Available online at www.sciencedirect.com

ScienceDirect

Information and Computation 204 (2006) 1756–1781

Information and Computation

www.elsevier.com/locate/ic

Eilenberg–Moore algebras for stochastic relations ☆

Ernst-Erich Doberkat *

Chair for Software Technology, University of Dortmund, Germany

Received 20 December 2004; revised 6 October 2005
Available online 11 October 2006

Abstract

We investigate the category of Eilenberg–Moore algebras for the Giry monad associated with stochastic relations over Polish spaces with continuous maps as morphisms. The algebras are identified as the positive convex structures on the base space. The forgetful functor assigning a positive convex structure the underlying Polish space has the stochastic powerdomain as its left adjoint.

MSC: 18C20; 68Q65; 68N30

Keywords: Stochastic relations; Stochastic powerdomain; Giry monad; Eilenberg–Moore algebras; Computation through monads; Convexity; Positive convex structures

1. Introduction

Modelling a computation through a monad (as suggested e.g., by Moggi [1]), one represents state transitions or the transformation from inputs a to outputs b through a morphism $a \rightarrow \mathbf{T}b$ with $\mathbf{T}$ as the functor underlying the monad. Working in a probabilistic setting, a state from a base space

☆ Research funded in part by Deutsche Forschungsgemeinschaft, Grant DO 263/8-1, Algebraische Eigenschaften stochastischer Relationen.
* Corresponding author.
E-mail address: doberkat@acm.org

0890-5401/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ic.2006.09.001

X is in this way associated with a subprobability distribution $K(x)$ on X. Here $K : X \to \mathbf{S}(X)$ is a morphism for the probability monad, in which the functor assigns a space its probabilities. But we now have only a distribution of the outputs, not the outputs proper. What is needed for this is a map $h : \mathbf{S}(X) \to X$ that would transform a distribution into a state proper.

Such a pair $\langle X, h\rangle$ is called an Eilenberg–Moore algebra (or simply an algebra) for this monad, whenever it is compatible with the monad. Structurally, these algebras help to construct an adjunction for which the monad is just the given one [2, Theorem VI.2.1]. In fact, this adjunction and the one constructed through the Kleisli category form in some sense the extreme points in a category of all adjunctions from which the given monad can be recovered [2, Theorem VI.5.3]. Thus it is of algebraic interest to identify these algebras in general, and in particular to the probability functor. It has as a Kleisli construction stochastic relations and is in this sense quite similar to the powerset functor; this is why we call this construction the stochastic powerdomain. For the powerset functor the algebras are completely characterized, but the stochastic side of the analogy is not explored fully yet. This paper proposes characterizations for these algebras under the assumption of continuity. We work in the category of Polish spaces (these spaces are explained in Section 2) with continuous maps as morphisms. In this category the algebras for the Giry monad are identified, and the category of all algebras is investigated. The natural approach is to think of these algebras in terms of an equivalence relation which may be thought to identify probability distributions, and to investigate either these relations or the partitions associated with them. These characterizations lead to the identification of the algebras as the positive convex structures on their base space. A similar result has been known for probability measures on compact Hausdorff spaces [3–5], but a full characterization of the subprobabilistic case on Polish spaces is new. With these results we are able to characterize the left adjoint to the forgetful functor that assigns each positive convex structure its underlying Polish space: the left adjoint is just the stochastic powerdomain, assigning a Polish space the space of all its subprobability measures together with a natural algebra, that comes with the Giry monad.

The paper contributes to the theory of stochastic relations by providing a characterization of the category of Eilenberg–Moore algebras for the subprobability functor in the category of Polish spaces with continuous mappings as morphisms, and by characterizing the left adjoint to the forgetful functor from above.

1.1. Organization

We define the objects we are dealing with in Section 2, in particular, the space of all subprobability measures on a Polish space is introduced together with the weak topology that renders it a Polish space. The Giry monad is also introduced. Section 4 is devoted to the characterization of the algebras for this monad through partitions, smooth equivalence relations, and positive convex structures (which are given a defining glance in Section 3). The category of algebras is shown to be isomorphic to these categories. Some examples are given in Section 5, indicating among others that the search for algebras in the—usually easily dealt with—finite case is somewhat hopeless. Section 6 identifies the subprobability functor together with the monad's multiplication as the left adjoint to the forgetful functor on the category of algebras. The case of probabilities is dealt with in a completely analogous manner, only that *positive convex* has to be replaced by *convex*. Section 7 has a brief look at related work, and finally Section 8 proposes further investigations along the lines developed here.

2. The Giry Monad

In this section, the constructions underlying the Giry monad are collected. We remind the reader of Polish spaces, of the topology of weak convergence on the space of all subprobabilities on a Polish space, and finally of the monad investigated by Giry.

Let X be a Polish space, i.e., a separable metric space for which a complete metric exists, and denote by $\mathbf{S}\,(X)$ the set of all subprobability measures on the Borel sets $\mathcal{B}(X)$ of X. The *weak topology* on $\mathbf{S}\,(X)$ is the smallest topology which makes $\tau \mapsto \int_X f\, \mathrm{d}\tau$ continuous, whenever $f \in \mathcal{C}(X) := \{g : X \to \mathbb{R} \mid g \text{ is bounded and continuous}\}$. It is well known that the discrete measures are dense, and that $\mathbf{S}\,(X)$ is a Polish space with this topology [6, Section II.6]. Let d be the metric on X, and put $d(x, A) := \inf\{d(x, y) \mid y \in A\}$ as the distance of $x \in X$ to the subset $A \subseteq X$, then the Prohorov metric $\mathbf{d}_P$ on $\mathbf{S}\,(X)$ is defined through

$$\mathbf{d}_P(\tau_1, \tau_2) := \inf\{\varepsilon > 0 \mid \forall A \in \mathcal{B}(X) : \tau_1(A) \leq \tau_2(A^\varepsilon) + \varepsilon \wedge \tau_2(A) \leq \tau_1(A^\varepsilon) + \varepsilon\}$$

with $A^\varepsilon := \{y \in X \mid d(y, A) < \varepsilon\}$ as the set of all elements of X having distance less that ε from A. This metric topologizes the topology of weak convergence, see [7, Theorem 6.8].

More explicitly, a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \in \mathbf{S}\,(X)$ converges to $\tau_0 \in \mathbf{S}\,(X)$ in this topology (indicated by $\tau_n \rightharpoonup_w \tau_0$) iff

$$\forall f \in \mathcal{C}(X) : \int_X f\, \mathrm{d}\tau_n \to \int_X f\, \mathrm{d}\tau_0$$

holds. The famous Portmanteau Theorem [6, II.6.1] states that this is equivalent to the condition

$$\liminf_{n\to\infty} \tau_n(G) \geq \tau_0(G)$$

whenever $G \subseteq X$ is an open set. We will assume throughout that $\mathbf{S}\,(X)$ is endowed with the weak topology.

Denote by $\mathfrak{Pol}$ the category of Polish spaces with continuous maps as morphisms. $\mathbf{S}$ assigns to each Polish space X the space of subprobability measures on X; if $f : X \to Y$ is a morphism in $\mathfrak{Pol}$, its image $\mathbf{S}\,(f) : \mathbf{S}\,(X) \to \mathbf{S}\,(Y)$ is defined through

$$\mathbf{S}\,(f)(\tau)(B) := \tau\left(f^{-1}[B]\right),$$

where $\tau \in \mathbf{S}\,(X)$ and $B \in \mathcal{B}(Y)$ is a Borel set. By virtue of the *Change of Variable Formula*

$$\int_Y g\, d\mathbf{S}\,(f)(\tau) = \int_X g \circ f\, \mathrm{d}\tau$$

it is easy to see that $\mathbf{S}\,(f)$ is continuous. Thus $\mathbf{S} : \mathfrak{Pol} \to \mathfrak{Pol}$ is functor.

Denote by $\mu_X : \mathbf{S}\,(\mathbf{S}\,(X)) \to \mathbf{S}\,(X)$ the map

$$\mu_X(M)(A) := \int_{\mathbf{S}\,(X)} \tau(A)\, M(\mathrm{d}\tau)$$

which assigns to each subprobability measure M on the Borel sets of $\mathbf{S}\ (X)$ a subprobability measure $\mu_X(M)$ on the Borel sets of X. Thus $\mu_X(M)(A)$ averages over the subprobabilities for A using measure M. It is immediate that $\mu_X(\delta_\tau) = \tau$, where $\delta_\tau \in \mathbf{S}\ (\mathbf{S}\ (X))$ is the *Dirac measure* on $\tau \in \mathbf{S}\ (X)$, thus

$$\delta_\tau(A) := \begin{cases} 1, & \text{if } \tau \in A, \\ 0, & \text{if } \tau \notin A. \end{cases}$$

Standard arguments show that

$$\int_X f \, \mathrm{d}\mu_X(M) = \int_{\mathbf{S}\ (X)} \left(\int_X f \, \mathrm{d}\tau \right) M(\mathrm{d}\tau)$$

for each measurable and bounded map $f : X \to \mathbb{R}$.

The map μ_X is a morphism in $\mathfrak{Pol}$, as the following Lemma shows.

Lemma 2.1. $\mu_X : \mathbf{S}\ (\mathbf{S}\ (X)) \to \mathbf{S}\ (X)$ *is continuous.*

Proof . Let $(M_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{S}\ (\mathbf{S}\ (X))$ with $M_n \rightharpoonup_w M_0$, then we get for $f \in \mathcal{C}(X)$ through the Change of Variable Formula, and because

$$\tau \mapsto \int_X f \, \mathrm{d}\tau$$

is a member of $\mathcal{C}(\mathbf{S}\ (X))$, the following chain

$$\begin{aligned} \int_X f \, \mathrm{d}\mu_X(M_n) &= \int_{\mathbf{S}\ (X)} \left(\int_X f \, \mathrm{d}\tau \right) M_n(\mathrm{d}\tau) \\ &\to \int_{\mathbf{S}\ (X)} \left(\int_X f \, \mathrm{d}\tau \right) M_0(\mathrm{d}\tau) \\ &= \int_X f \, \mathrm{d}\mu_X(M_0). \end{aligned}$$

Thus $\mu_X(M_n) \rightharpoonup_w \mu_X(M_0)$ is established, as desired. $\square$

The argumentation in [8] shows that $\mu : \mathbf{S}^2 \stackrel{\bullet}{\to} \mathbf{S}$ is a natural transformation. Together with $\eta_X : X \to \mathbf{S}\ (X)$, which assigns to each $x \in X$ the Dirac measure δ_x on x, and which is a natural transformation $\eta : \mathbb{1} \stackrel{\bullet}{\to} \mathbf{S}$, the triplet $\langle \mathbf{S}, \eta, \mu \rangle$ forms a monad [8]. It was originally proposed and investigated by Giry and will be referred to as the *Giry monad*. This means that these diagrams commute in the category of endofunctors of $\mathfrak{Pol}$ with natural transformations as morphisms:

$$\begin{tikzcd} \mathbf{S}^3 \arrow[r, "\mathbf{S}\mu"] \arrow[d, "\mu\mathbf{S}"'] & \mathbf{S}^2 \arrow[d, "\mu"] \\ \mathbf{S}^2 \arrow[r, "\mu"'] & \mathbf{S} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{S} \arrow[r, "\eta\mathbf{S}"] \arrow[dr, "id"'] & \mathbf{S}^2 \arrow[d, "\mu"] & \mathbf{S} \arrow[l, "\mathbf{S}\eta"'] \arrow[dl, "id"] \\ & \mathbf{S} & \end{tikzcd}$$

Definition 2.1. A stochastic relation $K : X \rightsquigarrow Y$ between the Polish spaces X and Y is a Kleisli morphism for the Giry monad.

Equivalently, a stochastic relation $K : X \rightsquigarrow Y$ may be represented as a map $K : X \to \mathbf{S}\,(Y)$ with the following properties:

(1) $x \mapsto K(x)$ is (weakly) continuous, so that $x \mapsto \int_Y f \; \mathrm{d}K(x)$ defines a continuous map on X, whenever $f : Y \to \mathbb{R}$ is continuous and bounded,
(2) $B \mapsto K(x)(B)$ constitutes a subprobability measure on the Borel sets $\mathcal{B}(Y)$ of Y for each $x \in X$.

The composition $L \circ K$ of stochastic relations $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ is the Kleisli product associated with the monad, thus

$$(L \circ K)(x)(C) = \int_Y L(y)(C)\; K(x)(\mathrm{d}y)$$

holds for $x \in X$ and $C \in \mathcal{B}(Z)$ (we use the composition symbol $\circ$ both for the product in the base category and for the Kleisli product, since no confusion can arise).

If the probability measures $\mathbf{P}\,(X)$ on X are considered, then we obtain the monad $\langle \mathbf{P}, \eta, \mu \rangle$ which is actually the monad that was investigated by Giry. We will concentrate on the subprobability functor $\mathbf{S}$ with occasional sidelong glances to the probability functor. The former one is a bit more convenient to work with because $\mathbf{S}\,(X)$ is positive convex.

3. Positive convex structures

Suppose the Polish space X is embedded into a vector space V over the reals as a positive convex structure. This means that, if $x_1, \ldots, x_k \in X$, $\langle \alpha_1, \ldots, \alpha_k \rangle \in \Omega$, then $\sum_{i=1}^{k} \alpha_i \cdot x_i \in X$. Here, we have put

$$\Omega := \left\{ \langle \alpha_1, \ldots, \alpha_k \rangle \mid k \in \mathbb{N}, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i \leq 1 \right\}$$

for the rest of the paper, the elements of Ω being called *positive convex tuples* or simply *positive convex*. In addition, forming positive convex combinations should be compatible with the topological structure on X, so it should be continuous. This means of course that $x_{i,n} \to x_{i,0}$ and $\alpha_n \to \alpha_0$ with $\alpha_0, \alpha_n \in \Omega$ together imply $\sum_{i=1}^{k} \alpha_{i,n} \cdot x_{i,n} \to \sum_{i=1}^{k} \alpha_{i,0} \cdot x_{i,0}$. These requirements are quite like those for a topological vector space, postulating continuity of addition and scalar multiplication. This meets the intuition about positive convexity, but it has the drawback that we have to look for the vector space V into which X to embed. It has the additional shortcoming that once we did identify V, the positive convex structure on X is fixed through the vector space, but we will see soon that we need some flexibility. Consequently, we propose an abstract description of positive convexity, much in the spirit of Pumplün's approach [9]. Thus the essential properties (for us, that is) of positive convexity are described intrinsically for X without having to resort to a vector space. This leads to the definition of a positive convex structure.

Definition 3.1. A *positive convex structure* $\mathcal{P}$ on the Polish space X has for each $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$ a continuous map $\alpha_{\mathcal{P}} : X^n \to X$ which we write as

$$\alpha_{\mathcal{P}}(x_1, \ldots, x_n) = \sum_{1 \le i \le n}^{\mathcal{P}} \alpha_i \cdot x_i,$$

such that

(1) $\sum_{1 \le i \le n}^{\mathcal{P}} \delta_{i,k} \cdot x_i = x_k$, where $\delta_{i,j}$ is Kronecker's δ (thus $\delta_{i,j} = 1$ if $i = j$, and $\delta_{i,j} = 0$, otherwise),
(2) the identity

$$\sum_{1 \le i \le n}^{\mathcal{P}} \alpha_i \cdot \left(\sum_{1 \le k \le m}^{\mathcal{P}} \beta_{i,k} \cdot x_k \right) = \sum_{1 \le k \le m}^{\mathcal{P}} \left(\sum_{1 \le i \le n}^{\mathcal{P}} \alpha_i \beta_{i,k} \right) \cdot x_k$$

holds whenever $\langle \alpha_1, \ldots, \alpha_n \rangle, \langle \beta_{i,k}, \ldots, \beta_{i,k} \rangle \in \Omega, 1 \le i \le n$.

Thus we will use freely the notation from vector spaces, omitting in particular the explicit reference to the structure whenever possible. Hence simple addition $\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2$ will be written rather than $\sum_{1 \le i \le 2}^{\mathcal{P}} \alpha_i \cdot x_i$, with the understanding that it refers to a fixed positive convex structure $\mathcal{P}$ on X.

It can be shown [9] that for a positive convex structure the usual rules for manipulating sums in vector spaces apply, e.g., the law of associativity,

$$(\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2) + \alpha_3 \cdot x_3 = \alpha_1 \cdot x_1 + (\alpha_2 \cdot x_2 + \alpha_3 \cdot x_3)$$

or

$$\sum_{i=1}^{n} \alpha_i \cdot x_i = \sum_{i=1, \alpha_i \neq 0}^{n} \alpha_i \cdot x_i.$$

Nevertheless, care should be observed, for of course not all rules apply: we cannot in general conclude $x = x'$ from $\alpha \cdot x = \alpha \cdot x'$, even if $\alpha \neq 0$.

A morphism $\theta : \langle X_1, \mathcal{P}_1 \rangle \to \langle X_2, \mathcal{P}_2 \rangle$ between continuous positive convex structures is a continuous map $\theta : X_1 \to X_2$ such that

$$\theta \left(\sum_{1 \le i \le n}^{\mathcal{P}_1} \alpha_i \cdot x_i \right) = \sum_{1 \le i \le n}^{\mathcal{P}_2} \alpha_i \cdot \theta(x_i)$$

holds for $x_1, \ldots, x_n \in X_1$ and $\langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$. In analogy to linear algebra, θ will be called an *affine* map. Positive convex structures with their morphisms form a category $\mathfrak{StrConv}$.

4. Characterizing the Algebras

The Eilenberg–Moore algebras are represented through partitions and through smooth equivalence relations, both on the respective space of subprobability measures. We first deal with partitions and investigate the partition induced by an algebra. This leads to a necessary and sufficient condition for a partition to be generated from an algebra which in turn can be used for characterizing the category of these algebras by introducing a suitable notion of morphisms for partitions. The second representation capitalizes on the fact that equivalence relations induced by continuous maps (as special cases of Borel measurable maps) have some rather convenient properties in terms of measurability. This is used for an alternative description of the category of all algebras.

4.1. Algebras

An *Eilenberg–Moore algebra* $\langle X, h\rangle$ for the Giry monad is an object X in $\mathfrak{Pol}$ together with a morphism $h : \mathbf{S}\,(X) \to X$ such that the following diagrams commute

$$\begin{tikzcd} \mathbf{S}\,(\mathbf{S}\,(X)) \arrow[r, "\mathbf{S}(h)"] \arrow[d, "\mu_X"'] & \mathbf{S}\,(X) \arrow[d, "h"] \\ \mathbf{S}\,(X) \arrow[r, "h"'] & X \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "\eta_X"] \arrow[dr, "id_X"'] & \mathbf{S}\,(X) \arrow[d, "h"] \\ & X \end{tikzcd}$$

When talking about algebras, we refer always to Eilenberg–Moore algebras for the Giry monad, unless otherwise indicated. An *algebra morphism* $f : \langle X, h\rangle \to \langle X', h'\rangle$ between the algebras $\langle X, h\rangle$ and $\langle X', h'\rangle$ is a continuous map $f : X \to X'$ which makes the diagram

$$\begin{tikzcd} \mathbf{S}\,(X) \arrow[r, "h"] \arrow[d, "\mathbf{S}(f)"'] & X \arrow[d, "f"] \\ \mathbf{S}\,(X') \arrow[r, "h'"'] & X' \end{tikzcd}$$

commute. Algebras together with their morphisms form a category $\mathfrak{Alg}$. This construction is discussed for monads in general in [2, Chapter IV.2].

Remark 4.1. Looking aside, we mention briefly a well-known monad in the category $\mathfrak{Set}$ of sets with maps as morphisms. The functor $\mathcal{P}$ assigns each set A its power set $\mathcal{P}\,(A)$, and if $f : A \to B$ is a map, $\mathcal{P}\,(f) : \mathcal{P}\,(A) \to \mathcal{P}\,(B)$ assigns each subset $A_0 \subseteq A$ its image $f\,[A_0]$, thus $\mathcal{P}\,(f)\,(A_0) = f\,[A_0]$. Define the natural transformation $\mu : \mathcal{P}^2 \overset{\bullet}{\to} \mathcal{P}$ through

$$\mu_A : \mathcal{P}\,(\mathcal{P}\,(A)) \ni M \mapsto \bigcup M \in \mathcal{P}\,(A),$$

and $\eta : \mathbb{1} \overset{\bullet}{\to} \mathcal{P}$ through $\eta_X : x \mapsto \{x\}$, then the triplet $\langle \mathcal{P}, \eta, \mu\rangle$ forms a monad (the *Manes monad*). It is well known that the algebras for this monad may be identified with the complete sup-semi lattices [2, Exercise VI.2.1].

For the rest of this paper each free occurrence of X refers to a Polish space.

We need some elementary properties for later reference. They are collected in the next Lemma.

Lemma 4.1.

(1) *Let* $f : A \to B$ *be a map between the Polish spaces* A *and* B, *and let* $\tau = \alpha_1 \cdot \tau_1 + \ldots + \alpha_n \cdot \tau_n$ *be a positive convex combination of subprobability measures* $\tau_1, \ldots, \tau_n \in \mathbf{S}(A)$. *Then* $\mathbf{S}(f)(\tau) = \alpha_1 \cdot \mathbf{S}(f)(\tau_1) + \ldots + \alpha_n \cdot \mathbf{S}(f)(\tau_n)$.
(2) *Let* $M = \alpha_1 \cdot M_1 + \ldots + \alpha_n \cdot M_n$ *be the positive convex combination of* $M_1, \ldots, M_n \in \mathbf{S}(\mathbf{S}(X))$. *Then* $\mu_X(M) = \alpha_1 \cdot \mu_X(M_1) + \ldots + \alpha_n \cdot \mu_X(M_n)$.

Proof . $\mathbf{S}(f) : \mathbf{S}(A) \to \mathbf{S}(B)$ and $\mu_X : \mathbf{S}(\mathbf{S}(X)) \to \mathbf{S}(X)$ both are affine maps. This follows immediately from the respective definitions. □

4.2. Positive Convex Partitions

We will show in this section that an algebra may be characterized in the way its fibres, i.e., the inverse images of points, partition the domain $\mathbf{S}(X)$. The point of interest here is that those partitions are positive convex and take closed values, they have an additional property due to continuity. This yields necessary and sufficient conditions for the characterization of partitions spawned by these algebras. A characterization of the morphisms in the category of all algebras is also derived.

Assume that the pair $\langle X, h\rangle$ is an algebra, and define for each $x \in X$

$$G_h(x) := \{\tau \in \mathbf{S}(X) \mid h(\tau) = x\} \left(= h^{-1}[\{x\}]\right).$$

Then $G_h(x) \neq \emptyset$ for all $x \in X$: because $h(\delta_x) = x, h$ is onto. The algebra h will be characterized through properties of the set-valued map G_h. Define the *weak inverse* $\exists R$ for a set-valued map $R : X \to \mathcal{P}(Y) \setminus \{\emptyset\}$ with non-empty images through

$$\exists R(W) := \{x \in X \mid R(x) \cap W \neq \emptyset\}.$$

for $W \subseteq Y$. If Y is a topological space, if R takes closed values, and if $\exists R(W)$ is compact in X whenever $W \subseteq Y$ is compact, then R is called *k-upper-semicontinuous* (abbreviated as k.u.s.c.). If Y is compact, this is the usual notion of upper-semicontinuity (cf. [10, Section 5.1]).

The importance of being k.u.s.c. becomes clear at once from

Lemma 4.2. *Let* $f : A \to B$ *be a surjective map between the Polish spaces* A *and* B, *and put* $G_f(b) := f^{-1}[\{b\}]$ *for* $b \in B$. *Then* f *is continuous iff* G_f *is k.u.s.c.*

Proof . A direct calculation for the weak inverse shows $\exists G_f(A_0) = f[A_0]$ for each subset $A_0 \subseteq A$. The assertion now follows from the well-known fact that a map between metric spaces is continuous iff it maps compact sets to compact sets. □

Applying this observation to the set-valued map G_h, we obtain

Proposition 4.1. *The set-valued map* $x \mapsto G_h(x)$ *has the following properties*:

(1) $\delta_x \in G_h(x)$ *holds for each* $x \in X$.
(2) $\mathcal{G}_h := \{G_h(x) \mid x \in X\}$ *is a partition of* $\mathbf{S}(X)$ *into closed and positive convex sets.*

(3) *$x \mapsto G_h(x)$ is k.u.s.c.*

(4) *Let $\sim_h$ be the equivalence relation on $\mathbf{S}\ (X)$ induced by the partition $\mathcal{G}_h$. If $\tau_i \sim_h \tau_i'$ $(1 \le i \le n)$, then*

$$(\alpha_1 \cdot \tau_1 + \ldots + \alpha_n \cdot \tau_n) \sim_h \left(\alpha_1 \cdot \tau_1' + \ldots + \alpha_n \cdot \tau_n'\right)$$

for the positive convex coefficients $\langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$.

Proof . Because $\{x\}$ is closed, and h is continuous, $G_h(x) = h^{-1}[\{x\}]$ is a closed subset of $\mathbf{S}\ (X)$. Because h is onto, every G_h takes non-empty values; it is clear that $\{G_h(x) \mid x \in X\}$ forms a partition of $\mathbf{S}\ (X)$. Because h is continuous, G_h is k.u.s.c. by Lemma 4.2. Positive convexity will follow immediately from part 4.

Assume that $h(\tau_i) = h(\tau_i') = x_i$ $(1 \le i \le n)$ and observe that $h(\delta_x) = x$ holds for all $x \in X$. Using Lemma 4.1, we get

$$\begin{aligned} h(\alpha_1 \cdot \tau_1 + \ldots + \alpha_n \cdot \tau_n) &= (h \circ \mu_X)\,(\alpha_1 \cdot \delta_{\tau_1} + \ldots + \alpha_n \cdot \delta_{\tau_n}) \\ &= (h \circ \mathbf{S}\ (h))(\alpha_1 \cdot \delta_{\tau_1} + \ldots + \alpha_n \cdot \delta_{\tau_n}) \\ &= h\left(\alpha_1 \cdot \delta_{h(\tau_1)} + \ldots + \alpha_n \cdot \delta_{h(\tau_n)}\right) \\ &= h\left(\alpha_1 \cdot \delta_{x_1} + \ldots + \alpha_n \cdot \delta_{x_n}\right). \end{aligned}$$

In a similar way, $h(\alpha_1 \cdot \tau_1' + \ldots \alpha_n \cdot \tau_n') = h\left(\alpha_1 \cdot \delta_{x_1} + \ldots \alpha_n \cdot \delta_{x_n}\right)$ is obtained. This implies the assertion. $\square$

Thus $\mathcal{G}_h$ is invariant under taking positive convex combinations. It is a positive convex partition in the sense of the following definition.

Definition 4.1. An equivalence relation ρ on $\mathbf{S}\ (X)$ is said to be *positive convex* iff $\tau_i\ \rho\ \tau_i'$ for $1 \le i \le n$ and $\langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega$ together imply

$$(\alpha_1 \cdot \tau_1 + \ldots + \alpha_n \cdot \tau_n)\ \rho\ \left(\alpha_1 \cdot \tau_1' + \ldots + \alpha_n \cdot \tau_n'\right)$$

for each $n \in \mathbb{N}$. A partition of $\mathbf{S}\ (X)$ is called *positive convex* iff its associated equivalence relation is.

Note that the elements of a positive convex partition form positive convex sets. The converse to Proposition 4.1 characterizes algebras

Proposition 4.2. *Assume $\mathcal{G} = \{G(x) \mid x \in X\}$ is a positive convex partition of $\mathbf{S}\ (X)$ into closed sets indexed by X such that $\delta_x \in G(x)$ for each $x \in X$, and such that $x \mapsto G(x)$ is k.u.s.c. Define $h : \mathbf{S}\ (X) \to X$ through $h(\tau) = x$ iff $\tau \in G(x)$. Then $\langle X, h \rangle$ is an algebra for the Giry monad.*

Proof . 1. It is clear that h is well defined and surjective, and that $\exists G(F) = h\,[F]$ holds for each subset $F \subseteq \mathbf{S}\ (X)$. Thus $h\,[K]$ is compact whenever K is compact, because G is k.u.s.c. Thus h is continuous by Lemma 4.2.

2. An easy induction establishes that h respects positive convex combinations: if $h(\tau_i) = h(\tau_i')$ for $i = 1, \dots, n$, and if $\alpha_1, \dots, \alpha_n$ are positive convex coefficients, then

$$h\left(\sum_{i=1}^{n} \alpha_i \cdot \tau_i\right) = h\left(\sum_{i=1}^{n} \alpha_i \cdot \tau_i'\right).$$

We claim that $(h \circ \mu_X)(M) = (h \circ \mathbf{S}\,(h))(M)$ holds for each *discrete* $M \in \mathbf{S}\,(\mathbf{S}\,(X))$. In fact, let

$$M = \sum_{i=1}^{n} \alpha_i \cdot \delta_{\tau_i}$$

be such a discrete measure, then Lemma 4.1 implies that

$$\mu_X(M) = \sum_{i=1}^{n} \alpha_i \cdot \tau_i,$$

thus

$$(h \circ \mu_X)(M) = h\left(\sum_{i=1}^{n} \alpha_i \cdot \tau_i\right) = h\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{h(\tau_i)}\right) = (h \circ \mathbf{S}\,(h))(M),$$

because we know also from Lemma 4.1 that

$$\mathbf{S}\,(h)(M) = \sum_{i=1}^{n} \alpha_i \cdot \delta_{h(\tau_i)}$$

holds.

3. Since the discrete measures are dense in the weak topology [6, Theorem II.6.3], we find for $M_0 \in \mathbf{S}\,(\mathbf{S}\,(X))$ a sequence $(M_n)_{n \in \mathbb{N}}$ of discrete measures M_n with $M_n \rightharpoonup_w M_0$. Consequently, we get from the continuity of both h and μ_X (Lemma 2.1) together with the continuity of $\mathbf{S}\,(h)$

$$(h \circ \mu_X)(M_0) = \lim_{n\to\infty} (h \circ \mu_X)(M_n) = \lim_{n\to\infty} (h \circ \mathbf{S}\,(h))(M_n) = (h \circ \mathbf{S}\,(h))(M_0).$$

This proves the claim. □

We have established

Proposition 4.3. *The algebras $\langle X, h\rangle$ for the Giry monad for Polish spaces X are exactly the positive convex k.u.s.c. partitions $\{G(x) \mid x \in X\}$ into closed subsets of $\mathbf{S}\,(X)$ such that $\delta_x \in G(x)$ for all $x \in X$ holds.*

We characterize the category $\mathfrak{Alg}$ of all algebras for the Giry monad. To this end we package the properties of partitions representing algebras into the notion of a G-partition. They will form the objects of category $\mathfrak{GPart}$.

Definition 4.2. $\mathcal{G}$ is called a *G-partition for X* iff

(1) $\mathcal{G} = \{G(x) \mid x \in X\}$ is a positive convex partition for $\mathbf{S}(X)$ into closed sets indexed by X,
(2) $\delta_x \in G(x)$ holds for all $x \in X$,
(3) the set-valued map $x \mapsto G(x)$ is k.u.s.c.

Define the objects of category $\mathfrak{GPart}$ as pairs $\langle X, \mathcal{G}\rangle$ where X is a Polish space, and $\mathcal{G}$ is a G-partition for X. A morphism f between $\mathcal{G}$ and $\mathcal{G}'$ will map elements of $G(x)$ to $G'(f(x))$ through its associated map $\mathbf{S}(f)$. Thus an element $\tau \in G(x)$ will correspond to an element $\mathbf{S}(f)(\tau) \in G'(f(x))$.

Definition 4.3. A morphism $f : \langle X, \mathcal{G}\rangle \to \langle X', \mathcal{G}'\rangle$ for $\mathfrak{GPart}$ is a continuous map $f : X \to X'$ such that $G(x) \subseteq \mathbf{S}(f)^{-1}\left[G'(f(x))\right]$ holds for each $x \in X$.

Define the functor $F : \mathfrak{Alg} \to \mathfrak{GPart}$ by associating with each algebra $\langle X, h\rangle$ its Giry partition $F(X,h)$ according to Proposition 4.3. Assume that $f : \langle X, h\rangle \to \langle X', h'\rangle$ is a morphism in $\mathfrak{Alg}$, and let $\mathcal{G} = \{G(x) \mid x \in X\}$ and $\mathcal{G}' = \{G'(x') \mid x' \in X'\}$ be the corresponding partitions. Then the properties of an algebra morphism yield

$$\begin{aligned} \tau \in \mathbf{S}(f)^{-1}\left[G'(f(x))\right] &\Leftrightarrow \mathbf{S}(f)(\tau) \in G'(f(x)) \\ &\Leftrightarrow (h' \circ \mathbf{S}(f))(\tau) = f(x) \\ &\Leftrightarrow (f \circ h)(\tau) = f(x). \end{aligned}$$

Thus $\tau \in \mathbf{S}(f)^{-1}\left[G'(f(x))\right]$, provided $\tau \in G(x)$. Hence f is a morphism in $\mathfrak{GPart}$ between $F(X,h)$ and $F(X',h')$. Conversely, let $f : \langle X, \mathcal{G}\rangle \to \langle X', \mathcal{G}'\rangle$ be a morphism in $\mathfrak{GPart}$ with $\langle X, \mathcal{G}\rangle = F(X,h)$ and $\langle X', \mathcal{G}'\rangle = F(X,h')$. Then

$$\begin{aligned} h(\tau) = x &\Leftrightarrow \tau \in G(x) \\ &\Rightarrow \mathbf{S}(f)(\tau) \in G'(f(x)) \\ &\Leftrightarrow h'(\mathbf{S}(f)(\tau)) = f(x), \end{aligned}$$

thus $h' \circ \mathbf{S}(f) = f \circ h$ is inferred. Hence f constitutes a morphism in category $\mathfrak{Alg}$.
Summarizing, we have shown

Proposition 4.4. *The category* $\mathfrak{Alg}$ *of algebras for the Giry monad is isomorphic to the category* $\mathfrak{GPart}$ *of G-partitions.*

4.3. Smooth Relations

The characterization of algebras so far encoded the crucial properties into a partition of $\mathbf{S}(X)$, thus indirectly into an equivalence relation on that space. We can move directly to a particular class of these relations when looking at an alternative characterization of the algebras through smooth equivalence relations. Hence we characterize algebras in terms of the kernel for the associated map, and we show how from such a partition an algebra may be gained. In contrast to the characterization in section 4.2 that started from the fibres $h^{-1}[\{x\}]$ we study here the kernel of h, i.e., the set $\{\langle \tau, \tau'\rangle \mid h(\tau) = h(\tau')\}$. The characterization is interesting in its own right and permits another characterization of morphisms for algebras, it will also help in giving an intrinsic characterization of algebras in terms of convex structures.

Definition 4.4. An equivalence relation ρ on a Polish space A is called *smooth* iff there exists a Polish space B and a Borel measurable map $f : A \to B$ such that

$$a_1 \,\rho\, a_2 \Leftrightarrow f(a_1) = f(a_2)$$

holds, thus ρ is just the kernel of f.

Smooth equivalence relations are a helpful tool in the theory of Borel sets [10] (where they are called sometimes *countably generated*: ρ is smooth iff there exists a countable family $(A_n)_{n\in\mathbb{N}}$ of Borel sets such that $x \,\rho\, x'$ iff $\forall n \in \mathbb{N} : [x \in A_n \Leftrightarrow x' \in A_n]$, see [10, Exercise 5.1.10]). They have some interesting properties that have been capitalized upon in the theory of labelled Markov transition processes [11] and stochastic relations [12].

Some basic notations and constructions first: denote for an equivalence relation ρ on A by A/ρ the factor space, i.e., the set of all equivalence classes $[a]_\rho$, and by

$$\varepsilon_\rho : A \to A/\rho$$

the canonical projection. If A is a Polish space, then let $\mathcal{T}/\rho$ be the final topology on A/ρ with respect to the given topology and ε_ρ, i.e., the largest topology on A/ρ which makes ε_ρ continuous. Clearly a map $g : A/\rho \to B$ for a topological space B is continuous with respect to $\mathcal{T}/\rho$ iff $g \circ \varepsilon_\rho : A \to B$ is continuous w.r.t. the given topologies. We will need this observation in the proof of Proposition 4.5.

Now let $\langle X, h\rangle$ be an algebra for the Giry monad. Obviously

$$\tau_1 \,\rho_h\, \tau_2 \Leftrightarrow h(\tau_1) = h(\tau_2)$$

defines a smooth equivalence relation on the Polish space $\mathbf{S}\,(X)$. Its properties are summarized in

Proposition 4.5. *The equivalence relation ρ_h is positive convex, each equivalence class $[\tau]_{\rho_h}$ is closed and positive convex, and the factor space $\mathbf{S}\,(X)/\rho_h$ is homeomorphic to X when the former is endowed with the topology $\mathcal{T}/\rho_h$.*

Proof . 1. Positive convexity of ρ_h follows from the properties of h exactly as in the proof of Proposition 4.1, from this, positive convexity of the classes is also inferred. Continuity of h implies that the classes are closed sets.

2. Define $\chi_h([\tau]_{\rho_h}) := h(\tau)$ for $\tau \in \mathbf{S}\,(X)$. Then $\chi_h : \mathbf{S}\,(X)/\rho_h \to X$ is well defined and a bijection. Let $G \subseteq X$ be an open set, then $\varepsilon_{\rho_h}^{-1}\left[\chi_h^{-1}[G]\right] = h^{-1}[G]$. Because $\mathcal{T}/\rho_h$ is the largest topology on $\mathbf{S}\,(X)/\rho_h$ that renders ε_{ρ_h} continuous, and because $h^{-1}[G] \subseteq \mathbf{S}\,(X)$ is open by assumption, we infer that $\chi_h^{-1}[G]$ is $\mathcal{T}/\rho_h$-open. Thus χ_h is continuous. On the other hand, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in X converging to $x_0 \in X$, then $\delta_{x_n} \rightharpoonup_w \delta_{x_0}$ in $\mathbf{S}\,(X)$, thus $\left[\delta_{x_n}\right]_{\rho_h} \to \left[\delta_{x_0}\right]_{\rho_h}$ in $\mathcal{T}/\rho_h$ by construction. Consequently χ_h^{-1} is also continuous. $\square$

Thus each algebra induces a G-triplet in the following sense.

Definition 4.5. A *G-triplet* $\langle X, \rho, \chi\rangle$ is a Polish space X with a smooth and positive convex equivalence relation ρ on $\mathbf{S}\,(X)$ such that $\chi : \mathbf{S}\,(X)/\rho \to X$ is a homeomorphism with $\chi([\delta_x]_\rho) = x$ for all $x \in X$. Here $\mathbf{S}\,(X)/\rho$ carries the final topology with respect to the weak topology on $\mathbf{S}\,(X)$ and ε_ρ.

Now assume that a G-triplet $\langle X, \rho, \chi\rangle$ is given. Define $h(\tau) := \chi([\tau]_\rho)$ for $\tau \in \mathbf{S}\,(X)$. Then $\langle X, h\rangle$ is an algebra for the Giry monad: $h(\delta_x) = x$ follows from the assumption, and because $h = \chi \circ \varepsilon_\rho$, holds, the map h is continuous. An argument very similar to that used in the proof of Proposition 4.2 shows that $h \circ \mu_X = h \circ \mathbf{S}\,(h)$ holds; this is so since ρ is assumed to be positive convex.

Definition 4.6. The continuous map $f : X \to X'$ between the Polish spaces X and X' constitutes a *G-triplet morphism* $f : \langle X, \rho, \chi\rangle \to \langle X', \rho', \chi'\rangle$ iff these conditions hold

(1) $\tau\,\rho\,\tau'$ implies $\mathbf{S}\,(f)(\tau)\,\rho'\,\mathbf{S}\,(f)\tau'$,
(2) the diagram

$$\begin{tikzcd} \mathbf{S}\,(X)/\rho \arrow[r, "\mathbf{S}(f)_{\rho,\rho'}"] \arrow[d, "\chi"'] & \mathbf{S}\,(X')/\rho' \arrow[d, "\chi'"] \\ X \arrow[r, "f"'] & X \end{tikzcd}$$

commutes, where $\mathbf{S}\,(f)_{\rho,\rho'}\left([\tau]_\rho\right) := \left[\mathbf{S}\,(f)(\tau)\right]_{\rho'}$.

G-triplets with their morphisms form a category $\mathfrak{GTrip}$.

Lemma 4.3. *Each algebra morphism $f : \langle X, h\rangle \to \langle X', h'\rangle$ induces a G-triplet morphism $f : \langle X, \rho_h, \chi_h\rangle \to \langle X', \rho_{h'}, \chi_{h'}\rangle$.*

Proof . 1. It is an easy calculation to show that $\tau\,\rho_h\,\tau'$ implies $\mathbf{S}\,(f)(\tau)\,\rho_{h'}\,\mathbf{S}\,(f)(\tau)$. This is so because f is a morphism for the algebras.

2. Since for each $\tau \in \mathbf{S}\,(X)$ there exists $x \in X$ such that $[\tau]_{\rho_h} = [\delta_x]_{\rho_h}$ (in fact, $h(\tau)$ would do, because $h(\tau) = h\left(\delta_{h(\tau)}\right)$, as shown above), it is enough to demonstrate that

$$\chi'_{h'}\left(\mathbf{S}\,(f)_{\rho_h,\rho'_{h'}}\left([\delta_x]_{\rho_h}\right)\right) = f\left(\chi_h\left([\delta_x]_{\rho_h}\right)\right)$$

is true for each $x \in X$. Because $\mathbf{S}\,(f)(\delta_x) = \delta_{f(x)}$, a little computation shows that both sides of the above equation boil down to $f(x)$. □

The morphisms between G-triplets are just the morphisms between algebras (when we forget that these games play in different categories).

Proposition 4.6. *Let $f : \langle X, \rho, \chi\rangle \to \langle X', \rho', \chi'\rangle$ be a morphism between G-triplets, and let $\langle X, h\rangle$, respectively, $\langle X', h'\rangle$ be the associated algebras. Then $f : \langle X, h\rangle \to \langle X', h'\rangle$ is an algebra morphism.*

Proof . Given $\tau \in \mathbf{S}\,(X)$ we have to show that $(f \circ h)(\tau)$ equals $(h' \circ \mathbf{S}\,(f))(\tau)$. Since $h(\tau) = \chi([\tau]_\rho)$, we obtain

$$(f \circ h)(\tau) = f\left(\chi([\tau]_\rho)\right)$$

$$= \chi'\left(\mathbf{S}\,(f)_{\rho,\rho'}([\tau]_\rho)\right)$$
$$= \chi'\left(\left[\mathbf{S}\,(f)(\tau)\right]_{\rho'}\right)$$
$$= (h' \circ \mathbf{S}\,(f))(\tau). \qquad \square$$

Putting all these constructions with their properties together, we obtain

Proposition 4.7. *The category* $\mathfrak{Alg}$ *of algebras for the Giry monad is isomorphic to the category* $\mathfrak{GTrip}$ *of G-triplets.*

The probabilistic case requires a separate discussion. It is treated similarly. We define an equivalence relation ρ on $\mathbf{P}(X)$ to be convex iff for each $n \in \mathbb{N}$ the conditions $\tau_i\ \rho\ \tau_i'$ for $1 \leq i \leq n$ and $\langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega_{\mathbf{c}}$ together imply $\left(\sum_{i=1}^n \alpha_i \cdot \tau_i\right)\ \rho\ \left(\sum_{i=1}^n \alpha_i \cdot \tau_i'\right)$, where

$$\Omega_{\mathbf{c}} := \{\langle \alpha_1, \ldots, \alpha_k \rangle \mid \alpha_i \geq 0, \alpha_1 + \ldots + \alpha_k = 1\}$$

are all convex coefficients. Then the ρ-classes form convex subsets of $\mathbf{P}(X)$. We introduce *PG-triplets* $\langle X, \rho, \chi \rangle$ for a Polish space X, a smooth convex equivalence relation ρ and a homeomorphism $\chi : \mathbf{P}(X)/\rho \to X$ with $\chi\left([\delta_x]_\rho\right) = x$ for all $x \in X$. A continuous map $f : X \to X'$ then is a PG-triplet morphism $\langle X, \rho, \chi \rangle \to \langle X', \rho', \chi' \rangle$ iff

(1) $\tau\ \rho\ \tau' \Rightarrow \mathbf{P}(f)(\tau) = \mathbf{P}(f)(\tau')$,
(2) $\chi' \circ \mathbf{P}(f)_{\rho,\rho'} = f \circ \chi$

Here $\mathbf{P}(f)_{\rho,\rho'}$ is defined in analogy to $\mathbf{S}\,(f)_{\rho,\rho'}$ in Definition 4.6 as

$$\mathbf{P}(f)_{\rho,\rho'}\left([\tau]_\rho\right) := \left[\mathbf{P}(f)(\tau)\right]_{\rho'}.$$

We see then that each algebra morphism $f : \langle X, h \rangle \to \langle X', h' \rangle$ induces a PG-triplet morphism $f : \langle X, \rho_h, \chi_h \rangle \to \langle X', \rho_{h'}, \chi_{h'} \rangle$, and vice versa. The reader is invited to fill in the details.
Summarizing, this yields

Proposition 4.8. *The category of algebras for the Giry monad for the probability functor is isomorphic to the full subcategory of G-triplets* $\langle X, \rho, \chi \rangle$ *with a smooth and convex equivalence relation such that* $\chi : \mathbf{P}(X)/\rho \to X$ *is a homeomorphism.*

We will show now that $\mathfrak{StrConv}$ is isomorphic to $\mathfrak{Alg}$.

4.4. Positive Convex Structures

The algebras can also be described without having to resort to $\mathbf{S}\,(X)$ by an intrinsic characterization through positive convex structures, and their morphisms as the affine maps on these structures.

This characterization is comparable to the one given by Manes for the power set monad, cp. Remark 4.1.

Lemma 4.4. *Given an algebra* $\langle X, h\rangle$, *define for* $x_1, \ldots, x_n \in X$ *and the positive convex coefficients* $\langle \alpha_1, \ldots, \alpha_n\rangle \in \Omega$

$$\sum_{i=1}^{n} \alpha_i \cdot x_i := h\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}\right),$$

then this defines a positive convex structure on X.

Proof . 1. Because

$$h\left(\sum_{i=1}^{n} \delta_{i,j} \cdot \delta_{x_i}\right) = h(\delta_{x_j}) = x_j,$$

property 1 in Definition 3.1 is satisfied.

2. Proving property 2, we resort to the properties of algebras, and of a monad

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i \cdot \left(\sum_{k=1}^{m} \beta_{i,k} \cdot x_k\right) &= h\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{\sum_{k=1}^{m} \beta_{i,k} \cdot x_k}\right) &(1)\\
&= h\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{h\left(\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}\right)}\right) &(2)\\
&= h\left(\sum_{i=1}^{n} \alpha_i \cdot \mathbf{S}\ (h)\left(\delta_{\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}}\right)\right) &(3)\\
&= (h \circ \mathbf{S}\ (h))\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}}\right) &(4)\\
&= (h \circ \mu_X)\left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}}\right) &(5)\\
&= h\left(\sum_{i=1}^{n} \alpha_i \cdot \mu_X\left(\delta_{\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}}\right)\right) &(6)\\
&= h\left(\sum_{i=1}^{n} \alpha_i \cdot \left(\sum_{k=1}^{m} \beta_{i,k} \cdot \delta_{x_k}\right)\right) &(7)\\
&= h\left(\sum_{k=1}^{m} \left(\sum_{i=1}^{n} \alpha_i \cdot \beta_{i,k}\right) \delta_{x_k}\right) &(8)\\
&= \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \alpha_i \cdot \beta_{i,k}\right) x_k &(9)
\end{aligned}$$

The equations (1) and (2) reflect the definition of the structure, equation (3) applies $\delta_{h(\tau)} = \mathbf{S}\ (h)(\delta_\tau)$, equation (4) uses the linearity of $\mathbf{S}\ (h)$ according to Lemma 4.1, equation (5) is due to h being an algebra, and now we are winding down. Equation (6) uses Lemma 4.1 again, this time for μ_X, equation (7) uses that $\mu_X \circ \delta_\tau = \tau$, equation (8) is just rearranging terms, and equation (9) is the definition again. $\square$

Let conversely such a positive convex structure be given. We show that we can define a G-triplet from it. Let

$$\mathcal{T}_X := \left\{ \sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i} \mid n \in \mathbb{N}, x_1, \ldots, x_n \in X, \langle \alpha_1, \ldots, \alpha_n \rangle \in \Omega \right\},$$

then $\mathcal{T}_X$ is dense in $\mathbf{S}\ (X)$. Put

$$h_0 \left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i} \right) := \sum_{i=1}^{n} \alpha_i \cdot x_i,$$

then $h_0 : \mathcal{T}_X \to X$ well defined. This is so since

$$\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i} = \sum_{j=1}^{m} \alpha'_j \cdot \delta_{x'_j}$$

implies that

$$\sum_{i=1, \alpha_i \neq 0}^{n} \alpha_i \cdot \delta_{x_i} = \sum_{j=1, \alpha'_j \neq 0}^{m} \alpha'_j \cdot \delta_{x'_j},$$

hence given i with $\alpha_i \neq 0$ there exists j with $\alpha'_j \neq 0$ such that $x_i = x'_j$ and vice versa. Consequently,

$$\sum_{i=1}^{n} \alpha_i \cdot x_i = \sum_{i=1, \alpha_i \neq 0}^{n} \alpha_i \cdot x_i = \sum_{j=1, \alpha'_j \neq 0}^{n} \alpha'_j \cdot x'_j = \sum_{j=1}^{n} \alpha'_j \cdot x'_j$$

is inferred from the properties of positive convex structures.

The map h_0 is uniformly continuous, because

$$\mathrm{d} \left(h_0 \left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i} \right), h_0 \left(\sum_{j=1}^{m} \beta_j \cdot \delta_{y_j} \right) \right) \leq \mathbf{d}_P \left(\sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}, \sum_{j=1}^{m} \beta_j \cdot \delta_{y_j} \right).$$

This is immediate from the definition of $\mathbf{d}_P$. Note that we need uniform continuity here, because otherwise a unique, continuous extension from the dense subset of discrete measures to all measures cannot be guaranteed.

Define ρ_0 as the kernel of h_0, then ρ_0 is a smooth equivalence relation on $\mathcal{T}_X$, and it is not difficult to see that the set of topological closures $\left\{ \left([t]_{\rho_0}\right)^{\mathsf{cl}} \mid t \in \mathcal{T}_X \right\}$ forms a partition of $\mathbf{S}\ (X)$:

(1) the closures of different equivalence classes are disjoint,

(2) given $\tau \in \mathbf{S}\ (X)$, one can find a sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathcal{T}_X$ with $t_n \rightharpoonup_w \tau$. Since X is Polish, in particular complete, the sequence $(h_0(t_n))_{n\in\mathbb{N}}$ converges to some t_0, and because h_0 is uniformly continuous, one concludes that $\tau \in \left([t_0]_{\rho_0}\right)^{\mathsf{cl}}$. Thus each member of $\mathbf{S}\ (X)$ is in some class.

This yields an equivalence relation ρ on $\mathbf{S}\ (X)$. Uniform continuity of h_0 gives a unique continuous extension h of h_0 to $\mathbf{S}\ (X)$, thus ρ equals the kernel of h, hence ρ is a smooth equivalence relation, and it is evidently positive convex. Defining on $\mathbf{S}\ (X)/\rho$ the metric

$$D([\tau_1]_\rho, [\tau_2]_\rho) := d(h(\tau_1), h(\tau_2)),$$

it is rather immediate that

(1) the metric space $(\mathbf{S}\ (X)/\rho, D)$ is homeomorphic to X with metric d,
(2) the topology induced by the metric D is just the final topology with respect to the weak topology on $\mathbf{S}\ (X)$ and ε_ρ.

It is clear that each affine and continuous map between positive convex structures gives rise to a morphisms between the corresponding G-triplets, and vice versa.

Thus we have established

Proposition 4.9. *The category* $\mathfrak{Alg}$ *of algebras for the Giry monad is isomorphic to the category* $\mathfrak{StrConv}$ *of positive convex structures with continuous affine maps as morphisms.*

For the probability functor we again mirror the development, but this time we need not go into details. We obtain eventually for the category $\mathfrak{pAlg}$ of algebras for the Giry monad, when restricted to the probability functor (with the obvious necessary adjustments made for morphisms).

Proposition 4.10. *The category* $\mathfrak{pAlg}$ *of algebras for the Giry monad for the probability functor is isomorphic to the full subcategory of continuous convex structures.*

This characterization has been known for the probability functor in the case that X is a compact Hausdorff space [3, 2.14] (the attribution to Swirszcz's work [4] in [3] is slightly unclear). The methods for the proof are, however, rather different: the compact case makes essential use of the right adjoint of the probability functor as a functor between the respective categories of compact Hausdorff spaces and compact convex sets. This adjoint is not yet characterized fully in the present situation. Thus Corollary 4.10 generalizes the characterization to Polish spaces.

5. Examples

This section illustrates the concept and proposes some examples by looking at some well-known situations, thus most of this section is not really new, probably apart from the proposed point of view. We first show that the monad carries for each Polish space an instance of an algebra with it. Then we prove that in the finite case an algebra exists only in the case of a singleton set. Finally a geometrically oriented example is discussed by investigating the barycenter of a probability in a compact and convex subset of $\mathbb{R}^n$. In each case it turns out that the convex structure associated with the algebra is the natural one, i.e., the one provided through the existing geometric structure.

5.1. Monad multiplication

We show that $\langle \mathbf{S}\ (X), \mu_X \rangle$ is an algebra whenever X is a Polish space. This is not only interesting in its own right, it shows moreover that each Polish space is associated in a natural fashion with a strongly convex structure. This association is actually more than meets the eye: we will show in section 6 that $X \mapsto \langle \mathbf{S}\ (X), \mu_X \rangle$ is the object part of the left adjoint to the forgetful functor $\mathfrak{Alg} \to \mathfrak{Pol}$. In analogy to [13], the pair $\langle \mathbf{S}\ (X), \mu_X \rangle$ is referred to (informally) as the *stochastic powerdomain* for Polish space X, because it exhibits properties similar to the probabilistic powerdomains studied in [13], and to the plain powerdomain that is based on the power set functor.

Example 5.1. The pair $\langle \mathbf{S}\ (X), \mu_X \rangle$ is always an algebra. We know from Lemma 2.1 that $\mu_X : \mathbf{S}\ (\mathbf{S}\ (X)) \to \mathbf{S}\ (X)$ is continuous. Because $\langle \mathbf{S}, \eta, \mu \rangle$ is a monad, the natural transformation $\mu : \mathbf{S}^2 \stackrel{\bullet}{\to} \mathbf{S}$ satisfies

$$\mu \circ \mathbf{S}\mu = \mu \circ \mu\mathbf{S}$$

in the category of functors with natural transformations as morphisms, see the diagram at the end of section 2. Since $(\mathbf{S} \circ \mu)_X = \mathbf{S}\ (\mu_X)$ and $(\mu \circ \mathbf{S})_X = \mu_{\mathbf{S}\ (X)}$, this translates to

$$\mu_X \circ \mathbf{S}\ (\mu_X) = \mu_X \circ \mu_{\mathbf{S}\ (X)}.$$

Because the equation $\mu_X \circ \eta_{\mathbf{S}\ (X)} = id_{\mathbf{S}\ (X)}$ is easily established through a simple computation, the defining diagrams are commutative.

Since

$$\mu_X(\alpha_1 \cdot \tau_1 + \ldots + \alpha_n \cdot \tau_n) = \alpha_1 \cdot \mu_X(\tau_1) + \ldots + \alpha_n \cdot \mu_X(\tau_n),$$

the positive convex structure induced on $\mathbf{S}\ (X)$ by this algebra is the natural one.

5.2. The finite case

The finite case can easily be characterized: there are no algebras for $\{1, \ldots, n\}$ unless $n = 1$. This will be shown now. As a byproduct we obtain a simple geometric description as a necessary condition for the existence of algebras.

We need a wee bit elementary topology for this.

Definition 5.1. A metric space A is called *connected* iff the decomposition $A = A_1 \cup A_2$ with disjoint open sets A_1, A_2 implies $A_1 = \emptyset$ or $A_2 = \emptyset$.

Thus a connected space cannot be decomposed into two non-trivial open sets. The connected subspaces of the real line $\mathbb{R}$ are just the open, half-open or closed finite or infinite intervals. The rational numbers $\mathbb{Q}$ are not connected. A subset $\emptyset \neq A \subseteq \mathbb{N}$ of the natural numbers which carries the discrete topology (because we assume that it is a Polish space) is connected as a subspace iff $A = \{n\}$ for some $n \in \mathbb{N}$.

The following facts about connected spaces are well known, see for example [14, Chapter 6.1], or any other standard reference to set-theoretic topology.

Lemma 5.1. *Let A be a metric space.*

(1) If A is connected, and $f : A \to B$ is a continuous and surjective map to another metric space B, then B is connected.
(2) If two arbitrary points in A can be joined through a connected subspace of A, then A is connected.

This has as a consequence

Corollary 5.1. *If $\langle X, h\rangle$ is an algebra for the Giry monad, then X is connected.*

Proof . If $\tau_1, \tau_2 \in \mathbf{S}\ (X)$ are arbitrary probability measures on X, then the line segment $\{c \cdot \tau_1 + (1-c) \cdot \tau_2 \mid 0 \leq c \leq 1\}$ is a connected subspace which joins τ_1 and τ_2. This is so because it is the image of the connected unit interval $[0,1]$ under the continuous map $c \mapsto c \cdot \tau_1 + (1-c) \cdot \tau_2$. Thus $\mathbf{S}\ (X)$ is connected by Lemma 5.1. Since h is onto, its image X is connected. □

Consequently it is hopeless to search for algebras for, say, the natural numbers or a non-trivial subset of it:

Corollary 5.2. *A subspace $A \subseteq \mathbb{N}$ has algebras for the Giry monad iff A is a singleton set.*

Proof . It is clear that a singleton set has an algebra. Conversely, if A has an algebra, then A is connected by Corollary 5.1and this can only be the case when A is a singleton. □

The next example deals with the unit interval:

Example 5.2. The map

$$h : \mathbf{S}\ ([0,1]) \ni \tau \mapsto \int_0^1 t\ \tau(dt) \in [0,1]$$

defines an algebra $\langle [0,1], h\rangle$. In fact, $h(\tau) \in [0,1]$ because τ is a subprobability measure. It is clear that $h(\delta_x) = x$ holds, and—by the very definition of the weak topology—that $\tau \mapsto h(\tau)$ is continuous. Thus by Proposition 4.2 it remains to show that the partition induced by h is positive convex. This is a fairly simple calculation. Consequently, the partition induced by h is a G-partition, showing that h is indeed the morphism part of an algebra.

It is not difficult to see that the positive convex structure induced on $[0,1]$ is the natural one.

This is the only algebra that has an integral representation through Lebesgue measure: suppose that

$$h^*(\tau) = \int_0^1 f(t)\ \tau(dt)$$

for some continuous f. Then $h^*(\delta_x) = f(x)$, from which $f(x) = x$ is inferred for each $x \in [0,1]$.

5.3. Barycenter

The final example has a more geometric touch to it and deals only with the probabilistic case. We work with bounded and closed subsets of some Euclidean space and show that the construction of a barycenter yields an algebra. Fix $X \subseteq \mathbb{R}^n$ as a bounded, closed and convex subset of the Euclidean space $\mathbb{R}^n$ (for example, X could be a closed ball or a cube in $\mathbb{R}^n$).

Denote for two vectors $x, x' \in \mathbb{R}^n$ by

$$x \bigstar x' := \sum_{i=1}^{n} x_i \cdot x'_i$$

their inner product. Then $\lambda x.x \bigstar x'$ constitutes a continuous linear map on $\mathbb{R}^n$ for fixed x'. In fact, each linear functional on $\mathbb{R}^n$ can be represented in this way.

Definition 5.2. The vector $x^* \in \mathbb{R}^n$ is called a *barycenter* of the probability measure $\tau \in \mathbf{P}(X)$ iff

$$x \bigstar x^* = \int_X x \bigstar y \; \tau(\mathrm{d}y)$$

holds for each $x \in X$.

Because X is compact, the integrand is bounded on X, thus the integral is always finite.

We collect some basic facts about barycenters and refer the reader to [15] for details.

Lemma 5.2. *The barycenter of $\tau \in \mathbf{P}(X)$ exists, it is uniquely determined, and it is an element of X.*

Proof . Once we know that the barycenter exists, uniqueness follows from the well-known fact that the linear functionals on $\mathbb{R}^n$ separate points. Existence of the barycenter is established in [15, Theorem 461 E], its membership in X follows from [15, Theorem 461 H]. □

These preparations help in establishing that the barycenter constitutes an algebra:

Proposition 5.1. *Let $b(\tau)$ be the barycenter of $\tau \in \mathbf{P}(X)$. Then $\langle X, b\rangle$ is an algebra for the Giry monad.*

Proof . 1. $b : \mathbf{P}(X) \to X$ is well defined by Lemma 5.2. From the uniqueness of the barycenter it is clear that $b(\delta_x) = x$ holds for each $x \in X$.

2. Assume that $(\tau_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbf{P}(X)$ with $\tau_n \rightharpoonup_w \tau_0$. Put $x_n^* := b(\tau_n)$ as the barycenter of τ_n, then $(x_n^*)_{n\in\mathbb{N}}$ is a sequence in the compact set X, thus has a convergent subsequence (which we take w.l.g. as the sequence itself). Let x_0^* be its limit. Then we have for all $x \in X$:

$$x \bigstar x_n^* = \int_X x \bigstar y \; \tau_n(\mathrm{d}y) \to \int_X x \bigstar y \; \tau_0(\mathrm{d}y) = x \bigstar x_0^*$$

Hence b is continuous.

3. It remains to show that the partition induced by b is convex. This, however, follows immediately from the linearity of $y \mapsto \lambda x.x \bigstar y$. □

Calculating the convex structure for b, we infer from affinity of the integral as a function of the measure and from

$$x\bigstar b(\tau) = \int_X x\bigstar y\ \tau(\mathrm{d}y)$$

that $(0 \leq c \leq 1, \tau_i \in \mathbf{P}(X))$

$$b(c \cdot \tau_1 + (1-c) \cdot \tau_2) = c \cdot b(\tau_1) + (1-c) \cdot b(\tau_2)$$

that the convex structure induced by b is the natural one.

It should be mentioned that this example can be generalized considerably to metrizable topological vector spaces. The terminological effort is, however, somewhat heavy, and the example remains essentially the same. Thus we refrain from a more general discussion.

Although the characterization of algebras in terms of positive convex structures yields a somewhat uniform approach, it becomes clear from these examples that the specific instances of the algebras provide a rather colorful picture unified only through the common abstract treatment.

6. The Left Adjoint

The identification of the algebras for the Giry monad and the observation from Example 5.1 that $\langle \mathbf{S}(X), \mu_X \rangle$ is always an algebra puts us in a position where we are able to identify the left adjoint for the forgetful functor $\mathbf{U} : \mathfrak{Alg} \to \mathfrak{Pol}$. Define $\mathbf{L}(X) := \langle \mathbf{S}(X), \mu_X \rangle$, for a Polish space X, hence $\mathbf{L}(X)$ is the stochastic powerdomain associated with X, for the continuous map $f : X \to Y$ put $\mathbf{L}(f) := \mathbf{S}(f)$, then we know from Example 5.1 that $\mathbf{L}(X)$ is an algebra. From Lemma 4.1, part 4.1, we see that $\mathbf{L}(f) : \mathbf{L}(X) \to \mathbf{L}(Y)$, is a morphism in $\mathfrak{Alg}$, and since $\mu : \mathbf{S}^2 \xrightarrow{\bullet} \mathbf{S}$ is a natural transformation, $\mathbf{L}(f)$ is an algebra morphism. Thus $\mathbf{L} : \mathfrak{Pol} \to \mathfrak{Alg}$ is a functor.

We will write as usual $\mathfrak{C}(a, b)$ for the morphisms $a \to b$ in category $\mathfrak{C}$.

Lemma 6.1. *Let $\theta : \mathbf{L}(X) \to \langle Y, h \rangle$ be a morphism in $\mathfrak{Alg}$, and put $\Theta(\theta)(x) := \theta(\delta_x)$.*
This defines a bijection

$$\Theta : \mathfrak{Alg}(\mathbf{L}(X), \langle Y, h \rangle) \to \mathfrak{Pol}(X, Y).$$

Proof . 1. Since $x \mapsto \delta_x$ defines a continuous map $X \to \mathbf{S}(X)$, and since the morphisms in $\mathfrak{Alg}$ are continuous as well, $\Theta(\theta) \in \mathfrak{Pol}(X, Y)$ whenever $\theta \in \mathfrak{Alg}(\mathbf{L}(X), \langle Y, h \rangle)$.

2. Now suppose that $\Theta(\theta_1)(x) = \Theta(\theta_2)(x)$, holds for all $x \in X$, thus $\theta_1(\delta_x) = \theta_2(\delta_x)$, for all $x \in X$. Let $\tau = \sum_{i=1}^m \alpha_i \cdot \delta_{x_i}$, be a discrete subprobability measure, then

$$\begin{aligned}\theta_1(\tau) &= \theta_1\left(\sum_{i=1}^m \alpha_i \cdot \delta_{x_i}\right)\\ &= \sum_{1 \leq i \leq m}^{\mathcal{P}} \alpha_i \cdot \theta_1(\delta_{x_i})\end{aligned}$$

$$
\begin{aligned}
&= \sum_{1 \le i \le m}^{\mathcal{P}} \alpha_i \cdot \theta_2(\delta_{x_i}) \\
&= \theta_2(\tau).
\end{aligned}
$$

Here $\mathcal{P}$ is the positive convex structure associated with the algebra $\langle Y, h\rangle$ by Proposition 4.9. Thus θ_1 agrees with θ_2 on all discrete measures. Since these measures are dense in the weak topology, and since θ_1 as well as θ_2 is continuous, we may conclude that $\theta_1(\tau) = \theta_2(\tau)$ holds for all $\tau \in \mathbf{S}\,(X)$. Thus Θ is injective.

3. Let $f : X \to Y$ be continuous, and put $\widetilde{\theta} := h \circ \mathbf{S}\,(f)$, the composition being formed in $\mathfrak{Alg}$. We claim that $\widetilde{\theta} \in \mathfrak{Alg}(\mathbf{L}\,(X)\,, \langle Y, h\rangle)$.

In fact, consider the diagram

$$
\begin{tikzcd}
\mathbf{S}\,(\mathbf{S}\,(X)) \arrow[r, "\mathbf{S}(\widetilde{\theta})"] \arrow[d, "\mu_X"'] & \mathbf{S}\,(Y) \arrow[d, "h"] \\
\mathbf{S}\,(X) \arrow[r, "\widetilde{\theta}"'] & Y
\end{tikzcd}
$$

We have

$$
\begin{aligned}
h \circ \mathbf{S}\,(\widetilde{\theta}), &= h \circ \mathbf{S}\,(h) \circ \mathbf{S}\,(\mathbf{S}\,(f\,)), \\
&= h \circ \mu_Y \circ \mathbf{S}\,(\mathbf{S}\,(f\,)) \text{ (because } \langle Y, h\rangle \text{ is an algebra)} \\
&= h \circ \mathbf{S}\,(f\,) \circ \mu_X \text{ (since } \mathbf{S}\,(f\,) \text{ is an } \mathfrak{Alg}\text{-morphism)} \\
&= \widetilde{\theta} \circ \mu_X
\end{aligned}
$$

which implies that the diagram is commutative, establishing the claim. Since for each $x \in X$

$$
\Theta(\widetilde{\theta})(x) = h(\mathbf{S}\,(f)(\delta_x)) = h(\delta_{f(x)}) = f(x)
$$

we conclude that $\Theta(\widetilde{\theta}) = f$, thus Θ is onto. $\square$

In order to establish the properties of an adjunction, we need to establish the naturalness of $\Theta = \Theta_{X,\langle Y,h\rangle}$, see [2, p. 80]. This means that we have to establish the commutativity of the diagrams below, given $f \in \mathfrak{Alg}(\langle Y, h\rangle, \langle Y', h'\rangle)$ and $g \in \mathfrak{Pol}(X', X)$.

The first diagram takes care of the hom-set functor $\mathfrak{Alg}(\mathbf{L}\,(X)\,, \cdot)$:

$$
\begin{tikzcd}
\mathfrak{Alg}(\mathbf{L}\,(X)\,, \langle Y, h\rangle) \arrow[r, "\Theta"] \arrow[d, "f_*"'] & \mathfrak{Pol}(X, Y) \arrow[d, "\mathbf{U}(f)_*"] \\
\mathfrak{Alg}(\mathbf{L}\,(X)\,, \langle Y', h'\rangle) \arrow[r, "\Theta"'] & \mathfrak{Pol}(X, Y')
\end{tikzcd}
$$

with

$$
f_* : \mathfrak{Alg}(\mathbf{L}\,(X)\,, \langle Y, h\rangle) \ni \theta \mapsto f \circ \theta \in \mathfrak{Alg}(\mathbf{L}\,(X)\,, \langle Y', h'\rangle)
$$

as composition from the left, similarly $\mathbf{U}(f\,)_*$. We see

$$
\mathbf{U}(f\,)_*\,(\Theta(\theta))(x) = (f \circ \Theta(\theta))(x) = f(\Theta(\theta)(x)) = f(\theta(\delta_x)),
$$

and

$$\Theta(f_*(\theta))(x) = f_*(\theta)(\delta_x) = f(\theta(\delta_x)),$$

hence the diagram commutes. The second diagram takes care of the contravariant hom-functor $\mathfrak{Alg}(\cdot, \langle Y, h\rangle)$:

$$\begin{array}{ccc} \mathfrak{Alg}(\mathbf{L}(X), \langle Y, h\rangle) & \xrightarrow{\Theta} & \mathfrak{Pol}(X, Y) \\ \downarrow{\scriptstyle \mathbf{L}(f)^*} & & \downarrow{\scriptstyle f^*} \\ \mathfrak{Alg}(\mathbf{L}(X'), \langle Y, h\rangle) & \xrightarrow[\Theta]{} & \mathfrak{Pol}(X', Y) \end{array}$$

Here

$$f^* : \mathfrak{Pol}(X, Y) \ni g \mapsto g \circ f \in \mathfrak{Pol}(X', Y)$$

is composition from the right, similarly for $\mathbf{L}(f)^*$. Because

$$f^*(\Theta(\theta))(x') = (\Theta(\theta) \circ f)(x') = \theta(\delta_{f(x')}),$$

and since

$$\Theta(\mathbf{L}(f)^*(\theta))(x') = \mathbf{L}(f)^*(\theta)(\delta_{x'}) = (\theta \circ \mathbf{S}(f))(\delta_{x'}) = \theta(\delta_{f(x')})$$

we see that this diagram commutes as well.

Summarizing, we have established:

Proposition 6.1. *The functor* $\mathbf{L} : \mathfrak{Pol} \to \mathfrak{Alg}$ *with* $\mathbf{L}(X) := \langle \mathbf{S}(X), \mu_X\rangle$ *and* $\mathbf{L}(f) := \mathbf{S}(f)$ *is left adjoint to the forgetful functor* $\mathbf{U} : \mathfrak{Alg} \to \mathfrak{Pol}$.

The probabilistic case is dealt with using the same arguments. The only place where the difference between subprobability measures and probability measures comes formally into the discussion is in the proof of Lemma 6.1. Proving surjectivity of Θ one has to take a convex combination of discrete measures, rather than a positive convex combination, as in the proof above. With this minor change all proofs carry over verbatim. We obtain for the category $\mathfrak{pAlg}$ of algebras for the probabilistic version of the Giry monad (the category has been introduced in Proposition 4.10).

Proposition 6.2. *The functor* $\mathbf{L}_{\text{prob}} : \mathfrak{Pol} \to \mathfrak{pAlg}$ *with* $\mathbf{L}_{\text{prob}}(X) := \langle \mathbf{P}(X), \mu_X\rangle$ *and* $\mathbf{L}_{\text{prob}}(f) := \mathbf{P}(f)$ *is left adjoint to the forgetful functor* $\mathbf{U} : \mathfrak{pAlg} \to \mathfrak{Pol}$.

Hence the forgetful functor on the algebras for the Giry monad has the stochastic powerdomain as its left adjoint. This emphasizes the close ties between positive convex, respectively, convex structures and probabilities and sheds further light on these functors. It also adds a formal underpinning to the intuitive understanding prevailing in Computer Science which often expresses the probability of an outcome as a convex combination of all the possible outcomes, see e.g., [16,13] for accounts in different fields. The interplay between convexity and probability is strikingly present in Heckmann's work [17] (in fact, he often interchanges both), but surprisingly not made explicit.

7. Related Work

The monad on which the present investigation is based was originally proposed and investigated by Giry [8] in an approach to provide a categorical foundation of Probability Theory. The functor on which it is based assigns each measurable space all probabilities defined on its σ-algebra, it is somewhat similar to the functor assigning each set its power set on which the monad investigated by Manes is based. While the Kleisli construction for the latter one leads to relations based on sets, it leads for the former one to stochastic relations as a similar relational construction. This point of view was stressed again by Panagaden in [18] when pointing out similarities between set based and probability based relations. It was extended further in [19]. In [20] this aspect is elaborated in depth by showing how a software architecture can be modelled using a monad as the basic computational model; the monad is shown to subsume both the Manes and the Giry monad as special cases. Stochastic relations turned out to be a fruitful field for investigations [19,21,12] in particular in such areas as labelled Markov transition systems and modelling stochastically algebraic aspects of modal and temporal logic [22]. Heckmann [17] discusses different approaches to probabilistic domains and addresses the question of the equivalence of different axiomatizations. His most powerful theory is called *Multiple Choice with Divergence*, and the axioms come very close to the axioms for positive convex sets—he even uses the notion of a formal linear combination $\sum_{i=1}^{n} p_i \cdot x_i$. The theory is related to the probabilistic powerdomains investigated by Jones and Plotkin [23,13]. In contrast to the work reported about in the present paper, however, no connection is drawn by Heckmann to probabilistic scenarios based on other than finite sets, and no relation to categorical constructions is attempted.

Another aspect of the Giry monad is dealt with in [24]: the discussion in the present paper uses the Prohorov metric for a metric on the space of subprobability measures, but it is fairly possible to work with the Kantorovich metric; van Breugel shows that this metric is particularly suited to model probabilistic nondeterminism, because it is robust to small changes. It may be interesting to pin down specifically which effect the choice of the metric has for the Giry monad and its associated structures.

The investigation of Eilenberg–Moore algebras for the probability functor using convexity arguments has been advocated e.g., in [25,9,5], pioneering work having been reported in [26,4], see [3]. It is clearly intimately connected with the question of identifying the adjoints for this functor, which are not yet completely known in the category of Polish spaces with continuous or with measurable maps as morphisms.

8. Further Work

We characterize the algebras for the Giry monad which assigns each Polish space its space of probabilities. The morphisms in this category are continuous maps between Polish spaces. Continuity is technically crucial for the argumentation. This approach will not work for general Borel measurable maps serving as morphisms between Polish spaces (although these maps are fairly interesting from the point of view of applications), thus a more general characterization for these algebras is desirable.

Continuity plays also a crucial role in some of the examples that are discussed. Through the geometric argument of connectedness we could show that for the discrete case no algebras exist, except in the very trivial case of a one point space. This argument also does work only when the morphisms involved are continuous. So it is desirable to find algebras for the general case of Borel maps over finite domains (probably they do not exist usually there either: one would also like to know that). The last example hints at a connection between these algebras and barycenters for compact convex sets in topological vector spaces. It ends here where the fun begins there, viz., when looking at Choquet's theory of integral representations [27]. There is room for further work exploring this avenue. The examples show that the world of algebras for this monad is quite colorfully polymorphous.

The most interesting question, however, addresses the expansion of the characterization given here for Borel measurable maps which are based on Polish spaces, or, going one crucial step further, on analytic ones. This goes hand in hand with the request for identifying adjoints for the probability functor (or its close cousin, the subprobability functor).

Acknowledgments

Dieter Pumplün's suggestions to investigate convex structures, and his general advice on algebras are highly appreciated. The referees provided very careful and detailed comments as well as constructive suggestions, for which I am very grateful.

References

[1] E. Moggi, Notions of computation and monads, Inform. Comput. 93 (1991) 55–92.
[2] S. MacLane, Categories for the Working Mathematician, Graduate Texts in Mathematics, Springer-Verlag, Berlin, 1997.
[3] V.V. Fedorchuk, Probability measures in topology, Russ. Math. Surv. (1991) 45–93 (Uspekhi Mat. Nauk 46:1 (1991), 41–80).
[4] T. Swirszcz, Monadic functors and convexity, Bull. Acad. Pol. Sci. Ser. Sci. Math. Astronom. Phys. 22 (1974) 39–42.
[5] O.R. Nykyforchyn, Semiconvex compacta, Comment. Math. Univ. Carolinae 38 (3) (1997) 761–774.
[6] K.R. Parthasarathy, Probability Measures on Metric Spaces, Academic Press, New York, 1967.
[7] P. Billingsley, Probability and Measure, John Wiley and Sons, New York, 1995.
[8] M. Giry, A categorical approach to probability theory, in: Categorical Aspects of Topology and Analysis, No. 915 in Lect. Notes Math., Springer-Verlag, Berlin, 1981, pp. 68–85.
[9] D. Pumplün, Positively convex modules and ordered normed linear spaces, J. Convex Anal. 10 (1) (2003) 109–127.
[10] S.M. Srivastava, A Course on Borel Sets, Graduate Texts in Mathematics, Springer-Verlag, Berlin, 1998.
[11] J. Desharnais, A. Edalat, P. Panangaden, Bisimulation of labelled Markov-processes, Inform. Comput. 179 (2) (2002) 163–193.
[12] E.-E. Doberkat, Stochastic relations: congruences, bisimulations and the Hennessy–Milner theorem, SIAM J. Comput. 35 (3) (2006) 590–626.
[13] C. Jones, G. Plotkin, A probabilistic powerdomain of evaluations, in: Proc. LICS'89, IEEE Computer Society Press, 1989, pp. 186–195.
[14] R. Engelking, General Topology, Revised and Completed Edition, vol. 6 of Sigma Series in Pure Mathematics, Heldermann-Verlag, Berlin, 1989.
[15] D.H. Fremlin, Measure Theory—vol. 4: Topological Measure Spaces, Torres Fremlin, Colchester, 2003.

[16] C. Morgan, A. McIver, K. Seidel, Probabilistic predicate transformers, ACM Trans. Prog. Lang. Syst. 18 (3) (1996) 325–353.
[17] R. Heckmann, Probabilistic domains, in: S. Tison (Ed.), Proc. 19th Int. Colloquium on Trees in Algebra and Programming, No. 787 in Lect. Notes Comp. Sci., Springer-Verlag, Berlin, 1994, pp. 142–156.
[18] P. Panangaden, Probabilistic relations, in: C. Baier, M. Huth, M. Kwiatkowska, M. Ryan (Eds.), Proc. PROBMIV, 1998, pp. 59–74.
[19] S. Abramsky, R. Blute, P. Panangaden, Nuclear and trace ideal in tensored *-categories, J. Pure Appl. Algebra 143 (1–3) (1999) 3–47.
[20] E.-E. Doberkat, Pipelines: Modelling a software architecture through relations, Acta Inform. 40 (2003) 37–79.
[21] E.-E. Doberkat, Semi-pullbacks and bisimulations in categories of stochastic relations, in: Proc. ICALP'03, vol. 2719 of Lect. Notes Comp. Sci., Springer-Verlag, Berlin, 2003, pp. 996–1007.
[22] E.-E. Doberkat, The Hennessy–Milner equivalence for continuous-times stochastic logic with mu-operator, J. Appl. Logic (2006) (in print).
[23] C. Jones, Probabilistic non-determinism, Ph.D. thesis, Univ. Edinburgh, 1989.
[24] F. van Breugel, Robust Models of Probabilistic Nondeterminism, Tech. Rep., Department of Computer Science, York University, Toronto, January 2005.
[25] D. Pumplün, Eilenberg–Moore algebras revisited, Tech. Rep. 29, Fachbereich Mathematik und Informatik, Fern-Universität Hagen, 1988.
[26] Z. Semadeni, Monads and their Eilenberg–Moore algebras in functional analysis, Queen's Papers Pure Appl. Math. 33.
[27] B. Fuchssteiner, W. Lusky, Convex Cones, vol. 56 of North-Holland Mathematical Studies, North-Holland Publishing Company, Amsterdam, 1981.