The Giry monad as a codensity monad

Kirk Sturtz

July 16, 2014

Abstract

We show the Giry monad is naturally isomorphic to a submonad of a double dualization monad on the category of measurable spaces. This shows a probability measure is equivalent to a weakly averaging affine functional taking values in the unit interval. This submonad arises canonically as the density monad of a functor from the category of convex spaces to the category of measurable spaces.

Keywords: Giry monad, categorical probability, codensity monad, convex spaces, double dualization monad

1 Introduction

In the paper Codensity and the ultrafilter monad [Leinster, 2013] the relationship between ultrafilters on a set $X$ and 2-generalized elements of $X$ was expounded. This relationship clarified that the ultrafilter monad $\mathcal{U}$ on the category of sets, $\text{Set}$, is a subfunctor of the double dualization monad $\text{Set}(2^\bullet, 2)$ where we use the exponent notation, $2^X = \text{Set}(X, 2)$ for every object $X$ in $\text{Set}$, for brevity. In this situation an ultrafilter $F \in \mathcal{U}(X)$ can be seen as a functional mapping parts of $X$ into 2 which satisfies the condition of being a finitely additive probability measure. Thus an ultrafilter is a primitive sort of probability measure which is “deterministic” as it assumes only the values 0 or 1. This construction can be generalized so that the primitive sort of probability measure becomes a true probability measure assuming values in the interval $I = [0, 1]$. In this introductory section we outline this procedure.

By replacing the category of sets with the category of measurable spaces $\text{Meas}$ and the object 2 by the measurable space $I$, endowed with the Borel $\sigma$-algebra generated by the open intervals, one obtains an analogous situation, $\mathcal{P} \hookrightarrow \text{Meas}(I^\bullet, I)$, of a submonad $\mathcal{P}$ of the double dualization monad $\text{Meas}(I^\bullet, I)$. Here $I^X = \text{Meas}(X, I)$ for all measurable spaces $X$. To define this (sub)monad $\mathcal{P}$ componentwise we first recall that a functional $G : I^X \to I$ is weakly averaging when, for all constant functions $\pi \in I^X$ with value $u = \pi(x) \in I$ for all $x \in X$, $G$ satisfies $G(\pi) = u$. Also, as $I$ has a natural convex structure defined by $u + \alpha v = \alpha u + (1 - \alpha)v$ for all $\alpha \in I$ the function space $I^X$ has a convex structure defined on it pointwise. Thus for $f, g \in I^X$ the “convex sum” $f +_\alpha g$ is
defined pointwise for each $x \in X$ by $f(x) +_{\alpha} g(x)$ making $f +_{\alpha} g \in I^X$. The functional $G : I^X \to I$ is affine\footnote{The set of all affine morphisms $\text{Meas}(I^X, I)$ is also an object in the category of convex spaces, $\text{Cvx}$, where the morphisms have been called affine, affine linear, convex linear, as well as convex.} when $G(f +_{\alpha} g) = G(f) +_{\alpha} G(g)$ for all $f, g \in I^X$ and all $\alpha \in I$. The monad $\mathcal{P}$ is defined on component $X$ by

$$\mathcal{P}(X) = \{ I^X \xrightarrow{G} I \mid G \text{ is weakly averaging and affine} \}.$$ 

We show this monad $\mathcal{P}$ is naturally isomorphic to the Giry monad $\mathcal{G}$. For $\text{Cvx}$ the category of convex spaces and affine morphisms, we construct a functor $\iota : \text{Cvx} \to \text{Meas}$ and show that the right Kan extension of $\iota$ along $\iota$ is the monad $\mathcal{P}$. This functor $\iota$ is itself a subfunctor of a double dualization functor using the unit interval $I$ which is an object in both of the categories, $\text{Cvx}$ and $\text{Meas}$.

While the monad $\mathcal{P}$ (or $\mathcal{G}$) can be viewed as a functor $\mathcal{P} : \text{Meas} \to \text{Cvx}$, because $\mathcal{P}(X)$ has a natural convex structure, obtaining a right adjoint $\iota : \text{Cvx} \to \text{Meas}$ to $\mathcal{P}$ would require a “barycenter natural transformation” $\epsilon : \mathcal{P} \circ \iota \to \text{Id}_{\text{Cvx}}$ as the counit. Presently no such construction is known or even whether such a construction is possible. However if we consider the situation of $\mathcal{P}$ as a codensity monad of a functor $\iota : \text{Cvx} \to \text{Meas}$ then the problem reduces to finding a natural transformation $\epsilon : \mathcal{P} \circ \iota \to \iota$ in $\text{Meas}$ which we show does exist and follows innately when viewed in terms of hom sets. As a consequence of this construction $\mathcal{P}$ is the right Kan extension of $\iota$ along itself. This is equivalent to saying the codensity monad of $\iota$ is $\mathcal{P}$.

Without imposing additional restrictions on the measurable spaces, such as requiring Polish Spaces \cite{Doberkat,2004}, the further characterization of the Giry monad beyond the standard monad characterizations has not been established. The possibility that the Giry monad arises canonically as a codensity monad appears first in the work of Leinster who specifically poses this question. Once the functor $\iota : \text{Cvx} \to \text{Meas}$ is established the fact that $\mathcal{P}$ is the codensity monad of this functor follows using the fact $\text{Meas}$ is complete and has a SMCC structure.

This paper is organized to sequentially show

(i) Both $\text{Cvx}$ and $\text{Meas}$ are symmetric monoidal closed categories (SMCC).
(ii) For every measurable space $X$ the set of weakly averaging affine morphisms $\mathcal{P}(X)$ is isomorphic to $\mathcal{G}(X)$ as convex spaces.
(iii) This isomorphism of convex spaces extends to an isomorphism of monads, $\mathcal{G} \cong \mathcal{P}$.
(iv) There exist a functor $\iota : \text{Cvx} \to \text{Meas}$.
(v) The monad $\mathcal{P}$ is the right Kan extension of $\iota$ along $\iota$.

An appendix provides additional information of the relationship between filters, ultrafilters, and weakly averaging affine functionals, and shows that proper filters $\mathcal{F}$ on a set $X$ correspond bijectively to weakly averaging affine functionals $\hat{\mathcal{F}} \in \text{Set}(2^X,2)$. This suggest a weakly averaging affine functional $G \in \mathcal{P}(X) \subset \text{Meas}(I^X, I)$, which bijectively corresponds to probability measure on $X$, can be seen as a generalization of a proper filter on a measurable space $X$ which, loosely speaking, one might call an $I$-valued proper filter on $X$. 

\footnote{The set of all affine morphisms $\text{Meas}(I^X, I)$ is also an object in the category of convex spaces, $\text{Cvx}$, where the morphisms have been called affine, affine linear, convex linear, as well as convex.}
Notation  Unless specifically defined otherwise, the symbols \( X \) and \( Y \) always denote measurable spaces while \( A \) and \( B \) always denote convex spaces. The symbol \( I \) denoting the unit interval is of course both a measurable space and a convex space. In the last section where a slice category is also used, in addition to the categories \( \text{Meas} \) and \( \text{Cvx} \), it is convenient to use the notation “\( X \in \text{ob} \ C \)” to denote an object in the category \( C \) and “\( f \in \text{ar} \ C \)” to denote an arrow in the category \( C \). For an object \( X \) in any category the identity arrow on \( X \) is denoted \( \text{id}_X \). The notation \( \mathbf{u} \) is used to denote a constant function with value \( u \) lying in the codomain of the function \( \mathbf{u} \).

2  The categories of interest

The two main categories of interest are \( \text{Meas} \) and \( \text{Cvx} \). While most of the categorical properties of \( \text{Meas} \) are well known the fact that \( \text{Meas} \) is a SMCC is apparently not well known and hence we give an overview of this fact.\(^2\) We first provide a brief summary of \( \text{Cvx} \) which is also a SMCC [Meng, 1987], and provide a brief overview of that construction. A more detailed description of the category of convex spaces can be found in [Fritz, 2009] who provides definitions with numerous examples and highlights the difference between geometric and combinatorial convex spaces.

2.1  The category of convex spaces

A convex space \((A, +)\) consist of a set \( A \) and a function

\[
A \times A \times I \longrightarrow A \\
(a_1, a_2, r) \mapsto a_1 +_r a_2
\]

satisfying the following axioms

1. \( a_1 +_0 a_2 = a_2 \)
2. \( a +_0 a = a \)
3. \( a_1 +_r a_2 = a_2 +_1 -r a_1 \)
4. \( (a_1 +_p a_2) +_q a_3 = a_1 +_p (a_2 +_r a_3) \) for 
   \[
   r = \begin{cases} 
   \frac{(1-p)q}{1-(pq)} & \text{if } pq \neq 1 \\
   \text{arbitrary} & \text{if } p = q = 1
   \end{cases}
   \]

The convex structure of the convex space \( I \) is defined by \( a_1 +_r a_2 = ra_1 + (1-r)a_2 \) for all \( a_1, a_2, r \in I \).

An affine morphism of convex spaces \( f : (A, +) \rightarrow (B, \oplus) \) satisfies

\[
f(a_1 +_r a_2) = f(a_1) \oplus_r f(a_2).
\]

These objects and morphisms determine the category of convex spaces \( \text{Cvx} \). If \( A \) and \( B \) are convex spaces we denote the set of all affine morphisms from \( A \) to \( B \) by \( \text{Cvx}(A, B) \).

\(^2\)We are not aware of this fact in the literature though it would be surprising that it is not known as its construction is similar to that used in topology.
2.2 The symmetric monoidal closed structure of \textbf{Cvx}

The unit of the SMCC structure on \textbf{Cvx} is the object 1 = \{\star\} with the only possible convex structure. The construction of the tensor product and function spaces in \textbf{Cvx} is virtually identical to the construction employed in the category \textbf{R-Mod}. Hence we limit ourself to reminding the reader of the basic construction.

The tensor product of two convex spaces $A \otimes B$ is obtained by taking the free convex structure on $A \times B$ and then taking the smallest congruence relation on this set such that
\[
\sum_{i=1}^{n} \alpha_i(a_i, b) \equiv (\sum_{i=1}^{n} \alpha_i a_i, b), \quad \text{and} \quad \sum_{i=1}^{n} \alpha_i(a, b_i) \equiv (a, \sum_{i=1}^{n} \alpha_i b_i).
\]

This tensor product $A \otimes B$ is universal in the sense that if $C$ is any convex space and $f : A \times B \to C$ is a bi-affine function (affine in each variable), then there exist a unique affine morphism $\hat{f}$ such that the diagram
\[
\begin{array}{ccc}
A \times B & \longrightarrow & A \otimes B \\
\downarrow f & & \downarrow \hat{f} \\
C & \longleftarrow &
\end{array}
\]
commutes.

The function space $B^A = \text{Cvx}(A, B)$ is defined pointwise. If $f, g \in \text{Cvx}(A, B)$ then $(f + \alpha g)(a) = f(a) + \alpha g(a)$.

Using these definitions the defining property of a closed monoidal category, $- \otimes B \vdash -$ for all convex spaces $B$, follows. The symmetry follows from the construction of the tensor product $\otimes$.

2.3 The symmetric monoidal closed structure of \textbf{Meas}

Throughout this section, $X$ and $Y$ denote measurable spaces. The category \textbf{Meas} is a SMCC with the tensor product $X \otimes Y$ defined by the coinduced (final) \(\sigma\)-algebra such that all the graph functions
\[
\Gamma_f : X \longrightarrow X \times Y \\
: x \mapsto (x, f(x))
\]
for $f : X \to Y$ a measurable function, as well as the graph functions
\[
\Gamma_g : Y \longrightarrow X \times Y \\
: y \mapsto (g(y), y)
\]
for $g : Y \to X$ a measurable function, are measurable.
Let $Y^X$ denote the set of all measurable functions from $X$ to $Y$ endowed with the $\sigma$-algebra induced by the set of all point evaluation maps

$$\begin{align*}
Y^X & \xrightarrow{ev_x} Y \\
\Gamma f \upharpoonright & \mapsto f(x)
\end{align*}$$

where the notation $\Gamma f \upharpoonright$ is used to distinguish between the measurable function $f : X \to Y$ and the point $\Gamma f \upharpoonright : 1 \to Y^X$ of the function space $Y^X$. After showing the SMCC structure we drop the distinction as it is common practice to let the context define which arrow we are referring to.

Because the $\sigma$-algebra structure on tensor product spaces is defined such that the graph functions are all measurable, it follows in particular the constant graph functions $\Gamma_x : X \to X \otimes Y^X$ sending $x \mapsto (x, \Gamma f \upharpoonright)$ are measurable.

Define the evaluation function

$$X \otimes Y^X \xrightarrow{ev_{X,Y}} Y$$

and observe that for every $\Gamma f \upharpoonright \in Y^X$ the right hand diagram in the $\text{Meas}$ diagrams

$$\begin{align*}
\begin{array}{ccc}
Y^X & \xrightarrow{ev_{X,Y}} & Y \\
\Gamma f \upharpoonright & \simeq & \text{Id}_X \otimes \Gamma f \upharpoonright \\
1 & \xrightarrow{\Gamma_x} & X \simeq X \otimes 1 \\
\end{array}
\end{align*}$$

is commutative as a set mapping, $f = ev_{X,Y} \circ \Gamma f \upharpoonright$. By rotating the above diagram and also considering the constant graph functions $\Gamma_x$ the right hand side of the diagram

$$\begin{align*}
\begin{array}{ccc}
X & \xrightarrow{\Gamma f \upharpoonright} & X \otimes Y^X \\
& \xrightarrow{\Gamma_x} & Y^X \\
& \xleftarrow{ev_{X,Y}} & \\
& \xleftarrow{ev_x} & Y
\end{array}
\end{align*}$$

also commutes for every $x \in X$. Since $f$ and $\Gamma f \upharpoonright$ are measurable, as are $ev_x$ and $\Gamma_x$, it follows by the elementary result on coinduced $\sigma$-algebras

**Lemma 2.1.** Let the $\sigma$-algebra of $Y$ be coinduced by a collection of maps $\{f_i : X_i \to Y\}_{i \in I}$. Then a function $g : Y \to Z$ is measurable if and only if the composition $g \circ f_i$ is measurable for each $i \in I$.

that $ev_{X,Y}$ is measurable because the graph functions generate the $\sigma$-algebra of $X \otimes Y^X$.

More generally, given any measurable function $f : X \otimes Z \to Y$ there exists a unique measurable map $\tilde{f} : Z \to Y^X$ defined by $\tilde{f}(z) = \Gamma f(\cdot, z) : 1 \to Y^X$ where $f(\cdot, z) : X \to Y$
sends \( x \mapsto f(x, z) \). This map \( \tilde{f} \) is measurable because the \( \sigma \)-algebra is generated by the point evaluation maps \( ev_x \) and the diagram

\[
\begin{array}{ccc}
Y^X & \xrightarrow{ev_x} & Y \\
\downarrow{\tilde{f}} & & \downarrow{f} \\
Z & \xrightarrow{\Gamma_x} & X \otimes Z
\end{array}
\]

commutes so that \( \tilde{f}^{-1}(ev_x^{-1}(B)) = (f \circ \Gamma_x)^{-1}(B) \in \Sigma_Z \).

Conversely given any measurable map \( g : Z \to Y^X \) it follows the composite \( ev_{X,Y} \circ (Id_X \otimes g) \) is a measurable map. This determines a bijective correspondence

\[
\text{Meas}(X \otimes Z, Y) \cong \text{Meas}(X, Y^Z).
\]

### Double dual mappings

Recall \( I = [0, 1] \) with the Borel \( \sigma \)-algebra generated by the open intervals.

**Lemma 2.2.** Given any measurable space \( X \) the double dual mapping

\[
\begin{array}{ccc}
X & \xrightarrow{\eta_X} & I^{I^X} \\
\downarrow{x} & & \downarrow{ev_x} \\
I^X & \xrightarrow{I} & I
\end{array}
\]

is a measurable function.

**Proof.** Since the functions \( \{ ev_f \}_{f \in \text{Meas}(X, I)} \) generate \( \Sigma_{I^{I^X}} \) it suffices to show that \( \eta_X^{-1}(ev_f^{-1}(U)) \in \Sigma_X \) for \( U \in \Sigma_I \). But this set is just \( f^{-1}(U) \) which is measurable since \( f \) is measurable. \( \square \)

### 3 The submonad of the double dualization monad

Using the SMCC category structure we have the double dualization monad \( \text{Meas}(I^\bullet, I) \) on \( \text{Meas} \) specified by

\[
\begin{array}{ccc}
\text{Meas}(I^\bullet, I) & :_{ob} & X \to \text{Meas}(I^X, I) \\
:_{ar} & \begin{array}{c} \xrightarrow{f} \end{array} & \begin{array}{c} \xrightarrow{\text{Meas}(I^I, I)} \end{array} \text{Meas}(I^Y, I)
\end{array}
\]

where \( \text{Meas}(I^I, I)(G) = G \circ I^f \) is the pushforward of \( G \) by \( f \),

\(^3\)In this diagram and those to follow we abuse notation following the doctrine of expressing the mapping into a function space not as the name of an element, like \( ev_x \in I^{I^X} \) for the given map \( \eta_X(x) \), but rather as the morphism corresponding to the named element. The dashed arrow notation is employed to make it easier to read given the multiple arrows involved.
The double dualization monad, similar to any double dualization monad on a SMCC, has the unit \( \eta \) and counit \( \mu \) given componentwise by

\[
\begin{array}{ccc}
X \xrightarrow{\eta_X} \text{Meas}(I^X, I) & \quad & \text{Meas}(I^{\text{Meas}(I^X, I)}, I) \xrightarrow{\mu_X} \text{Meas}(I^X, I) \\
x \mapsto I^X & \quad & \mu_X(Q) \mapsto I^X
\end{array}
\]

\[
\begin{array}{ccc}
Q \xrightarrow{\mu_X(Q)} I & \quad & \mu_X(Q) \mapsto Q(ev_f)
\end{array}
\]

For any \( y \in Y \) let \( \overline{y} : X \to Y \) denote the constant map with value \( y \). A map \( G : Y^X \to Y \), in any category with a terminal object and function spaces, is called **weakly averaging** if it satisfies the condition \( G(\overline{y}) = y \) for all \( y \in Y \) \(^4\). Because \( I \) has a convex structure the space \( I^X \) of measurable functions has a convex structure associated with it defined pointwise by \((f + \alpha g)(x) = f(x) + \alpha g(x)\) for all \( f, g \in I^X \).

Define the submonad \( \mathcal{P} \hookrightarrow \text{Meas}(I^\bullet, I) \) componentwise by

\[
\mathcal{P}(X) = \{G \in \text{Meas}(I^X, I) \mid G \text{ is affine and weakly averaging}\}.
\]

This submonad has the same unit (with the reduced codomain) and same counit (with reduced domain and codomain) as \( \text{Meas}(I^\bullet, I) \) because for any measurable space \( X \) all the evaluation maps are affine and weakly averaging by the pointwise convex structure on \( I^X \).

### 4 Probability measures as weakly averaging affine functionals

Throughout this section, as well as subsequent sections, let \( X \) denote a measurable space. For any subset \( S \) of \( X \) we denote its complement by \( S^c \).

**Lemma 4.1.** For \( G \in \mathcal{P}(X) \) and \( \chi_S, \chi_T : X \to I \) the characteristic functions with \( S, T \in \Sigma_X \) it follows

---

\(^4\)This terminology is taken from *Sets for Mathematics* [Lawvere and Rosebrugh, 2005] who specifically address the double dualization process and subfunctors thereof into objects with extra structure.
(i) $G(\chi_X) = 1$ and $G(\chi_\emptyset) = 0$

(ii) $G(\chi_{S^c}) = 1 - G(\chi_S)$

(iii) $G(\chi_{S \cap T}) + G(\chi_{S \cup T}) = G(\chi_S) + G(\chi_T)$

(iv) If $S \subseteq T$ then $G(\chi_S) \leq G(\chi_T)$

(v) If $\{S_i\}_{i=1}^\infty$ is a disjoint cover of $S$ by measurable sets then

$$G(\chi_S) = \lim_{N \to \infty} \{\sum_{i=1}^N G(\chi_{S_i})\}$$

(vi) For any $\alpha \in I$, $G(\alpha f) = \alpha G(f)$

Proof. (i) Since $\chi_X$ and $\chi_\emptyset$ are constant functions the result follows from the weakly
averaging condition. (ii) Consider the constant function

$$\frac{1}{2} \chi_S + \frac{1}{2} \chi_{S^c} : X \to I.$$ Since $G \in \mathcal{P}(X)$ it follows

$$\frac{1}{2} = G\left(\frac{1}{2} \chi_S + \frac{1}{2} \chi_{S^c}\right) = \frac{1}{2} (G(\chi_S) + G(\chi_{S^c}))$$

which implies $G(\chi_S) + G(\chi_{S^c}) = 1$ and hence the result. (iii) This is a consequence of the
observation that for all $S, T \in \Sigma_X$ the equation

$$\frac{1}{2} \chi_{S \cap T} + \frac{1}{2} \chi_{S \cup T} = \frac{1}{2} \chi_S + \frac{1}{2} \chi_T$$

holds and both the left and right terms are measurable functions $X \to I$. Applying $G$
to both sides of this expression gives the result. (iv) Apply the weakly averaging affine
morphism $G$ to both sides of the equation

$$\frac{1}{2} \chi_T + \frac{1}{2} \chi_\emptyset = \frac{1}{2} \chi_S + \frac{1}{2} \chi_{T \cap S^c}$$

and use the condition $G(\chi_{T \cap S^c}) \geq 0$. (v) By part (iii) $G(\chi_{S_1 \cup S_2}) = G(\chi_{S_1}) + G(\chi_{S_2})$ since
the $S_i$ are disjoint. Iterating this gives

$$G(\chi_{\bigcup_{i=1}^N S_i}) = \sum_{i=1}^N G(\chi_{S_i})$$

and because $\bigcup_{i=1}^N S_i \subset S$ by (iv) it follows $\sum_{i=1}^N G(\chi_{S_i}) \leq G(\chi_S)$. Because $\{\sum_{i=1}^N G(\chi_{S_i})\}_{N=1}^\infty$
is a monotone increasing sequence bounded above by $G(\chi_S)$ the result follows. (vi) If $\alpha \in I$
then for $f : X \to I$ it follows that $\alpha f + (1 - \alpha)\chi_{\emptyset} : X \to I$ so using the fact $\chi_{\emptyset}$ is the additive identity element of $I^X$

$$G(\alpha f) = G(\alpha f + (1 - \alpha)\chi_{\emptyset}) = \alpha G(f) + (1 - \alpha)G(\chi_{\emptyset}) = \alpha G(f).$$

\[\square\]

**Lemma 4.2.** Every simple measurable function $f : X \to I$ can be written as a convex sum, $f = \sum_{i=1}^n a_i \chi_{S_i}$ with $\sum_{i=1}^n a_i = 1$.

**Proof.** We can assume the simple measurable function $f = \sum_{i=1}^n a_i \chi_{S_i}$ is written with pairwise disjoint measurable sets $\{S_i\}_{i=1}^n$ and has increasing coefficients, $a_1 \leq a_2, \ldots \leq a_n$.

This sum can be rewritten as the “telescoping” function

$$f = a_1\chi_{\bigcup_{i=1}^n S_i} + (a_2 - a_1)\chi_{\bigcup_{i=2}^n S_i} + \ldots + (a_j - a_{j-1})\chi_{\bigcup_{i=j}^n S_i} + \ldots + (a_n - a_{n-1})\chi_{S_n} + (1 - a_n)\chi_{\emptyset}$$

which satisfies the condition that the sum of the coefficients is one. \[\square\]

As every measurable function $f : X \to I$ can be written as the limit of a sequence of simple functions, $f = \lim_{M \to \infty} \{f_N\}_{N=1}^M$ with $0 \leq f_N \leq f$ for all $N$, it follows for $\tilde{f}_N$ the simple function $f_N$ expressed as a convex sum using Lemma 4.2 that $G(f) = \lim_{M \to \infty} \{G(\tilde{f}_N)\}_{N=1}^M$.

**Lemma 4.3.** There exist an isomorphism of convex spaces

$$\mathcal{P}(X) \xrightarrow{\phi} \mathcal{G}(X)$$

$$G \quad \mapsto \quad \nu_G$$

where $\nu_G(S) = G(\chi_S)$ for all $S \in \Sigma_X$.

**Proof.** The verification that $\nu_G$ defines a probability measure follows directly from the definition of $\nu_G$ in terms of $G$ and the characteristic functions by applying Lemma 4.1.

The inverse of $\phi$ maps $P \in \mathcal{G}(X)$ to $\hat{P} : I^X \to I$ which is defined on the characteristic functions of the space $I^X$ by $\hat{P}(\chi_S) = P(S)$. Since every measurable function $f : X \to I$ is a limit of sequence of monotone increasing simple functions $\{f_N\}_{N=1}^\infty$ define

$$\hat{P}(f) = \lim_{N \to \infty} \{\hat{P}(\tilde{f}_N)\}$$

where, for $\tilde{f}_N$ the simple function $f_N = \sum_{i=1}^{m(N)} \alpha_{i,N} \chi_{S_{i,N}}$ expressed as a convex sum with the integer $m(N)$ depending upon $N$, $\hat{P}(\tilde{f}_N) = \sum_{i=1}^{m(N)} \alpha_{i,N} \hat{P}(\chi_{S_{i,N}})$. By the construction $\hat{P}$ is a weakly averaging affine morphism.

These two constructions are inverse to each other. The isomorphism is affine because the convex structures on both convex spaces are defined pointwise. \[\square\]

**Theorem 4.4.** The isomorphism of convex spaces in Lemma 4.3 extends to a natural isomorphism of monads $\phi : \mathcal{P} \to \mathcal{G}$. 
Proof. First we show that $\phi_X : \mathcal{P}(X) \to \mathcal{G}(X)$ is an isomorphism of measurable spaces which requires showing $\mathcal{P}(X)$ with its subspace $\sigma$-algebra is isomorphic to the $\sigma$-algebra on $\mathcal{G}(X)$. Recall that the Giry monad is endowed with the smallest $\sigma$-algebra such that each of the evaluation maps $ev_S : \mathcal{G}(X) \to I$ sending a probability measure $P \mapsto P(S)$ is measurable, for every measurable set $S$ in $X$. On the other hand the function space $I^X$ has the smallest $\sigma$-algebra such that each of the evaluation maps $ev_f : I^X \to I$ are measurable for every measurable function $f : X \to I$, so for $U \in \mathcal{B}_I$ it follows the set

$$ev_f^{-1}(U) = \{ I^X \xrightarrow{G} I \mid G(f) \in U \}$$

is measurable in $I^X$ and sets of this form, as $f$ varies over $I^X$ and $U$ varies over $\mathcal{B}_I$ form a generating set for the $\sigma$-algebra on $I^X$. Being more economical it suffices to take the generating set on the characteristic functions $f = \chi_S$ for all $S \in \Sigma_X$. Restriction of the $\sigma$-algebra generated by these elements $\{ ev_{\chi_S}^{-1}(U) \}_{S \in \Sigma_X, U \in \Sigma_I}$ to the subset $\mathcal{P}(X)$ gives the $\sigma$-algebra on $\mathcal{P}(X)$. Under the mapping $\phi$ the generating set elements in (1) get mapped to the subsets of $\mathcal{G}(X)$ corresponding to the preimage of the diagonal map in the diagram

$$\begin{array}{ccc}
\mathcal{P}(X) & \xrightarrow{ev_{\chi_S}} & I \\
\phi_X \downarrow & & \downarrow ev_S \\
\mathcal{G}(X) & \xrightarrow{\int_X \chi_S d_\mu} & ev_G
\end{array}$$

which are the generating elements for the $\sigma$-algebra of $\mathcal{G}(X)$. The converse then follows similarly mapping the generating elements of $\mathcal{G}(X)$ to the generating elements of $\mathcal{P}(X)$.

For $f : X \to Y$ the measurable function $\mathcal{P}(f)$ is just the pushforward map shown in Diagram 1, restricted to appropriate domain and codomain, which coincides with the Giry monad definition because for all measurable $S \in \Sigma_Y$

$$(G \circ I^f)(\chi_S) = G(\chi_{f^{-1}(S)}) \quad \text{where} \quad (G \circ I^f)(\chi_S) = G(f^{-1}(S))$$

The map $\phi$ is a natural transformation as the Meas diagram

$$\begin{array}{ccc}
X & \xrightarrow{\mathcal{P}(f)} & \mathcal{G}(Y) \\
\phi_X \downarrow & & \phi_Y \\
\mathcal{P}(X) & \xrightarrow{\phi_X} & \mathcal{G}(X) \\
\downarrow f & & \downarrow \mathcal{G}(f) \\
Y & \xrightarrow{\mathcal{P}(Y)} & \mathcal{G}(Y)
\end{array}$$

Diagram 2. The naturality of $\phi$. 


commutes because for all $G \in \mathcal{P}(X)$

$$G(f)(\phi_X(G)) = G(f)(\nu_G) = \nu_G f^{-1} = \phi_Y(G \circ I^f) \quad \text{using (2)}$$

This natural transformation has the inverse natural transformation specified in Theorem 4.3.

The natural isomorphism $\phi : \mathcal{P} \rightarrow \mathcal{G}$ is a morphism of monads as it makes the two requisite diagrams commute. Recalling the unit of the Giry monad is defined by $\eta'_X(x) = \delta_x$ while the counit is specified by $\mu'_X(Q)(S) = \int_{q \in \mathcal{G}(X)} q(S) dQ$ for all $S \in \Sigma_X$, the commutativity of the left diagram follows from

$$(\phi_X(\eta_X(x))) (S) = (\phi_X(ev_x)) (S) = \nu_{ev_x}(S) = ev_x(\chi_S) = \delta_x(S) = \eta'_X(x)(S)$$

while the commutativity of the right diagram follows from the east-south path giving

$$(\phi_X(\mu_X(Q))) (S) = (\phi_X(\mu_X(Q))) (S) = \nu_{\mu_X(Q)}(S) = \mu_X(Q)(\chi_S) = Q(ev_{\chi_S})$$

while the south-east path yields the same value because

$$(\mu'_X \circ (\phi \circ \phi)_{\mathcal{P}(X)})(S) = \mu'_X(\phi_g(X)(Q \circ I^{\phi_X}))(S) = \mu'_X(\nu_{QoI^{\phi_X}})(S) = \int_{q \in \mathcal{G}(X)} ev_S(q) d
\nu_{QoI^{\phi_X}} = \int_{p \in \mathcal{P}(X)} (ev_S \circ \phi_X)(p) d\nu_Q = Q(ev_{\chi_S})$$
5 Constructing the functor $\iota : \text{Cvx} \to \text{Meas}$

For $A$ a convex space endow the set of functions $I^A = \text{Set}(A, I)$ with the $\sigma$-algebra generated by all the evaluation maps $ev_a : I^A \to I$ mapping $h \mapsto h(a)$ for every $a \in A$ and every function $h : A \to I$. Give $\text{Cvx}(A, I) \subset I^A$ the subspace $\sigma$-algebra. Let $\text{Cvx}_w(A, I)$ denote the set of weakly averaging affine functions from $A$ to $I$ and endow $\text{Cvx}_w(\text{Cvx}(A, I), I)$ with the $\sigma$-algebra generated by the evaluation maps

$$\text{Cvx}_w(\text{Cvx}(A, I), I) \xrightarrow{ev_h} I$$

for all $h \in \text{Cvx}(A, I)$\footnote{The set $\text{Cvx}(A, I)$ is nonempty because every constant function is affine.} For $k : A \to B$ an affine morphism of convex spaces

$$\text{Cvx}_w(\text{Cvx}(A, I), I) \xrightarrow{\text{Cvx}_w(\text{Cvx}(k, I), I)} \text{Cvx}_w(\text{Cvx}(B, I), I) \xrightarrow{G \circ I^k}$$

where $I^k : I^B \to I^A$ is the map defined on all $g \in I^B$ by $I^k(g) = g \circ k$. It is easy to verify

$$\iota(\bullet) \overset{\text{def}}{=} \text{Cvx}_w(\text{Cvx}(\bullet, I), I) : \text{Cvx} \to \text{Meas}$$

is functorial.

**Lemma 5.1.** For $k : A \to B$ and $g : B \to I$ affine morphisms in $\text{Cvx}$ the $\text{Meas}$ diagram

![Diagram](image)

commutes.

**Proof.** For all $K \in \iota(A)$

$$ev_{g \circ k}(K) = K(g \circ k) = (K \circ I^k)(g) = ev_g(K \circ I^k) = (ev_g \circ \iota(k))(K)$$
6 The Giry monad as a codensity monad

The functor \( \iota : \text{Cvx} \to \text{Meas} \) induces the functor

\[
\begin{align*}
\text{Meas}^\iota : \text{Meas} & \to \text{Meas}^\iota \to \text{Meas}^{\text{Cvx}} \\
\text{ob} & : F \mapsto F \circ \iota \\
\text{ar} & : F \xrightarrow{\alpha} G \mapsto F \circ \iota \xrightarrow{\alpha \iota} G \circ \iota 
\end{align*}
\]

and a universal arrow from the functor \( \text{Meas}^\iota \) to the object \( \iota \in \text{ob} \) \( \text{Meas}^{\text{Cvx}} \) is called the right Kan extension of \( \iota \) along \( \iota \), or more succinctly, the codensity monad of \( \iota \). Like any universal arrow the right Kan extension is a pair \((R^\iota, \epsilon)\) where \( R^\iota \in \text{ob} \) \( \text{Meas}^{\text{Meas}} \) and \( \epsilon : R^\iota \circ \iota \Rightarrow \iota \) is the universal arrow such that if \( \alpha : S \circ \iota \Rightarrow \iota \) then there exist a unique adjunct \( \overline{\alpha} : S \Rightarrow R^\iota \) such that the diagram on the right in

\[
\begin{array}{ccc}
R^\iota & \longrightarrow & \iota \\
\downarrow \overline{\alpha} & & \overline{\alpha}_\iota \downarrow & \\
S & \Rightarrow & S \circ \iota \\
\alpha & \Rightarrow & \alpha \iota
\end{array}
\]

in \( \text{Meas}^{\text{Meas}} \) in \( \text{Meas}^{\text{Cvx}} \)

Diagram 3. The codensity monad of \( \iota \) as a universal arrow.

commutes, and conversely given \( \overline{\alpha} \) there exist a unique arrow \( \alpha \) making the diagram on the right commute. The property of being a codensity monad of \( \iota \) can equivalently be expressed in terms of the diagram

\[
\begin{array}{ccc}
\text{Cvx} & \xrightarrow{\iota} & \text{Meas} \\
\downarrow \iota & & \downarrow \epsilon \\
& & \text{Meas} \\
\overline{\alpha} & \Rightarrow & R^\iota \\
\overline{\alpha}_\iota & \Rightarrow & S \circ \iota \\
\end{array}
\]

which indicates the fact that there exist a natural transformation \( \epsilon : R^\iota \circ \iota \Rightarrow \iota \) such that if \( (S, \alpha) \) also satisfies \( \alpha : S \circ \iota \Rightarrow \iota \) then there exist a unique natural transformation \( \overline{\alpha} : S \Rightarrow R^\iota \) such that \( \alpha = \epsilon \circ \overline{\alpha}_\iota \).

As \( \text{Meas} \) is complete the codensity monad \( R^\iota \) can be constructed pointwise \[\text{MacLane, 1971} \text{ Theorem 1, page 233} \] using the slice category \( (X \downarrow \iota) \) of objects under \( X \in \text{ob} \) \( \text{Meas} \) which has the objects and arrows
The objects of \((X_\downarrow \mu)\) are \((f, A)\), and the arrows are \(k \mapsto k\) where \(A, B \in \text{ob } \text{Cvx}\), \(f, g \in \text{ar } \text{Meas}\) and \(k \in \text{ar } \text{Cvx}\). There is a projection functor \(Q : (X_\downarrow \mu) \to \text{Cvx}\) mapping the objects \((f, A) \mapsto A\) and arrows \(k \mapsto k\) which when composed with \(\iota\) yields a composite functor whose limit

\[
\lim_{\leftarrow} \left( (X_\downarrow \mu) \xrightarrow{Q} \text{Cvx} \xrightarrow{\iota} \text{Meas} \right)
\]

we claim is precisely \(\mathcal{P}(X)\). Towards this end we require the following construction.

Given the object \(f : X \to \iota(A)\) in \((X_\downarrow \mu)\), since \(\iota(A)\) is a subobject \(\iota(A) \hookrightarrow I^{\text{Cvx}(A,I)}\) in \text{Meas} by the SMCC structure of \text{Meas} \(f\) determines a map

\[
\text{Cvx}(A, I) \xrightarrow{\hat{f}} I^{X} \quad \hat{f}[h](x) = f(x)[h] \quad \forall h \in \text{Cvx}(A, I), \forall x \in X
\]

where the notation ""[h]"" is used to emphasize that the argument is itself a function and to avoid excessive parentheses. By the definition of \(\hat{f}\) and the fact \(f(x) \in \iota(A)\) it follows

**Lemma 6.1.** Given the object \(f : X \to \iota(A)\) in \((X_\downarrow \mu)\) the map \(\hat{f}\) defined by (4) is a weakly averaging affine map.

Given any \(G \in \mathcal{P}(X)\) its composite with \(\hat{f}\) gives the “pushforward” map

\[
\begin{array}{ccc}
\text{Cvx}(A, I) & \xrightarrow{\hat{f}} & I^{X} \\
\downarrow & & \downarrow G \\
I & \xrightarrow{G} & I
\end{array}
\]

which is a weakly averaging affine map because the components defining it are and hence \(G \circ \hat{f} \in \iota(A)\).

**Theorem 6.2.** For each \(X \in \text{ob } \text{Meas}\), \(\mathcal{P}(X) = \lim_{\leftarrow} \left( (X_\downarrow \mu) \xrightarrow{Q} \text{Cvx} \xrightarrow{\iota} \text{Meas} \right)\) with the natural transformation \(\lambda\) of the cone \((\mathcal{P}(X), \lambda)\) over \(\iota \circ Q\) specified by

\[
\begin{array}{ccc}
\mathcal{P}(X) & \xrightarrow{\lambda_f} & \iota(A) \\
\downarrow & & \downarrow \iota(k) \\
I^{X} & \xrightarrow{G} & \text{Cvx}(A, I) \xrightarrow{G \circ \hat{f}} I
\end{array}
\]

Diagram 4. The slice category \((X_\downarrow \mu)\).
for every \( f \in \text{Meas}(X, \iota(A)) \), every \( A \in \text{ob} \ Cvx \) and where \( \hat{f} \) is defined by \([\frac{4}{4}]\). Each \( \theta \in \text{Meas}(Y, X) \) induces a unique arrow

\[
P(\theta) : \lim(\iota \circ Q) \to \lim(\iota \circ Q')
\]

commuting with the limiting cones, where \( Q' : (Y \downarrow \iota) \to Cvx \) is the projection functor. This construction makes \( P : \text{Meas} \to \text{Meas} \) functorial, and by defining \( \epsilon : P \circ \iota \to \iota \) componentwise by

\[
\epsilon_A = \lambda_{\text{id}_{\iota(A)}} \quad \text{for all } A \in \text{ob} \ Cvx
\]

the pair \( (P, \epsilon) \) is the codensity monad of \( \iota \).

**Proof.** The proof is broken into multiple parts which are denoted using italicized headings. \((P(X), \lambda)\) as a limit cone over \( \iota \circ Q \)

Given the \((X \downarrow \iota)\) arrows

\[
\begin{array}{ccc}
X & \xrightarrow{f} & \iota(A) \\
\downarrow{g} & & \downarrow{\iota(k)} \\
\iota(B) & \xrightarrow{\iota(k)} & B
\end{array}
\]

in \((X \downarrow \iota)\) in \( Cvx \)

\[
\begin{array}{ccc}
\lambda_f & \xrightarrow{\text{Cvx}(A, I)} & G \circ \hat{f} \\
\downarrow{G \circ \hat{f}} & & \downarrow{\iota(k)} \\
\text{Cvx}(B, I) & \xrightarrow{\text{Cvx}(B, I)} & I
\end{array}
\]

it follows \( \hat{g} = \hat{f} \circ k \) which makes \((P(X), \lambda)\) a cone over \( \iota \circ Q \) because for every \( G \in P(X) \) the diagram

\[
\begin{array}{ccc}
\lambda_f & \xrightarrow{\text{Cvx}(A, I)} & G \circ \hat{f} \\
\downarrow{G \circ \hat{f}} & & \downarrow{\iota(k)} \\
\text{Cvx}(B, I) & \xrightarrow{\text{Cvx}(B, I)} & I
\end{array}
\]

commutes.
Now suppose that \((Z, \omega)\) is also a cone over the functor \(\iota \circ Q\). We must show there exist a unique arrow \(\eta\) making the diagram commute. The commutativity of the outer path implies

\[
\omega_g(z)[h] = \omega_f(z)[h \circ k] \quad \forall h \in \text{Cvx}(B, I), \forall z \in Z.
\]  

(5)

Fix an element element \(z \in Z\). To satisfy the required commutativity condition \(\omega_f(z) = \eta(z) \circ \lambda_f\) for all objects \(f : X \to \iota(A)\) in \((X \downarrow \mu)\) it is necessary and sufficient that the function \(\eta(z)\) satisfy

\[
\begin{align}
(a) \quad & \eta(z)\left[\hat{f}[h]\right] = \omega_f(z)[h] \quad \forall h \in \text{Cvx}(A, I), \forall f \in \text{Meas}(X, \iota(A)), \\
(b) \quad & \eta(z) \in \mathcal{P}(X)
\end{align}
\]

(6)

The first condition can be used to define \(\eta(z)\) because every \(\gamma \in \text{Meas}(X, I)\) determines an object \(\gamma' : X \to \iota(I)\) in \((X \downarrow \mu)\) specified by

\[
\gamma'(x)[h] = h(\gamma(x)) \quad \forall h \in \text{Cvx}(I, I)
\]

(7)

which in turn, via the construction in (4), determines the map \(\hat{\gamma}' : \text{Cvx}(I, I) \to I^X\) in \(\text{Meas}\) specified by

\[
\hat{\gamma}'[h](x) = \gamma'(x)[h] = h(\gamma(x)) \quad \forall x \in X.
\]

(8)

This map \(\hat{\gamma}'\) is also a weakly averaging affine function because \(h\) is affine, \(\gamma(x)\) is weakly averaging for all \(x \in X\), and \(\text{Cvx}(I, I)\) has the convex structure defined pointwise on \(I\), i.e., \(\text{Cvx}(I, I) \subset I^I\) and the convex structure on \(I^I\) is defined pointwise which restricts to \(\text{Cvx}(I, I)\). Since \(\hat{\gamma}'[id_I](x) = \gamma'(x)[id_I] = \gamma(x)\) for all \(x \in X\) it follows \(\gamma = \hat{\gamma}'[id_I]\) and we can use condition (6a) to define \(\eta(z)\) as

\[
\eta(z)[\gamma] = \eta(z)\left[\hat{\gamma}'[id_I]\right] \overset{\text{def}}{=} \omega_{\gamma'}(z)[id_I] \quad \forall z \in Z, \forall \gamma \in \text{Meas}(X, I)
\]

(9)

or equivalently, for every \(z \in Z\) the map \(\eta(z)\) is defined as a function by the commutativity of the diagram

\[
\begin{array}{ccc}
\text{Cvx}(I, I) & \xrightarrow{\hat{\gamma}'} & I^X \\
\downarrow \eta(z) & & \downarrow \eta(z) \circ \hat{\gamma}' \overset{\text{def}}{=} \omega_{\gamma'}(z) \\
I & \xrightarrow{\eta(z)} & I
\end{array}
\]
for every $\gamma \in \text{Meas}(X, I)$.

We now proceed to verify condition (6b) by showing $\eta(z)$ as defined by (9) is both a weakly averaging and an affine function.

**Weakly averaging condition**

Let $\pi \in \text{Meas}(X, I)$ be a constant function with value $u \in I$. Let $k \in \text{Cvx}(I, I)$ be the constant map $k = u$ where we retain the symbol “$k$” to avoid confusion between the two constant functions with value $u$. For any $\gamma \in \text{Meas}(X, I)$ the diagram on the left in

\[
\begin{array}{ccc}
X & \xrightarrow{\gamma'} & \text{Meas}(\gamma', I) \\
\downarrow{\pi} & & \downarrow{\text{id}_I} \\
\text{Meas}(\pi, I) & \xrightarrow{\text{id}_I} & \text{Meas}(\pi, I)
\end{array}
\]

commutes because, for all $x \in X$, the composite map $\iota(k)(\gamma'(x))$ is the pushforward map

\[
\begin{array}{ccc}
\text{Cvx}(I, I) & \xrightarrow{\gamma'(x)} & I \\
\downarrow{\pi(x)} & & \downarrow{\text{id}_I} \\
\text{Cvx}(k, I) & \xrightarrow{h \circ k} & \text{Cvx}(I, I)
\end{array}
\]

\[
\gamma'(x)(h \circ k) = h(k(\gamma(x))) = h(u)
\]

Given the cone $(Z, \omega)$ over the functor $\iota \circ Q$ it follows the diagram on the right in Diagram 5 also commutes which says

\[
\omega_{\pi'}(z)[id_I] = (\iota(k) \circ \omega_{\gamma'})(z)[id_I] = \omega_{\gamma'}(z)[k] = \omega_{\gamma'}(z)[\pi] = u
\]

where the last equality follows because $\omega_{\gamma'}(z)$ is weakly averaging. This shows that

\[
\eta(z)[\pi] = \omega_{\pi'}(z)[id_I] = u
\]

which proves $\eta(z)$ is weakly averaging.
**Affine condition**

Consider the object $I \times I$ in CVx which has a convex structure defined componentwise by
\[
\sum_{i=1}^{n} r_i (u_i, v_i) = \left( \sum_{i=1}^{n} r_i u_i, \sum_{i=1}^{n} r_i v_i \right) \text{ where } \sum_{i=1}^{n} r_i = 1, \forall r_i \in I.
\]

Let $\alpha \in I$. The map $\pi_1 + \alpha \pi_2 : I \times I \to I$ defined by $\pi_1 + \alpha \pi_2 : (u, v) \mapsto u + \alpha v$ is affine because
\[
(\pi_1 + \alpha \pi_2) \left( \sum_{i=1}^{n} r_i (u_i, v_i) \right) = \alpha \sum_{i=1}^{n} r_i u_i + (1 - \alpha) \sum_{i=1}^{n} r_i v_i
= \sum_{i=1}^{n} r_i (\alpha u_i) + \sum_{i=1}^{n} r_i ((1 - \alpha) v_i)
= \sum_{i=1}^{n} r_i (u_i + \alpha v_i)
= \sum_{i=1}^{n} r_i ((\pi_1 + \alpha \pi_2)(u_i, v_i))
\]

Moreover this map $\pi_1 + \alpha \pi_2$ is also measurable with $I \times I$ having the product $\sigma$-algebra.

For $\gamma_1, \gamma_2 \in \text{Meas}(X, I)$ we obtain the induced maps $\gamma_1', \gamma_2' \in \text{Meas}(X, \iota(I))$ by the construction given in (7) and the diagram on the left in

\[
\begin{array}{ccc}
X & \xrightarrow{\gamma_1} & I(\iota(I)) \\
\downarrow{\gamma_1'} & & \downarrow{\iota(\pi_1)} \\
X & \xrightarrow{\iota(I)} & I(\iota(I))
\end{array}
\]

and the diagram on the right in

\[
\begin{array}{ccc}
Z & \xrightarrow{\omega_{(\gamma_1, \gamma_2)'}} & I(\iota(I)) \\
\downarrow{\omega_{\gamma_1'}} & & \downarrow{\iota(\pi_1)} \\
Z & \xrightarrow{\omega_{\gamma_2'}} & I(\iota(I))
\end{array}
\]

in $(X, \mu)$, in $\text{Meas}$

Diagram 6. Determining the affine property.

commutes. Hence, given the cone $(Z, \omega)$ over the functor $\iota \circ Q$, the diagram on the right also commutes which yields the two equations
\[
\omega_{(\gamma_1, \gamma_2)'}(z)[\pi_1] = \omega_{\gamma_1'}(z)[id_I] \quad \text{and} \quad \omega_{(\gamma_1, \gamma_2)'}(z)[\pi_2] = \omega_{\gamma_2'}(z)[id_I]
\]
for all $z \in Z$. As $\omega_{(\gamma_1, \gamma_2)'}(z)$ is affine it follows, using the two above equations, that
\[
\omega_{(\gamma_1, \gamma_2)'}(z)[\pi_1 + \alpha \pi_2] = \omega_{\gamma_1'}(z)[id_I] + \alpha \omega_{\gamma_2'}(z)[id_I] \quad \forall z \in Z \tag{10}
\]

---

6Here it is more convenient, and standard, to use define the convex structure of $I \times I$ using the representation of the form $\sum_{i=1}^{n} r_i u_i$ with the relation $\sum_{i=1}^{n} r_i = 1$ rather than the free representation $(\ldots (u_1 + s_1 u_2) + s_2 u_3) + \ldots + s_{n-1} u_n)$, where the elements $s_i \in I$ for all $i = 1, \ldots, n - 1$ determine the coefficients $r_i$ and vice versa as these two representations are, assuming the $r_i \neq 0$, easily computed recursively.
Now using the commutative diagram on the left in

\[
\begin{array}{ccc}
X \xrightarrow{(\gamma_1, \gamma_2)'} & \iota(I \times I) & \xrightarrow{\omega(\gamma_1, \gamma_2)'} & Z \\
\downarrow \iota(I) & & \downarrow \omega(\gamma_1 + \alpha \gamma_2) & \downarrow \iota(I) \\
\iota(I) & & \iota(I) \\
\end{array}
\]

in \((X \downarrow \iota)\) in \textbf{Meas}

it follows the diagram on the right also must commute which gives the equation

\[
\omega(\gamma_1 + \alpha \gamma_2)'(z)[id_I] = \omega(\gamma_1, \gamma_2)'(z)[\pi_1 + \alpha \pi_2]. \tag{11}
\]

Combining the last two results and using the definition of \(\eta(z)\) it follows

\[
\eta(z)(\gamma_1 + \alpha \gamma_2) = \omega(\gamma_1 + \alpha \gamma_2)'(z)[id_I] \quad \text{by def. of } \eta(z)(\gamma_1 + \alpha \gamma_2)
\]

by (11)

\[
= \omega(\gamma_1, \gamma_2)'(z)[\pi_1 + \alpha \pi_2] \quad \text{by (11)}
\]

by (11)

\[
= \omega_1'(z)[id_I] + \alpha \omega_2'(z)[id_I] \quad \text{by (10)}
\]

by def. of \(\eta(z)(\gamma_1)\) and \(\eta(z)(\gamma_2)\)

which shows \(\eta(z)\) is affine for all \(z \in Z\).

\textbf{Functoriality of } \mathcal{P}

The object \(X\) with all the \((X \downarrow \iota)\) arrows gives a cone over the functor \(\iota \circ Q\) and the unique arrow \(\eta : X \to \mathcal{P}(X)\) making the diagram

\[
\begin{array}{ccc}
X \xrightarrow{\eta} & \mathcal{P}(X) & \xrightarrow{\iota(A)} \iota(A) \\
\downarrow f & & \downarrow \iota(k) \\
\mathcal{P}(X) & \xrightarrow{\lambda_f} & \iota(A) \\
\downarrow g & & \downarrow k \\
B & \xrightarrow{\lambda_g} & \iota(B) \\
\end{array}
\]

in \textbf{Meas} in \textbf{Cvx}

Diagram 7. The unit of the monad \(\mathcal{P}\) at component \(X\).

commute is precisely the unit of the monad \(\mathcal{P}\) at the component \(X\), \(\eta = \eta_X\), because

\[
(\lambda_g \circ \eta_X)(x) = ev_x \circ \hat{g}
= g(x)
= f(x) \circ k \quad \text{because } g = \iota(k) \circ f.
= ev_x \circ \hat{f} \circ k
= (\iota(k) \circ \lambda_f \circ \eta_X)(x)
\]
Consequently, by precomposition of the cone with vertex \( X \) shown in Diagram 7, each \( \theta \in \text{Meas}(Y, X) \) induces a cone with vertex \( Y \) over \( \iota \circ Q \) and hence uniquely determines an arrow

\[
\mathcal{P}(\theta) : \lim(\iota \circ Q) \longrightarrow \lim(\iota \circ Q')
\]

where \( Q' : (Y \downarrow \iota) \to \text{Cvx} \) is the projection functor, making \( \mathcal{P} \) functorial in the above construction which coincides with the previously defined operation of \( \mathcal{P} \) on \( \text{Meas} \) arrows, i.e., as the pushforward map.

Having established that for each measurable space \( X \) the pair \( (\mathcal{P}(X), \{\lambda_g\}_{g \in \text{ob}(X)}) \) forms a limiting cone over the functor \( \iota \circ Q \) the rest of the proof now follows the proof of [MacLane, 1971, Theorem 1, p233] verbatim for the general construction of the pointwise right Kan extension of \( \iota \) along \( \iota \). We give the proof showing the naturality of \( \epsilon \), which expands upon the proof given by MacLane, and refer the reader to MacLanes proof that if \( S : \text{Meas} \to \text{Meas} \) is another functor with \( \alpha : S \circ \iota \to \iota \) a natural transformation then it corresponds bijectively with a natural transformation \( \overline{\alpha} : S \to \mathcal{P} \). This result simply depends upon \( (\mathcal{P}(X), \{\lambda_g\}_{g \in \text{ob}(X)}) \) being a limiting cone over the functor \( \iota \circ Q \) and the functoriality of \( \mathcal{P} \).

Defining the universal arrow \( \epsilon \)

Let \( A \) be an object of \( \text{Cvx} \). The identity map \( \hat{id}_{\iota(A)} : \iota(A) \to \iota(A) \) is an object in the slice category \( (\iota(A) \downarrow \iota) \) and, just as we defined in (4), using the SMCC structure of \( \text{Meas} \) we obtain a measurable map

\[
\begin{array}{c}
\text{Cvx}(A, I) \\
\downarrow \hat{id}_{\iota(A)} \\
A \\
\hline
\iota(A) \\
\downarrow \iota \\
K(h) \\
\uparrow \\
\hat{id}_{\iota(A)}(h) \\
\iota(A) \\
\downarrow \\
I \\
\end{array}
\]

from which we see \( \hat{id}_{\iota(A)}(h) = ev_h \), the evaluation map at \( h \). Corresponding to this object in the slice category \( (\iota(A) \downarrow \iota) \) there is the component map \( \lambda_{\hat{id}_{\iota(A)}} : \mathcal{P}(\iota(A)) \to \iota(A) \) of the natural transformation \( \lambda \), of the universal cone \( (\mathcal{P}(\iota(A)), \lambda) \) over the functor \( \iota \circ Q'' \), where \( Q'' : (\iota(A) \downarrow \iota) \to \text{Cvx} \) is the projection functor. The universal arrow \( \epsilon \) is defined componentwise at \( A \) by

\[
\begin{array}{c}
\mathcal{P}(\iota(A)) \\
\downarrow \epsilon_A = \lambda_{\hat{id}_{\iota(A)}} \\
\iota(A) \\
\downarrow G \\
\text{Cvx}(A, I) \\
\downarrow \hat{id}_{\iota(A)} \\
I \iota(A) \\
\downarrow G \\
I \\
\end{array}
\]

hence for all \( G \in \mathcal{P}(\iota(A)) \) the map \( \epsilon_A(G) \) is specified by \( \epsilon_A(G)[h] = G(ev_h) \) for all \( h \in \text{Cvx}(A, I) \), and \( \epsilon_A \) is a measurable weakly averaging affine map because both component maps are measurable weakly averaging affine maps.
Naturality of $\epsilon$

For $k : A \to B$ a Conv morphism the naturality of $\epsilon$ requires the Meas diagram

\[
\begin{array}{cccc}
\mathcal{P}(\iota(A)) & \xrightarrow{\epsilon_A} & \iota(A) \\
\mathcal{P}(\iota(B)) & \xrightarrow{\epsilon_B} & \iota(B)
\end{array}
\quad
\begin{array}{cccc}
\mathcal{P}(\iota(k)) & \xrightarrow{\iota(k)} & \iota(k) \\
G & \xrightarrow{G \circ \iota(k)} & G \circ \iota(k)
\end{array}
\]

To commute. Evaluating the two expressions at the bottom right in the diagram at the affine morphism $h : B \to I$ gives

\[
((\iota(k) \circ \epsilon_A)(G))[h] = (\epsilon_A(G) \circ I^k)[h] = \epsilon_A(G)[h \circ k] = G(ev_{h \circ k}) = G(ev_h \circ \iota(k)) \text{ by Lemma 5.1}
\]

and hence $\epsilon$ is a natural transformation.

Appendix. Filters as weakly averaging affine functionals

Let $2 = \{0, 1\}$ be the totally ordered set with the natural order $0 < 1$. For $X$ any set let $P(X)$ denote the powerset of $X$. This set $P(X)$ is partially ordered by the inclusion relation $\subset$.

A filter $\mathcal{F}$ on a set $X$ is a $\land$ semi-lattice homomorphism

\[
(P(X), \subset) \xrightarrow{\mathcal{F}} (2, <)
\]

where $\mathcal{F}(X) = 1$. Hereafter, for brevity, we drop the explicit reference to the relations $\subset$ and $<$ in the above definition and simply write $\mathcal{F} : P(X) \to 2$ to denote a filter. A filter $\mathcal{F}$ on $X$ is often viewed as the subset $\mathcal{F}^{-1}(\{1\}) \subseteq P(X)$ which is (i) nonempty, (ii) upward closed and (iii) closed under finite intersections. The nonempty condition is equivalent to $\mathcal{F}(X) = 1$. A proper filter is a filter for which $\mathcal{F}(\emptyset) = 0$ which is equivalent
to $\mathcal{F}^{-1}(\{1\}) \neq P(X)$. Consequently a proper filter on a set $X$ is equivalent to a ∧ semi-
lattice homomorphism $\mathcal{F}: P(X) \to 2$ which preserves both the top and bottom elements
of $(P(X), \subset)$.

Define a convex structure on $2$ by

$$0 +_\alpha 1 = \begin{cases} 0 & \text{iff } \alpha \in (0, 1] \\ 1 & \text{otherwise} \end{cases}$$

and the convex structure on $2^X$, the space of characteristic functions of $X$, is determined
pointwise by the convex structure of $2$ so $\chi_A \leq \chi_B$ iff $A \subset B$.

**Theorem.** Let $X$ be any object in $\textbf{Set}$. There is a bijective correspondence between proper
filters $\mathcal{F}$ on $X$ and weakly averaging affine functionals $\hat{\mathcal{F}} \in \textbf{Set}(2^X, 2)$.

**Proof.** Given a filter $\mathcal{F}: P(X) \to 2$ on $X$ we obtain a weakly averaging affine functional
$\hat{\mathcal{F}} \in \textbf{Set}(2^X, 2)$ by defining $\hat{\mathcal{F}}(\chi_A) = \mathcal{F}(A)$. By the convex structure of $2^X$ it follows
$\chi_A +_\alpha \chi_B = \chi_{A\cap B}$ and hence the affine property

$$\hat{\mathcal{F}}(\chi_A +_\alpha \chi_B) = \hat{\mathcal{F}}(\chi_A) +_\alpha \hat{\mathcal{F}}(\chi_B) \quad \forall \alpha \in (0, 1) \tag{12}$$

follows directly from the fact $\mathcal{F}$ is a filter. As $\hat{\mathcal{F}}(\chi_X) = \mathcal{F}(X) = 1$ and $\hat{\mathcal{F}}(\chi_\emptyset) = \mathcal{F}(\emptyset) = 0$ the weakly averaging property is satisfied and therefore $\hat{\mathcal{F}} \in \textbf{Set}(2^X, 2)$ is a weakly
averaging affine functional.

Conversely given a weakly averaging affine functional $\hat{\mathcal{F}} \in \textbf{Set}(2^X, 2)$ by defining
$\mathcal{F}(A) = \hat{\mathcal{F}}(\chi_A)$ for every $A \subset X$ it follows by (12) that by taking the meet operation $\land$
as $+_\alpha$ for any $\alpha \in (0, 1)$ gives

$$\mathcal{F}(A \cap B) = \mathcal{F}(A) \land \mathcal{F}(B)$$

and hence $\hat{\mathcal{F}}$ determines a proper filter. These two constructions are inverse to each other
yielding a bijective correspondence. \qed

An ultrafilter is a lattice homomorphism $\mathcal{F}: P(X) \to 2$ and a proper lattice homo-
morphism satisfies the finite additivity condition

$$\mathcal{F}(A \cap B) + \mathcal{F}(A \cup B) = \mathcal{F}(A) + \mathcal{F}(B)$$

and hence is a primitive sort of probability measure, as it is a deterministic probability
measure assuming values of only 0 or 1. By changing the base from 2 to $I$ and consid-
ering measurable functions $X \to I$ (or take the discrete $\sigma$-algebra making all functions
measurable) we are led to Lemma 4.1 and the bijective correspondence of Lemma 4.3.
References

[Doberkat, 2004] E.E. Doberkat, Eilenberg-Moore algebras for stochastic relations. Information and Computation, 204 (2006), 1756-1781.

[Fritz, 2009] T. Fritz, Convex Spaces I: Definitions and Examples. http://arxiv.org/abs/0903.5522, 2009.

[Giry, 1982] M. Giry, A categorical approach to probability theory, in Categorical Aspects of Topology and Analysis, Vol. 915, 68-85, Springer-Verlag, 1982.

[Lawvere and Rosebrugh, 2005] F.W. Lawvere and R. Rosebrugh, Sets for Mathematics, Cambridge University Press, 2005.

[Leinster, 2013] T. Leinster, Codensity and the ultrafilter monad. Theory and Applications of Categories, Vol. 28, No. 13, 332 - 370 (2013).

[Lindner, 1979] H. Lindner, Affine parts of monads, Vol. 33, (1979), 437-443.

[MacLane, 1971] S. MacLane, Categories for the Working Mathematician, Springer-Verlag, 1971.

[Meng, 1987] X. Q. Meng, Categories of Convex Sets and of Metric Spaces, with Applications to Stochastic Programming. Dissertation, SUNY Buffalo, 1987.

Kirk Sturtz
Universal Mathematics
Dayton, OH USA

kirksturtz@UniversalMath.com