Weakly-affine monads

Paolo Perrone
Department of Computer Science, University of Oxford

Fabio Gadducci
Department of Computer Science, University of Pisa, Pisa, IT

Davide Trotta
Department of Computer Science, University of Pisa, Pisa, IT

Abstract

To be written.

2012 ACM Subject Classification

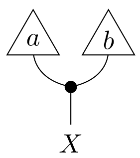
Keywords and phrases string diagrams, gs-monoidal categoriess

Digital Object Identifier 10.4230/LIPIcs...

1 Weakly Markov categories

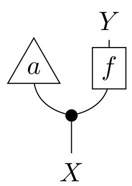
Let $\mathcal{C}$ be a GS-category. For every object X, the set $\mathcal{C}(X, I)$ has a canonical commutative monoid structure as follows: the monoidal unit is the discard map $X \to I$, and given $a, b : X \to I$, their product ab is given by copying, as follows.

How to call them? effects? co-states?



If a morphism $f : X \to Y$ is copyable and discardable, precomposition with f induces a morphism of monoids $\mathcal{C}(Y, I) \to \mathcal{C}(X, I)$.

The monoid $\mathcal{C}(X, I)$ acts on the set $\mathcal{C}(X, Y)$: given $a : X \to I$ and $f : X \to Y$, $a \cdot f$ is given as follows,



and the product $(f, g) \mapsto (f \otimes g) \circ \mathrm{copy}_X$ is equivariant for this action.

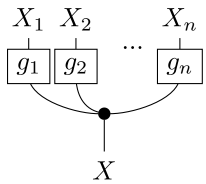
▶ **Definition 1.1.** A GS-category $\mathcal{C}$ is called *weakly Markov* if for every object X, the monoid $\mathcal{C}(X, I)$ is a group.

Every Markov category is weakly Markov: for each X, the monoid $\mathcal{C}(X, I)$ is the trivial group.

▶ **Definition 1.2.** Given two parallel morphisms $f, g : X \to Y$ in a weakly Markov GS-category $\mathcal{C}$, we say that f and g are *equivalent*, and write $f \sim g$, if they lie in the same orbit for the action of $\mathcal{C}(X, I)$.

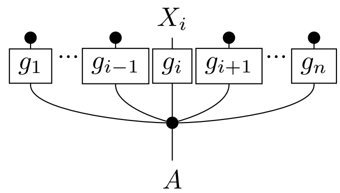
▶ **Definition 1.3.** A morphism $f : A \to X_1 \otimes \cdots \otimes X_n$ in a A GS-category $\mathcal{C}$ is said to exhibit *conditional independence of the X_i given A* if and only if it can be expressed as a product of the following form.

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



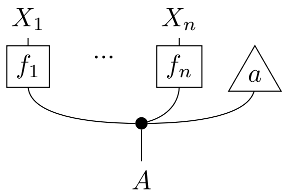
▶ **Proposition 1.4.** *Let $f : A \to X_1 \otimes \cdots \otimes X_n$ be a morphism in a GS-category $\mathcal{C}$. Then f exhibits conditional independence of the X_i given A if and only if it is equivalent to the product of all its marginals.*

Proof. Denote the marginals of f by $f_1, \ldots, f_n$. Suppose that f is a product as in Definition 1.3. For each $i = 1, \ldots, n$, by marginalizing, we get that f_i is equal to the following.



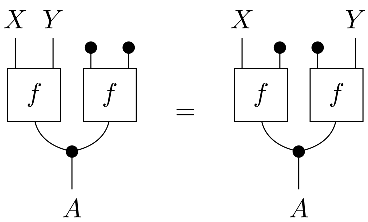
Therefore for each i we have that $f_i \sim g_i$.

Conversely, suppose that f is equivalent to the product of its marginals, i.e. that there exists $a : X \to I$ such that f is equal to the following.



One can then choose $g_i = f_i$ for all $i < n$, and $g_n = a \cdot f_n$. ◀

▶ **Remark 1.5.** For $n = 2$, a morphism $f : A \to X \otimes Y$ in a weakly Markov GS-category $\mathcal{C}$ exhibits conditional independence of X and Y given A if and only if the following equation holds.



▶ **Proposition 1.6.** *Let $\mathcal{C}$ be a weakly Markov GS-category. Let $f : A \to X \otimes Y \otimes Z$ be a morphism exhibiting conditional independence of $X \otimes Y$ (jointly) and Z given A, as well as conditional independence of X and $Y \otimes Z$ given A. Then f exhibits conditional independence of X, Y and Z given A.*

Proof. By marginalizing out X, we have that f_{YZ} exhibits conditional independence of Y and Z given A. Since by hypothesis f exhibits conditional independence of X and $Y \otimes Z$ given A, by Proposition 1.4 we have that f is equivalent to the product of f_X and f_{YZ}. But, again by Proposition 1.4, f_{YZ} is equivalent to the product of f_Y and f_Z, so we have that f is equivalent to the product of all its marginals. Using Proposition 1.4 in the other direction, thie means that f exhibits conditional independence of X, Y and Z given A. ◀

Let's now consider the case where the GS structure comes from a commutative monad on a cartesian monoidal category $\mathcal{D}$. In this case, the monoid structure of Kleisli morphisms $X \to 1$ comes from the following canonical internal monoid structure of $T1$ in $\mathcal{D}$, given by

$$1 \xrightarrow{\eta} T1, \qquad T1 \times T1 \xrightarrow{c} T(1 \times 1) \xrightarrow{\cong} T1.$$

The monoid structure of Kleisli morphisms $X \to 1$ is now given as follows. The unit is

$$X \xrightarrow{!} 1 \xrightarrow{\eta} T1,$$

and the multiplication of the morphisms $f, g : X \to T1$ is

$$X \xrightarrow{\text{copy}} X \times X \xrightarrow{f \times g} T1 \times T1 \xrightarrow{c} T(1 \times 1) \xrightarrow{\cong} T1.$$

▶ **Proposition 1.7.** *Let $\mathcal{D}$ be a cartesian monoidal category, and let T be a commutative monad on $\mathcal{D}$. The Kleisli category of T is weakly Markov if and only if $T1$ with its canonical internal monoid structure is a group.*

mi sembra che la proposizione 1.7 funzioni in generale per con $\mathcal{D}$ GS, no?

Proof. First, suppose that $T1$ is an internal group, and denote by $\iota : T1 \to T1$ its inversion map. The inverse of the morphism $f : X \to T1$ in $\mathrm{Kl}_T(X, 1)$ is given by $\iota \circ f$: indeed, the following diagram commutes,

$$\begin{tikzcd}
X \arrow[r, "\text{copy}"] \arrow[d, "f"'] & X \times X \arrow[d, "f \times f"'] \arrow[dr, "f \times (\iota \circ f)"] & & \\
T1 \arrow[r, "\text{copy}"] \arrow[d, "!"'] & T1 \times T1 \arrow[r, "\mathrm{id} \times \iota"'] & T1 \times T1 \arrow[r, "c"] & T(1 \times 1) \arrow[d, "\cong"] \\
1 \arrow[rrr, "\eta"] & & & T1
\end{tikzcd}$$

where the bottom rectangle commutes since ι is the inversion map for $T1$. The analogous diagram with $\iota \times \mathrm{id}$ in place of $\mathrm{id} \times \iota$ commutes analogously.

Conversely, suppose that for every X, the monoid structure on $\mathrm{Kl}_T(X, 1)$ has inverses. Then in particular we can take $X = T1$, and the inverse of the Kleisli morphism $\mathrm{id} : T1 \to T1$ is an inversion map for $T1$. ◀

This feels vaguely like Yoneda, but in monoidal sauce. Can't make it precise for now, though.

2 Introduction

For context:

▶ **Proposition 2.1.** *A monoid $(M, \cdot, 1)$ is a group if and only if the associativity square*

$$\begin{tikzcd}
M \times M \times M \arrow[r, "\cdot \times \mathrm{id}"] \arrow[d, "\mathrm{id} \times \cdot"] & M \times M \arrow[d, "\cdot"] \\
M \times M \arrow[r, "\cdot"] & M
\end{tikzcd} \tag{1}$$

is a pullback.

Proof. The square (1) is a pullback, both of sets and of groups, if and only if given $a, g, h, c \in M$ such that $ag = hc$, there exists a unique $b \in M$ such that $g = bc$ and $h = ab$. First, suppose that g is a group. The only possible choice of b is

$$b = a^{-1}h = gc^{-1},$$

which is unique by uniqueness of inverses.

Conversely, suppose that (1) is a pullback. We can set $g, h = e$ and $c = a$ so that $ae = ea = a$. Instantiating the pullback property, there is a unique b such that $ab = e$ and $ba = e$, that is, $b = a^{-1}$. ◀

Recall that a monoidal functor generalizes a monoid object (in turn, generalizing a monoid). Similarly, a *weakly affine monoidal functor* generalizes a group in the sense of the proposition above.

▶ **Proposition 2.2.** *Let $(G, \cdot, 1)$ be a group and let X be a set. A function $\alpha : M \times X \to X$ determines a left action if and only if the square*

$$\begin{tikzcd} G \times G \times X \arrow[r, "\cdot \times \mathrm{id}"] \arrow[d, "\mathrm{id} \times \alpha"] & M \times X \arrow[d, "\alpha"] \\ G \times X \arrow[r, "\alpha"] & X \end{tikzcd} \tag{2}$$

commutes and it is a pullback.

Proof. By definition, the square (2) commutes if and only if α and $\cdot$ are compatible. Now we show that the commutative square (2) is a pullback if and only if α satisfies the identity axiom, i.e. $\alpha(e, x) = x$ for every x in X. Now, if (2) is a pullback, then there exists a function $\beta : X \to G$ such that the diagram

$$\begin{tikzcd} X \arrow[dr, "\langle e!, \beta, \mathrm{id} \rangle"] \arrow[drr, bend left, "\langle e!, \mathrm{id} \rangle"] \arrow[ddr, bend right, "\langle e!, \mathrm{id} \rangle"'] & & \\ & G \times G \times X \arrow[r, "id \times \alpha"] \arrow[d, "\cdot \times id"'] & G \times X \arrow[d, "\alpha"] \\ & G \times X \arrow[r, "\alpha"'] & X \end{tikzcd}$$

commutes, where $e! : X \to G$ is the function assigning the identity element e to every element x of X. Now, since the left triangle commutes, then we have that $e = e \cdot \beta(x)$ for every x of X, i.e. $\beta(x) = e$ for every x of X. Now, since the right triangle commutes, we can conclude that $\alpha(\beta(x), x) = \alpha(e, x) = x$ for every x in X.

Now we show that $\alpha(e, x) = x$ implies that the commutative square (2) is a pullback. Let us consider a set Y and the functions $\langle f_1, f_2 \rangle : Y \to G \times X$ and $\langle g_1, g_2 \rangle : Y \to G \times X$ such that $\alpha(f_1(y), f_2(y)) = \alpha(g_1(y), g_2(y))$. By applying $\alpha(f_1(y)^{-1}, -)$ to both sides, and then combining the compatibility of α with the assumption that $\alpha(e, x) = x$, we can conclude that $f_2(y) = \alpha(f_1(y)^{-1} \cdot g_1(y), g_2(y))$. Therefore, we can conclude that the diagram

$$
\begin{tikzcd}
Y \arrow[dr, "\langle f_1,\gamma,g_2\rangle"] \arrow[drr, bend left, "\langle f_1,f_2\rangle"] \arrow[ddr, bend right, "\langle g_1,g_2\rangle"'] & & \\
& M \times M \times X \arrow[r, "id\times\alpha"] \arrow[d, "\cdot\times id"'] & M \times X \arrow[d, "\alpha"] \\
& M \times X \arrow[r, "\alpha"'] & X
\end{tikzcd}
$$

commutes, where the function $\gamma : Y \to M$ is defined by $\gamma(y) := f_1^{-1}(y) \cdot g_1(y)$. By the unicity of the inverse in a group, this function is also unique, and hence we can conclude that the commutative square (2) is a pullback. ◀

3 Weakly-affine monads

Nome da scegliere e valutare se dare la def per una arbitraria gs

▶ **Definition 3.1.** Let T be a commutative monad on a category $\mathcal{A}$ with finite products. A triple (X, Y, Z) of objects of $\mathcal{A}$ is said to be **TBA** if the commutative square

$$
\begin{tikzcd}
T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"] \arrow[d, "c_{X,Y} \times \mathrm{id}"'] & T(X) \times T(Y \times Z) \arrow[d, "c_{X,Y\times Z}"] \\
T(X \times Y) \times T(Z) \arrow[r, "c_{X\times Y,Z}"] & T(X \times Y \times Z)
\end{tikzcd}
$$

is a pullback.

esempi?

▶ **Definition 3.2.** Let T be a commutative monad on a category $\mathcal{A}$ with finite products. We say that the monad T is **weakly affine** if the following associativity diagram is a pullback for every X, Y, Z in $\mathcal{A}$:

$$
\begin{tikzcd}
T(X) \times T(Y) \times T(Z) \arrow[r, "\mathrm{id} \times c_{Y,Z}"] \arrow[d, "c_{X,Y} \times \mathrm{id}"'] & T(X) \times T(Y \times Z) \arrow[d, "c_{X,Y\times Z}"] \\
T(X \times Y) \times T(Z) \arrow[r, "c_{X\times Y,Z}"] & T(X \times Y \times Z)
\end{tikzcd}
\tag{3}
$$

Let us derive some general properties before looking at examples. It is a standard fact that for any commutative monad T, the composite arrow

$$T(1) \times T(1) \xrightarrow{c_{1,1}} T(1 \times 1) \xrightarrow{\cong} T(1)$$

equips $T(1)$ with the structure of a commutative monoid internal to $\mathcal{A}$ with unit $\eta_1 : 1 \to T(1)$.

▶ **Lemma 3.3.** *If T is weakly affine, then $T(1)$ is a group.*

Proof. If T is weakly affine, then taking $X = Y = Z = 1$ in (3) shows that this monoid must be an abelian group: assuming that $\times$ is a strict monoidal structure for simplicity, we obtain a unique arrow $\iota\colon T(1) \to T(1)$ such that the diagram

$$\begin{tikzcd}
T(1) \arrow[dr, "{(\mathrm{id},\iota,\mathrm{id})}"] \arrow[drr, bend left, "{(\mathrm{id},\eta_1!)}"] \arrow[ddr, bend right, "{(\eta_1!,\mathrm{id})}"'] & & \\
& T(1)\times T(1)\times T(1) \arrow[r, "\mathrm{id}\times c_{1,1}"] \arrow[d, "c_{1,1}\times \mathrm{id}"'] & T(1)\times T(1) \arrow[d, "c_{1,1}"] \\
& T(1)\times T(1) \arrow[r, "c_{1,1}"] & T(1)
\end{tikzcd}$$

and the commutativity shows that ι satisfies the equations making it the inversion map for a group structure. ◀

▶ **Proposition 3.4.** *If T is weakly affine, then for every object X, the morphism $c_{1,X}$: $T(1) \times T(X) \to T(X)$ determines a (left) group action.*

Proof. The compatibility axiom follows from the fact that the diagram

$$\begin{tikzcd}
T(1)\times T(1)\times T(X) \arrow[r, "\mathrm{id}\times c_{1,X}"] \arrow[d, "c_{1,1}\times \mathrm{id}"'] & T(1)\times T(X) \arrow[d, "c_{1,X}"] \\
T(1)\times T(X) \arrow[r, "c_{1,X}"'] & T(X)
\end{tikzcd}$$

commutes for every strong and commutative monad. Moreover, following the same proof used for Proposition 2.2, we can conclude that the identity axiom is satisfied since T is weakly affine. In particular, because $T(1)$ is a group by Lemma 3.3, and the previous square is a pullbackc (by definition of weakly affine monad). ◀

▶ **Proposition 3.5.** *Let T be a weakly affine monad. If the diagram*

$$\begin{tikzcd}
T(1) \arrow[r, "\mathrm{id}"] \arrow[d, "\iota"'] & T(1) \arrow[d, "\eta_{T1}"] \\
T(1) \arrow[r, "T(\eta_1)"'] & T^2(1)
\end{tikzcd}$$

commutes, then $T^2(1) \cong T(1)$ in $\mathcal{A}$.

Proof. To prove the result it is enough to show that $T(1) \cong 1$ in the Kleisli category $\mathcal{A}_T$. We know from Lemma that $T(1)$ is a group in $\mathcal{A}$, where the arrow $\eta_1\colon 1 \to T(1)$ is the unit of the group, and $\iota\colon T(1) \to T(1)$ is the inversion map. Therefore, we have that the composition $\iota\eta_1\colon 1 \to T(1)$ has to be equal to η_1. Therefore, we can consider the arrows $1 \to T(1)$ and $T(1) \to 1$ in the Kleisli category $\mathcal{A}_T$ given by $T(\eta_1)\eta_1$ and ι respectively. The composition $T(\eta_1)\eta_1$ with ι in $\mathcal{A}_T$ is given by $\mu_{1,1}T(\iota)T(\eta_1)\eta_1$. Employing the naturality of η_1 and the fact that $\iota\eta_1 = \eta_1$, it is direct to check that $\mu T(\iota)T(\eta_1)\eta_1 = \eta_1$, that is the identity $1 \to 1$ in $\mathcal{A}_T$. Now to show that the other composition gives the identity on $T(1)$ in $\mathcal{A}_T$, it is enough to show that $T(\eta_1)\iota = \eta_{T(1)}$, but this follows by hypothesis. ◀

(Paolo) Credo che $T(\eta_1)\iota \neq \eta_{T(1)}$ nell'esempio delle misure nonzero. Per ogni x in $(0,\infty) = T1$ abbiamo che $\eta_{T(1)}(x) = \delta_x$ (delta di Dirac), mentre $T\eta_1(\iota(x)) = T\eta_1(1/x) = 1/x\,\delta_1$.

▶ **Corollary 3.6.** *Let T be a weakly affine monad. If the diagram*

$$\begin{tikzcd} T(1) \arrow[r, "\mathrm{id}"] \arrow[d, "\iota"'] & T(1) \arrow[d, "\eta_{T1}"] \\ T(1) \arrow[r, "T(\eta_1)"'] & T^2(1) \end{tikzcd}$$

commutes, then $T(1)$ is an idempotent group, namely $\iota = \mathrm{id}_{T1}$.

Proof. By Lemma 3.3 we have that $T(1)$ is a group. If $\eta_{T1} = T(\eta_1)\iota$, then we can apply the multiplication of the monad to both sides, obtaining $\iota = \mathrm{id}_{T1}$. ◀

The following result shows that weak affinity occurs frequently. Recall that a strong monad $T\colon \mathcal{A} \to \mathcal{A}$ on a category $\mathcal{A}$ with finite products is **affine** if $T(1) \cong 1$ (see also Remark **??**). Three relevant examples of affine monads are the distribution monad on **Set** (for discrete probability), the Giry monad on the category of measurable spaces (for measure-theoretic probability, see Examples **??** and **??**), and the expectation monad, see [**?**].

▶ **Proposition 3.7.** *Let T be a commutative monad on a category $\mathcal{A}$ with finite limits. If T is affine, then it is weakly affine.*

Proof. Let $m_{X,Y} : T(X \times Y) \longrightarrow TX \times TY$ be the arrow defined as the pairing of $T(\pi_1)$ and $T(\pi_2)$. Then it is known that T is affine if and only if $m_{X,Y} c_{X,Y} = \mathrm{id}_{TX \times TY}$ [**?**, Lemma 4.2(i)].[1] In particular, $c_{X,Y}$ is a split mono and therefore mono.

To show that (3) is a pullback, we prove the universal property starting with a diagram

$$\begin{tikzcd} A \arrow[drrr, bend left, "{(f_1,f_2)}"] \arrow[dr, dashed, "\exists!"] \arrow[ddr, bend right, "{(g_1,g_2)}"'] & & & \\ & TX \times TY \times TZ \arrow[rr, "\mathrm{id} \times c_{Y,Z}"] \arrow[d, "c_{X,Y} \times \mathrm{id}"'] & & TX \times T(Y \times Z) \arrow[d, "c_{X,Y \times Z}"] \\ & T(X \times Y) \times TZ \arrow[rr, "c_{X \times Y,Z}"] & & T(X \times Y \times Z) \end{tikzcd} \tag{4}$$

where the dashed arrow will be constructed; its uniqueness is clear since $\mathrm{id} \times c_{Y,Z}$ and $c_{X,Y} \times \mathrm{id}$ are mono, so it remains to prove existence. Taking the unlabelled arrows to be (induced by) product projections, we have the commutative diagram

$$\begin{tikzcd} A \arrow[r, "{(g_1,g_2)}"] \arrow[d, "{(f_1,f_2)}"'] & T(X \times Y) \times TZ \arrow[r] & T(X \times Y \times Z) \arrow[d] \\ TX \times T(Y \times Z) \arrow[urr, "c_{X,Y \times Z}" description] \arrow[rr] & & T(Y \times Z) \end{tikzcd}$$

where the upper left triangle commutes by assumption, and the lower right triangle commutes by naturality of c with respect to the unique arrow $X \to 1$ together with $T1 \cong 1$ and the fact that $c_{1,Y \times Z}$ is a coherence isomorphism. By the naturality of c, f_2 can be written as the composite

$$A \xrightarrow{(g_1,g_2)} T(X \times Y) \times TZ \longrightarrow TY \times TZ \xrightarrow{c_{Y,Z}} T(Y \times Z).$$

[1] For probability monads, this equation can be interpreted as stating that the marginals of a product distribution are the original factors [**?**].

By analogous reasoning, we identify g_1 with the composite

$$A \xrightarrow{(f_1,f_2)} TX \times T(Y \times Z) \longrightarrow TX \times TY \xrightarrow{c_{X,Y}} T(X \times Y).$$

Getting back to (4), we take the dashed arrow to be the arrow whose component on TX is given by f_1, on TZ by g_2, and on TY by the diagonal in the diagram

$$\begin{tikzcd} A \arrow[r, "f_2"] \arrow[d, "g_1"'] & T(Y \times Z) \arrow[d] \\ T(X \times Y) \arrow[r] & TY \end{tikzcd}$$

which commutes for similar reasons as above. The fact that this arrow recovers the f_2 component after composition with $\mathrm{id} \times c_{Y,Z}$ and the g_1 component after composition with $c_{X,Y} \times \mathrm{id}$ follows by the expressions for f_2 and g_1 derived above. The fact that it recovers f_1 and g_2 is by construction.

◀

▶ **Remark 3.8.** We are not aware of any relation between weakly affine monads in our sense and Jacobs' *strongly affine* monads [**?**], other than the fact that strongly affine implies affine implies weakly affine.

▶ **Example 3.9.** We present a family of examples of commutative monads that are weakly affine but not affine. Let A be an abelian group (written multiplicatively). Then the functor $T_A \coloneqq A \times -$ on **Set** has a canonical structure of commutative monad, where the lax structure components $c_{X,Y}$ are given by multiplying elements in A while carrying the elements of X and Y along.

Since $T_A \cong A$, the monad T_A is clearly not affine unless A is the trivial group. However, T_A is always weakly affine. Indeed, in order to show that (3) is a pullback, it suffices to show that the associativity square of A

$$\begin{tikzcd} A \times A \times A \arrow[r, "\mathrm{id} \times \cdot"] \arrow[d, "\cdot \times \mathrm{id}"] & A \times A \arrow[d, "\cdot"] \\ A \times A \arrow[r, "\cdot"] & A \end{tikzcd}$$

is a pullback. Using element-wise reasoning, this amounts to showing that the system of equations $ax = c$ and $xb = d$ has a solution for $x \in A$ if and only if $cb = ad$, and in this case the solution is unique. But this is indeed the case with $x = a^{-1}c = db^{-1}$. (Note that this argument does not even require A to be abelian, but we need to require this in order for T_A to be commutative.)

▶ **Example 3.10.** Many monads in categorical measure theory are weakly affine but not affine. Let e.g. $M^* : \mathbf{Set} \to \mathbf{Set}$ be the monad assigning to every set the set of finitely supported discrete *nonzero* measures on M^*, or equivalently let $M^*(X)$ for any set X be the set of nonzero finitely supported functions $X \to [0, \infty)$. The monad structure is defined in terms of the same formulas as for the distribution monad on **Set** and the components $c_{X,Y}$ are also given by the formation of product measures, or equivalently point-wise products of functions $X \to [0, \infty)$.

Since $M^*1 \cong (0, \infty)$, this monad is clearly not affine. However, it is weakly affine, and we limit ourselves to a sketch of the proof. Indeed to prove that (3) is a pullback, we again reason in terms of elements. If all measures are normalised, then we are back in the situation of the distribution monad, which is affine and the claim follows. In the general case, one can reduce to the normalised case by showing that the normalisation of the desired element of $M^*(Y)$ is uniquely determined. This works in the same way as in Example 3.9 with $A = (0, \infty)$.

On the other hand, if the zero measure is included, then we obtain a commutative monad M which can be seen as the monad of semimodules for the semiring of nonnegative reals. Since $M1 \cong [0, \infty)$ is not a group under multiplication, M is not weakly affine.

The previous two examples and Lemma 3 suggest the following problem.

▶ **Problem 3.11.** Let T be a commutative monoid such that $T(1)$ is an abelian group. Does it follow that T is weakly affine?