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

[Electronic Notes in Theoretical Computer Science 341 (2018) 121–149](https://doi.org/10.1016/j.entcs.2018.11.007)

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

Bimonoidal Structure of Probability Monads

# Tobias Fritz[a](#_bookmark0)*,*[1](#_bookmark1) and Paolo Perrone[a](#_bookmark0)*,*[2](#_bookmark1)

a *Max Planck Institute for Mathematics in the Sciences Leipzig, Germany*

Abstract

We give a conceptual treatment of the notion of joints, marginals, and independence in the setting of cate- gorical probability. This is achieved by endowing the usual probability monads (like the Giry monad) with a monoidal and an opmonoidal structure, mutually compatible (i.e. a bimonoidal structure). If the under-

lying monoidal category is cartesian monoidal, a bimonoidal structure is given uniquely by a commutative strength. However, if the underlying monoidal category is not cartesian monoidal, a strength is not enough to guarantee all the desired properties of joints and marginals. A bimonoidal structure is then the correct requirement for the more general case.

We explain the theory and the operational interpretation, with the help of the graphical calculus for monoidal categories. We give a definition of stochastic independence based on the bimonoidal structure, compatible with the intuition and with other approaches in the literature for cartesian monoidal categories. We then show as an example that the Kantorovich monad on the category of complete metric spaces is a bimonoidal monad for a non-cartesian monoidal structure.

*Keywords:* Monoidal monads, probabilistic powerdomain, strong monads, randomness, stochastic correlation, Kantorovich duality, Wasserstein distance.

# Introduction

The standard way to treat randomness categorically is via a *probability monad*, of which classic examples are the Giry monad [[7]](#_bookmark38) and the probabilistic powerdomain [[10]](#_bookmark41). The interpretation is the following: let C be a category whose objects we think of as spaces of possible values that a variable may assume. A probability monad *P* on C makes it possible to talk about random variables on objects *X* ∈ C, or equivalently random elements of *X*: an element *p* ∈ *PX* specifies the *law* of a random variable on *X*.

A central theme of probability theory is that random variables can form joints and marginals. For this to make sense in C, we need C to be a monoidal category,

1 Email: [fritz@mis.mpg.de](mailto:fritz@mis.mpg.de)

2 Email: [perrone@mis.mpg.de](mailto:perrone@mis.mpg.de)

<https://doi.org/10.1016/j.entcs.2018.11.007>

1571-0661/© 2018 The Author(s). Published by Elsevier B.V.

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

and we need *P* to interact well with the monoidal structure. We argue that this interaction is best modelled in terms of a bimonoidal structure.

A first structure which links a monad with the tensor product in a category is that of a *strength*. A strength for a probability monad is a natural map *X* ⊗ *PY* → *P* (*X* ⊗ *Y* ), whose interpretation is the following: an element of *X* and a random element of *Y* determine uniquely a random element of *X* ⊗ *Y* which has the correct marginals, and whose randomness is all in the *Y* component. In the language of probability theory, (*x, q*) ∈ *X* ⊗ *PY* defines the product distribution of *δx* and *q* on *X* ⊗ *Y* . In the literature, the operational meaning of a strength for a monad, which includes the usage in probability, is well explained in [[16]](#_bookmark47), and in [[10]](#_bookmark41) for the case of the probabilistic powerdomain. A compendium of probability monads appearing in the literature, with information about their strength, can be found in [[9]](#_bookmark40).

The monoidal structure can be thought of as a refinement of the idea of strength. The basic idea is that given two probability measures *p* ∈ *PX* and *q* ∈ *PY* , one can canonically define a probability measure *p* ⊗∇ *q* ∈ *P* (*X* ⊗ *Y* ), the “product distribution” [3](#_bookmark2) . This is not the only possible joint distribution that *p* and *q* have, but it can be obtained without additional knowledge (of their correlation). When a strength satisfies suitable symmetry conditions (commutative strength) it defines automatically a monoidal structure [[11,](#_bookmark42)[8]](#_bookmark39).

An opmonoidal structure formalizes the dual intuition, namely that given a joint probability distribution *r* ∈ *P* (*X* ⊗ *Y* ) we canonically have the marginals on *PX* and *PY* as well. A bimonoidal structure is a compatible way of combining the two structures, in a way consistent with the usual properties of products and marginals in probability. When the underlying category is cartesian monoidal, then *P* is auto- matically opmonoidal. In this case, we show that if *P* carries a monoidal structure, then it is automatically bimonoidal. Therefore a commutative strong monad on a cartesian monoidal category is canonically bimonoidal. This is for example the case of the probabilistic powerdomain [[10]](#_bookmark41). We argue that the bimonoidal structure is the structure of relevance for probability theory: if the underlying category is not cartesian monoidal, or the strength is not commutative, then one cannot talk about joints and marginals in the usual way just by having a strong monad.

However, not every probability monad in the literature is bimonoidal, not even strong; a famous counterexample is in [[17]](#_bookmark48). While a non-bimonoidal probability monad could be of use in measure theory to talk about spaces of measures, it would be far from applications to probability, since it would not permit talking about concepts like stochastic independence and correlation, which in probability theory play a central role. We thus want to argue that in order for a monad to *really* count as a probability monad, it should be a bimonoidal monad.

In Section [2](#_bookmark3) we describe the setting of semicartesian monoidal categories and affine monads, which we argue is the one of relevance for classical probability theory. In such a setting, we will represent the concepts using a graphical calculus analogous

3 Our reason for denoting it by *p* ⊗∇ *q* rather than by *p* ⊗ *q* is that we want to interpret *p* : 1 → *PX* and *q* : 1 → *PY* as morphisms, so that *p* ⊗ *q* : 1 ⊗ 1 → *PX* ⊗ *PY* is not yet the product distribution. Rather, one needs to compose *p* ⊗ *q* with the monoidal structure ∇ : *PX* ⊗ *PY* → *P* (*X* ⊗ *Y* ), which is the subject of the present paper, see Section [3.2.](#_bookmark9)

to that of [[14]](#_bookmark45), presented in [2.1.](#_bookmark5) In Section [3](#_bookmark6) we will sketch the basic theory and interpretation of a bimonoidal structure for probability monads, using the graphical calculus. The same definitions in terms of commutative diagrams can be found in Appendix [A.](#_bookmark53) In [3.1,](#_bookmark8) we will show how this permits to talk about functions between products of random variables. In [3.2,](#_bookmark9) we show how to define a category of probability spaces from a probability monad, in such a way that the monoidal structure is inherited. This permits to connect with other treatments of stochastic independence in the literature. In [3.3](#_bookmark10) we will see in more detail why this formalism generalizes the strength of probability monads on cartesian monoidal categories. In Section [4,](#_bookmark14) we give a notion of stochastic independence based on the bimonoidal structure of the monad, and show that it satisfies some of the intuitively expected properties. In [4.1](#_bookmark18) we show that, if the base category is cartesian monoidal, our definition agrees with the one given by Franz [[5]](#_bookmark32), and it is compatible with the definition of independence structure given by Simpson [[18]](#_bookmark49). Finally, in Section [5](#_bookmark20) we will give a nontrivial example of a bimonoidal monad, the Kantorovich monad on complete metric spaces [[19,](#_bookmark50)[6]](#_bookmark37). The precise proofs and calculations of the statements of Section [5](#_bookmark20) can be found in Appendix [B.](#_bookmark55)

# Semicartesian monoidal categories and affine monads

By definition, a *semicartesian monoidal category* is a monoidal category in which the monoidal unit 1 is a terminal object. For probability theory, this is a very appealing feature of a category, because such an object can be interpreted as a trivial space, having only one possible state. In other words, the object 1 would

have the property that for every object *X*, *X* ⊗ 1 ∼= *X* (monoidal unit), so that tensoring with 1 does not increase the number of possible states, and moreover there

is a unique map ! : *X* → 1 (terminal object), which we can think of as “forgetting the state of *X*”. Cartesian monoidal categories are in particular semicartesian. Not every monoidal category of interest in probability theory is cartesian, but most of them are semicartesian (in particular, all the ones listed in [[9]](#_bookmark40)).

Semicartesian monoidal categories have another appealing feature for probabil- ity: every tensor product space comes equipped with natural projections onto its factors:

*X* ⊗ *Y X* ⊗ *Y*

id⊗!

!⊗id

∼=

*X* ⊗ 1 *X,*

∼=

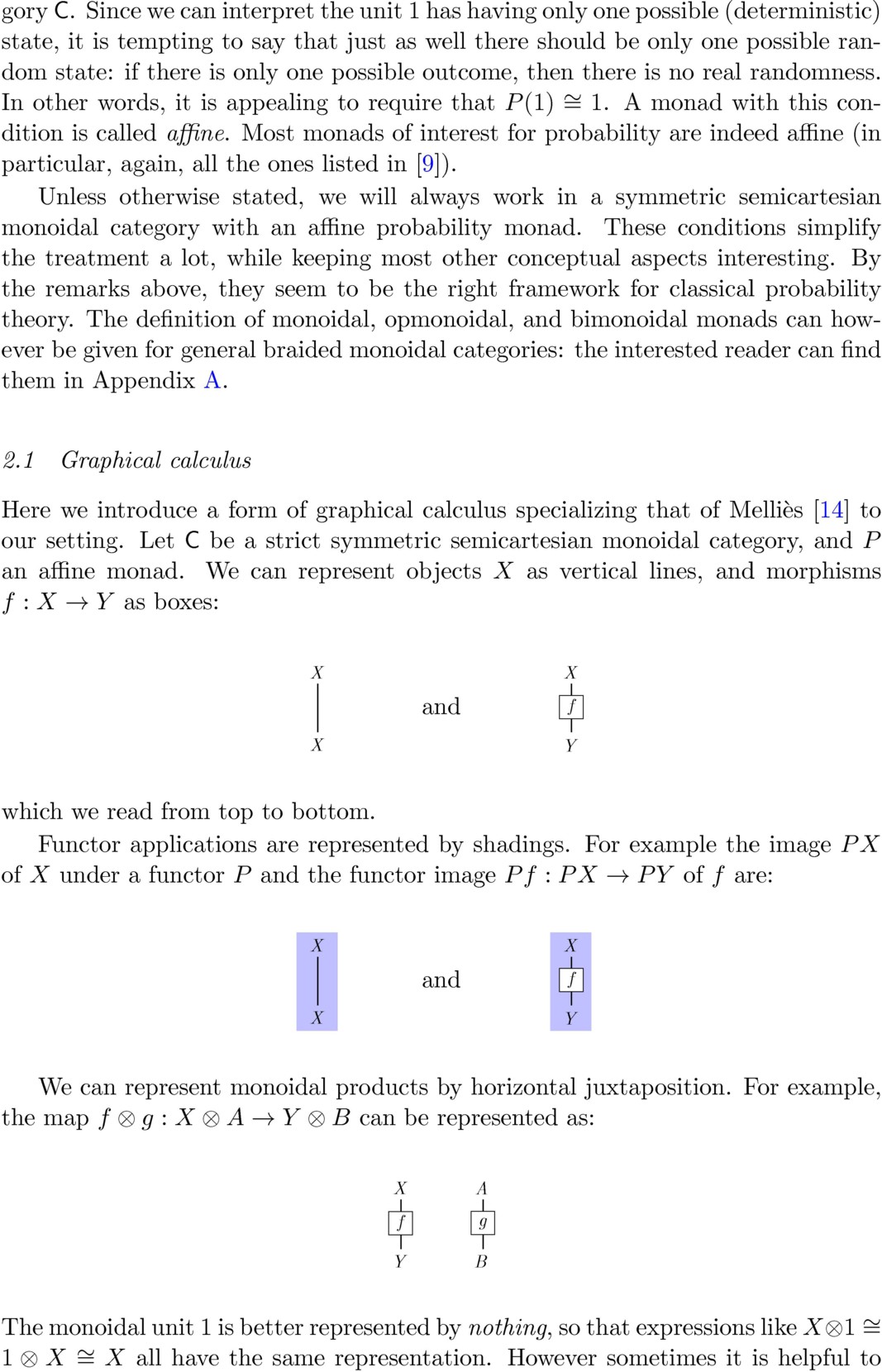
1 ⊗ *Y Y,*

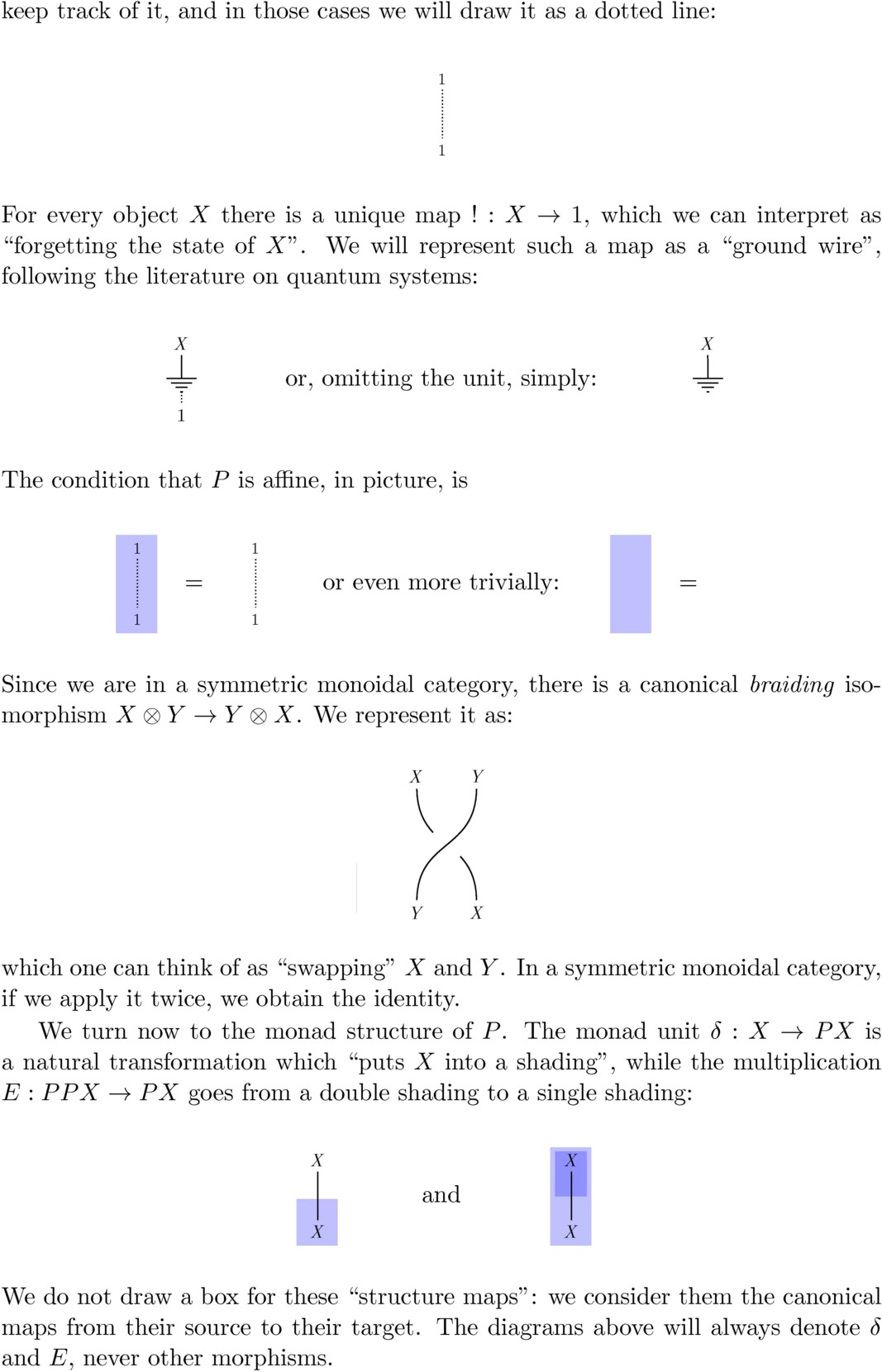
which satisfy the universal property of the product projections if and only if the

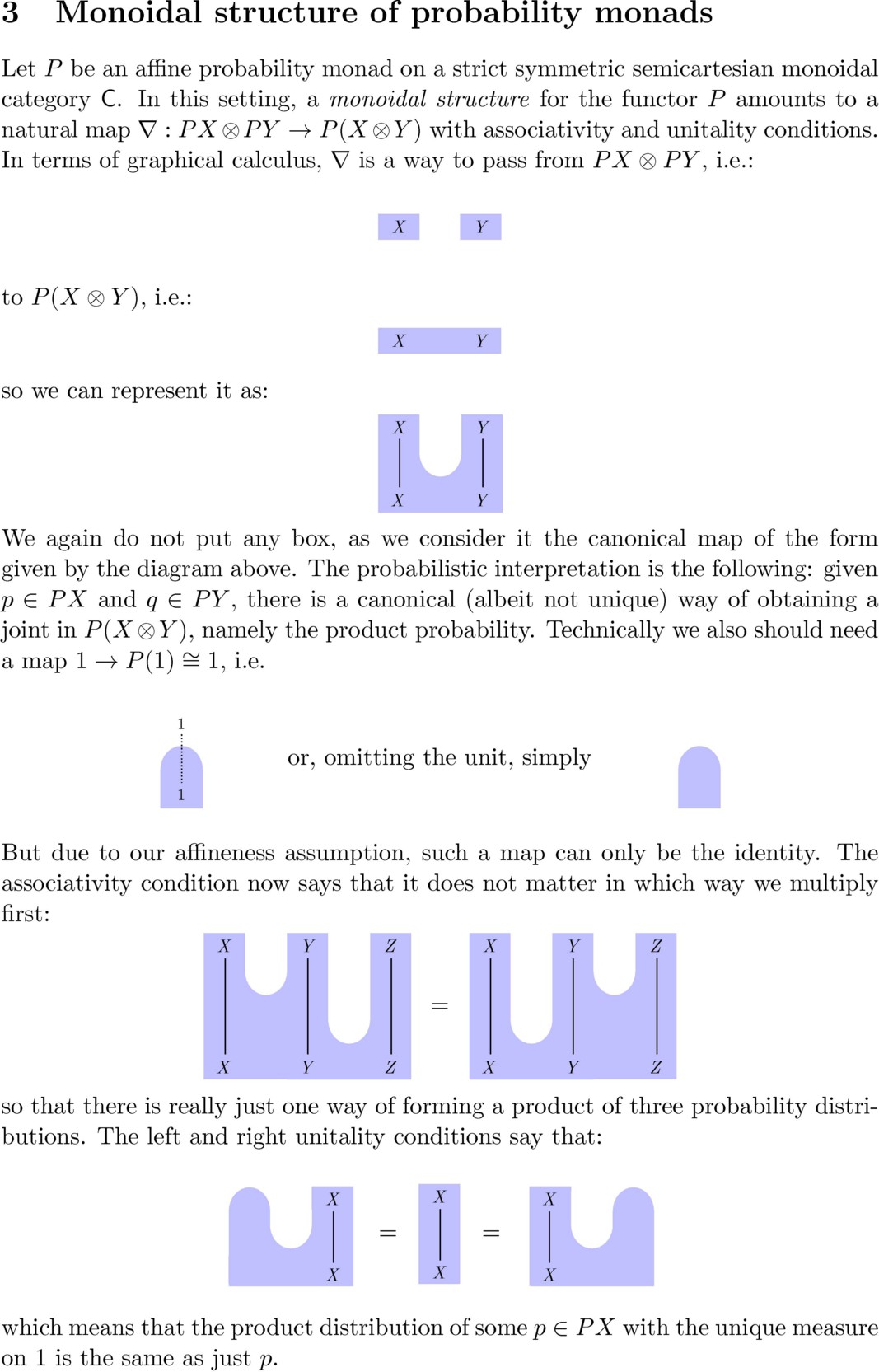
category is cartesian monoidal. These maps are important in probability theory, because they give the *marginals*. Since these projections are automatically natural in *X* and *Y* , a semicartesian monoidal category is always equivalently a *tensor category with projections* in the sense of [[5,](#_bookmark32) Definition 3.3]; see [[13]](#_bookmark44) for more background.

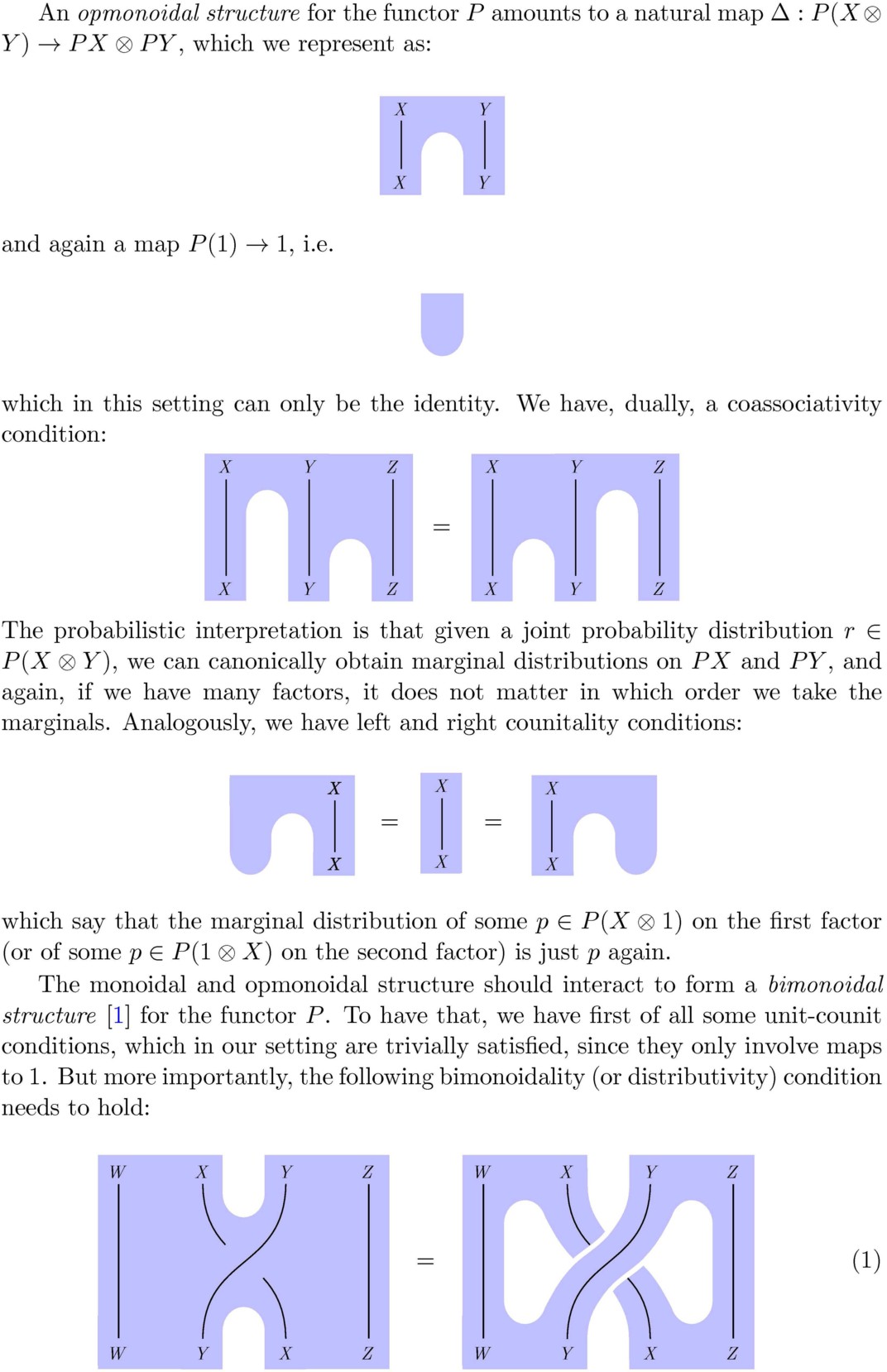
Suppose now that *P* is a probability monad [4](#_bookmark4) on a semicartesian monoidal cate-

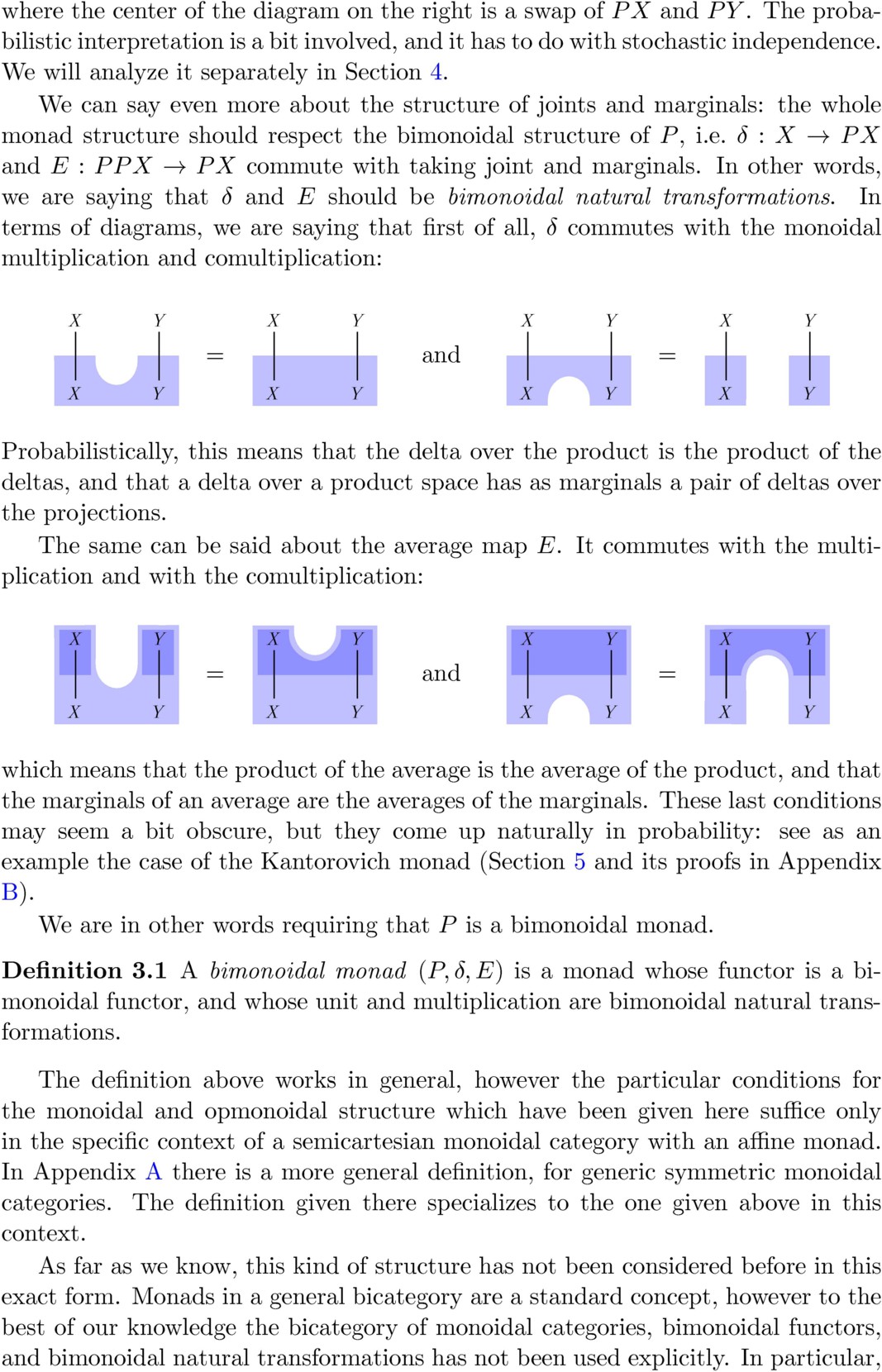
4 In this work, “probability monad” is not a technical term: any monad could be in principle considered a probability monad. We merely use this term in order to indicate our intended interpretation in terms of randomness, as in the case of the Giry monad or the probabilistic powerdomain.

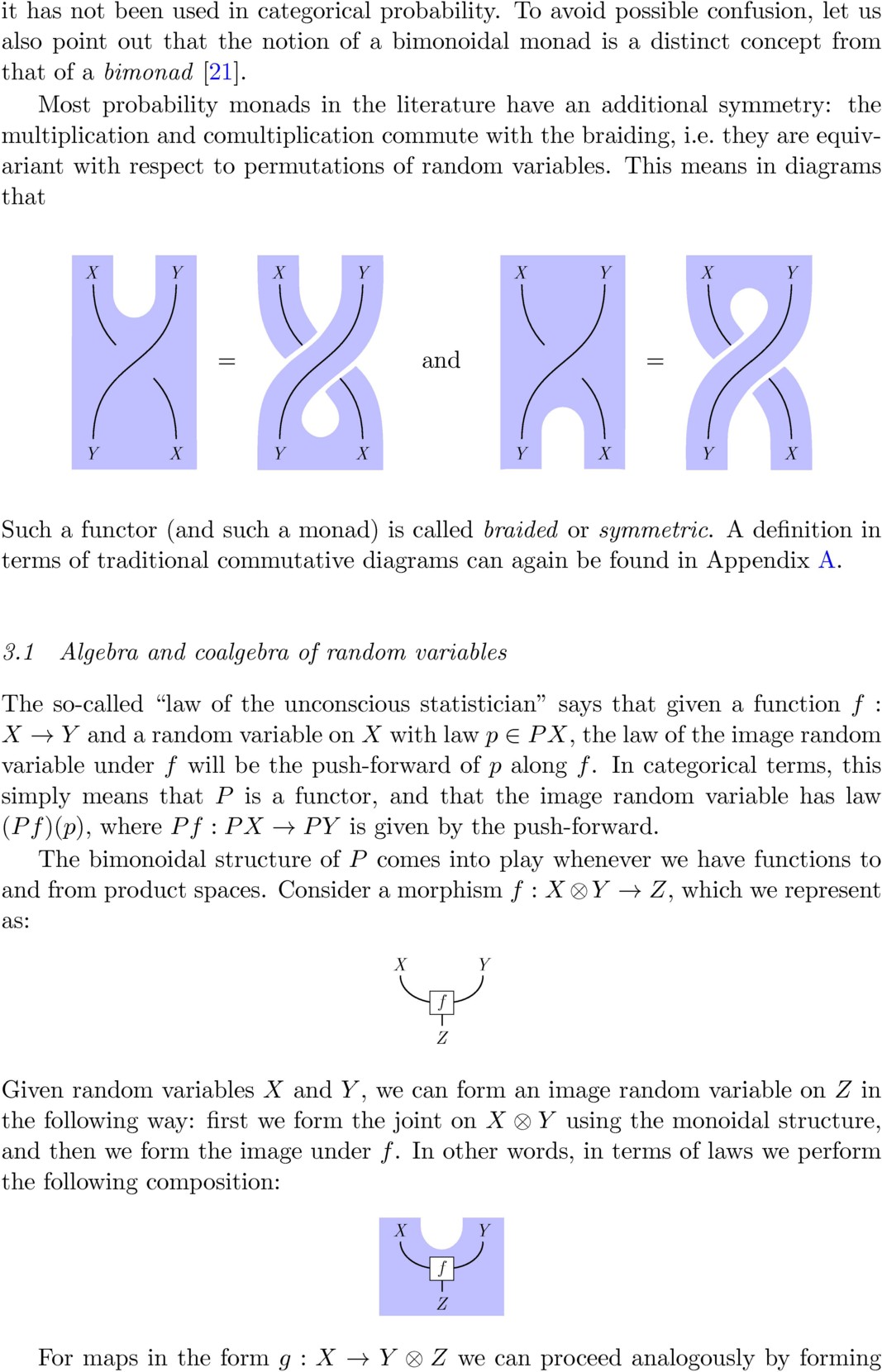


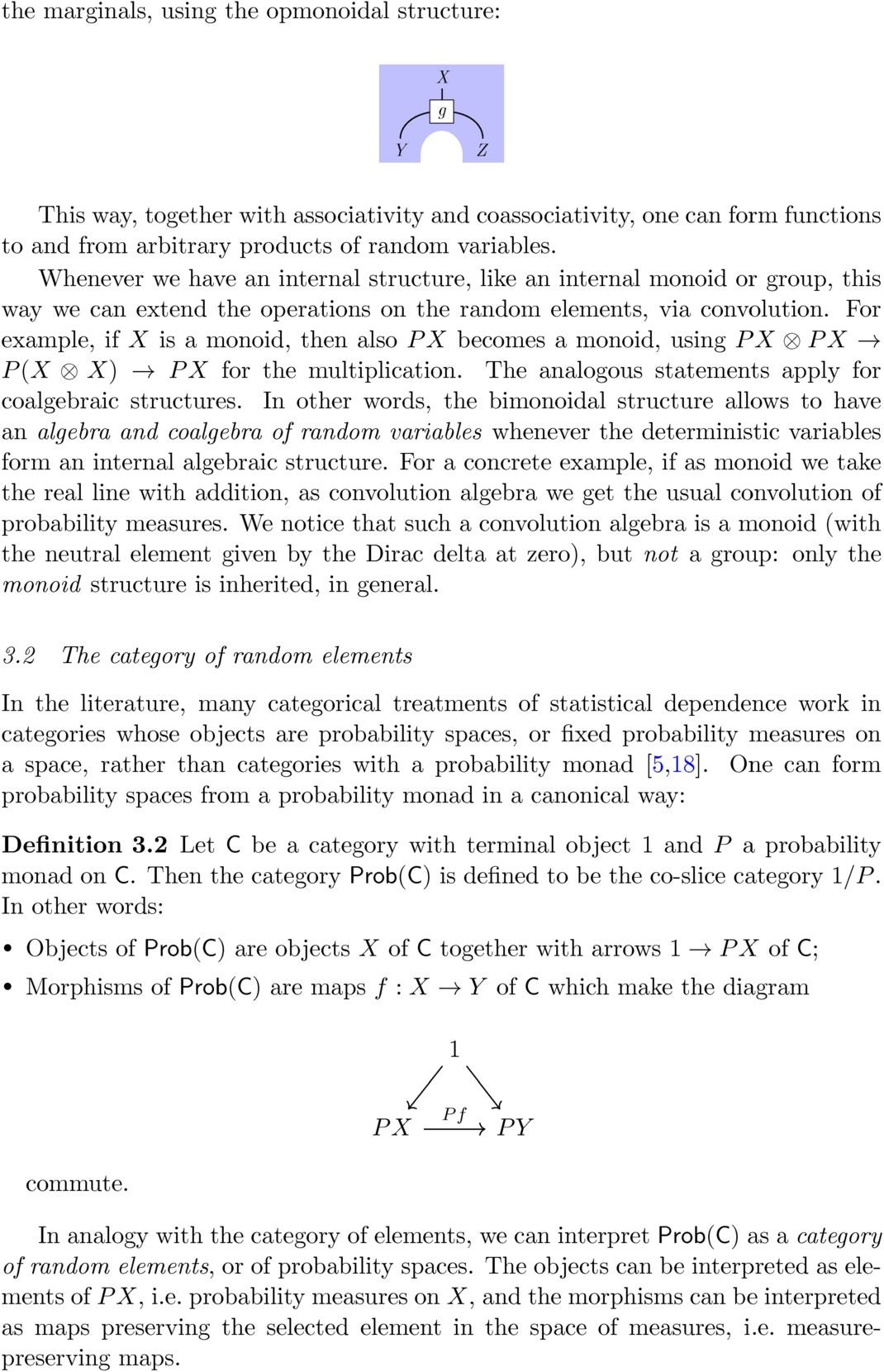












Under some mild assumptions, if C has a semicartesian monoidal structure we can transfer that structure to the category of random elements, with a construction analogous to that of Section [3.1.](#_bookmark8)

probability monad on C with monoidal structure ∇. We define the following Definition 3.3 Let C be a semicartesian monoidal category and *P* an affine monoidal structure on Prob(C): given *p* : 1 → *PX* and *q* : 1 → *PY* , we define *p* ⊗∇ *q* : 1 → *P* (*X* ⊗ *Y* ) to be the composition:

1 ∼= 1 ⊗ 1

*p*⊗*q*

*PX* ⊗ *PY*

∇ *P* (*X* ⊗ *Y* )*.*

and for morphisms we proceed analogously.

This way (Prob(C)*,* ⊗∇) is a semicartesian monoidal category, with the unit 1 → 1 isomorphic to the terminal object. In other words, it is always a tensor category with projections in the sense of [[5]](#_bookmark32), generalizing the construction given

in Section 3.1 therein (in which the base category Meas is cartesian monoidal). In general (and in all interesting cases in the literature), Prob(C) equipped with this monoidal structure is *not cartesian monoidal*, not even if C is: the product probability does not satisfy the universal property of a categorical product (see for example [[5]](#_bookmark32) for a discussion on this). [5](#_bookmark11)

Some of the upcoming results will refer to Prob(C), whose objects we also call

notation *p* ⊗∇ *q* for the product probability. *laws*, as they generalize laws of random variables. In particular we will use the

*3.3 Bimonoidal monads on a cartesian monoidal category*

Suppose now that the monoidal structure of C is *cartesian* monoidal, i.e. that the monoidal product is given by the categorical product (so, in particular, C is semi-

cartesian). The projection maps *π*1 : *X* × *Y* → *X* and *π*2 : *X* × *Y* → *Y* now satisfy a universal property. Let’s now apply *P* , so that we get maps *Pπ*1 : *P* (*X* × *Y* ) → *PX* and *Pπ*2 : *P* (*X* × *Y* ) → *PY* . By the universal property of the product, there is then a *unique* map *P* (*X* × *Y* ) → *PX* × *PY* compatible with the projections, i.e. making the following diagram commute:

*Pπ*1

*PX π*1

*P* (*X* × *Y* )

*PX* × *PY*

*Pπ*2

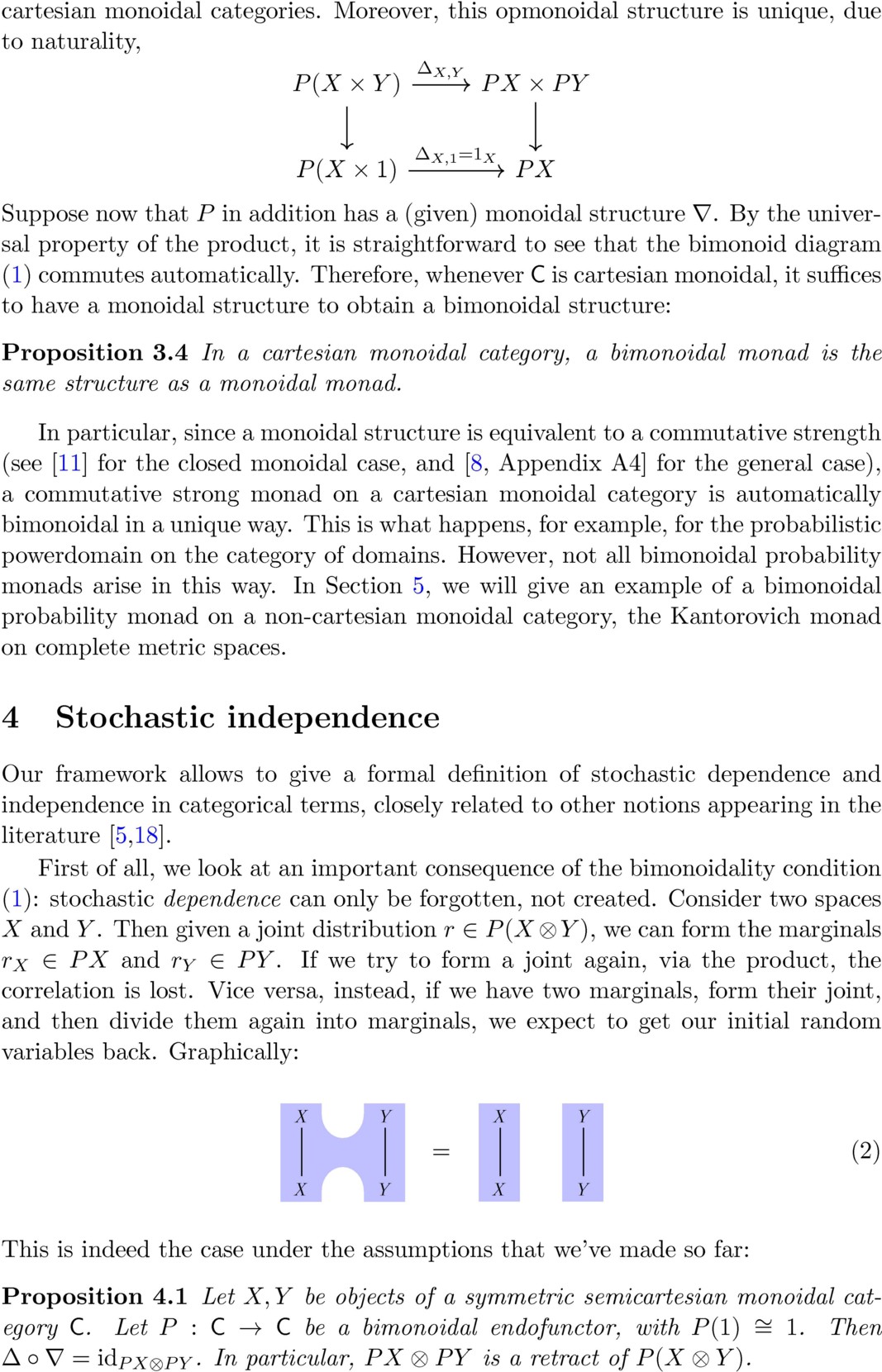
*π*2 *PY*

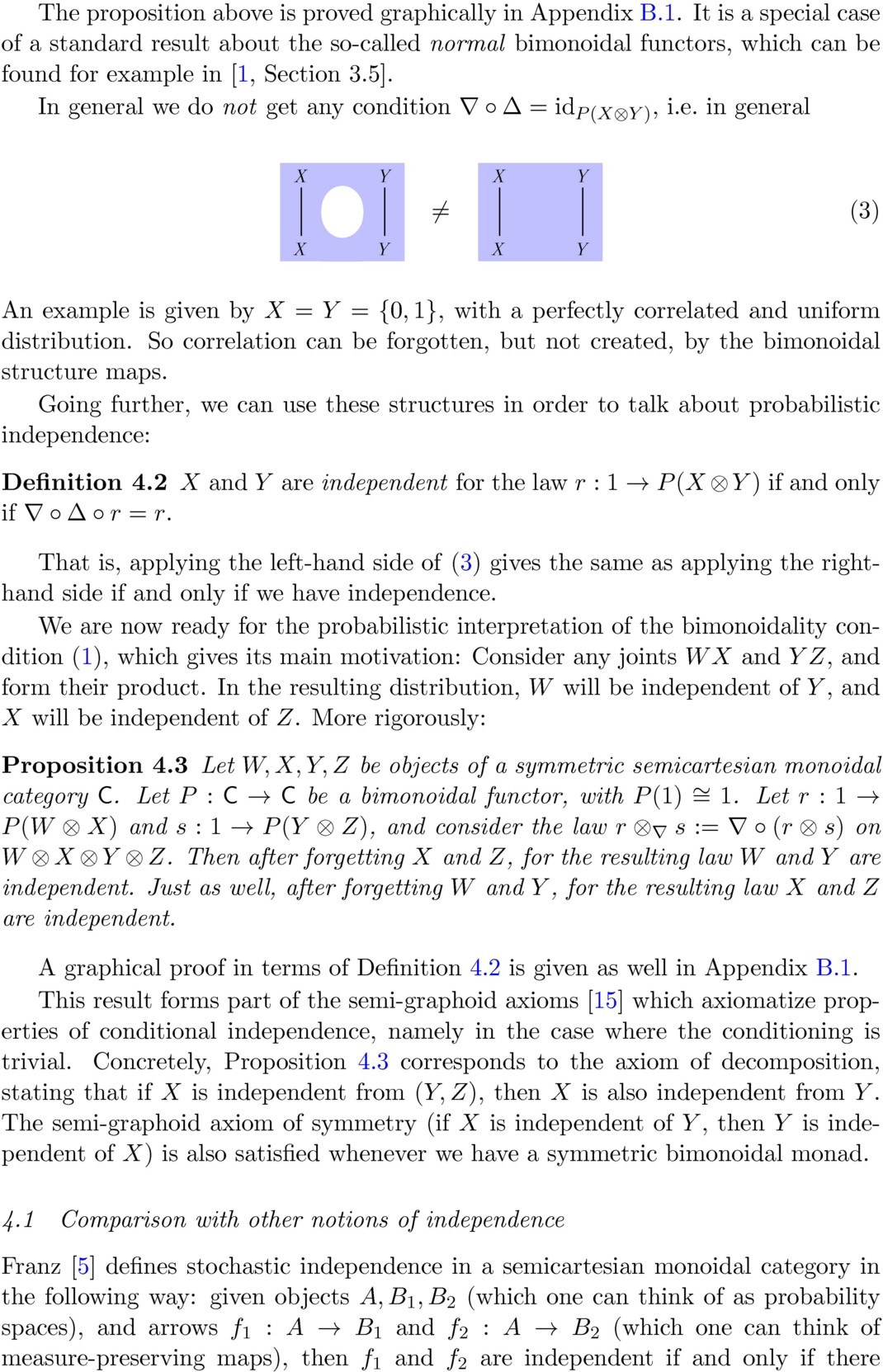
This gives a natural map ∆ : *P* (*X* × *Y* ) → *PX* × *PY* . Such a map exists and is

unique for any (finite) number of factors, so it is automatically associative. There-

fore *P* has a canonical opmonoidal structure. This is true for all functors *P* between

5 The intuitive idea that “the product probability has the same information as the pair of marginals” can be made rigorous in a different manner, see Section [4](#_bookmark14).





exists *h* : *A* → *B*1 ⊗ *B*2 making this diagram commute:

*A*

*f*1 *f*2

*h*

(4)

*B*1 *π*1

*B*1 ⊗ *B*2

*π*2 *B*2

where *π*1*, π*2 are the projections of the tensor product. He then proves [[5,](#_bookmark32) Proposi- tion 3.5] that in the category Prob of (traditional) probability spaces, this notion of independence is equivalent to the standard one of probability theory. We propose a generalization of that result, which holds for categories of random elements obtained by generic *cartesian* monoidal categories.

Proposition 4.4 *Let* C *be a cartesian monoidal category and P an affine bi- monoidal probability monad. Consider a law s* : 1 → *PA and maps f*1 : *A* → *B*1 *and f*2 : *A* → *B*2*. Then f*1 *and f*2 *are independent in the sense of Franz [*[*5*](#_bookmark32)*] if and only if B*1 *and B*2 *are independent for the law P* (*f*1*, f*2) ◦ *s in the sense of Definition* [*4.2*](#_bookmark16)*.*

So in the case of cartesian monoidal base categories, the two approaches agree. The proof can be found in Appendix [B.2,](#_bookmark58) and goes along the lines of the proof of [[5,](#_bookmark32) Proposition 3.5].

Simpson [[18]](#_bookmark49) defines an *independence structure* as a certain collection of multi- spans that contains the singleton families. Given again a *cartesian* monoidal cate- gory C and an affine monad *P* on *C*, and given a finite multispan {*fi* : *A* → *Bi*}*i*∈*I* in C, we can form a multispan in the category Prob(C) by precomposing with a law *r* : 1 → *PA*. We can call such a resulting multispan *independent*, in analogy with Definition [4.2,](#_bookmark16) iff

∇*I* ◦ ∆*I* ◦ *P* ((*fi*)*i*∈*I* ) ◦ *r* = *P* ((*fi*)*i*∈*I* ) ◦ *r,*

where (*f* ) : *A* → *B* is the tupling of the *f* given by the cartesian monoidal structure, and ∇ and ∆ are the maps respectively (*PB* ) → *P* ( *B* ) and *P* ( *i Bi*) → *i*(*PBi*) obtained by iterating respectively ∇ and ∆ (by associativity and coassociativity, the resulting maps are unique). Independent multispans defined in this way form then an independence structure in the sense of [[18,](#_bookmark49) Definition 2.1], in a way analogous to Examples 2.1 and 2.25 therein: they are closed with respect to multispan composition, and to forming subfamilies. Therefore, again in the case of a cartesian monoidal base category, our definition is compatible with Simpson’s approach.

*i i*∈*I* Q*i*∈*I i* *i*

*I I* Q*i i* Q*i i*

Q Q

# 5 Bimonoidal structure of the Kantorovich monad

The Kantorovich monad is a probability monad on complete metric spaces. It was first defined by van Breugel for compact and for complete 1-bounded metric spaces [[19]](#_bookmark50). We will use here the definitions and results of [[6]](#_bookmark37), which work for all complete metric spaces.

Consider the category CMet whose:

* Objects are complete metric spaces;
* Morphisms are short maps, i.e. functions *f* : *X* → *Y* such that

*d* *f* (*x*)*, f* (*x*′) ≤ *d*(*x, x*′)

for all *x, x*′ ∈ *X*;

* As monoidal structure, we define *X* ⊗ *Y* to be the set *X* × *Y* , with the metric:

*d* (*x, y*)*,* (*x*′*, y*′) := *d*(*x, x*′) + *d*(*y, y*′)*.*

This category can be thought of as a category of enriched categories and functors [[12,](#_bookmark43) Section 2], and the monoidal structure is closed but not cartesian. Further motivation for the choice of this category is given in [[6]](#_bookmark37). In particular, by choosing as morphisms the short maps, one can obtain *PX* as a colimit of spaces of empirical distributions of finite sequences [[6,](#_bookmark37) Section 3], which would not be possible if one allowed for more general morphisms (like continuous or Lipschitz functions).

We recall the basic definitions of [[6]](#_bookmark37).

Definition 5.1 Let *X* be a complete metric space.

* A Radon probability measure *p* on *X* is said to have *finite first moment* if for every short map *f* : *X* → R,

∫

*f dp <* ∞*.*

*X*

Every such probability measure can be specified uniquely by its integration against short maps to R: the set of such measures can be identified with the set of positive, Scott-continuous linear functionals on the space of Lipschitz func- tions on *X*. Hence, in the following, we explicitly construct such measures by specifying their action on short maps.

* The *Kantorovich-Wasserstein* space *PX* is the space of all Radon probability measures on *X* with finite first moment, equipped with the metric:

*d*(*p, q*) := sup ∫ *f dp* − ∫

*f dq* *,*

*f* :*X*→R *X* *X*

where the supremum ranges over all short maps *X* → R. With this metric, *PX*

is itself a complete metric space.

* Given *f* : *X* → *Y* , we define *Pf* : *PX* → *PY* as the map assigning to *p* ∈ *PX* its push-forward measure (*Pf* )(*p*) := *f*∗*p* ∈ *PY* . The latter is defined by saying that for all *g* : *Y* → R short,

∫*Y g d*(*f*∗*p*) := ∫*X g* ◦ *f dp.*

*f*∗*p* also has finite first moment, and this assignment makes *P* into a functor.

A concise treatment of Wasserstein spaces can be found in [[2]](#_bookmark34) and a more com- prehensive one in [[20]](#_bookmark51). For the basic measure-theoretic setting, we refer the reader to [[3,](#_bookmark35)[4]](#_bookmark36).

The functor *P* admits a monad structure, with the unit *δ* : *X* → *PX* given by the Dirac distributions

∫

*f* (*y*) *d*(*δ*(*x*))(*y*) := *f* (*x*)*,*

*X*

and the multiplication *E* : *PPX* → *PX* given by forming the expected or average distribution,

∫

∫

∫

*f d*(*Eµ*) :=

*X PX*

*X*

*f* (*x*) *dp*(*x*) *dµ*(*p*)*.*

We can now define product joints and marginals, which will equip *P* with a bimonoidal structure.

Definition 5.2 Let *p* ∈ *PX, q* ∈ *PY* . We denote *p* ⊗∇ *q* the joint probability measure on *X* ⊗ *Y* defined by:

∫ ∫

*f* (*x, y*) *d*(*p* ⊗∇ *q*)(*x, y*) := *f* (*x, y*) *dp*(*x*) *dq*(*y*)*.*

*X*⊗*Y X*⊗*Y*

Let now *r* ∈ *P* (*X* ⊗ *Y* ). We denote (*rX* ) the marginal probability on *X* defined

by:

∫

∫

*f* (*x*) *drX* (*x*) := *f* (*x*) *dr*(*x, y*)*.*

*X X*⊗*Y*

The marginal on *Y* is defined analogously.

It is straightforward to check that the functionals defined in Definition [5.2](#_bookmark21) are positive, linear, and Scott-continuous, therefore they specify uniquely Radon prob- ability measures of finite first moment.

In the rest of this section we will show that the joints and marginals in Definition

[5.2](#_bookmark21) equip the Kantorovich monad on CMet with a bimonoidal monad structure (Theorem [5.15](#_bookmark31)). The proofs with the actual calculations are in Appendix [B.](#_bookmark55)

We will prove now that the product joint construction equips *P* with a monoidal structure.

Definition 5.3 Let *X, Y* ∈ CMet. We define the map ∇ : *PX* ⊗ *PY* → *P* (*X* ⊗ *Y* ) as mapping (*p, q*) ∈ *PX* ⊗ *PY* to the joint *p* ⊗∇ *q* ∈ *P* (*X* ⊗ *Y* ).

Proposition 5.4 ∇ : *PX* ⊗ *PY* → *P* (*X* ⊗ *Y* ) *is short.*

Therefore, ∇ is a morphism of CMet. This would not be the case if we took as monoidal structure for CMet the cartesian product: for the product metric, ∇ is Lipschitz, but in general not short. The fact that ∇ equips *P* with a monoidal structure now follows directly from the naturality and associativity of the product probability construction (as sketched in Section [3](#_bookmark6)). In other words, the proofs of the next three statements (see Appendix [B.3](#_bookmark57)) can be adapted to most other categorical contexts in which the map ∇ is of a similar form.

Proposition 5.5 ∇ : *PX* ⊗ *PY* → *P* (*X* ⊗ *Y* ) *is natural in X and Y .* Proposition 5.6 (*P,* id1*,* ∇) *is a symmetric lax monoidal functor* CMet → CMet*.* Proposition 5.7 (*P, δ, E*) *is a symmetric monoidal monad.*

We know that a monoidal monad is the same as a commutative monad, and therefore obtain:

Corollary 5.8 *P is a commutative strong monad, with strength X* ⊗ *PY* → *P* (*X* ⊗

*Y* ) *given by:*

(*x, q*) '→ *δx* ⊗∇ *q* ∈ *P* (*X* ⊗ *Y* )*.*

We now turn to the analogous statements for the marginals, and show that they equip *P* with an opmonoidal structure.

Definition 5.9 Let *X, Y* ∈ CMet. We define the map ∆ : *P* (*X* ⊗ *Y* ) → *PX* ⊗ *PY*

as mapping *r* ∈ *P* (*X* ⊗ *Y* ) to the pair of marginals (*rX, rY* ) ∈ *PX* ⊗ *PY* .

Proposition 5.10 ∆ : *P* (*X* ⊗ *Y* ) → *PX* ⊗ *PY is short.*

Therefore ∆ is a morphism of CMet. Again, the following statements follow just from the properties of marginals, and their proofs (see Appendix [B.4](#_bookmark59)) can be adapted to most other categorical contexts provided that ∆ is of a similar form.

Proposition 5.11 ∆ : *P* (*X* ⊗ *Y* ) → *PX* ⊗ *PY is natural in X, Y .*

Proposition 5.12 *The marginal map together with the trivial counitor defines a symmetric oplax monoidal functor* (*P,* id1*,* ∆)*.*

Proposition 5.13 (*P, δ, E*) *is a symmetric opmonoidal monad.*

The lax and oplax monoidal structure interact to give a bimonoidal structure. The following statements also follow just from the properties of joints and marginals.

Proposition 5.14 *P is a symmetric bilax monoidal functor.*

The main result then just follows as a corollary:

Theorem 5.15 *The Kantorovich monad is a symmetric bimonoidal monad, with monoidal structure given by the product joint, and opmonoidal structure given by the marginals.*

By Proposition [4.1,](#_bookmark13) we therefore have:

Corollary 5.16 ∆*X,Y* ◦ ∇*X,Y* = id*P X*⊗*P Y . Therefore, the inclusion* ∇ *of product measures into general joints, is an isometric embedding for the Kantorovich metric, and its image is a retract of the space of all joints.*

# Acknowledgement

The authors would like to thank the anonymous reviewers for the very helpful comments.

# References

1. Aguiar, M. and S. Mahajan, “Monoidal Functors, Species and Hopf Algebras,” AMS, 2010.
2. Basso, G., *A* *Hitchhiker’s guide to Wasserstein distances* (2015), [Available at: http://n.ethz.ch.](http://n.ethz.ch/~gbasso/download/A%20Hitchhikers%20guide%20to%20Wasserstein/A%20Hitchhikers%20guide%20to%20Wasserstein.pdf)
3. Bogachev, V. I., “Measure Theory,” Springer, 2000.
4. Edgar, G. A., “Integral, Probability, and Fractal Measures,” Springer, 1998.
5. Franz, U., *What is stochastic independence?*, Non-commutativity, infinite-dimensionality and probability at the crossroads. Proceedings of the RIMS workshop on infinite-dimensional analysis and quantum probability. (2001).
6. Fritz, T. and P. Perrone, *A Probability Monad as the Colimit of Finite Powers* (2017), submitted. [arXiv:1712.05363.](https://arxiv.org/abs/1712.05363)
7. Giry, M., *A* *Categorical Approach to Probability Theory*, in: *Categorical aspects of topology and analysis*, Lecture Notes in Mathematics 915, 1982 .
8. Golbeaut-Larrecq, J., S. Lasota and D. Nowak, *Logical relations for monadic types*, Mathematical Structures in Computer Science 18 (2008), pp. 1169–1217.
9. Jacobs, B., *From Probability Monads to Commutative Effectuses*, Journ. of Logical and Algebraic Methods in Programming (2017), in press. Available at <http://www.cs.ru.nl/B.Jacobs/PAPERS/probability-monads.pdf>.
10. Jones, C. and J. D. Plotkin, *A Probabilistic Powerdomain of Evaluations*, Proceedings of the Fourth Annual Symposium of Logics in Computer Science (1989).
11. Kock, A., *St**rong functors and monoidal monads*, Arch. Math. (Basel) 23 (1972), pp. 113–120.
12. Lawvere, W., *Metric spaces, generalized logic and closed categories*, Rendiconti del seminario matematico e fisico di Milano 43 (1973).

URL <http://www.tac.mta.ca/tac/reprints/articles/1/tr1abs.html>

1. Leinster, T., *Monoidal categories with projections* (2016), *n*-Category Caf¨e blog post, [https://golem.ph.utexas.edu/category/2011/05/an operadic introduction to en.html.](https://golem.ph.utexas.edu/category/2016/08/monoidal_categories_with_proje.html)
2. Melli`es, P.-A., *Functorial Boxes in String Diagrams*, International Workshop on Computer Science Logic (2006).
3. Pearl, J. and A. Paz, *Graphoids: a graph-based logic for reasoning about relevance relations*, UCLA Computer Science Dept., Technical Report 850038 (1985).
4. Plotkin, G. D. and J. Power, *Notion of Computation Determine Monads* (2002).
5. Sato, T., *The Giry Monad is not strong for the canonical symmetric monoidal closed structure on Meas* (2016), [arXiv:1612.05939.](https://arxiv.org/abs/1612.05939)

URL <http://arxiv.org/abs/1612.05939>

1. Simpson, A., *Category-theoretic structure for independence and conditional independence*, Electronic Notes in Theoretical Computer Science 336 (2018), pp. 281–297.
2. van Breugel, F., *The Metric Monad for Probabilistic Nondeterminism* (2005), [Available at](http://www.cse.yorku.ca/~franck/research/drafts/monad.pdf) [http://www.cse.yorku.ca](http://www.cse.yorku.ca/~franck/research/drafts/monad.pdf).
3. Villani, C., “Optimal transport: old and new,” Grundlehren der mathematischen Wissenschaften 338, Springer, 2009.
4. Willerton, S., *A diagrammatic approach to Hopf monads*, Arabian Journal of Science and Engineering C 33 (2008).

# A Monoidal, opmonoidal and bimonoidal monads

We recall the definition of the different monoidal structures for a functor, for the case of braided (including symmetric) monoidal categories. For more results and more general definitions, we refer to [[1]](#_bookmark33).

Let (C*,* ⊗) and (D*,* ⊗) be braided monoidal categories.

Definition A.1 A *lax monoidal functor* (C*,* ⊗) → (D*,* ⊗) is a triple (*F, η,* ∇), such that:

1. *F* : *C* → *D* is a functor;
2. The “unit” *η* : 1D → *F* (1C) is a morphism of D;
3. The “multiplication” ∇ : *F* (−)⊗*F* (−) ⇒ *F* (−⊗−) is a natural transformation of functors C×C → D;
4. The following “associativity” diagram commutes for every *X, Y, Z* in C:

∼=

(*FX* ⊗ *FY* ) ⊗ *FZ*

∇*X,Y* ⊗id

*FX* ⊗ (*FY* ⊗ *FZ*)

id⊗∇*Y,Z*

*F* (*X* ⊗ *Y* ) ⊗ *FZ FX* ⊗ *F* (*Y* ⊗ *Z*)

∇*X*⊗*Y,Z*

∼=

*F* ((*X* ⊗ *Y* ) ⊗ *Z*)

∇*X,Y* ⊗*Z*

*F* (*X* ⊗ (*Y* ⊗ *Z*))

1. The following “unitality” diagrams commute for every *X* in C:

1D ⊗ *FX*

∼=

*FX*

*η*⊗id

∼=

*F* (1C) ⊗ *FX*

∇1C *,X*

*F* (1C ⊗ *X*)

*FX* ⊗ 1D

∼=

*FX*

id⊗*η*

∼=

*FX* ⊗ *F* (1C)

∇*X,*1C

*F* (*X* ⊗ 1C)

We say that (*F, η,* ∇) is also *braided*, or *symmetric* if C is symmetric, if in addition the multiplication commutes with the braiding:

*FX* ⊗ *FY*

∼=

*FY* ⊗ *FX*

∇

∼=

*F* (*X* ⊗ *Y* )

∇

*F* (*Y* ⊗ *X*)

Definition A.2 Let (*F, ηF ,* ∇*F* ) and (*G, ηG,* ∇*G*) be lax monoidal functors (C*,* ⊗) → (D*,* ⊗). A *lax monoidal natural transformation*, or just *monoidal natural transformation* when it’s clear from the context, is a natural transformation *α* : *F* ⇒ *G* which is compatible with the unit and multiplication map. In particular, the following diagrams must commute (for all *X, Y* ∈ C):

1D *F* (1C)

*ηF*

*ηG*

*FX* ⊗ *FY*

∇*F*

*F* (*X* ⊗ *Y* )

*α*1C

*αX* ⊗*αY*

*αX*⊗*Y*

*G*(1C)

*GX* ⊗ *GY*

∇*G*

*G*(*X* ⊗ *Y* )

Definition A.3 An *oplax monoidal functor* (C*,* ⊗) → (D*,* ⊗) is a triple (*F, ǫ,* ∆), such that:

1. *F* : *C* → *D* is a functor;
2. The “counit” *ǫ* : *F* (1C) → 1D is a morphism of D;
3. The “comultiplication” ∆ : *F* (− ⊗ −) ⇒ *F* (−) ⊗ *F* (−) is a natural transformation of functors

C × C → D;

1. The following “coassociativity” diagram commutes for every *X, Y, Z* in C:

∼=

*F* ((*X* ⊗ *Y* ) ⊗ *Z*)

*F* (*X* ⊗ (*Y* ⊗ *Z*))

∆*X*⊗*Y,Z* ∆*X,Y* ⊗*Z*

*F* (*X* ⊗ *Y* ) ⊗ *FZ FX* ⊗ *F* (*Y* ⊗ *Z*)

∆*X,Y* ⊗id

(*FX* ⊗ *FY* ) ⊗ *FZ*

id⊗∆*Y,Z*

∼=

*FX* ⊗ (*FY* ⊗ *FZ*)

1. The following “counitality” diagrams commute for every *X* in C:

∆1C *,X*

*F* (1C ⊗ *X*)  *F* (1C) ⊗ *FX*

∆*X,*1C

*F* (*X* ⊗ 1C)  *FX* ⊗ *F* (1C)

∼=

*FX* ∼=

*ǫ*⊗id

1D ⊗ *FX*

∼=

*FX* ∼=

id⊗*ǫ*

*FX* ⊗ 1D

We say that (*F, ǫ,* ∆) is also *braided*, or *symmetric* if C is symmetric, if in addition the comultiplication commutes with the braiding:

∼=

*F* (*X* ⊗ *Y* )

*F* (*Y* ⊗ *X*)

∆

*FX* ⊗ *FY*

∆

∼=

*FY* ⊗ *FX*

Definition A.4 Let (*F, ǫF ,* ∆*F* ) and (*G, ǫG,* ∆*G*) be oplax monoidal functors (C*,* ⊗) → (D*,* ⊗). An *oplax monoidal natural transformation*, or just *monoidal natural transformation* when it’s clear from the context, is a natural transformation *α* : *F* ⇒ *G* which is compatible with the counit and comultiplication map. In particular, the following diagrams must commute (for all *X, Y* ∈ C):

1D *F* (1C)

*ǫF*

*ǫG*

*FX* ⊗ *FY*

∆*F*

*F* (*X* ⊗ *Y* )

*α*1C

*αX* ⊗*αY*

*αX*⊗*Y*

*G*(1C)

*GX* ⊗ *GY*

∆*G*

*G*(*X* ⊗ *Y* )

Definition A.5 A *bilax monoidal functor* (C*,* ⊗) → (D*,* ⊗) is a “quintuplet” (*F, η,* ∇*, ǫ,* ∆) such that:

1. (*F, η,* ∇) : (C*,* ⊗) → (D*,* ⊗) is a lax monoidal functor;
2. (*F, ǫ,* ∆) : (C*,* ⊗) → (D*,* ⊗) is an oplax monoidal functor;
3. The following “bimonoidality” diagram commutes:

*F* (*W* ⊗ *X*) ⊗ *F* (*Y* ⊗ *Z*)

∇*W* ⊗*X,Y* ⊗*Z*

∆*W,X* ⊗∆*Y,Z*

*F* (*W* ⊗ *X* ⊗ *Y* ⊗ *Z*) *F* (*W* ) ⊗ *F* (*X*) ⊗ *F* (*Y* ) ⊗ *F* (*Z*)

∼= ∼= (A.1)

*F* (*W* ⊗ *Y* ⊗ *X* ⊗ *Z*) *F* (*W* ) ⊗ *F* (*Y* ) ⊗ *F* (*X*) ⊗ *F* (*Z*)

∆*W* ⊗*Y,X*⊗*Z*

∇*W,Y* ⊗∇*X,Z*

*F* (*W* ⊗ *Y* ) ⊗ *F* (*X* ⊗ *Z*)

1. The following three “unit/counit” diagrams commute:

*η*

1

*η*

*F* (1)

*ǫ*

1

∼=

1 1 ⊗ 1

*F* (1)

*η*⊗*η*

∼=

*F* (1 ⊗ 1)

∆1*,*1

*F* (1) ⊗ *F* (1)

∼=

1 ⊗ 1

*F* (1)

*ǫ*⊗*ǫ*

1

*ǫ*

∼=

*F* (1 ⊗ 1)

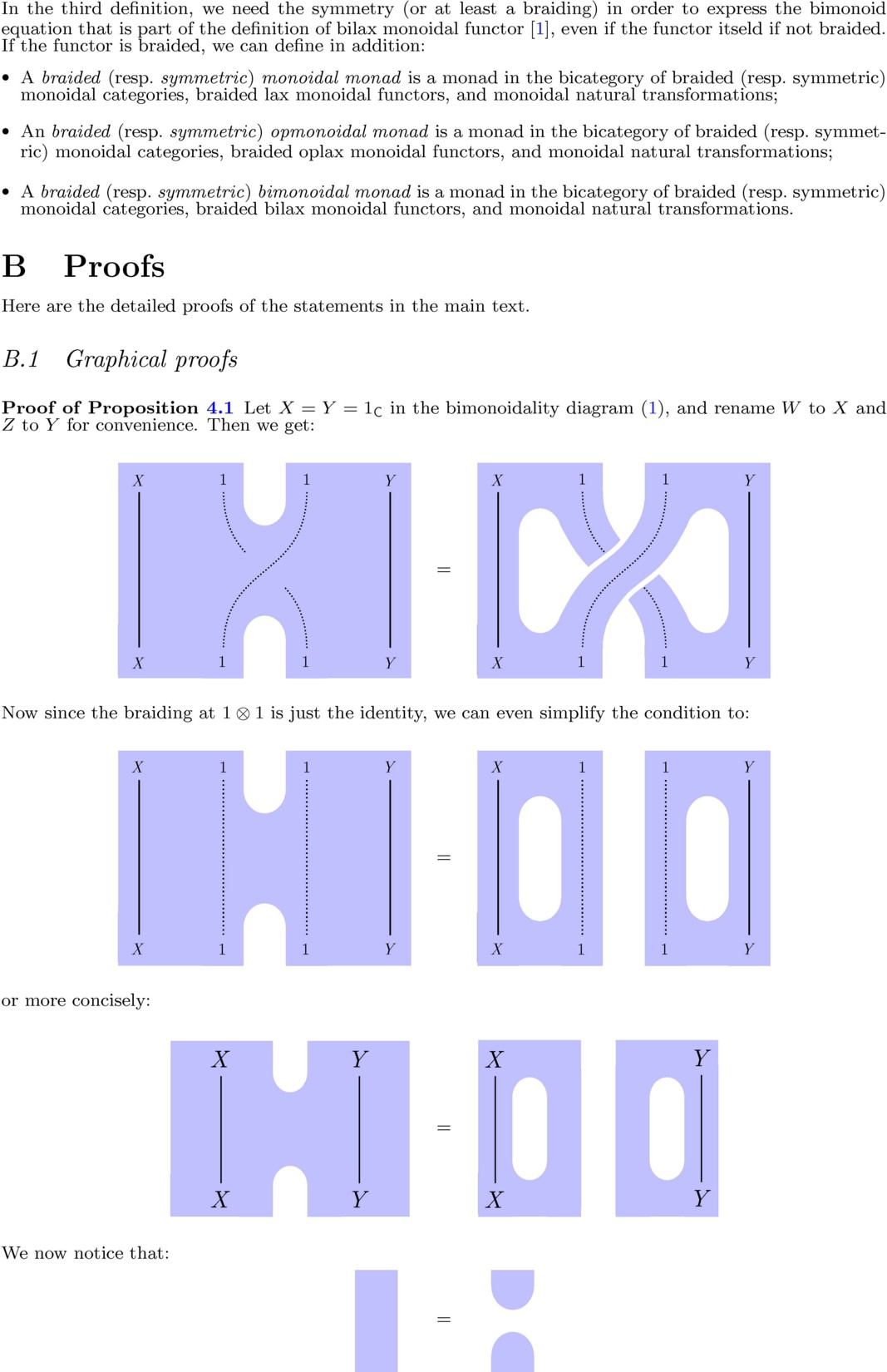
∇1*,*1

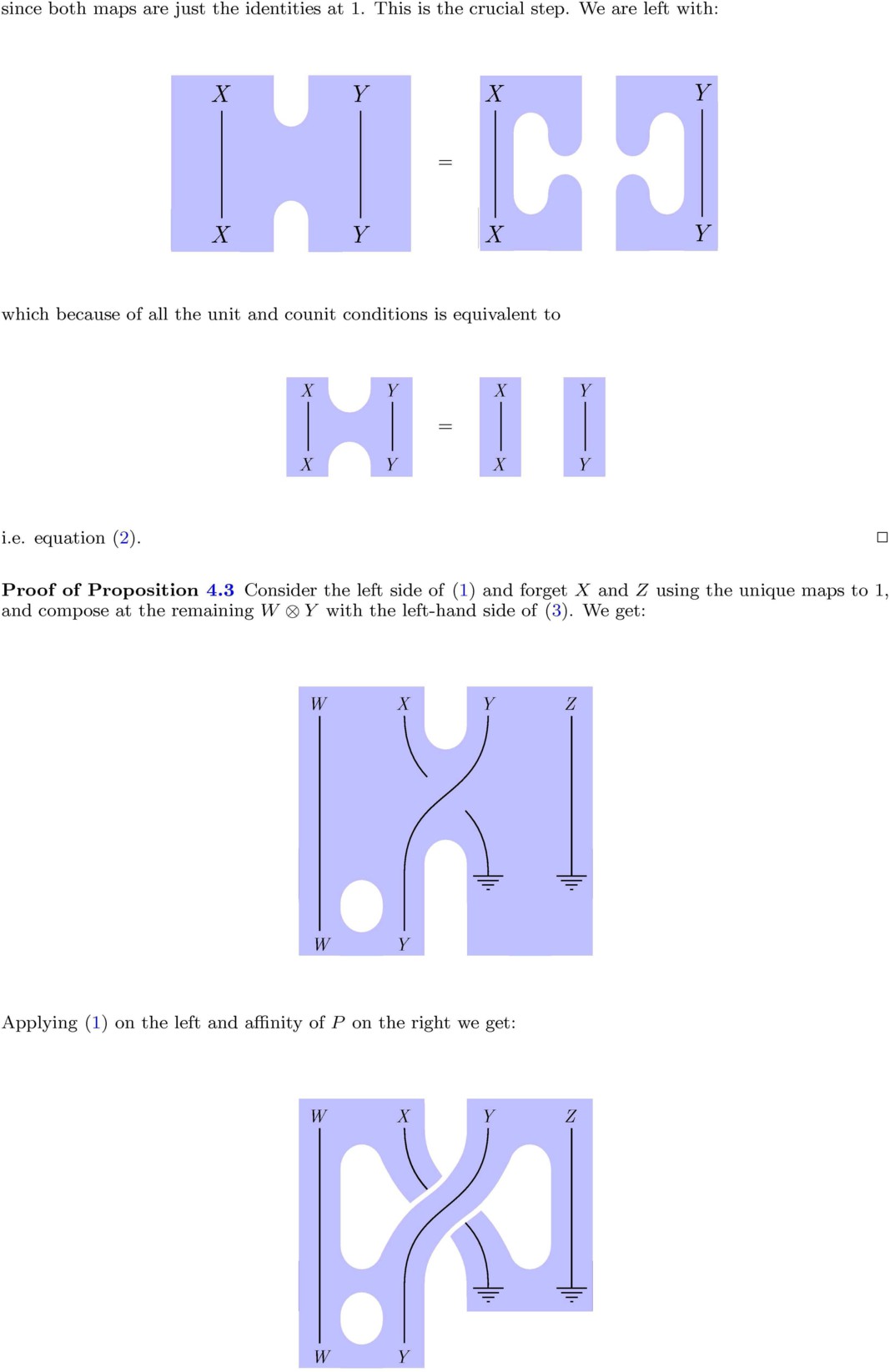
*F* (1) ⊗ *F* (1)

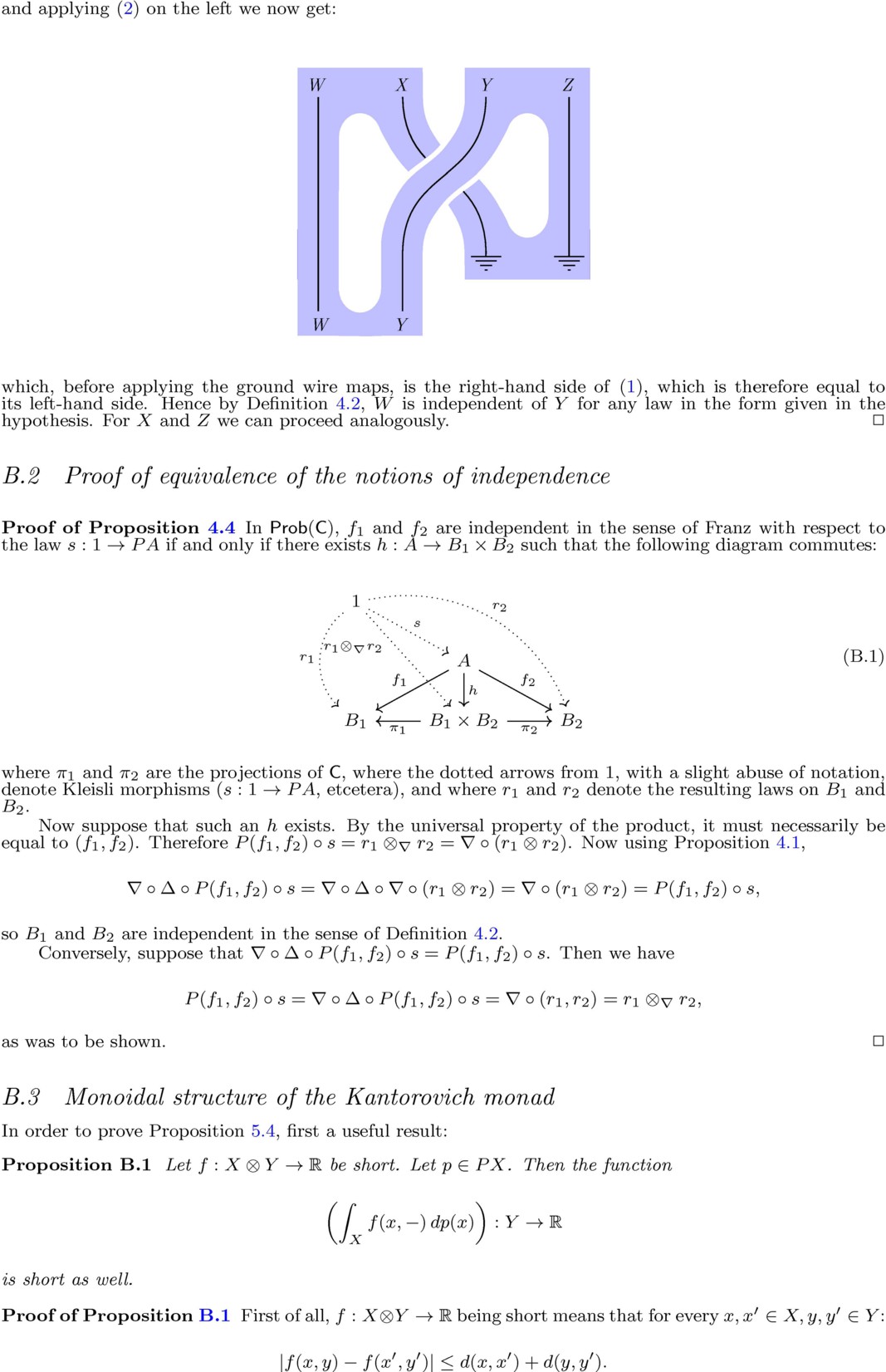
Definition A.6 Let (*F, ǫF ,* ∆*F* ) and (*G, ǫG,* ∆*G*) be bilax monoidal functors (C*,* ⊗) → (D*,* ⊗). A *bilax monoidal natural transformation*, or just *monoidal natural transformation* when it’s clear from the context, is a natural transformation *α* : *F* ⇒ *G* which is a lax and oplax natural transformation.

Definition A.7 Now, we define:

* A *monoidal monad* is a monad in the bicategory of monoidal categories, lax monoidal functors, and monoidal natural transformations;
* An *opmonoidal monad* is a monad in the bicategory of monoidal categories, oplax monoidal functors, and monoidal natural transformations;
* A *bimonoidal monad* is a monad in the bicategory of braided monoidal categories, bilax monoidal functors, and monoidal natural transformations.







Now:

∫*X*

*f* (*x, y*) *dp*(*x*) −

∫

*X*

′

*f* (*x, y* ) *dp*(*x*)

∫ *f* (*x, y*) − *f* (*x, y*′) *dp*(*x*)

=

∫

*X*

′

≤ *f* (*x, y*) − *f* (*x, y* ) *dp*(*x*)

≤ *d*(*x, x*) + *d*(*y, y*′) *dp*(*x*)

*X*

∫

∫

*X*

= *d*(*y, y*′) *dp*(*x*)

*X*

= *d*(*y, y*′)*.*

2

Proof of Proposition [5.4](#_bookmark22) To prove that ∇ it is short, let *p, p*′ ∈ *PX, q, q*′ ∈ *PY* . Then

*d* ∇(*p, q*)*,* ∇(*p*′*, q*′)

= *d* *p* ⊗∇ *q, p*′ ⊗∇ *q*′

∫

= sup

*f* :*X*⊗*Y* →R

= sup

*f* :*X*⊗*Y* →R

= sup

*f* :*X*⊗*Y* →R

= sup

*f* :*X*⊗*Y* →R

*X*⊗*Y*

*X*⊗*Y*

∫

∫

*X*⊗*Y*

∫ ∫

*X*

*f* (*x, y*) *d*(*p* ⊗∇ *q* − *p*′ ⊗∇ *q*′)(*x, y*)

*f* (*x, y*) *d* *p* ⊗∇ *q* − *p*′ ⊗∇ *q* + *p*′ ⊗∇ *q* − *p*′ ⊗∇ *q*′ (*x, y*)

*f* (*x, y*) *d* (*p* − *p*′) ⊗ *q* + *p*′ ⊗∇ (*q* − *q*′) (*x, y*)

*f* (*x, y*) *dq*(*y*) *d*(*p* − *p*′)(*x*)

*Y*

+ ∫ ∫ *f* (*x, y*) *dp*′(*x*) *d*(*q* − *q*′)(*y*)

*Y* *X*

≤ sup

∫

*g*:*X*→R *X*

*g*(*x*)*d*(*p* − *p*′)(*x*) + sup

*h*:*Y* →R *Y*

∫

*h*(*y*)*d*(*q* − *q*′)(*y*)

= *d*(*p, p*′) + *d*(*q, q*′)

= *d* (*p, q*)*,* (*p*′*, q*′) *,*

where by replacing the partial integral of *f* by *g* we have used Proposition [B.1.](#_bookmark56) 2

Proof of Proposition [5.5](#_bookmark23) By symmetry, it suffices to show naturality in *X*. Let *f* : *X* → *Z*. We need to show that this diagram commutes:

*PX* ⊗ *PY*

*f*∗⊗id

*PZ* ⊗ *PY*

∇*X,Y*

∇*Z,Y*

*P* (*X* ⊗ *Y* )

(*f* ⊗id)∗

*P* (*Z* ⊗ *Y* )

Now let *p* ∈ *PX, q* ∈ *PY* , and *g* : *Z* ⊗ *Y* → R. Then

*Z*⊗*Y*

∫

*f* (*z, y*) *d* (*f* ⊗ id)∗∇*X,Y* (*p, q*) (*z, y*) = ∫

*X*⊗*Y*

∫

*g*(*f* (*x*)*, y*) *d*(∇*X,Y* (*p, q*))(*x, y*)

= *g*(*f* (*x*)*, y*) *dp*(*x*) *dq*(*y*)

*X*⊗*Y*

∫

= *g*(*z, y*) *d*(*f*∗ *p*)(*z*) *dq*(*y*)

*Z*⊗*Y*

=

∫

*Z*⊗*Y*

∫

*g*(*z, y*) *d* (*f*∗ *p*) ⊗ *q* (*z, y*)

= *g*(*z, y*) *d* ∇*Z,Y* ◦ (*f*∗ ⊗ id)(*p, q*) (*z, y*)*.*

*Z*⊗*Y*

2

Proof of Proposition [5.6](#_bookmark24) Since both maps are natural, we only need to check the coherence diagrams. Since the unitor is just the identity at the terminal object, the unit diagrams commute. The associativity diagram at each *X, Y, Z*

id⊗∇*Y,Z*

*PX* ⊗ *PY* ⊗ *PZ * *PX* ⊗ *P* (*Y* ⊗ *Z*)

∇*X,Y* ⊗id

∇*X,Y* ⊗*Z*

∇*X*⊗*Y,Z*

*P* (*X* ⊗ *Y* ) ⊗ *PZ * *P* (*X* ⊗ *Y* ⊗ *Z*)

gives for (*p, q, r*) ∈ *PX* ⊗ *PY* ⊗ *PZ* on one path

(*p, q, r*) '→ (*p* ⊗∇ *q, r*) '→ (*p* ⊗∇ *q*) ⊗∇ *r,*

and on the other path

(*p, q, r*) '→ (*p, q* ⊗∇ *r*) '→ *p* ⊗∇ (*q* ⊗∇ *r*)*.*

The product of probability distributions is now associative, as a simple calculation can show.

The symmetry condition is straightforward. 2

Proof of Proposition [5.7](#_bookmark25) We know that (*P,* id1*,* ∇) is a lax monoidal functor. We need to check now that *δ* and *E* are monoidal natural transformations. Again we only need to show the commutativity with the multiplication, since the unitor is trivial. For *δ* : idCMet ⇒ *P* we need to check that this diagram commute for each *X, Y* :

*X* ⊗ *Y*

*δ*⊗*δ*

*δ*

*PX* ⊗ *PY*

∇*X,Y*

*P* (*X* ⊗ *Y* )

which means that for each *x* ∈ *X, y* ∈ *Y δx* ⊗∇ *δy* = *δ*(*x,y*), which is easy to check (the delta over the product is the product of the deltas). For *E* : *PP* ⇒ *P* we first need to find the multiplication map

2

∇

*X,Y*

: *PPX* ⊗ *PPY* → *PP* (*X* ⊗ *Y* ) (the unit is just twice the deltas, and the unit diagram again trivially

commutes). This map is given by

∇*P X,P Y* (∇*X,Y* )∗

*P* (*PX*) ⊗ *P* (*PY* )  *P* (*PX* ⊗ *PY* )  *P* (*P* (*X* ⊗ *Y* ))

and more explicitly, if *µ* ∈ *PPX, ν* ∈ *PPY* , and *f* : *P* (*X* × *Y* ) → R,

*P* (*X*⊗*Y* )

∫

*f* (*r*) *d* ∇2

(*µ, ν*) (*r*) = ∫

*P* (*X*⊗*Y* )

∫

*f* (*r*) *d* (∇*X,Y* )∗ ◦ ∇*P X,P Y* (*µ, ν*) (*r*)

=

*P* (*X*⊗*Y* )

*f* (*r*) *d* (∇*X,Y* )∗(*µ* ⊗∇ *ν*) (*r*)

*X,Y*

= ∫*P X*⊗*P Y f* (∇*X,Y* (*p, q*)) *d*(*µ* ⊗∇ *ν*)(*p, q*)

= ∫*P X*⊗*P Y f* (*p* ⊗∇ *q*) *dµ*(*p*) *dν*(*q*)*.*

Now we have to check that this map makes this multiplication diagram commute:

*PPX* ⊗ *PPY*

2

∇

*X,Y*

*PP* (*X* ⊗ *Y* )

*EX* ⊗*EY*

*EX*⊗*Y*

*PX* ⊗ *PY*

∇*X,Y*

*P* (*X* ⊗ *Y* )

Now let *µ* ∈ *PPX, ν* ∈ *PPY* , and *g* : *X* × *Y* → R. We have, using the formula for ∇2 found above,

*X*⊗*Y*

∫

*g*(*x, y*) *d* ∇*X,Y* ◦ (*EX , EY* )(*µ, ν*) (*x, y*) =

=

∫

*X*⊗*Y*

*g*(*x, y*) *d* ∇*X,Y* (*Eµ, Eν*) (*x, y*)

=

∫

*X*⊗*Y*

∫ ∫

*g*(*x, y*) *d* *Eµ* ⊗∇ *Eν* (*x, y*)

= *g*(*x, y*) *dp*(*x*) *dq*(*y*) *dµ*(*p*) *dν*(*q*)

*P X*⊗*P Y X*⊗*Y*

∫ ∫

= *g*(*x, y*) *d*(*p* ⊗∇ *q*)(*x, y*) *dµ*(*p*) *dν*(*q*)

*P X*⊗*P Y X*⊗*Y*

*X,Y*

=

∫

*P* (*X*×*Y* )

*X*⊗*Y*

∫

*g*(*x, y*) *dr*(*x, y*) *d*(∇2

(*µ, ν*))(*r*)

=

∫

*X*⊗*Y*

*g*(*x, y*) *d* *EX*⊗*Y* ◦ ∇2

(*µ, ν*) (*x, y*)*.*

Therefore the diagram commutes, and (*P, δ, E*) is a monoidal monad. 2

*X,Y*

* 1. *Opmonoidal structure of the Kantorovich monad*

Just as in the case of joints, to prove the Proposition [5.10](#_bookmark26) we first prove the following useful result.

Proposition B.2 *Let f* : *X* → R *and g* : *Y* → R *be short. Then* (*f* + *g*) : *X* ⊗ *Y* → R *given by*

(*x, y*) '→ *f* (*x*) + *g*(*y*) *is short.*

Proof of Proposition [B.2](#_bookmark60) Let *x, x*′ ∈ *X* and *y, y* ∈ *Y* . Then

|*f* (*x*) + *g*(*y*) − *f* (*x*′) − *f* (*y*′)| ≤ |*f* (*x*) − *f* (*x*′)| + |*g*(*y*) − *g*(*y*′)|

≤ *d*(*x, x*′) + *d*(*y, y*′) = *d* (*x, y*)*,* (*x*′*, y*′) *.*

2

Proof of Proposition [5.10](#_bookmark26) To prove that ∆ is short, let *p, q* ∈ *P* (*X* ⊗ *Y* ), and denote *pX , pY , qX , qY*

their marginals. Then:

*d* ∆(*p*)*,* ∆(*q*) = *d* (*pX , pY* )*,* (*qX , qY* ) = *d*(*pX , qX* ) + *d*(*pY , qY* )

∫

∫

= sup

*f* :*X*→R *X*

= sup ∫

*f* (*x*) *d*(*pX* − *qX* )(*x*) + sup

*g*:*Y* →R *Y*

*f* (*x*) *d*(*p* − *q*)(*x, y*) + sup ∫

*g*(*y*) *d*(*pY* − *qY* )(*y*)

*g*(*y*) *d*(*p* − *q*)(*x, y*)

*f* :*X*→R

*X*⊗*Y*

*g*:*Y* →R

*X*⊗*Y*

= sup sup

∫

*f* :*X*→R *g*:*Y* →R

*X*⊗*Y*

*f* (*x*) + *g*(*y*) *d*(*p* − *q*)(*x, y*)

≤ sup

*h*:*X*⊗*Y* →R

*h*(*x, y*) *d*(*p* − *q*)(*x, y*)

= *dP* (*X*⊗*Y* )(*p, q*)*,*

where by replacing *f* + *g* with *h* we have used Proposition [B.2.](#_bookmark60) 2

Proof of Proposition [5.11](#_bookmark27) By symmetry, it suffices to show naturality in *X*. Let *f* : *X* → *Z*. We need to show that this diagram commutes:

∆*X,Y*

*P* (*X* ⊗ *Y* )

(*f* ⊗id)∗

∆*Z,Y*

*P* (*Z* ⊗ *Y* )

*PX* ⊗ *PY*

*f*∗⊗id

*PZ* ⊗ *PY*

Let now *p* ∈ *P* (*X* ⊗ *Y* ). We have to prove that:

∆*Z,Y* ◦ (*f* ⊗ id)∗*p* = (*f*∗ ⊗ id) ◦ ∆*X,Y* (*p*)*.*

On one hand:

(*f*∗ ⊗ id) ◦ ∆*X,Y* (*p*) = (*f*∗ ⊗ id)(*pX , pY* )

= (*f*∗*pX , pY* )*.* On the other hand, let *h* : *Z* → R and *g* : *Y* → R be short. Then:

∫ ∫

*h*(*z*) *d*(((*f* ⊗ id)∗*p*)*Z* )(*z*) = *h*(*z*) *d*((*f* ⊗ id)∗*p*)(*z, y*)

*Z Z*⊗*Y*

∫

= *h*(*f* (*x*)) *dp*(*x, y*)

*X*⊗*Y*

∫

= *h*(*f* (*x*)) *dpX* (*x*)

*X*

∫

= *h*(*z*) *d*(*f*∗*pX* )(*x*)*,*

*Z*

and:

∫ ∫

*g*(*y*) *d*(((*f* ⊗ *id*)∗*p*)*Y* )(*y*) = *g*(*y*) *d*((*f* ⊗ *id*)∗*p*)(*z, y*)

*Y Z*⊗*Y*

∫

= *g*(*y*) *dp*(*x, y*)

*X*⊗*Y*

∫

= *g*(*y*) *dpY* (*y*)*,*

*Y*

so the two components are again (*f*∗*pX , pY* ). 2

Proof of Proposition [5.12](#_bookmark28) We already have naturality of the maps, and the counitor is trivial, we just have to check coassociativity. Namely, that the following diagrams commutes for each *X, Y, Z*:

*P* (*X* ⊗ *Y* ⊗ *Z*)

∆*X,Y* ⊗*Z*

*P* (*X*) ⊗ *P* (*Y* ⊗ *Z*)

∆*X*⊗*Y,Z*

id⊗∆*Y* ⊗*Z*

*P* (*X* ⊗ *Y* ) ⊗ *P* (*Z*)

∆*X*⊗*Y* ⊗id

*P* (*X*) ⊗ *P* (*Y* ) ⊗ *P* (*Z*)

Now given *p* ∈ *P* (*X* ⊗ *Y* ⊗ *Z*), we get:

(∆*X*⊗*Y* ⊗ id) ◦ ∆*X*⊗*Y,Z* (*p*) = (∆*X*⊗*Y* ⊗ id)(*pXY , pZ* ) = (*pX , pY , pZ* )*,*

and:

(id ⊗ ∆*Y* ⊗*Z* ) ◦ ∆*X,Y* ⊗*Z* (*p*) = (id ⊗ ∆*Y* ⊗*Z* )(*pX , pY Z* ) = (*pX , pY , pZ* )*,*

since there is only one way of forming marginals.

The symmetry condition is again straightforward. 2

Proof of Proposition [5.13](#_bookmark29) We know that (*P,* id1*,* ∆) is an oplax monoidal functor. We need to check now that *δ* and *E* are comonoidal natural transformations. Again we only need to show the commutativity with the comultiplication, since the counitor is trivial. For *δ* : idCMet ⇒ *P* we need to check that this diagram commute for each *X, Y* :

*X* ⊗ *Y*

*δ*

*δ*⊗*δ*

*P* (*X* ⊗ *Y* )

∆*X,Y*

*PX* ⊗ *PY*

which means that for each *x* ∈ *X, y* ∈ *Y* , (*δ*(*x,y*))*X* = *δx* and (*δ*(*x,y*))*Y* = *δy*, which is again easy to check (the marginals of a delta are the deltas at the projections). For *E* : *PP* ⇒ *P* we first need to find the

comultiplication map ∆2

*X,Y*

: *PP* (*X* ⊗ *Y* ) → *PPX* ⊗ *PPY* (the unit is just twice the deltas, and the unit

diagram again trivially commutes). This map is given by:

(∆*XY* )∗ ∆*P X,P Y*

*P* (*P* (*X* ⊗ *Y* )) *P* (*PX* ⊗ *PY* )  *P* (*PX*) ⊗ *P* (*PY* )

and more explicitly, if *µ* ∈ *P* (*P* (*X* ⊗ *Y* )), and *f* : *PX* → R and *g* : *PY* → R are short:

*f* (*p*) *d* ((∆*XY* )∗*µ*)*P X* (*p*) = ∫

∫

∫

*P X*

*P X*⊗*P Y*

*f* (*p*) *d* ((∆*XY* )∗*µ*)*P X* (*p, q*)

since *g* only depends on *PX*, and analogously:

∫

= *f* (*rX* ) *dµ*(*r*)

*P* (*X*⊗*Y* )

*g*(*q*) *d* ((∆*XY* )∗*µ*)*P Y* (*q*) = ∫

*P Y*

*P* (*X*⊗*Y* )

*f* (*rY* ) *dµ*(*r*)*.*

We have to check that this map makes this multiplication diagram commute:

*PP* (*X* ⊗ *Y* )

2

∆*X,Y*

*PPX* ⊗ *PPY*

*EX*⊗*Y*

*EX* ⊗*EY*

*P* (*X* ⊗ *Y* )

∆*X,Y*

*PX* ⊗ *PY*

Now let *µ* ∈ *P* (*P* (*X* ⊗ *Y* )), and *f* : *X* → R and *g* : *Y* → R short. We have, using the formula for ∆2 found above:

∫

∫

*f* (*x*) *d* ((*EX*⊗*Y µ*)*X* ) (*x*) =

*X X*⊗*Y*

*f* (*x*) *d* (*EX*⊗*Y µ*) (*x, y*)

= *f* (*x*) *dr*(*x, y*) *dµ*(*r*)

∫ ∫

*P* (*X*⊗*Y* ) *X*⊗*Y*

∫ ∫

= *f* (*x*) *d*(*rX* )(*x*) *dµ*(*r*)

*P* (*X*⊗*Y* ) *X*

=

∫

*P X*⊗*P Y*

∫ *f* (*x*) *dp*(*x*) *d* (∆*XY* )∗*µ* (*p, q*)

*X*

=

∫

∫

*P X* *X*

*f* (*x*) *dp*(*x*) *d* ((∆*XY* )∗*µ*)*P X* (*p*)

and analogously:

= ∫ *f* (*x*) *d* *EX* ((∆*XY* )∗*µ*)*P X* (*x*)*,*

*X*

which means:

*g*(*y*) *d* ((*EX*⊗*Y µ*)*Y* ) (*y*) =

*Y Y*

∫

∫

*f* (*y*) *d* *EY* ((∆*XY* )∗*µ*)*P Y* (*y*)*,*

∆*X,Y* ◦ *EX*⊗*Y µ* = (*EX* ⊗ *EY* ) ◦ ∆*P X,P Y* (∆*XY* )∗*µ*)

= (*EX* ⊗ *EY* ) ◦ (∆*P X,P Y* ◦ (∆*XY* )∗)*µ*

= (*EX* ⊗ *EY* ) ◦ ∆2 *µ.*

*X,Y*

Therefore the diagram commutes, and (*P, δ, E*) is an opmonoidal monad. 2

* 1. *Bimonoidal structure of the Kantorovich monad*

Proof of Proposition [5.14](#_bookmark30) We already know that *P* is lax and oplax. We only need to check the compatibility diagrams between the two structures. The unit diagrams are trivial, because the unitors are trivial. The bimonoidality diagram:

*P* (*W* ⊗ *X*) ⊗ *P* (*Y* ⊗ *Z*)

∇*W* ⊗*X,Y* ⊗*Z*

∆*W,X* ⊗∆*Y,Z*

*P* (*W* ⊗ *X* ⊗ *Y* ⊗ *Z*) *P* (*W* ) ⊗ *P* (*X*) ⊗ *P* (*Y* ) ⊗ *P* (*Z*)

∼= ∼=

*P* (*W* ⊗ *Y* ⊗ *X* ⊗ *Z*) *P* (*W* ) ⊗ *P* (*Y* ) ⊗ *P* (*X*) ⊗ *P* (*Z*)

∆*W* ⊗*Y,X*⊗*Z*

∇*W,Y* ⊗∇*X,Z*

*P* (*W* ⊗ *Y* ) ⊗ *P* (*X* ⊗ *Z*)

says that given *p* ∈ *P* (*W* ⊗ *X*)*, q* ∈ *P* (*Y* ⊗ *Z*):

∆*W* ⊗*Y,X*⊗*Z* ◦ ∇*W* ⊗*X,Y* ⊗*Z* (*p, q*) = (∇*W,Y* ⊗ ∇*X,Z* ) ◦ (∆*W,X* ⊗ ∆*Y,Z* )(*p, q*) Now on one hand:

(∇*W,Y* ⊗ ∇*X,Z* ) ◦ (∆*W,X* ⊗ ∆*Y,Z* )(*p, q*) = (∇*W,Y* ⊗ ∇*X,Z* )(*pW , pX , qY , qZ* )

= (*pW* ⊗∇ *qY , pX* ⊗∇ *qZ* )*.*

On the other hand:

∆*W* ⊗*Y,X*⊗*Z* ◦ ∇*W* ⊗*X,Y* ⊗*Z* (*p, q*) = ∆*W* ⊗*Y,X*⊗*Z* (*p* ⊗∇ *q*)*.*

The marginal of *p* ⊗ *q* on *W* ⊗ *Y* is, by Fubini’s theorem, let *f* : *W* ⊗ *Y* → R:

∫ ∫

*f* (*w, y*) *d*((*p* ⊗∇ *q*)*W Y* )(*w, y*) = *f* (*w, y*) *d*(*p* ⊗∇ *q*)(*w, x, y, z*)

*W* ⊗*Y W* ⊗*X*⊗*Y* ⊗*Z*

∫

= *f* (*w, y*) *dp*(*w, x*) *dq*(*y, z*)

*W* ⊗*X*⊗*Y* ⊗*Z*

∫

= *f* (*w, y*) *dpW* (*w*) *dqY* (*y*)

*W* ⊗*Y*

∫

= *f* (*w, y*) *d*(*pW* ⊗∇ *qY* )(*w, y*)*,*

*W* ⊗*Y*

and similarly the marginal on *X* ⊗ *Z* is given by *pX* ⊗∇ *qZ* . In other words, if the pairs are independent, the components from different pairs are also independent. It follows that *P* is bilax monoidal. 2