that *P* (*A*)= *A fdx*. In this case *P* is said to have *density f* with respect to the Lebesgue measure. *f* is sometimes noted *dP* , where *λ* denotes Lebesgue. For *P, Q ∈ G*(R) with respective densities w.r.t. Lebesgue *fP , fQ*, the convolution product of *P* and *Q* is defined to be the measure *P ∗ Q* having density *fP ∗Q*(*x*)= R *fP* (*x*)*fQ*(*x− t*)*dt* (see Kallenberg [[22](#_bookmark58)], Lemma 1.28).

∫

∫

*dλ*

## Finitely supported measures

When *X* is a finite, discrete space such that *X* = *{x*1*,..., xn}*, *G*(*X*) is in bijec- tion with the simplex Δ*n ⊆* R*n*, where Δ*n* = *{*(*p*1*,..., pn*) *⊆* R*n | pi ≥* 0*, pi* = 1*}*. Notice that Δ*n* is an *n −* 1 dimensional space. *M* (*X*) corresponds to the positive

Σ

orthant, noted R*n*

*≥*0

. Since for *X* finite *G*(*X*) is (topologically) a subspace of a fi-

nite dimensional vector space, it is homeomorphic to Δ*n ∩* R*n* while the topology

of *M* (*X*) corresponds to that of R*n ∩* R*n*. If we note the *n*-element set, we in

*≥*0

particular have the trivial identities *M* ( )= R*n*

*≥*0

and *M* ( ) *× M* ( )= *M* ( + ).

* 1. *Projective limits of topological spaces*

Many of our theorems will deal with spaces obtained as *projective limits* (also known as *inverse limits* or *coﬁltered limits*) of topological spaces. These topological pro- jective limits are defined as adequate topologisations of projective limits in **Set**, the usual category of sets and functions.

Let (*I, ≤*) be a directed partially ordered set seen as a category and let *D* : *Iop →* **Set** be a cofiltered **Set** diagram. The projective limit of *D* is a terminal cone (lim *D, πi*) over *D* where lim *D* is the set

lim *D* , *{x | D*(*i ≤ j*)(*πj*(*x*)) = *πi*(*x*)*}⊆ D*(*i*)

*i*

and the *πi* : *j D*(*j*) *→ D*(*i*) are the canonical projections. Notice that *D* is con- travariant from *I* to **Set**. As emphasised in the definition, lim *D* is the subset of the cartesian product *i D*(*i*) containing all sequences of elements that respect the constraints imposed by the diagram *D*. The elements of lim *D* are called *threads* and the maps *D*(*i ≤ j*) : *D*(*j*) *→ D*(*i*) are the *bonding maps*. Of course, lim *D* can be empty (see [[34](#_bookmark70)] for a short example). A sufficient condition to ensure non-emptiness of the limit is to consider functors *D* where *I* is countable and the bonding maps are surjective [[6](#_bookmark45)]. As a convenience, we will note those bonding maps as *πij* , *D*(*i ≤ j*), and we write countable cofiltered surjective diagrams *ccd* for short.

Writing *U* : **Top** *→* **Set** for the underlying set functor, cofiltered limits in **Top** for diagrams *D* : *Iop →* **Top** are obtained by endowing the **Set** limit of *U ◦ D* with the initial topology for the canonical projections *{πi}i∈I* . The following useful additional fact follows by considering lim *D* as the intersection of the (closed) subsets of *i D*(*i*) satisfying *D*(*i ≤ j*)(*πj*(*x*)) = *πi*(*x*) for all pairs (*i, j*) s.t. *i ≤ j*.

**Lemma 2.4** *([*[*11*](#_bookmark42)*], Ch. 1, §8.2, Corollaire 2)* lim *D is a closed subset of* *i D*(*i*)*.*

* 1. *The Bochner extension theorem*

The construction of a stochastic process given a system of consistent finite- dimensional marginals is an important tool in probability theory, a classical example being the construction of the Brownian motion using the *Kolmogorov extension the- orem* [[25](#_bookmark61)]. Besides Kolmogorov’s there are many other variants, collectively called *Bochner extension theorem* [[24](#_bookmark60)]. They differ in the amount of structure of the space over which probabilities are considered (measurable, topological or vector spaces) – and we will make crucial use of the Bochner extension theorem for Polish spaces, which admits a particularly elegant presentation.

**Theorem 2.5** *For all D a* ccd *in* **Pol***, G*(lim *D*) *∼*= lim *G ◦ D. We denote by* bcn : lim *G ◦ D → G*(lim *D*) *this homeomorphism.*

In words, the Bochner extension theorem states that any projective family of probabilities that satisfy the diagram constraints (elements of lim *G ◦ D*) can be uniquely lifted to a probability over the limit space (elements of *G*(lim *D*)) – and

what’s more, this extension is a homeomorphism! This presentation of the Bochner extension seems not to be well-known: a similar statement is given in Metivier ([[24](#_bookmark60)], Theorem 5.5) in the case of locally compact spaces, which intersects but does not include Polish spaces; Fedorchuk proves the continuity of *G* on the class of compact Hausdorff spaces in [[15](#_bookmark51)] while more recently Banakh [[4](#_bookmark40)] provides an

extension theorem in the more general setting of Tychonoff spaces, using properties of the Stone-Cˇech compactification.

# Zero-dimensional Polish spaces and their properties

It is natural in applications to consider finitary approximations of stochastic pro- cesses. Accordingly, the correctness of such approximations should correspond to some kind of limiting argument, stating that increasingly finer approximations yield in some suitable sense the original object. In view of the Bochner extension theo- rem, it suffices to consider as input a projective family of probabilities supported by the finitary approximants of the underlying space. However, the very same theorem tells us that we can only obtain by this means probabilities on a projective limit of finite spaces (also called *proﬁnite* spaces), a rather restrictive class:

**Proposition 3.1** *A space is a countable projective limit of ﬁnite discrete spaces if and only if it is a compact, zero-dimensional Polish space.*

The proof can be found under a slightly different terminology in Borceux & Janelidze [[10](#_bookmark46)], where it is shown that these spaces correspond to *Stone spaces* – indeed, profinite spaces are exactly the spaces homeomorphic to the Stone dual of their Boolean algebra of clopen sets! As the proof of this proposition is quite enlightening for the developments to come, we provide it here.

**Proof.** Let *D* : *Iop →* **Pol***fin* be a *ccd* of finite spaces. Polishness of lim *D* comes from the closure of **Pol** under countable limits. Finite spaces are compact and by Tychonoff’s theorem so is *i D*(*i*). Lemma [2.4](#_bookmark8) asserts that lim *D* is closed in this compact product, hence lim *D* is itself compact. Recall that lim *D* has the initial topology for the canonical projections maps *πi* : lim *D → D*(*i*), therefore a base of lim *D* is constituted of finite intersections of prebase opens *π—*1(*Xi*), for *Xi ⊆ D*(*i*). Since the *D*(*i*) are discrete, any of their subsets is clopen and so are the prebase opens; we conclude by noticing that a finite intersection of clopen sets is again clopen.

*i*

Conversely, let *Z* be a compact zero-dimensional Polish space. As *Z* is zero- dimensional Polish, its topology is generated by a countable base of clopen sets. Since *Z* is compact, each clopen can be written as a finite union of base clopens. Therefore its Boolean algebra of clopens *Clo*(*Z*) is also generated by the same countable base, and is itself countable. Note that *Clo*(*Z*) does not depend on the choice of the base! Let us consider the set *I*(*Z*) of all finite clopen partitions of

*Z*. For any *i ∈ I*(*Z*), there exist by assumption a continuous surjective quotient map *fi* : *Z → i*. Since *i* is discrete, the fibres of *fi* are clopen. Note that *I*(*Z*) is also countable. *I*(*Z*) is partially ordered by partition refinement: for all *i, j ∈*

*I*(*Z*), we write *i ≤ j* if there exists a surjective “bonding” map *fji* : *j → i* such that *fji ◦ fj* = *fi* (any such map, if it exists, is unique). *I*(*Z*) is also directed by considering pairwise intersections of the cells of any two partitions. The system of finite discrete quotients of *Z* together with the bonding maps *fji* clearly defines a *ccd* that we write *D* : *I*(*Z*)*op →* **Pol***fin*, mapping each element of *I*(*Z*) to itself and the partial order of refinement to the bonding maps. Therefore, there exists a limit cone (lim *D, πi*). By universality of this cone, there exists a unique continuous map *η* : *Z →* lim *D* s.t. *fi* = *πi ◦ η*. Let us show that *η* is an homeomorphism. As lim *D* and *Z* are both compact, it is enough to show that *η* is a bijection. Recall that *Clo*(*Z*) separates points (it contains a base for a Hausdorff topology) therefore for any *x /*= *y ∈ Z* we can exhibit two clopen cells separating them, implying that *η* is injective. Surjectivity of *η* is a consequence of that of the quotient and bonding maps. *2*

We denote by **Pol***cz* the full subcategory of **Pol** where objects are compact and zero-dimensional – by the previous proposition, these spaces are exactly the profinite Polish spaces. From the data of a projective system of finitely supported probabilities, Prop. [3.1](#_bookmark11) together with Bochner’s extension theorem (Thm. [2.5](#_bookmark10)) only allow us to obtain probabilities supported by such profinite spaces. Our extension of Dirichlet as a natural transformation from the setting of finite spaces to that of arbitrary Polish spaces must therefore imperatively bridge the gap from profinite spaces to arbitrary Polish spaces.

The solution we propose is mediated by zero-dimensional Polish spaces in a decisive way. More precisely, our construction can be framed as the iterative reduc- tion of the extension problem to increasingly smaller subcategories of **Pol** (depicted below): the (full) subcategory of zero-dimensional spaces **Pol***z*, that of compact zero-dimensional spaces **Pol***cz* and finally the subcategory **Pol***fin* of finite Polish spaces. The categorical setting is informally sketched in the following figure:

**Pol***fin*

*⊆*

*ω*

,*7*

**P**¸**o***8***l***cz*

*⊆*

*z*

,*7*

**P** **o***,***l***z*

*⊆*

**P** **o***,***l**

The two essential operations, highlighted in the figure above, are:

* the *zero-dimensionalisation Z*, which yields a zero-dimensional refinement of a Polish space for which a countable base of the topology has been chosen, and
* the *zero-dimensional Wallman compactiﬁcation ω*, which yields a compact zero- dimensional Polish space from a zero-dimensional one along, again, a choice of a base of clopens sets.

We attract the attention of the reader on the fact that these operations are a priori *not* functorial. However, as we shall see in the rest of this section, these operations exhibit powerful properties which are sufficient to proceed to the extension.

* 1. *Zero-dimensionalisation*

Zero-dimensionalisation takes as input a Polish space *X* along a choice of some countable base *F* for *X*. It produces a Polish zero-dimensional topology on the same underlying set as *X*, that we denote by *zF* (*X*).

**Proposition 3.2** *Let* (*X, TX* ) *be Polish and let F be a countable base for X. Let Boole*(*F*) *be the Boolean algebra generated by F. Let zF* (*X*) *be the space which admits Boole*(*F*) *as a base of its topology. zF* (*X*) *veriﬁes the following properties:*

* + 1. *zF* (*X*) *is Polish;*
    2. *zF* (*X*) *is zero-dimensional;*
    3. *the Borel sets are preserved: B*(*X*)= *B*(*zF* (*X*))*;*
    4. *the identity function idF* : *zF* (*X*) *→ X is continuous.*

In order to prove Prop. [3.2](#_bookmark12) we need some classical facts from descriptive set theory, taken verbatim from Kechris [[23](#_bookmark59)], Sec. 13:

**Lemma 3.3** *For any Polish space* (*X, TX* ) *and any closed set A, there exists a Polish topology TXA so that TX ⊆ TXA , A is clopen in TXA and B*(*TX* ) = *B*(*TXA* )*. Moreover, TX ∪ {O ∩ A | O ∈ TX} is a base of TXA .*

**Lemma 3.4** *Let* (*X, TX* ) *be Polish and let {TXn }n∈*N *be a family of Polish topologies on X, then the topology TX∞ generated by ∪nTXn is Polish. Moreover if ∀n, TXn ⊆ B*(*TX* )*, then B*(*TX∞* )= *B*(*TX* )*.*

**Proof. (Proposition** [**3.2**](#_bookmark12)**)** For each *On ∈ F*, let us denote *An* = *X \ On*. Consider the family of Polish topologies *TXAn n∈*N, as obtained using Lemma [3.3](#_bookmark13). Lemma [3.4](#_bookmark14) entails that the topology generated by *∪nTXAn* is Polish. Recall that each *TXAn* has base *TX ∪{O ∩ An | O ∈ TX}*. Closing *∪nTXAn* under finite intersections yields that the topology generated by *∪nTXAn* has base *TX ∪ {O ∩ C | O ∈ TX,C ∈ Boole*(*F*)*}*. Since *F* is a base of *TX* and *F ⊆ Boole*(*F*), an equivalent base of the topology generated by *∪nTXAn* is *Boole*(*F*). By definition, we deduce that the topology of *zF* (*X*) is generated by *∪nTXAn* .

}

1. Lemma [3.4](#_bookmark14) entails that the resulting space is indeed Polish. An equivalent base to *TX ∪ TX|Fc* is *F ∪ F|Fc* and the elements of this base are clopen, hence the

*δ δ*

resulting space is also zero-dimensional.

1. Zero-dimensionality is a trivial consequence of taking a Boolean algebra as a base.
2. Preservation of Borel sets is a further consequence of Lemma [3.4](#_bookmark14).
3. Continuity of the identity is a trivial consequence of the fact that *zF* (*X*) is finer than *X*.

*2*

To the best of our knowledge, we can’t do away with the dependency on *F*: one can exhibit a Polish space *X* with two distinct bases *F, G* such that *zF* (*X*) */*= *zG* (*X*).

Despite this apparent lack of canonicity, any Polish topology is entirely determined by its collection of zero-dimensional refinements [6](#_bookmark18) :

**Theorem 3.5** *Any Polish space X has the ﬁnal topology for the family*

*{idF* : *zF* (*X*) *→ X}F of all the (continuous) identity maps from its zero- dimensionalisations, where F ranges over all the countable bases of X.*

The proof of this theorem relies on the following lemma.

**Lemma 3.6** *Let X bea Polish space and* (*xn*)*n∈*N *→ x a convergent sequence in X. Let F be a countable base for X.* (*xn*)*n∈*N *converges to x in zF* (*X*) *if x /∈ ∪O∈F ∂O.*

**Proof.** Recall that a countable base of *zF* (*X*) is *F ∪ F|Fc*. Assume *x* is not in the boundary of any element *O ∈ F*. Let *U* be a basic open neighbourhood of *x* in *zF* (*X*). If *U ∈F* then it is trivial to exhibit the convergence property by referring

*δ*

to the topology of *X* only. If not, we have that *U* = *O ∩ D*, where *D* = *∩n X \ Oi*,

*i*=1

*Oi ∈ F*; in other terms *x ∈* (*X \ ∪n*

*i*=1

*Oi*) *∩ O*. Observe that since the *Oi* are open,

*Oi* = *Oi∪∂Oi* – therefore, using the initial assumption, we have *x ∈* (*X\∪n*

*i*=1

*Oi*)*∩O*,

which is an open set in *X*. The result follows. *2*

**Proof. (Theorem** [**3.5**](#_bookmark15)**)** It suffices to prove that for all topological space *Y* , a function *f* : *X → Y* is continuous if and only if *f ◦ idF* : *zF* (*X*) *→ Y* is continuous for all countable base *F*. The forward implication is trivial. Assume that for all countable base *F*, *f◦idB* : *zF* (*X*) *→ Y* is continuous. Consider a converging sequence (*xn*)*n∈*N *→ x* in *X*. It is sufficient to exhibit *one* space *zF* (*X*) where this sequence also converges. Lemma [3.6](#_bookmark16) gives as a sufficient criterion that *x* does not belong to *∂O* for any *O ∈ F*. Let us build such a base. Consider a dense set *D* of *X*. Let *d* : *X*2 *→* [0*,* 1] be some metric that completely metrises *X*. Without loss of generality, assume *x ∈ D*. Write *rn* , *d*(*x, dn*) for *dn ∈D \ {x}*. For all *n*, take the family of open balls centred on each *dn* with rational radii strictly below *rn*, e.g. *rn/*3. Since *diam*(*B*(*dn, rn/*3)) = *diam*(*B*(*dn, rn/*3)), *x /∈ ∂B*(*dn, r*) for *r < rn/*3. This family still constitutes a neighbourhood base. The countable union of countable sets is countable, therefore it constitutes a countable base of *X*. *2*

Notice that the topologies of *G*(*X*) and *G*(*zF* (*X*)) might be different, and there is in general no continuous map from *G*(*X*) to *G*(*zF* (*X*)). It should also be emphasised that the “zero-dimensionalisation” of a Polish space is not an innocent operation: for instance if *X* is compact and non-zero-dimensional then *zF* (*X*) will *never* be compact! However, we have the following powerful analogue to Thm. [3.5](#_bookmark15):

**Theorem 3.7** *For X Polish, G*(*X*) *has the ﬁnal topology for the family of identity maps {G*(*id**F* ): *G*(*zF* (*X*)) *→ G*(*X*)*}F where F ranges over countable bases of X.*

**Proof.** As before, it is sufficient to prove that a map *f* : *G*(*X*) *→ Y* is continuous if and only if all precompositions *f ◦ G*(*idF* ) are continuous. If *f* is continuous then the composites clearly also are. Let us consider the reverse implication and suppose that all composites are continuous. Let (*Pn*)*n∈*N *-G*(*X*) *P* be a sequence

6 We mention this fact *en passant* but do not use it in the following developments.

of probabilities converging weakly to *P* in *G*(*X*). It is sufficient to exhibit one *F*

s.t. *Pn - P* in *G*(*zF* (*X*)). Let us recall the following theorem ([[7](#_bookmark43)], Theorem 2.2): *For any Y Polish, let U be a subset of B*(*Y* ) *such that (i) U is closed under ﬁnite intersections (ii) each open set in X is a ﬁnite or countable union of elements in*

*U. If Pn*(*A*) *→ P* (*A*) *for all A in U, then Pn -Y P* . Recall that *Boole*(*F*) is a base of *zF* (*X*). This base trivially verifies condition (i) of the previous theorem. It is therefore sufficient to build a base *F* of *X* such that condition (ii) is verified, i.e. *Pn*(*A*) *→ P* (*A*) for all *A ∈ Boole*(*F*). Observe that the *P* -continuity sets in *X* form a Boolean algebra ([[27](#_bookmark63)], Lemma 6.4). It then suffices to form a base of *X* included in the Boolean algebra of continuity sets of *X*, which is always possible: for any point *x ∈ X*, there can at most be countably many radii *ϵ* s.t. the open ball *B*(*x, ϵ*) has a boundary with strictly positive mass. *2*

* 1. *Zero-dimensional Wallman compactiﬁcations*

Compactifications are topological operations embedding topological spaces into compact spaces. Common examples are the Alexandrov one-point compactifica- tion (for locally compact spaces) of the Stone-Cˇech compactification for Tychonoff spaces. In most settings, this embedding is also required to be dense. By choos-

ing the compactification carefuly, one can preserve some relevant properties of the starting space – in our case, Polishness and zero-dimensionality.

A well-behaved class of compactifications (that includes Alexandrov and Stone- Cˇech as special cases) is that of Wallman compactifications. The general method by which one obtains such a compactification from a given topological space *X* can be decomposed in two steps:

* + 1. one first selects a suitable sublattice of the lattice of open sets of *X* (a *Wallman base*);
    2. then, one topologises (in a standard way) the space of maximal ideals of that particular sublattice.

These compactifications are surveyed in Johnstone [[21](#_bookmark57)] and (less abstractly) in Beck- enstein et al. [[5](#_bookmark41)]. Van Mill [[30](#_bookmark66)] provides some facts on Wallman compactifications of separable metric spaces. An extensive topos-theoretic perspective is also given by Caramello [[12](#_bookmark47)]. In the remainder of this section, we present this compactification method and apply it to the case of Polish zero-dimensional space, yielding a zero- dimensional compactification that we denote by *ω*. We then highlight its connection with Prop. [3.1](#_bookmark11) and study those of its properties that are relevant to our goal.

## Spaces of maximal ideals

All the material here is standard from the litterature on Stone duality for dis- tributive lattices. See e.g. Johnstone [[21](#_bookmark57)] for more details.

**Proposition 3.8** *Let X be a set and L be a (distributive) sublattice of the lattice of subsets of X. The space* max(*L*) *has the set of maximal ideals of L as points and admits subsets of the form B*(*O*)= *{I ∈* max(*L*) *| O /∈ I} as a base. Moreover:*

1. max(*L*) *is T* 1 *and compact;*
2. *if L is furthermore* normal *as a lattice, i.e. if for all O*1*, O*2 *∈ L such that*

*O*1 *∪ O*2 = *X, there exists disjoint Oj , Oj*

*such that Oj*

*⊆ O*1*, Oj*

*⊆ O*2 *then*

max(*L*) *is Hausdorff.*

1 2 1 2

**Proof.** It suffices to check that the family *B*(*O*) where *O* ranges in *L* is indeed a base (i.e. closed under finite intersections). Maximal ideals are by definition proper. Since *L* is distributive, maximal ideals on *L* are moreover prime: for all *I ∈* max(*L*), *B, Bj ∈ L*, if *B ∩ Bj ∈ I* then either *B ∈ I* or *Bj ∈ I* ([[21](#_bookmark57)], I 2.4). Let *B, Bj ∈L* be given. We show *B*(*O*) *∩ B*(*Oj*)= *B*(*O ∩ Oj*). Consider *I ∈ B*(*O*) *∩ B*(*Oj*): we have *O /∈ I* and *Oj /∈ I*, therefore (by primality) *O∩Oj /∈ I*, which implies *I ∈ B*(*O∩Oj*). Conversely, if *I ∈ B*(*O ∩ Oj*) then *O ∩ Oj /∈ I*. Since ideals are downward closed, we must have *O /∈ I* and *Oj /∈ I*. For the proof of (*i*) and (*ii*), see [[21](#_bookmark57)], II resp. 3.5 and

3.6. *2*

## Wallman bases and compactifications

Wallman compactifications are defined as spaces of maximal ideals over Wallman bases, which are particular lattices that are also bases in the topological sense. Here, we will follow the definition given in [[21](#_bookmark57)]:

**Definition 3.9 ([**[**21**](#_bookmark57)**], IV 2.4)** Let *X* be a topological space and let *fX* be its lattice of open sets. A *Wallman base* is a sublattice of *fX* that is a base for *X* and which verifies:

For all *U ∈ fX* and *x ∈ U* , there exists a *V ∈ fX* such that *X* = *U ∪ V* and *x /∈ V* .

The following lemma is key in considering a space of maximal ideals over a Wallman base as a compactification (see **([**[**21**](#_bookmark57)**], IV 2.4)** for a proof):

**Lemma 3.10** *Let X be a topological space and let L be a Wallman base for X. ηL*(*x*)= *{O ∈L| x /∈ O} is a maximal ideal of L. Moreover, if X is T* 0 *then ηL is* *an embedding into* max(*L*)*.*

We are now in position to define Wallman compactifications:

**Definition 3.11** Let *X* be a *T* 0 space and *L* a Wallman base. We denote *ωL*(*X*)= max(*L*) the *Wallman compactiﬁcation* of *X* for *L*.

## Zero-dimensional compactifications

We will now show that taking the inverse limit of the finite partitions of a Polish zero-dimensional space (as in the proof of Prop. [3.1](#_bookmark11)) corresponds – when applied to a *non-compact* Polish zero-dimensional space – to a Wallman compactification of that space, which exhibits very good properties.

Consider a zero-dimensional Polish space *Z*. In opposition to the compact case, the Boolean algebra *Clo*(*Z*) of clopens of *Z* is not necessarily countably generated: we therefore consider partitions of *Z* taken in some countable Boolean sub-algebra *C ⊆ Clo*(*X*) such that *C* is a (topological) base for *Z*. Observe that such a base is

always trivially a normal Wallman base. In the following, we call such countable Boolean sub-algebras that generate the topology “Boolean bases”. We define:

C(*X*) , *{C | C* is a countable Boolean base of *X}*

We write *IC* (*Z*) for the directed partial order of clopen partitions of *Z* taken in *C ∈* C(*X*). Since *C* is countable, so is *IC* (*Z*). We recall that the construction of *IC* (*Z*) is described in the proof of Prop. [3.1](#_bookmark11).

**Proposition 3.12** *For C ∈* C(*Z*)*, let DC* : *Iop*(*Z*) *→* **Pol***fin be the diagram of ﬁnite clopen partitions of Z seen as discrete spaces, then* lim *DC is a zero-dimensional compactiﬁcation of Z homeomorphic to ωC* (*Z*)*.*

*C*

**Proof.** Existence and non-emptiness of lim *DC* stems from surjectivity of the bond- ing maps and countability of *C*. Note that lim *DC* is Polish. Zero-dimensionality is an hereditary property, so it only remains to exhibit an homeomorphism with *ωC* (*Z*). First, observe that since *C* is a Boolean algebra, maximal *C*-ideals are in one-to-one correspondence with maximal *C*-ultrafilters via the complement map: elements of lim *DC* correspond to *C*-filters, they are upward closed and codirected by intersec- tion. They are moreover maximal: for any *U ∈* lim *DC* and all *C ∈ C*, either *C ∈ U* or *Cc ∈ U* . A basic clopen in lim *DC* is of the form *π—*1(*C*) where *C ∈ i ∈ IC* , which correspond to the ultrafilter *{U ∈* lim *DC | C ∈ U}*. This in turns, through the nega- tion map, correspond to a basic clopen of *ωC* (*Z*) (see Prop. [3.8](#_bookmark19)). Every basic clopen of *ωC* (*Z*) similarly correspond to a basic clopen in lim *DC* . Therefore, the spaces are homeomorphic, from which we conclude that lim *DC* is a Polish zero-dimensional compactification. *2*

*i*

As *ωC* (*Z*) is always a profinite space, Prop. [3.1](#_bookmark11) ensures there always exists a cofiltered diagram *D* in **Pol***fin* such that lim *D ∼*= *ωC* (*Z*). We will switch from one point of view to the other freely. We should insist on the fact that our compactifica- tion is *not* the Stone-Cˇech compactification, as these are in general not metrisable (except when compactifying an already metrisable compact space, obviously). Take

for instance the discrete (hence zero-dimensional) Polish space N: *β*N has cardinality 22*ℵ*0 ([[33](#_bookmark69)], Theorem 3.2) while Polish spaces have cardinality at most 2*ℵ*0 . We would obtain Stone-Cˇech if we were to take the Wallman compactification over the full lattice of open sets, however. *ωC* (*Z*) enjoys a property reminiscent of Stone-Cˇech:

**Proposition 3.13** *Let Z be a Polish zero-dimensional space. For each continuous map f* : *Z → K to a compact zero-dimensional space K, there exists a Boolean base C and a continuous map ωC* (*f* ): *ωC* (*Z*) *→ K such that ωC* (*f* ) *◦ ηC* = *f, where ηC* : *Z → ωC* (*Z*) *is the embedding of Z into its compactiﬁcation.*

**Proof.** Prop. [3.1](#_bookmark11) entails that there exists a *ccd DK* : *Iop →* **Pol***fin* s.t. *K ∼*= lim *DK*, with limit cone (lim *DK, {πi* : lim *Dk → Dk*(*i*)*}i∈I* ). Note that by continuity of *πi ◦f* , each finite clopen partition of *K* induces a finite clopen partition of *X*. By choosing a Boolean base of clopens *C* of *Z* that contains *f—*1(*Clo*(*K*)), we can exhibit a compactification *ωC* (*Z*) with an associated cone (*ωC* (*Z*)*, {λi* : *ωC* (*Z*) *→ DK*(*i*)*}*) and therefore an unique map *ωC* (*f* ): *ωC* (*Z*) *→ K* such that *ωC* (*f* ) *◦ ηC* = *f* . *2*

**Corollary 3.14** *For any continuous f* : *Z → Zj between zero-dimensional spaces, there exists Boolean bases C, Cj of respectively Z and Zj such that there exists a map ωCC′* (*f* ): *ωC* (*Z*) *→ ωC′* (*Zj*) *verifying ωCC′* (*f* ) *◦ ηC* = *ηC′ ◦ f.*

Zero-dimensional Polish Wallman compactifications were considered in [[2](#_bookmark38)], which however does not state Prop. [3.13](#_bookmark22).

* 1. *Projective limit measures on zero-dimensional compactiﬁcation*

For *Z* Polish zero-dimensional, the developments of Sec. [3.2](#_bookmark21) allow us to map any measure in *G*(*Z*) to *G*(*ωC* (*Z*)) (for any choice of a Boolean base *C*) through *G*(*ηC* ). Crucially, thanks to Lemma [2.1](#_bookmark5) this is a faithful operation.

Therefore any measure on *Z* can be obtained, up to isomorphism, as a pro- jective limit of finitely supported measures. However, as pointed out before, the converse operation is the difficult one. Let *D* be a diagram such that *ωC* (*Z*) *∼*= lim *D* and *{Pi}i ∈* lim *G ◦ D* a projective family of finitely supported probabilities. There is in general no way to assert that the corresponding projective limit probability *P ∈ G*(*ωC* (*Z*)) obtained through the Bochner extension theorem restricts to *G*(*Z*). We delineate the conditions under which a probability can be restricted to a sub- space and propose a simplification of previous arguments (see [[26](#_bookmark62)]), based on the properties of the Giry monad. Note that the results to follow are not specific to zero-dimensional spaces.

Polish subspaces of Polish spaces are always *Gδ* sets (and conversely, see [[23](#_bookmark59)], 3.11), hence Borel sets. This allows for a simple restriction criterion.

**Proposition 3.15** *Let P ∈ M* (*Y* ) *be a ﬁnite measure on a Polish space Y and let X ⊆ Y be a Polish subspace (hence a Gδ in Y ). The restriction of P to X, deﬁned as the set function P|X* , (*B ∈ B*(*Y* ) *∩ X*) *'→ P* (*B*)*, veriﬁes P|X ∈ G*(*X*) *if and only if P* (*X*)= 1*.*

**Proof.** *P|X* is trivially a finite measure on the trace *σ*-algebra. We observe that *B*(*X*) = *B*(*Y* ) *∩ X*: this is a consequence of Theorem 15.1 in [[23](#_bookmark59)] (essentially, this follows from the Borel isomorphism theorem for Polish spaces), therefore *P|X ∈ M* (*X*). Since *P* (*X*) = 1 and *X ∈ B*(*Y* ), *P|X* (*X*) = 1 and *P|X ∈ G*(*X*). The converse is easy. *2*

This criterion lifts to “higher-order” probabilities, that is probabilities over spaces of probabilities, thanks to the multiplication of the Giry monad. The fol- lowing theorem states that such a higher order probability measure restricts to a subspace if and only if it restricts *in the mean*. This is essentially Theorem 1.1 in [[26](#_bookmark62)].

**Theorem 3.16** *For all X ⊆ Y Polish spaces and all P ∈ G*2(*Y* ) *we have P|G*(*X*) *∈*

*G*2(*X*) *if and only if* (*μY* (*P* ))*|X ∈ G*(*X*)*.*

**Proof.** The forward implication is trivial. By Lemma [2.1](#_bookmark5), *G*2(*X*) is a subspace of *G*2(*Y* ). By Prop. [3.15](#_bookmark24), it is sufficient to prove that *P* (*G*(*X*)) = 1. By assumption that *μ*(*P* )*|X ∈ G*(*X*) and Prop. [3.15](#_bookmark24), we have that *μ*(*P* )(*X*) = 1, which unfolds as

*G*(*Y* ) *EVXdP* = 1. So it suffices to prove that *G*(*Y* ) *EVXdP* =1 *⇒ P* (*G*(*X*)) = 1. Assume *P* (*G*(*X*)) *<* 1, then there must exist a Borel set *A ⊆ G*(*Y* ) *\ G*(*X*) with *P* (*A*) *>* 0. Any probability *p ∈ A* assigns positive measure to some Borel set *B ⊆ Y \ X*, therefore *A EVXdP <* 1. *2*

∫

∫ ∫

# The Dirichlet process

The Dirichlet process stands out among other Bayesian methods in that the prior and posterior distributions are second order probabilities, that is elements of *G*2(*X*). Learning becomes an operation of type *X → G*2(*X*) *→ G*2(*X*), mapping some ev- idence in *X* and a prior in *G*2(*X*) to a posterior in *G*2(*X*), and it can be proved that the second-order stochastic process induced by sampling from identically and independently distributed random variable will converge (in Kullback-Leibler diver- gence, hence in the weak topology [[18](#_bookmark54)]) to a singular distribution over the law of the target.

* 1. *The Dirichlet distribution*

For a fixed finite discrete space *X*, Dirichlet is a function *DX* : *M* +(*X*) *→ G*2(*X*), the parameter in *M* +(*X*) representing the initial prior as well as the degree of certainty about this prior (encoded in its total mass). As we highlight below, *DX* is *continuous* and verifies other properties, among which *naturality* and *normalisation*. Some of the material on the finitary Dirichlet distribution contained in this section can be found (presented differently) in e.g. [[17](#_bookmark53)]. In the following we take *X* = *{x*1*,.* *, xn}*

to be a finite discrete space of cardinality *n*.

## Definition of *D*X

For *Q ≡* (*q*1*,..., qn*) *∈ M* +(*X*), *DX* (*Q*) admits a (continuous) density *dX* (*Q*)

w.r.t. the *n −* 1 dimensional Lebesgue measure given by

Γ(Σ

*i qi*)

*qi—*1 ,

Σ ⎞*qn—*1

*dX* (*Q*)(*p*1*,..., pn—*1)

,

*i*

Γ(*qi*)

*pi*

1*≤i<n*

⎝1 *−*

1*≤i<n*

*pi*⎠

(1)

for all (*pi*)*i<n ∈* Δ*n ∩* (0*,* 1)*n—*1, with *dX* (*Q*) = 0 elsewhere. In Equation [1](#_bookmark26), Γ(*q*) is the continuous gamma function, verifying Γ(*n*) = (*n −* 1)! for *n* integer and defined as Γ(*t*)= *∞ xt—*1*e—xdx*. It is clear that for any bounded continuous function

∫

*X*

continuous in the weak topology.

*f ∈ C* (Δ

*b*

*X*

0

), *D*

*n*

*X* (*Q*)(*f* )= Δ*n*

∫ *fd*

*dxn* varies continuously with *Q*, therefore *D* is

## Γ representation

A useful alternative representation of *DX* (*Q*) relies on sampling from Γ distribu- tions. For parameters *q, r ∈* (0*, ∞*), Γ(*q, r*) (not to be mistaken with the Γ function!)

*rq* Γ(*q*)

is an element of *G*(0*, ∞*) with continuous density *f*

*q,r*

(*x*) = *xq−*1*e−x/r* . Using the

identity (0*, ∞*)*n* = *M* +(*X*) allows to write that for any *Q ≡* (*q*1*,..., qn*) *∈ M* +(*X*) and *R ≡* (*r*1*,..., rn*) *∈ M* +(*X*), the product probability *⊗*0*≤i≤n*Γ(*qi, ri*) belongs in *G*(*M* +(*X*)). Recall that *νX* is the normalisation map (Sec. [2.1](#_bookmark6)).

**Proposition 4.1 *(***Γ ***representation)***

*For all* (*qi*)1*≤i≤n ∈ M* +(*X*)*, dX* (*q*1*,..., qn*)= *d* (*G*(*νX* )(*⊗i*Γ(*qi,* 1)))*,*

*dλ*

*where dP*

*dλ*

*denotes the density of P with respect to Lebesgue.*

**Proof.** We write *xn* , 1*− i<n xi*. For any *Q ≡* (*q*1*,..., qn*) *∈ M* +(*X*), the function *νX* maps the half-line *{tQ | t ∈* R*>*0*}* to its unique point of intersection with Δ*n*. This yields a change of variables *T* (*t, q*1*/ |Q| ,..., qn—*1*/ |Q|*)= *Q* with determinant *tn—*1. Expressing the density of *⊗i*Γ(*qi,* 1) in these new coordinates and integrating over *t* to obtain the density at the point of intersection, one obtains:

Σ

*d ∞*

∫

*dλ* (*G*(*νX* )(*⊗i*Γ(*qi,* 1)))(*x*1*,..., xn—*1) = *dt*

(*tx* )*qi—*1*e—txi*

*t*

*i*

Γ(*q* )

*n—*1

0 *i≤n* *i*

⎡ *xqi—*1 ⎤ ∫ *∞* Σ

0

= ⎣

*i i≤n* Γ(*qi*)

⎦

*t*

*i qi—*1*e—tdt*

The integral identity *∞ xb—*1*e—ax* = *a—b*Γ(*b*) (obtained by integration by parts, or see [[20](#_bookmark56)] 3.381.4) yields the desired result. *2*

0

∫

A more detailed proof can be found in [[1](#_bookmark37)], Sec. 2.3 or [[3](#_bookmark39)], Chap. 27. For an extensive survey of other representations of Dirichlet, see [[17](#_bookmark53)].

## Extension of *D* to non-zero measures

Both the definition in Eq. [1](#_bookmark26) and the Γ representation have the drawback of being confined to strictly positive measures *M* +(*X*) (since Γ is not defined at 0). *M* + is not a functor unless one restricts to the category of Polish spaces and maps with dense ranges. This is a serious drawback, as strict positivity of a measure is a topological property: in particular, it is not preserved by zero-dimensionalisation of the underlying space! We therefore continuously extend *DX* to the better-behaved nonzero measures, so that it verifies for *Q ∈ M* +(*X \ {xi}*):

*DX* (*M∗* (*ei*)(*Q*)) , *G*2(*ei*)(*DX\{x }*(*Q*)) (2)

*i*

where *ei* : *X \ {xi} → X* is the inclusion (observe that *M∗* (*ei*)(*Q*)(*xi*) = 0). The following proposition asserts that this extension is well-defined as a continuous map.

**Proposition 4.2** *Let Q ∈ M* +(*X\{xi}*) *be given. Let δx*

*i*

*be the singular probability*

*measure at xi. Then DX* (*M∗* (*ei*)(*Q*)+ *ϵδx* ) *weakly converges to G*2(*ei*)(*DX\{x }*(*Q*))

*i* *i*

*as ϵ →* 0*.*

**Proof.** We consider w.l.o.g. the case *|X|* = 3, *X* = *{x*1*, x*2*, x*3*}*, *xi* = *x*3. Let

*f ∈ Ub*(*G*(*X*)) be given. Still w.l.o.g., take *f * = 1 and non-negative. In the interest

of conciseness, we write Γ*⊗*(*q*1*, q*2*, q*3) , Γ(*q*1*,* 1) *⊗* Γ(*q*2*,* 1) *⊗* Γ(*q*3*,* 1). Using the Γ representation, we have to prove:

*I*(R3*∗ , ϵ*) *≡* ∫ 3*∗* (*f◦νX* )*d*Γ*⊗*(*q*1*, q*2*, ϵ*) *−−→* ∫ 2*∗* (*f◦G*(*e*3)*◦νx*1*,x*2 )*d*Γ*⊗*(*q*1*, q*2) *≡ J*(R2*∗* )

*≥*0

R*≥*0

*є→*0

R*≥*0

*≥*0

Note that (*f ◦ νX* )(*p*1*, p*2*, p*3)= *f* (*p*1*/ i pi, p*2*/ i pi, p*3*/ i pi*) while (*f ◦ G*(*e*3) *◦ νx*1*,x*2 )(*p*1*, p*2)= *f* (*p*1*/*(*p*1+*p*2)*, p*2*/*(*p*1+*p*2)*,* 0). *νX* is uniformly continuous. Uniform continuity of *f ◦ νX* implies that for all *ϵ*1 *>* 0, there exists an *η >* 0 such that for all (*a, b*) *∈* R2*∗* , for all *c ∈* [0*, η*], *|*(*f ◦ νX* )(*a, b, c*) *−* (*f ◦ G*(*e*3) *◦ νx ,x* )(*a, b*)*| < ϵ*1.

Σ Σ Σ

*≥*0 1 2

Therefore,

*I*(R2*∗ ×* [0*, η*]*, ϵ*)

*≥*0

∫ 2*∗*

*≤ ϵ*1 +

R*≥*0*×*[0*,η*]

*f* (*p*1*/* Σ*i pi, p*2*/* Σ*i pi,* 0)*d*[Γ*⊗*(*q*1*, q*2*, ϵ*)](*p*1*, p*2*, p*3)

*≤ ϵ*1 + ∫ 2*∗ f* (*p*1*/* Σ*i pi, p*2*/* Σ*i pi,* 0) ∫[0*,η*] *d*Γ(*ϵ,* 1)(*p*3) *d*[Γ*⊗*(*q*1*, q*2)](*p*1*, p*2)

R*≥*0

*≤ ϵ*1 + *J*(R2*∗* )

*≥*0

One gets the symmetric inequality *J*(R2*∗* ) *− ϵ*1 *≤ I*(R2*∗*

*×* [0*, η*]*, ϵ*) by the exact

*≥*0 *≥*0

same process. As for the other half of the integral,

*I*(R2*∗*

*≥*0

∫

∫

*×* (*η, l*]*, ϵ*)

*≤*

*∗*

R2

∫*≥*0

(*η,l*]

*f* (*p*1*/*

Σ*i pi, p*2*/*

Σ*i pi, p*3*/*

Σ*i pi*)

*p*1*−ce−p*3

3 Γ(*є*) *d*Γ*⊗*(*q*1*, q*2)

3

*≤ ϵ*

2*∗* (*η,l*]

R

*≥*0

*f* (*p*1*/*

Σ*i pi, p*2*/*

Σ*i pi, p*3*/*

Σ*i pi*)

*p*1*−ce−p*3

Γ(1+*є*)

*dp*3 *d*Γ*⊗*(*q*1*, q*2)

*≤ ϵK*(*η, ϵ*)

∫

where we used the identity Γ(1+*ϵ*)= *ϵ*Γ(*ϵ*) on the third line and *K*(*η, ϵ*) is a bounded quantity. Therefore, *I*(R3*∗ , ϵ*) converges to *J*(R2*∗* ) as *ϵ →* 0. This concludes the

*≥*0 *≥*0

proof. *2*

## Naturality of *D*

As we are going to show, *D* is a natural transformation from *M∗* to *G*2. This property, usually called *aggregation* in the literature, is an indirect consequence of the closure under convolution of the Γ distribution, a property that we assume without proof:

Γ(*q*1*, r*) *∗* Γ(*q*2*, r*)= Γ(*q*1 + *q*2*, r*) (3)

A corollary of Eq. [3](#_bookmark29) is that the pushforward of a product of Γ distributions is a product of Γ distributions.

**Corollary 4.3** *Consider f* : *X → Y surjective for X, Y ﬁnite and discrete, then:*

*G*(*M* +(*f* ))(*⊗x∈X* Γ(*qx, r*)) = *⊗y∈Y* Γ(Σ*x∈f−*1(*y*) *qx, r*)

**Proof.** It suffices to consider the case of *|X|* = 2*, |Y |* = 1, in which case *M* +(*f* ) : *M* +(*X*) *→ M* +(*Y* ) is simply the addition and *G*(*M* +(*f* )) is by defini- tion the convolution! Eq. [3](#_bookmark29) then yields the claim. An induction on *|f—*1(*y*)*|* for each *y ∈ f* (*X*) allows to conclude in the general case. *2*

**Proposition 4.4** *D* , *X '→ DX is a natural transformation D* : *M∗ ⇒ G*2 *when*

*M∗ and G*2 *are restricted to* **Pol***fin.*

**Proof.** It suffices to show that for any *f* : *X → Y* a (continuous) function between finite discrete spaces, one has *G*2(*f* ) *◦ DX* = *DY ◦ M∗* (*f* ). First of all, Prop. [4.2](#_bookmark28) allows us to restrict our attention to *f* surjective and *Q* = *xi '→ qi ∈ M* +(*X*) strictly positive. Then:

(*DY ◦ M* +(*f* ))(*Q*) = *G*(*νY* )(*⊗y∈Y* Γ(Σ

*xi∈f*

*−*1(*y*)

*qi,* 1)) (Prop. [4.1](#_bookmark27))

= *G*(*νY* )((*G ◦ M* +)(*f* )(*⊗x ∈X* Γ(*qi,* 1))) (Cor. [4.3](#_bookmark30))

*i*

= *G*(*νY ◦ M* +(*f* ))(*⊗x ∈X* Γ(*qi,* 1))

*i*

= *G*(*G*(*f* ) *◦ νX* )(*⊗xi∈X* Γ(*qi,* 1)) (Lemma [2.3](#_bookmark7))

= (*G*2(*f* ) *◦ DX* )(*Q*) (Prop. [4.1](#_bookmark27))

*2*

**Example 4.5** Let us verify naturality on a simple case: take *X* = *{*1*,* 2*,* 3*}*, *Y* = *{*1*,* 2*}*, *f* (1) = 1*,f* (2) = *f* (3) = 2 and *Q* = *i '→ qi*. By definition, *G*(*f* )(*p*1*, p*2*, p*3)= (*p*1*, p*2+*p*3). *G*(*f* )*—*1(*p*1*, p*2) corresponds to the closed line segment

*{*(*p*1*, α, p*2 *− α*) *| α ∈* [0*, p*2]*}*, from which we deduce that

*G*2(*DX* (*Q*))(*B*) = ∫

*G*(*f* )*−*1(*B*)

∫

*dX* (*Q*)(*p*1*, p*2)*dp*1*dp*2

= ∫ *dp*1

∫*B*

1*—p*1

0

Γ(*q*1 + *q*2 + *q*3) *pq*1*—*1

*dX* (*Q*)(*p*1*, p*2)*dp*2

∫ 1*—p*1

*pq*2*—*1(1 *− p*

*— p* )*q*3*—*1*dp*

1 Γ(*q*1)Γ(*q*2)Γ(*q*3) 1 2

=

*dp*

*B*

0

1 2 2

The last integral can be carried out using the generalised binomial theorem on the rightmost term: the integral identity *w xa*(*w − x*)*bdx* = *wa*+*b*+1 Γ(*a*+1)Γ(*b*+1)

∫

0

holds ([[20](#_bookmark56)], 3.191.1). On the other hand, we have:

1

1

Γ(*a*+*b*+2)

(*D ◦ M∗* (*f* ))(*Q*)(*B*)= ∫

Γ(*q*1 + *q*2 + *q*3) *pq*1*—*1(1 *− p* )*q*2+*q*3*—*1*dp*

*Y*

So naturality holds.

*B* Γ(*q*1)Γ(*q*2 + *q*3) 1

## Normalisation

The Dirichlet distribution *DX* : *M* +(*X*) *→ G*2(*X*) obeys a consistency relation- ship called *normalisation*:

*μX ◦ DX* = *νX* (4)

*∗ DD*(*j*) 2*¸* ¸*,* 2

*M∗* (*πi* )

*M∗* (*πij* )

z

*M∗* (*πj* )

*M* ( *D*,(*j,*))

*G* (*D*,(*j,*)) *¸*

*ρj*

*G* (*πj* )

*M∗*(*ωC* (*Z*)) *∼*= *M∗*(lim *D*)

*u*  lim*¸*(*G*2 *◦ D*) *G*(bcn)*◦*bcn *G*2*¸*(lim *D*) *∼*= *G*2(*ω* (*Z*))

*G*2 (*πij* )

*C*

*ρi*

*˛ DD*(*i*) ,*7*

*M∗*(*D*(*i*)) *G*2*¸*(*D*(*i*)) ¸*,*

*G*2 (*πi* )

Fig. 1. Construction of *D*ˆ on *ωC*(*Z*)

Let us verify this identity for a parameter *Q* = (*q*1*,..., qn*) *∈* R*n*

*>*0

*⊆ M* +(*X*).

Observe that Eq. [4](#_bookmark31) holds when *|X|* = 2, as *DX* (*q*1*, q*2) degenerates to a *BetaX* (*q*1*, q*2)

distribution which is known to have mean ( *q*1 *,*  *q*2 ) (see eg [[3](#_bookmark39)], Sec. 16.5 for a

*q*1+*q*2 *q*1+*q*2

definition of the *Beta* distribution and the proof of this property).

In the case of *X* an arbitrary finite discrete space, let *fi* : *X → {xi, •}* be the lumping function verifying *fi*(*xi*)= *xi*, *fi*(*xj◦*=*i*)= *•*. By naturality of *μ* and *D* :

(*μX ◦ DX* )(*Q*)(*xi*) = (*μ{x ,•} ◦ G*2(*fi*) *◦ DX* )(*Q*)(*xi*)

*i*

= (*μ{x ,•} ◦ D{x ,•} ◦ M* +(*fi*))(*Q*)(*xi*)

*i* *i*

Σ *qi*

= (*μ{xi,•} ◦ Beta{xi,•}*)(*qi,*

*j◦*=*i qj*)(*xi*) = Σ

*j*

*q*

*j*

* 1. *Extension to zero-dimensional Polish spaces*

The finite support case is instructive but lacks generality. We proceed to the ex- tension of finitely supported Dirichlet distributions to Dirichlet processes supported by arbitrary zero-dimensional Polish spaces. Our construction preserves both nat- urality and continuity – in fact, it can be framed as the extension of the natural transformation *D* from **Pol***fin* to **Pol***z*, the full subcategory of zero-dimensional Polish spaces and continuous maps. In what follows, we denote by *F|*C : C *→* **Pol** the restriction of the domain of some endofunctor *F* : **Pol** *→* **Pol** to a subcategory C of **Pol**. When unambiguous, we drop this notation.

**Theorem 4.6** *There exists a unique (up to isomorphism) natural transformation*

*D*ˆ : *M|***Pol** *⇒ G*2*|***Pol** *such that D*ˆ *coincides with D on* **Pol***fin.*

*z z*

**Proof.** We prove existence, naturality and uniqueness.

**Existence.** For any given choice of a Boolean base *C*, let *ηC* : *Z → ωC* (*Z*) be the embedding of a Polish zero-dimensional space *Z* into its compactifica- tion *ωC* (*Z*) (Lemma [3.10](#_bookmark20)). *ωC* (*Z*) is compact zero-dimensional so by Prop. [3.1](#_bookmark11) there exists a *ccd* of finite spaces *D* such that *ωC* (*Z*) *∼*= lim *D*. Let us con-

struct *D*ˆ , the extension of Dirichlet to lim *D* (see Fig. [1](#_bookmark32)). Applying the func-

lim *D*

tor *M∗* yields a cone *C* = (*M∗* (lim *D*)*, {M∗* (*πi*): *M∗* (lim *D*) *→* (*M∗ ◦ D*)(*i*)*}i*). Applying the finitary Dirichlet *D* on the base of this cone yields a *ccd* in

*ν*lim *D*

*νD*(*j*)

*G*(*D*¸,(,*j,*)*,*)¸*¸*

*M∗* (*πij* )

*M∗* (*πi* )

z

*G*2 (*πij* )

*.G*2 (*πi* )

*G*(*D*¸(,*i,*))*¸*

z

*G*(lim,*D*)*¸*

*M∗*(*D*(*j*))

*μD*(*i*)

*DD*(*j*) *G*2*¸*(*D*(*j*))

*μ*lim *D*

*M∗* (*π* ) ,*,*

*j*

*νD*(*i*)

,*,*¸*¸G*2 (*πj* )

*M∗*(*Z*)

*M¸∗* (lim *D*)

*M* (*ηC* )

*∗*

*D*ˆlim *D*  *G*2*¸*(lim *D*)

*μD*(*j*)

*˛*

*M∗*(*D*(*i*))

*DD*(*i*)

*G*2*¸*(*D*(*i*))

Fig. 2. Commutation of normalisation and Dirichlet averaging

*G*2 *◦ D*, of which we take the limit, obtaining a terminal cone *T* = (lim *G*2 *◦*

*D,* *ρi* : lim *G*2 *◦ D → G*2(*D*(*i*))} ). By naturality of *D* , the cone *C* extends to a

cone *Cj* = (*M∗* (lim *D*)*,* *D*

*D*(*i*)

*i*

* *M∗*

(*πi*): *M*

*∗* (lim *D*)

*→* (*G*2 *◦ D*)(*i*)} ). By univer-

sality of *T* , there exists a unique morphism *u* : *M∗* (lim *D*) *→ G*2(lim *D*) map- ping *Cj* to *T* . The Bochner extension theorem (Thm [2.5](#_bookmark10)) yields an isomorphism *G*(bcn) *◦* bcn : lim *G*2 *◦ D → G*2(lim *D*) (the fact that *G*(bcn) is an isomorphism is a consequence of Lemma. [2.1](#_bookmark5)). This yields a morphism

*i*

*D*ˆlim *D* : *M∗* (lim *D*) *→ G*2(lim *D*)

ˆ

*D*

lim *D*

= *u ◦ G*(bcn) *◦* bcn

that trivially coincides with *D* when lim *D* happens to be finite. In order to conclude

the existence part of the extension, we need to show that *D*ˆ *◦M∗* (*ηC* ): *M∗* (*Z*) *→*

lim *D*

*G*2(lim *D*) actually ranges in *G*2(*ηC* (*Z*)) *⊆ G*2(lim *D*), after which we can set *D*ˆ*Z* ,

*D*ˆlim *D ◦ M∗* (*ηC* ). By Theorem [3.16](#_bookmark25), it suffices to check that for any *Q ∈ M∗* (*Z*),

(*μ ◦ D*ˆ *◦ M∗* (*η* ))(*Q*)

lim *D* lim *D C*

*ηC*(*Z*)

*∈ G*(*ηC* (*Z*))

which by Prop. [3.15](#_bookmark24) amounts to checking that this measure attributes full measure to *ηC* (*Z*. We take advantage of the *normalisation property* (Eq. [4](#_bookmark31)) of *D* . Thanks to this property and to the naturality of *μ*, the diagram in Fig. [2](#_bookmark34) commutes. The Bochner extension theorem entails the universality of the cone (*G*(lim *D*)*, {G*(*πi*)*}i*) at the top of the diagram, therefore commutation of the diagram in Fig. [2](#_bookmark34) entails the existence of a unique morphism from the cone (*M∗* (lim *D*)*, νD*(*i*)*○M∗* (*π* ) ) to

}

*i*

*i*

(*G*(lim *D*)*, {G*(*πi*)*}i*) (morphism represented as a dashed line in Fig. [2](#_bookmark34)). This mor-

phism is no other than the normalisation *ν*lim *D* : *M∗* (lim *D*) *→ G*(lim *D*). Therefore,

(*μ*lim *D ◦ D*ˆ

lim *D*

* + *M∗* (*ηC* ))(*Q*) = (*ν*

lim *D*

* *M∗* (*ηC* ))(*Q*)

*M∗* (*Z*)

*M∗*(*ηC*(*Z*))

*∗*

*M∗*(*f* ) *¸∗*

*M∗*(*ωCC′* (*f* ) ) *¸∗*

*M*

(*Zj*)

*M∗*(*ηC′* (*Z′*))

*j*

*M* (*ωC* (*Z*))

*M* (*ωC′* (*Z* ))

*∗ M∗*(*πi○ωCC′* (*f* ) ) *¸∗*

*D*ˆ *D*ˆ *′*

*M* (*ωC* (*Z*))

*M* (*DZ′* (*i*))

*ωC* (*Z*)

*G*2(*ωC* (*Z*))

*ωC′* (*Z* )

*C′* (*Zj*))

*G¸*2(*ω*

*G*2(*ωCC′* (*f* ))

ˆ

*ω* (*Z*)

*D*

*C*

*G*2(*ωC* (*Z*))

*D Z′* (*i*)

*D*

*DZ′* (*i*))

*G¸*2(

*G*2(*πi○ωCC′* (*f* ))

1. Reducing naturality to the case of com- pact zero-dim. Polish spaces.
2. Finitary case.

Trivially, *M∗* (*ηC* )(*Q*)(*Z \ ηC* (*Z*)) = 0, therefore (*ν*lim *D ◦ M∗* (*ηC* ))(*Q*) is concentrated on *ηC* (*Z*). Hence, up to isomorphism, *D*ˆ*Z* restricts to a morphism *D*ˆ*Z* : *M∗* (*Z*) *→ G*2(*Z*). This concludes the proof of existence.

**Naturality.** For any map *f* : *Z → Zj* between zero-dimensional Polish spaces,

we must prove *D*ˆ *′ ◦ M∗* (*f* ) = *G*2(*f* ) *◦ D*ˆ*Z* . By Corollary [3.14](#_bookmark23), we can reduce the

*Z*

task to the case of a morphism *ωCC′* (*f* ) : *ωC* (*Z*) *→ ωC′* (*Zj*) between compact zero-

dimensional spaces (see Fig. [3a](#_bookmark35)). It remains to prove *D*ˆ *′ ◦ M∗* (*ωCC′* (*f* )) =

*G*2(*ωCC′* (*f* )) *◦ D*ˆ

*ωC′* (*Z* )

. By Prop. [3.1](#_bookmark11), *ωC*(*Z*) *∼*= lim *DZ* and *ωC′* (*Zj*) *∼*= lim *DZ′*

*ωC*(*Z*)

where *DZ* and *DZ′* are their respective finite discrete quotient *ccd* s. Let us write

(*ωC′* (*Zj*)*, {πi* : *ωC′* (*Zj*) *→ DZ′* (*i*)*}i*) the terminal cone corresponding to *DZ′* . The universal property of this limit cone allows to reduce the problem to the commuta- tion of the diagram in Fig. [3b](#_bookmark35):

ˆ *′ ◦ M∗* (*f* )= *G*2(*f* ) *◦ D*ˆ*Z ⇔ ∀i, G*2(*πi*) *◦ D*ˆ *′ ◦ M∗* (*f* )= *G*2(*πi*) *◦ G*2(*f* ) *◦ D*ˆ*Z*

*D*

*Z*

*Z*

*⇔ ∀i, DD* (*i*) *◦ M∗* (*πi*) *◦ M∗* (*f* )= *G*2(*πi*) *◦ G*2(*f* ) *◦ D*ˆ*Z*

*Z′*

As already argued in the proof of Prop. [3.13](#_bookmark22), any finite discrete clopen partition of *ωC′* (*Zj*) induces a finite discrete clopen partition of *ωC*(*Z*) since the two spaces are related by the continuous function *ωCC′* (), therefore the diagram in Fig. [3b](#_bookmark35) commutes.

**Uniqueness.** Assume there exists two distinct natural transformations

*D*ˆ*, D*ˆ*j* :

*M∗* (*Z*) *→ G*2(*Z*) that coincide with *D* on finite spaces. It is clear that it is enough to exhibit a contradiction in the case of *Z* compact zero-dimensional Polish. We refer to Fig. [1](#_bookmark32) for the notations. Let *D* be a *ccd* of finite spaces such that *Z ∼*= lim *D*, with canonical projections *πi* : lim *D → D*(*i*). By assumption, there must exist a measure *Q ∈ M∗* (*Z*) such that *D*ˆ(*Q*) */*= *D*ˆ*j*(*Q*). But both *D*ˆ and *D*ˆ*j* verify

(by assumption of naturality and consistency with the finitary case) the equalities

*G*2(*πi*) *◦ D*ˆ*j* = *DD*(*i*) *◦ M∗* (*πi*)= *G*2(*πi*) *◦ D*ˆ*Z* for all *i*. Therefore, *Q* induces through

*Z*

*D*ˆ and *D*ˆ*j* the same projective family of finite-dimensional Dirichlet distributions

}

*DD*(*i*) *◦ M∗* (*πi*)(*Q*) , which yields (by unicity of extensions, see Theorem [2.5](#_bookmark10)) a contradiction. *2*

*i*

* 1. *Extension to arbitrary Polish spaces*

Let *X* be an arbitrary Polish space. As shown in Theorem [3.7](#_bookmark17), *G*(*X*) has the final topology for the family of identity maps *{G*(*idF* ): *G*(*zF* (*X*)) *→ G*(*X*)*}F* where *F* ranges over countable bases of *X*. In order to harness this theorem, we need the following fact:

**Lemma 4.7** *Let X be Polish and zF* (*X*)*, zG* (*X*) *be two zero-dimensional reﬁne- ments as constructed in Prop.* [*3.2*](#_bookmark12)*. Then DzF* (*X*) *and DzG* (*X*) *are equal in* **Set***.*

**Proof.** The set of countable bases of *X* is directed by union. Let us write *H≡ F∪G*. The (continuous) identity functions *idFH* : *zH*(*X*) *→ zF* (*X*) and *idGH* : *zH*(*X*) *→ zG* (*X*) lift to identity functions *G*2(*idFH*)*, G*2(*idGH*), and similarly for the functor *M∗* . Therefore, the commutation relation *G*2(*idFH*) *◦ Dz* (*X*) = *Dz* (*X*) *◦ M∗* (*idFH*) boils down in **Set** to the equality of *DzF* (*X*) and *DzG* (*X*) (and similarly for *G*). *2*

*H F*

Finally, we have:

**Theorem 4.8** *There exists a unique (up to isomorphism) natural transformation*

*D*ˆ : *M∗ ⇒ G*2 *such that D*ˆ *coincides with D on* **Pol***.*

**Proof.** Let *X* be a Polish space. Consider the family *{zF* (*X*)*}F* of its zero- dimensional refinements, as constructed in Prop. [3.2](#_bookmark12). For each *zF* (*X*), Theorem

[4.6](#_bookmark33) asserts the existence of a continuous Dirichlet map *D*ˆ

*zF* (*X*)

: *M∗* (*zF* (*X*)) *→*

*G*2(*zF* (*X*)), which extends by continuity of the identity and functoriality to a con- tinuous map

*G*2(*idF* ) *◦ D*ˆ : *M∗* (*zF* (*X*)) *→ G*2(*X*) (5)

*zF* (*X*)

By Lemma [4.7](#_bookmark36), all these maps coincide in **Set**. Theorem [3.7](#_bookmark17) allows to conclude. *2*

# 5 Conclusion

Our construction of the Dirichlet process in categorical style subsumes existing ones [[17](#_bookmark53),[26](#_bookmark62)] while establishing continuity and naturality. However, further work, which we intend to pursue right away, is required to consolidate our understanding of the finitary approximation framework we have built for higher-order probabil- ities. The Giry monad can be generalised from **Pol** to the category of Tychonoff spaces, however our construction relies heavily on the properties of Polish spaces: for instance we use the fact that zero-dimensional Polish spaces are Borel sets of their compactifications (Prop. [3.15](#_bookmark24)); the measurable selection theorem used in Lemma [2.1](#_bookmark5) also requires the spaces considered to be Polish. The process by which we rebuild Dirichlet relies on some simple properties of Γ distributions. Naturality is a conse- quence of closure of Γ under convolution (a particular case of *inﬁnite divisibility* also exhibited by e.g. normal distributions), and the fact that Dirichlet restricts to *X*, which is only a subset of its compactification *wX*, follows from the *normalisation* property (see *§*4.1). By axiomatising these properties, we can generalise our main result. However, it remains to be seen whether other interesting distributions on R*>*0 fit the conditions and generate Dirichlet-like processes.

Beyond the immediate questions above, we can return to the less immediate goals expounded on in the introduction, namely higher-order learning using uncertain chains of Dirichlet type. Any uncertain Markov chain *τ* , meaning a morphism *X → G*2(*X*) in **Pol**, can be post-composed with the multiplication of *G* to obtain the “mean” Markov chain of type *X → G*(*X*). We will investigate the case where *τ* takes values in Dirichlet processes -focusing on the tractable “uncertain chains of Dirichlet type”. Such chains can be decomposed as *α* : *X → G*(*X*) followed by the Dirichlet natural transformation *DX* : *M∗* (*X*) *→ G*2(*X*). The first component *α* is *τ* ’s parameterising chain. As *μ ◦ DX ◦ α* is the normalised version of *α*, *α* is again up to normalisation the mean chain of the uncertain *τ* . Our construction ensures that *τ* is continuous by construction. At this stage, it is already possible given *τ* : *X → G*2(*X*) to quantify the uncertainty at each point by considering moments of the “Kantorovich” random variable *Kx* , (*P ∈ G*(*X*)) *'→ dK*((*μ ◦ τ* )(*x*)*,P* ), where *KX* is defined over the probability triple (*G*(*X*)*,τ* (*x*)) and *dK* is metrises

*G*(*X*). The next step is to adapt the Bayesian learning scheme which in the discrete

case maps the *prior DX* (*Q*) to the *posterior DX* (*Q* + *s*), for *Q* in *M∗* (*X*) the current parameter, given *s* a multiset of observed values in *X* (seen as a counting measure). Via the projective limit construction, learning can be led at the level of behavioural approximants [[13](#_bookmark49)] and a subsidiary goal is to understand how the two levels relate. The second goal consists of in extending the probabilistic Kantorovich metric to uncertain chains (of this specific type) and understand its evolution under learning. Until now we assumed that the state of the system is fully observable, but the above questions should be developed as well in a broader context where the state is only partially and noisily so. In this setting, naturality of *D* might allow to compare uncertain chains defined on distinct state spaces by embedding them in some universal Polish space – giving a quantitative account of both the differences in their respective state spaces *and* their dynamics.

# Acknowledgement

We warmly thank Laurent Dufloux and Florence Clerc for their valuable insights, and the anonymous reviewers for their thorough and helpful work.

# References

1. Bela A Frigyik AK and MR Gupta. Introduction to the Dirichlet distribution and related processes. Technical report, Technical Report UWEETR-2010-0006, Department of Electrical Engineering, University of Washington., 2010.
2. F.G. Arenas and M.A. Snchez-Granero. Wallman compactification and zero-dimensionality.

*Divulgaciones Matemticas*, 7(2):151–155, 1999.

1. N. Balakrishnan and V.B. Nevzorov. *A Primer on Statistical Distributions*. Wiley, 2004.
2. Taras Banakh. The topology of spaces of probability measures, i: Functors *pτ* and *p*ˆ. *Matematychni Studii*, 1995.
3. Edward Beckenstein, Lawrence Narici, and Charles Suffel. *Topological algebras*. Mathematic studies. North-Holland, 1977.
4. George M. Bergman. Some empty inverse limits. [https://math.berkeley.edu/~gbergman/papers/](https://math.berkeley.edu/~gbergman/papers/unpub/) [unpub/](https://math.berkeley.edu/~gbergman/papers/unpub/).
5. Patrick Billingsley. *Convergence of Probability Measures*. Wiley, 1968.
6. Vladimir I Bogachev. *Measure Theory I*. Springer, 2006.
7. Vladimir I Bogachev. *Measure Theory II*. Springer, 2006.
8. Francis Borceux and George Janelidze. *Galois theories*. Number 72 in Cambridge studies in advanced mathematics. Cambridge University Press, 2001.
9. N. Bourbaki. *Elements de math´ematique. Topologie G´en´erale.* Springer, 1971.
10. O. Caramello. Gelfand spectra and Wallman compactifications. *ArXiv e-prints*, April 2012.
11. Philippe Chaput, Vincent Danos, Prakash Panangaden, and Gordon Plotkin. Approximating Markov Processes by averaging. *Journal of the ACM*, 61(1), January 2014. 45 pages.
12. Alexandre Donz´e and Oded Maler. Robust satisfaction of temporal logic over real-valued signals. In

*Formal modeling and analysis of timed systems*, pages 92–106. Springer Berlin Heidelberg, 2010.

1. V.V. Fedorchuk. Functors of probability measures in topological categories. *Journal of Mathematical Sciences*, 91(4):3157–3204, 1998.
2. Thomas S Ferguson. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, pages 209–230, 1973.
3. Subhashis Ghosal. The Dirichlet process, related priors and posterior asymptotics. In N. L. Hjort et al, editor, *Bayesian Nonparametrics*, pages 36–83. Cambridge University Press, 2010.
4. Alison L. Gibbs, Francis, and Edward Su. On choosing and bounding probability metrics. *Internat.* *Statist. Rev.*, pages 419–435, 2002.
5. M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects* *of Topology and Analysis*, number 915 in Lecture Notes In Mathematics, pages 68–85. Springer-Verlag, 1981.
6. I.S. Gradshteyn and D. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier Science, 2000.
7. P.T. Johnstone. *Stone Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1986.
8. Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.
9. Alexander S Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Text in Mathematics*. Springer, 1995.
10. Michel M´etivier. Limites projectives de mesures. martingales. applications. *Annali di Matematica*, 1963.
11. B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2003.
12. Peter Orbanz et al. Projective limit random probabilities on polish spaces. *Electronic Journal of Statistics*, 5:1354–1373, 2011.
13. K.R. Parthasarathy. *Probability Measures on Metric Spaces*. AMS Chelsea Publishing Series. Academic Press, 1972.
14. Evgeny B Stukalin, Hubert Phillips III, and Anatoly B Kolomeisky. Coupling of two motor proteins: a new motor can move faster. *Physical review letters*, 94(23):238101, 2005.
15. Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems.

*Theoretical Computer Science*, 331(1):115 – 142, 2005. Automata, Languages and Programming.

1. Jan Van Mill. *The infinite-dimensional topology of function spaces*, volume 64. Elsevier, 2002.
2. Cedric Villani. *Optimal transport, Old and New*. Grundlehren der mathematischen Wissenschaften. Springer, 2006.
3. DanielH. Wagner. Survey of measurable selection theorems: An update. In Dietrich Klzow, editor, *Measure Theory Oberwolfach 1979*, volume 794 of *Lecture Notes in Mathematics*, pages 176–219. Springer Berlin Heidelberg, 1980.
4. Russell C. Walker. *The Stone-Cˆech compactification*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, 1974.
5. William Waterhouse. An empty inverse limit. *Proceedings of the American Mathematical Society*, 36(2), 1972.
6. David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

# Topological and measurable spaces

We recall some basic facts about topological and measurable spaces.

* 1. *Topological spaces*

## Basic definitions

A *topological space* (*X, fX* ) is given by a set *X* and a set *fX* of *open subsets of X* such that *X ∈ fX* , *∅ ∈ fX* and *fX* is closed under arbitrary unions and finite intersections. A set is *closed* if its complement is open. It is *clopen* if it is both closed and open. A set of subsets *S ⊆ fX* is a *base* of *fX* whenever it is closed under finite intersections and its closure under arbitrary unions yields *fX* . A set of subsets *S ⊆ fX* is a *pre-base* of *fX* when its closure under finite intersections is a base of *fX* . The *closure* of a subset *A ⊆ X* is noted *A*, it is the smallest closed set containing *A*. Conversely, the *interior* of *A ⊆ X* is noted int (*A*), it is the largest open set contained in *A*. The *boundary* of *A* is noted *∂A* and is defined as *∂A* , *A \* int (*A*). A space (*X, fX* ) is *separable* if there exists a countable subset *D ⊆ X* that is *dense* in *X*, i.e. *D* = *X*. Except where it might lead to ambiguities, we will omit *fX* and write the space simply *X*. A map *f* : *X → Y* between two topological spaces is *continuous* if and only if for all *OY ∈ fY* , we have *f—*1(*OY* ) *∈ fX* . An *homeomorphism* is a bicontinuous bijection. *Y ⊆ X* is a *subspace* of *X* if its opens are of the form *O ∩ Y* , for *O ∈ fX* . Topological spaces with continuous maps form a category, noted **Top**.

## Initial and final topologies

Let *I* = *{fi* : *X →* (*Xi, fXi* )*}i* be a family of functions *fi* from a set *X* into topo- logical spaces (*Xi, fXi* ). The initial topology induced by *I* is the coarsest topology

on *X* making the *fi* continuous. If is defined as the topology *fI* generated by the sets

*f—*1(*O*) *| O ∈ fX* }. The final topology is defined dually, as the finest topology

*i*

*i*

*i*

on *X* making a family of functions *F* = *{fi* : (*Xi, fXi* ) *→ X}i* continuous. A subset

*O ⊆ X* is open if and only if *f—*1(*O*) *∈ fX* for all *i*. It is straightforward to check

*i*

*i*

that this defines a topology.

Limits and colimits in **Top** are defined by endowing them with resp. the initial and final topologies on the **Set** limits and colimits. In particular, the topological product is defined as the initial topology on the **Set** product w.r.t. the canonical projections.

## Metric and metrisable spaces

A *distance* function on a set *X* is a function *d* : *X × X →* [0*, ∞*] that obeys the following axioms, *∀x, y, z ∈ X*: (i) symmetry: *d*(*x, y*)= *d*(*y, x*), (ii) *d*(*x, y*) *≥* 0, *d*(*x, y*)=0 iff *x* = *y*, (iii) *d*(*x, y*) *≤ d*(*x, z*)+*d*(*z, y*). Any distance *d* on *X* induces the topology of a *metric space*, with base the open balls *B*(*x, ϵ*)= *{y | d*(*x, y*) *< ϵ}*, for all *x ∈ X*, *ϵ >* 0. A metric space is noted (*X, d*). A sequence of points (*xn ∈ X*)*n∈*N *converges* to a point *x ∈ X* if for all *ϵ >* 0, there exists a *N ∈* N such that for all *n ≥ N* , *d*(*xn, x*) *< ϵ*. A sequence is *Cauchy* if for all *ϵ >* 0, there exists an *N ∈ N* such that for all *m, n ≥ N* , *d*(*xm, xn*) *< ϵ*. A metric space is *complete* if

all Cauchy sequences converge. A space is *metrisable* if its topology is generated by some distance. A space is *completely metrisable* if its topology is generated by some distance that makes it *complete*. A *Polish space* is a separable, completely metrisable space. Polish spaces form a full subcategory of **Top**, noted **Pol**. **Pol** has all countable limits and all countable disjoint unions [[11](#_bookmark42)]. It includes the category of finite, discrete topological spaces **Pol***fin* as a full subcategory (a space *X* is discrete if *fX* = *℘*(*X*)).

## Separation conditions

A topological space *X* is *Hausdorff* if for any two points *x, y ∈ X* there exists disjoint open sets *Ox, Oy* such that *x ∈ Ox*, *y ∈ Oy*. In a Hausdorff space, all singletons are closed. *X* is *completely regular* if for any closed set *F ⊆ X* and any point *x ∈ X \ F* , there exists a continuous function *f* : *X →* R such that *f* (*x*)=0 and *f* (*y*) = 1 for all *y ∈ F* . A space is *Tychonoff* if it is completely regular and Hausdorff. All metrisable spaces are Tychonoff.

## Compactness

An *open cover* of a space *X* is a family *{Oi ∈ fX}i* of open subsets such that

*∪iOi* = *X*.A topological space is *compact* if any open cover of the space has a finite sub-cover. A subset of *X* is compact if it verifies this property. All finite subsets are compact. The continuous image of a compact set is compact. All spaces we consider will be Hausdorff, accordingly all compact spaces will be implicitly Hausdorff. All compact subspaces of Hausdorff spaces are closed. *Tychonoff* ’s theorem asserts that an arbitrary product of compact spaces is compact. A continuous bijection between compact (Hausdorff!) spaces is always a homeomorphism.

## Zero-dimensional spaces

A topological space is *zero-dimensional* if it has a base of clopen sets. The set of clopen sets of a space *X* is noted *C*(*X*). It is a Boolean algebra. One easily deduces that zero-dimensional Polish spaces have a countable base of clopen sets. Zero-dimensionality is a hereditary property and is preserved by subspaces.

## Compactifications

A *compactiﬁcation* of a (Tychonoff) topological space *X* is a compact space *Y* into which *X* embeds homeomorphically and such that the closure of *X* in *Y* is *Y* itself (a non-Tychonoff spaces need not embed in its compactification).

* 1. *Measurable spaces*

A *measurable space* (*X,* Σ*X* ) is a set *X* along a *σ*-algebra Σ*X* , that is a set of subsets of *X* closed under complements and countable unions that contain *X*. If

*X* is a topological space we note *B*(*X*) the *Borel σ*-algebra generated from its topology. A map *f* : (*X,* Σ*X* ) *→* (*Y,* Σ*Y* ) between measurable spaces is *measurable* if *f—*1(*A*) *∈* Σ*X* for all *A ∈* Σ*Y* . If *f* : *X → Y* is a continuous map, *f* is also

measurable between the corresponding Borel measure spaces. Borel measure spaces arising from Polish spaces verify the “Isomorphism theorem” [[23](#_bookmark59)]:

**Theorem A.1** *For all X and Y Polish spaces, B*(*X*) *∼*= *B*(*Y* ) *if and only if X and*

*Y have the same cardinality.*

# Proof of Theorem [2.5](#_bookmark10)

This proof is adapted from Metivier [[24](#_bookmark60)].

Let *D* : *Iop →* **Pol** be our *ccd*, with canonical projections *πi* : lim *D → D*(*i*). Let

*{Pi}i ∈* lim *G ◦ D* be given. We proceed to continuously extend this family to an element *P ∈ G*(lim *D*). Consider *A* = *∪i∈I π—*1(*B*) *| B ∈ B*(*D*(*i*)) . By directed- ness, *A* is an algebra of lim *D*-Borel sets. We define the set function *P*0 : *A→* [0*,* 1] by *P*0(*π—*1(*B*)) = *Pi*(*B*). Codirectedness of the family *{Pi}i* ensures that (i) *P*0 is consistent as a function and that (ii) *P*0 is finitely additive, therefore *P*0 is a *charge*. As *P*0 is finite, hence *σ*-finite, it is sufficient to exhibit that *P*0 is *σ*-additive on *A* and the Carath´eodory extension theorem ([[35](#_bookmark71)], Theorem 1.7) will yield the sought unique projective limit Borel measure. *σ*-additivity is equivalent to the implication

*i*

}

*i*

*∀n, P*0(*An*) *≥ δ ⇒ ∩nAn /*= *∅* for all *δ >* 0 and all decreasing sequence of Borel sets (*An*)*n∈*N ([[8](#_bookmark48)], Prop. 1.3.3).

Let (*An*)*n* be such a sequence. Each *An* is by construction of the form *An* =

*π—*1(*B∗*

) for some *i ∈ I*, where *B∗*

*∈ B*(*D*(*i*)). We map this sequence (*An*)

*i i,n i,n n*

to a family *{Bcn ∈ B*(*D*(*cn*))*}cn* of Borel sets indexed by an increasing sequence

*cn*

*n*

(*cn*)

*⊆ π*

*n∈*N

, cofinal in *I*, such that for all *n*, *An* = *π—*1(*Bc*

) and *Bc*

*n*+1

*—*1

*cncn*+1

(*Bcn* ).

The cofinal increasing sequence (*cn*)*n∈*N is constructed by induction on any fixed

enumeration of *I*. By construction, there is some *in ∈ I* for which *An* = *π—*1(*Bi* ).

*in*

*n*

By cofinality, there exists *cn ≥ in* and by measurability, *Bcn*

*—*1

*incn*

, *π*

(*Bin*

) is mea-

surable. By directedness, *An* = *π—*1(*Bc*

*cn*

*n*

). Now consider *m ≤ n* with *An ⊆ Am*. We

have *An* = *An ∩ Am* = *π—*1(*Bc*

) *∩ π—*1(*Bc*

). By directedness, *π—*1 = *π—*1

* *π—*1

*cn n*

*∩ π*

*cm m*

*cm cmcn cn*

therefore *An* = *π—*1(*Bc*

*cn*

*n*

*—*1

*cmcn*

(*Bcm*

)). For *n* fixed, this generalises to *An* =

*π—*1(*Bc*

) = *π—*1(*∩m≤nπ—*1

(*Bc*

)). Therefore, *Bc*

= *∩m≤n*+1*π—*1

(*Bc* ) =

*cn n cn*

*cmcn m*

*n*+1

*cmcn*+1 *m*

*∩m≤n*+1(*π—*1

*cncn*+1

*—*1

*cmcn*

* *π*

)(*Bcm*

*—*1

*cncn*+1

(*∩m≤nπ—*1

(*Bcm*

*—*1

*cncn*+1

(*Bcn* ).

We construct a nonempty (compact!) set *K* s.t. *K ⊆ An* for all *n*. By cofi-

) *⊆ π*

*cmcn*

)) = *π*

nality of (*cn*)*n∈*N, it is sufficient to construct of a family of non-empty compact sets *{Kcn ⊆ Bcn }n∈*N that is projective, i.e. verifying *πcncn*+1 (*Kcn*+1 ) = *Kcn* for all

*n*. Such a projective family of compact sets can in turn be obtained from a se-

quence of non-empty compact sets *Kj*

*cn*

*cn*+1

(*K*

*cn*

*n∈*N

(*K*

*cn*

*m*

verifying *Kj*

*—*1

*cncn*+1

*⊆ π*

*j* ). In-

deed, setting, for all *m*, *Kcm*

= *∩n≥mπc*

*mcn*

*j* ), we trivially have that *Kc* is

compact. As an intersection of a decreasing sequence of non-empty compact sets

(in a metrisable space), *Kcm* is also non-empty (this is Cantor’s intersection the-

(*K*

*cn*

*cn*

orem). Moreover, *Kcm*

= *∩n≥mπc*

*mcn*

*j* ) *⊇ ∩n≥m*+1(*πc*

*mcm*+1

* *πc*

*m*+1*cn*

)(*Kj* ) *⊇*

*πcmcm*+1

(*∩n≥m*+1*πc*

*m*+1*cn*

*j* )) = *πc*

*mcm*+1

(*Kc*

*m*+1

). To prove the reverse inclusion,

it suffices to show that for all *x ∈ Kcm*

(*K*

, *π*

*cn*

*—*1

*cmcm*+1

(*K*

*cmcm*+1

*cn*

(*x*) *∩ Kc*

*m*+1

*/*= *∅*. We have

*—*1

*π*

*cmcm*+1

(*x*) *∩Kc*

*m*+1

*—*1

*cmcm*+1

= *π*

(*x*) *∩* (*∩n≥m*+1*πc*

*m*+1*cn*

*j* )) = *∩n≥m*+1(*π—*1

(*x*) *∩*

*πcm*+1*cn*

(*K*

*j* )). Notice that *π—*1

(*x*) *∩ πc*

*m*+1*cn*

(*K*

*j* ) is compact for all *n*. Since

by definition *π—*1

(*K*

*cmcm*+1

(*x*) *⊆ Kc*

*m*+1

*⊆ πc*

*m*+1*cn*

*j* ) for all *n*, this intersection is

non-empty.

*cn*

*cn*

*cn*

We have reduced the goal to providing a sequence of non-empty compact sets

*K*

*cn*+1

(*K*

*cn*

*j ⊆ Bc*

*cmcm*+1

*cn*

*n*

*n∈*N

verifying *Kj*

*—*1

*cncn*+1

*⊆ π*

*j* ). Recall that *P* (*An*) *≥ δ >* 0 for

all *n*, which implies *Pcn* (*Bcn* ) *≥ δ >* 0 for all *n*. Finite Borel measures on Polish

spaces are *Radon*: for all *P ∈ G*(*X*) with *X* Polish, for all *B ∈ B*(*X*), *P* (*B*) = sup *{P* (*K*) *| K ⊆ B, K* compact*}* ([[7](#_bookmark43)], Theorem 1.4); therefore each *Pcn ∈ G*(*D*(*cn*))

is Radon. We build by induction a sequence *Kj*

*cn*

*\ K*

) *<*

*j*

*n∈*N

such that *Pcn*

(*Bcn cn*

Σ*n*  *є* and *Kj*

*k*=0 2*k*+1

*cn*+1

*cncn*+1

*cn*

*Kj*

*⊆ π—*1

(*Kj*

). For *n* = 0, we obtain *Kj*

verifying *Pc* (*Bc \*

*c*0

*c*0

0

0

) *< ϵ/*2 by application of the Radon property. Our inductive hypothesis consists

*j*

in the existence of a sequence

*Kck*

0*≤k≤n* having the aforementioned properties.

By assumption, *Bc*

*⊆ π—*1

(*Bc* ). We have *π—*1

(*Kj*

) *∩ Bc*

= *Bc ∩*

*n*+1

*cncn*+1 *k*

*cncn*+1 *cn*

*n*+1

*n*+1

*—*1

[*π*

(*K*

*cn*

*cncn*+1

(*Bcn*

*—*1

*cncn*+1

*j* )], therefore *Pc*

*n*+1

(*Bc*

*n*+1

*—*1

*cncn*+1

*\ π*

*j* )) *<* Σ*n*

*ϵ/*2*k*+1.

To conclude it suffices to pick, using the Radon property, *Kj*

) *\ π*

(*K*

*k*=0

*cn*+1

*cn*

s.t. *Pc*

*n*+1

((*Bc*

*n*+1 *∩*

*—*1 *j*

*π*

(*K*

)) *\ K*

*cncn*+1 *cn*

*j*

*cn*+1

) *< ϵ/*2*n*+2. One then has a sequence verifying all the required

properties – in particular, its elements are non-empty (since they have positive

*cn*

measure) and they verify *Kj*

*cn*+1

*—*1

*cncn*+1

*⊆ π*

(*K*

*j* ). This concludes the existence and

unicity of the measure associated to *{Pi}i ∈* lim *G ◦ D*.

We now prove that this extension is a homeomorphism. Observe that the maps *G*(*πi*): *G*(lim *D*) *→ G*(*D*(*i*)) define a cone over *G◦ D*, therefore there exists an uni- versal (continuous!) mediating map bcn*—*1 : *G*(lim *D*) *→* lim *G◦D*, which associates to *P ∈ G*(lim *D*) a projective system *{G*(*πi*)(*P* )*}i∈I* of probabilities. As Borel mea- sures are entirely specified by their values on open sets (Lemma 7.1.2, [[9](#_bookmark44)]), bcn*—*1 is injective. The uniqueness of the procedure described above ensures that *P* is pre- cisely the extension corresponding to *{G*(*πi*)(*P* )*}i∈I* , therefore bcn*—*1 is surjective. Let us prove continuity of bcn. Consider (*tn ∈* lim *G ◦ D*)*n∈*N a sequence converging to *t ∈* lim *G ◦ D*, i.e. for all *i ∈ I* that *tn*(*i*) weakly converges (in the sense of Sec. [2.1](#_bookmark4)) to *t*(*i*), which is equivalent by the “Portmanteau” theorem to strong conver- gence on *t*(*i*)-continuity sets for all *i*. Let us write (*Pn* = bcn(*tn*))*n ,P* = bcn(*t*) the projective limit measures of resp. (*tn*)*n , t*. We must prove *Pn - P* . It can be eas- ily verified, by commutation of the interior and closure operations with topological products, that for all *i*, if *B ∈ B*(*G*(*D*(*i*))) is a *t*(*i*)-continuity set then *π—*1(*B*) is a *P* -continuity set. Let *di* be a distance compatible with *D*(*i*), consider for *x ∈ D*(*i*) the neighbourhood *Ni,x*(*ϵ*)= *{y ∈ D*(*i*) *| di*(*x, y*)*}*. For distinct *ϵk*, *∂Ni,x*(*ϵk*) are dis- joint. Therefore there cannot be more than a countable family of *{ϵk >* 0*}k* such that *t*(*i*)(*∂Ni,x*(*ϵk*)) *>* 0. We deduce that each prebase open *π—*1(*Ni,x*(*ϵ*)) contains a continuity set. Since continuity sets form an algebra, Corollary 1 to Theorem 2.2 of [[7](#_bookmark43)] applies and we conclude that *Pn - P* . Therefore, bcn is a homeomorphism.

*i*

*i*