A categorical approach to the probability theory

Vladimir Voevodsky

Started Stage 1, February 3 2004; Stage 2, December 2 2004; Stage 3, October 2 2007;

Contents

1 Introduction

Let us look at the standard approach to mathematical modeling of a deterministic process. One starts with a set X and a family of maps $\phi_{t1,t2} : X \to X$ where $t1, t2$ are two numbers which are

whose points correspond to the possible states of the system in question. A change in the state of the system is modeled as a map from this set to itself. A "process" is usually a family of such maps – one for each interval $[t_0, t_1]$ of the line representing time, which satisfy the obvious composition condition for intervals of the form $[t_0, t_1]$, $[t_1, t_2]$ and $[t_0, t_2]$. For example, any (deterministic) computer program which takes t_0, t_1, and the state of the system at time t_0 as an input and produces the state of the system at time t_1 as an output defines a "process" in the sense specified above.

If the program we use is not deterministic but uses a random number generator to compute new values of the variables from the old ones it does not define such a process.

Consider now the case when we have a process whose computer model is based on a randomized algorithm to produce the new values of the variables from the old ones. As an example we may look at a simple population dynamics model where the the state of the system is determined by the number of organisms currently alive, time is discrete and to produce the state at the next moment of time our algorithm uses a random number generator to determine whether a given organism survives (with probability p) or dies (with probability $1 - p$).

Note that all the notions used in the mathematical description of a deterministic process naturally belong to the language of the category theory: we have a set X and a family of morphisms (maps) $f_{[t_0,t_1]} : X \to X$ satisfying the composition condition.

The stochastic category described below allows one to repeat the same description in a randomized case simply by replacing the category of sets with the stochastic category.

2 Stochastic categories

1 The category of measurable spaces

Let us first recall the following definition.

Definition 1.1 *A σ-algebra A on a set X is a collection of subsets of X satisfying the following conditions.*

1. *The empty subset is in A.*
2. *For a countable family U_i of elements of A one has $\cup_i U_i \in A$.*
3. *For U in A the complement $X \backslash U$ to U in X is in A.*

For a set of σ-algebras A_α on X the collection of subsets $\bigcap_\alpha A_\alpha$ is the largest σ-algebra contained in all A_α and we will write $\sum_\alpha A_\alpha$ for the smallest σ-algebra which contains all of the A_α.

Let $f : X \to Y$ be a map of sets. For a collection A of subsets of X we let $f(A)$ denote the collection of subsets U of Y such that $f^{-1}(U) \in A$. For a collection B of subsets of Y we let $f^{-1}(B)$ denote the collection of subsets of X of the form $f^{-1}(U)$ where $U \in B$. It is easily seen that if A (resp. B) is a σ-algebra then $f(A)$ (resp. $f^{-1}(B)$) is a σ-algebra.

Definition 1.2 *The category MS of measurable space is defined as follows:*

Objects of MS are measurable spaces i.e. pairs of the form (X, A) where X is a set and A is a σ-algebra of subsets of X.

Morphisms from (X, A) to (Y, B) are maps of sets $f : X \to Y$ such that for each $V \in B$ one has $f^{-1}(V) \in A$.

Compositions of morphisms and the identity morphisms correspond to the compositions of maps of sets and to the identity maps of sets.

The associativity of the composition and the defining property of the identity maps are obvious and therefore MS is indeed a category.

Sending (X, A) to X we get a functor from MS to the category $Sets$ of sets. This functor has two adjoints. The right adjoint sends X to $(X, \{\emptyset, X\})$ and the left adjoint to $(X, 2^X)$ where 2^X is the set of all subsets of X. We will say that a morphism in MS is surjective, injective or bijective if the morphism of the underlying sets has the corresponding property.

The measurable spaces $(\emptyset, \{\emptyset\})$ and $(pt, 2^{pt})$ give us an initial object and a final object of MS. To simplify the notation we will write $\emptyset$ instead of $(\emptyset, 2^{\emptyset})$ and pt instead of $(pt, 2^{pt})$.

Theorem 1.3 [**mscomplete**] *The category MS is a complete category i.e. any small diagram in MS has a limit.*

Proof: By [3, Theorem 1, p.113] it is sufficient to show that products and equalizers exist in MS.

Let $(X_\alpha, B_\alpha)_{\alpha\in A}$ be a family of measurable spaces. Set

$$\prod_\alpha (X_\alpha, B_\alpha) = (\prod_\alpha X_\alpha, \sum_\alpha pr_\alpha^{-1}(B_\alpha))$$

where pr_α is the projection from $\prod_\alpha X_\alpha$ to X_α. One verifies easily that this measurable space together with the obvious maps from it to (X_α, B_α) is indeed the product of the family $(X_\alpha)_{\alpha\in A}$ in MS.

Let $f, g : (X, A) \to (Y, B)$ be a pair of morphisms in MS. Consider the equalizer diagram in Sets corresponding to f and g

$$Z \xrightarrow{i} X \rightrightarrows Y$$

and define the equalizer of f and g in MS by the formula

$$[\textbf{eqdef}]eq(f, g) = (Z, i^{-1}(A)) \tag{1}$$

as in the case of the product one verifies easily that together with the obvious morphism $eq(f, g) \to X$ this measurable space is indeed the equalizer of the morphisms f and g in MS.

Remark 1.4 [**powerspace**] Let X be a set and (Y, B) a measure space. The product of as many copies of (Y, B) as there are elements in X can also be described in a slightly different way. Consider the set Y^X of all maps of sets from X to Y. For any V in B and any x in X let $A(x, V)$ be the set of all $g : X \to Y$ such that $g(x) \in V$. Let B^X be the σ-algebra on Y^X generated by the subsets $A(x, V)$. Then our product is given by $(Y, B)^X = (Y^X, B^X)$.

Theorem 1.5 *[***mscocomplete***] The category MS is co-complete i.e. any small diagram in MS has a colimit.*

Proof: By [3, Theorem 1, p.113] applied to the category MS^{op} it is sufficient to show that MS has coproducts and coequalizers.

Let $(X_\alpha, B_\alpha)_{\alpha\in A}$ be a family of measurable spaces. Set

$$\coprod_\alpha (X_\alpha, B_\alpha) = (\coprod_\alpha X_\alpha, \bigcap_\alpha i_\alpha(B_\alpha))$$

where i_α is the canonical map from X_α to $\coprod_\alpha X_\alpha$. One verifies easily that this measurable space together with the obvious maps to it from (X_α, B_α) is indeed the coproduct of the family $(X_\alpha)_{\alpha\in A}$ in MS.

Let $f, g : (X, A) \to (Y, B)$ be a pair of morphisms in MS. Consider the coequalizer diagram in Sets corresponding to f and g

$$X \rightrightarrows Y \xrightarrow{p} Z$$

and define the coequalizer of f and g in MS by the formula

$$[\textbf{coeqdef}]coeq(f, g) = (Z, p(B)). \tag{2}$$

As in the case of the coproduct one verifies easily that together with the obvious morphism $(Y, B) \to coeq(f, g)$ this measurable space is indeed the coequalizer of the morphisms f and g in MS.

Lemma 1.6 *[***epimono1***] A morphism $f : (X, A) \to (Y, B)$ in MS is an epimorphism (resp. a monomorphism) if and only if it is surjective (resp. injective).*

Proof: The 'if' part is obvious both for epimorphisms and for monomorphisms. Let us prove the 'only if' parts. Assume that f is a monomorphism. Then it is injective since otherwise there would be two different morphisms from the point *pt* to X whose compositions with f coincide. Assume that f is an epimorphism. Then it is surjective since otherwise there would be two different morphisms from Y to $(\{0, 1\}, 2^{\{0,1\}})$ whose compositions with f coincide.

Recall that a morphism $X \to Y$ is called an effective epimorphism if $X \times_Y X \rightrightarrows X \xrightarrow{f} Y$ is a coequalizer diagram and an effective monomorphism if it is an effective epimorphism in the opposite category.

Lemma 1.7 *[***epimono2***] A morphism $f : (X, A) \to (Y, B)$ in MS is an effective epimorphism iff it is an epimorphism and $B = f(A)$. It is an effective monomorphism iff it is a monomorphism and $A = f^{-1}(B)$.*

Proof: The statement for the epimorphisms follows from (2) and the statement for the monomorphisms from (1).

Example 1.8 [**bijective**] Let X be a set and $A_2 \subset A_1$ be two σ-algebras on X. Then the identity of X defines a bijective morphism $(X, A_1) \to (X, A_2)$. This morphism is an epimorphism and a monomorphism but unless $A_2 = A_1$ it is not an isomorphism.

Proposition 1.9 *[***epimono3***] For any morphism $f : (X, A) \to (Y, B)$ there exists a unique decomposition of the form $f = i \circ b \circ p$ where i is an effective monomorphism, b is a bijection and p is an effective epimorphism.*

Proof: Let $X \xrightarrow{p} Z \xrightarrow{i} Y$ be the decomposition of f into a surjection and an injection in the category of sets. It defines a decomposition of f in the category MS of the form

$$(X, A) \xrightarrow{p} (Z, p(A)) \xrightarrow{b} (Z, i^{-1}(B)) \xrightarrow{i} (Y, B)$$

which satisfies the conditions of the proposition by Lemmas 1.5 and 1.6. The uniqueness easily follows from the same two lemmas.

2 Category of kernels

We define the *category of Kernels* $\mathcal{K}$ as follows. Objects of $\mathcal{K}$ are pairs (X, A) where X is a set and A is a σ-algebra of subsets of X i.e. objects are measurable spaces. Morphisms in $\mathcal{K}$ are called *kernels*.

Definition 2.1 *[***d1***] A kernel $f = f(x, U)$ from (X, A) to (Y, B) is a function*

$$f(-, -) : X \times B \to [0, \infty]$$

such that for any $x \in X$ the function

$$f(x, -) : U \mapsto f(x, U)$$

is a measure on (Y, B) and for any $U \in B$ the function

$$f(-, U) : x \mapsto f(x, U)$$

is a measurable function on (X, A).

For a measure μ on (X, A), a measurable function f on the same space and a measurable subset Y of X we let

$$\int_Y f d\mu$$

denote the integral of f restricted to Y with respect to μ.

Lemma 2.2 *[**comp1**] Let f be a kernel $(X, A) \to (Y, B)$ and $g : Y \to [0, \infty]$ be a non-negative measurable function on Y. Then the function*

$$f^*(g) : x \mapsto \int_Y g df(x, -)$$

is a measurable function on (X, A).

Proof: Consider the class $\mathcal{C}$ of all g such that $f^*(g)$ is measurable. By definition of a kernel this class contains defining functions I_U of subsets U in B. Hence it contains all non-negative simple functions on (Y, B). The continuity property of the integral (e.g. [1, Th.15.1(iii),p.204]) implies that if $0 \leq g_n \uparrow g$ where g_n are in $\mathcal{C}$ then g is in $\mathcal{C}$. By [1, Th.13.5, p.185] the smallest class satisfying these two properties contains all measurable functions.

Now let $f : (X, A) \to (Y, B)$, $g : (Y, B) \to (Z, C)$ be two kernels. Consider the function on $X \times C$ of the form

$$[\textbf{comp2}](x, W) \mapsto \int_Y g(-, W) df(x, -) \tag{3}$$

This function is well defined since $g(-, W)$ is measurable. For each W it is a measurable function on (X, A) by Lemma 2.2. On the other hand for any x the function

$$W \mapsto \int_Y g(-, W) df(x, -)$$

is a measure on (Y, C) by the standard properties of the integral. Therefore, (3) defines a kernel from (X, A) to (Z, C) which we denote by $g \circ f$ and call the composition of f and g.

For every (X, A) the kernel Id which takes x to the measure δ_x concentrated in x is the identity morphism. The following three lemmas imply that our composition is associative and therefore measure spaces, kernels and compositions (3) define a category. We denote this category by $\mathcal{K}$ and call the category of kernels.

Lemma 2.3 *[**funcmes**] Let μ be a measure on (X, A) and $f : (X, A) \to (Y, B)$ a kernel. Then the function $f_*(\mu)$ on B of the form*

$$U \mapsto \int_X f(-, U) d\mu$$

is a measure on (Y, B).

Proof: Obvious.

Lemma 2.4 *[**tudysyudy**] Let $f : (X, A) \to (Y, B)$ be a kernel, μ a measure on (X, A) and g a measurable non-negative function on (Y, B). Then one has*

$$\int f^*(g) d\mu = \int g df_*(\mu)$$

Proof: If g is the simple function corresponding to a subset $U \in B$ then our equality holds by definitions. For a general g the result follows by the same continuity argument as in the proof of Lemma 2.2.

Lemma 2.5 *[**assos**] The composition of kernels defined by (3) is associative.*

Proof: It follows immediately from definitions and Lemma 2.4.

For a topological space X we will write simply X instead of the usual $(X, \mathcal{B})$ for the measure space with the underlying set X and the underlying σ-algebra the Borel σ-algebra on X. We will further consider sets as topological spaces with the discrete topology (all subsets are open). Combining these two conventions we will write X for the measure space with the underlying set X and the underlying σ-algebra of all subsets of X.

Example 2.6 *[**ex0**]*For any (X, A) there is a unique kernel from $\emptyset$ to (X, A). Therefore $\emptyset$ is the initial object of the category of kernels. Since there is a unique measure on $\emptyset$ there is also a unique kernel from any (X, A) to the empty set i.e. $\emptyset$ is also the final object.

Example 2.7 *[**ex1**]*We will denote the object of $\mathcal{K}$ corresponding to the one element set by $\mathbf{1}$. A morphism from $\mathbf{1}$ to (X, A) is the same as a measure on (X, A). A morphism from (X, A) to $\mathbf{1}$ is a non-negative measurable function on (X, A). In particular

$$[\mathbf{h11}]Hom(\mathbf{1}, \mathbf{1}) = \mathbf{R}_{\geq 0} \cup \{\infty\} \tag{4}$$

and for any (X, A) the composition pairing

$$Hom(\mathbf{1}, (X, A)) \times Hom((X, A), \mathbf{1}) \to Hom(\mathbf{1}, \mathbf{1})$$

takes (μ, f) to $\int f\mu$. Note that the composition on (4) is of the form $(a, b) \mapsto ab$ where $0\infty = \infty 0 = 0$ as is usually assumed in measure theory.

Example 2.8 [**matrixex**] Let $\mathbf{n}$ be the measure space with the underlying set $\{1, \ldots, n\}$ and the σ-algebra of all subsets. Then $Hom(\mathbf{n}, \mathbf{n})$ is the set of $n \times n$ matrices with entries from $[0, \infty]$. The composition is given by the product of matrices.

Example 2.9 [**ex0new**] Let (X, A) be a measurable space and f a non-negative measurable function on it. Then the mapping which sends a point x of X to the measure $f(x)\delta_x$ is a kernel which we denote I_f. If $\mu : \mathbf{1} \to (X, A)$ is a measure on (X, A) the the composition $I_f \circ \mu$ is the 'product measure' which sends $Y \in A$ to $\int_Y f d\mu$. We will denote this measure by $f * \mu$.

Let (X, A), (Y, B) be measurable spaces and let $f : X \to Y$ be a measurable map. Sending $x \in X$ to the measure $\delta_{f(x)}$ on Y concentrated in $f(y)$ defines a morphism from (X, A) to (X, B) in $\mathcal{K}$. To verify the integrability condition note that for a subset U in Y the function $x \mapsto \delta_{f(x)}(U)$ is the characteristic function of the subset $f^{-1}(U)$. Hence the second condition of Definition 2.1 is equivalent to the condition that f is measurable. This construction defines a functor from the category of measurable spaces and measurable maps to the category of kernels. To distinguish morphisms in $\mathcal{K}$ which correspond to maps of measure spaces from the general morphisms we will call the former *deterministic* morphisms.

Example 2.10 *[**ex5**]*Let $\mu : \mathbf{1} \to (X, A)$ be a measure on (X, A) and $f : (X, A) \to (Y, B)$ a measurable map considered as a kernel. Then $f \circ \mu = f_*(\mu)$ is the "direct image" of μ with respect to f.

Example 2.11 [**retract**]Let (X, A) be a measure set and (U, A_U) be a measurable subset of X considered with the induced σ-algebra. Then the embedding $(U, A_U) \to (X, A)$ can be split by a projection p where $p(x, -)$ is zero for $x \in X - U$ and is the measure concentrated in x for $x \in U$. Hence any measurable subset (including the empty one) of a measure space is canonically a retract of this space in $\mathcal{K}$.

The functor from the category of measurable spaces to $\mathcal{K}$ does not reflect isomorphisms i.e. some morphisms of measurable spaces may become isomorphisms when considered in $\mathcal{K}$. Let (Y, B) be a measurable space and $f : X \to Y$ a be any surjection of sets. Let $f^{-1}(B)$ be the σ-algebra on X which consists of subsets of the form $f^{-1}(U)$ for $U \in B$. Then measures on $(X, f^{-1}(B))$ are in one-to-one correspondence with measures on (Y, B). In particular for each point $b \in B$ we have a measure f_b on $(X, f^{-1}(A))$ corresponding to the delta measure δ_b on (Y, B). Sending b to f_b gives us a kernel $(Y, B) \to (X, f^{-1}(A))$ and one verifies easily that it is inverse to the obvious kernel $(X, f^{-1}(B)) \to (Y, B)$. Hence, from the point of view of the category of kernels, the measurable spaces (Y, B) and $(X, f^{-1}(B))$ are indistinguishable.

For measure spaces (X, A), (X', A') the measure space $(X \coprod X', A \coprod A')$ is easily seen to be both a product and a coproduct of (X, A) and (X', A') in $\mathcal{K}$. Together with Example 2.6 it shows that $\mathcal{K}$ has both finite products and finite coproducts which coincide. For any two objects the set of morphisms between them is an abelian semi-group and moreover a "module" over $\mathbf{R}_+ \cup \{\infty\}$. However (since we do not allow negative measures) morphisms can not be subtracted and therefore $\mathcal{K}$ is not an additive category.

Lemma 2.12 [**prcopr**] *Let (X_α, A_α) be a family of measure spaces. The measure space $(\coprod X_\alpha, \sum A_\alpha)$ (resp. $(\coprod X_\alpha, \cap A_\alpha)$) is the product (resp. the coproduct) of the family (X_α, A_α) in $\mathcal{K}$.*

Proof: ???

The families $\sum A_\alpha$ and $\cap A_\alpha$ coincide if our family is finite or countable but are different in general. In particular the countable products and coproducts in $\mathcal{K}$ coincide.

Example 2.13 [**prcopr2**]The set of natural numbers $\mathbf{N}$ considered with the σ-algebra of all subsets is both the product and the coproduct of a countable number of copies of $\mathbf{1}$. The sets $Hom_{\mathcal{K}}(\mathbf{N}, \mathbf{1})$ and $Hom_{\mathcal{K}}(\mathbf{1}, \mathbf{N})$ can both be identified with the set $[0, \infty]^{\mathbf{N}}$ of infinite sequences of (extended) non-negative real numbers.

Lemma 2.14 [**l1**] *Let G be a finite group of measurable automorphisms of a measure space (X, A). Then the measure space $(X/G, A^G)$ is the categorical quotient of (X, A) in $\mathcal{K}$ with respect to the action of G.*

Proof: ???

3 Category of bounded kernels

A kernel $f : (X, A) \to (Y, B)$ is called bounded if the function

$$\beta_f : x \mapsto f(x, Y)$$

is a bounded function on X. Note that this condition means in particular that β_f takes only finite values i.e. that for any x the measure $f(x, -)$ on (Y, B) is finite. The composition of bounded kernels is bounded and therefore measure spaces and bounded kernels form a subcategory $\mathcal{K}^b$ in $\mathcal{K}$ called the category of bounded kernels.

For (X, A), (X', A') consider the measure space $(X \times X', A \times A')$ where $A \times A'$ is the σ-algebra generated by $U \times V$ with $U \in A$ and $V \in A'$. If $f : (X, A) \to (Y, B)$ and $f' : (X', A') \to (Y', B')$ are bounded kernels define $f \times f'$ as the family which takes (x, x') to the product measure $f(x, -) \times f'(x', -)$ on $Y \times Y'$. Standard results about products of finite measures imply that $f \times f'$ is a bounded kernel. One can easily see that this construction defines a symmetric monoidal structure on $\mathcal{K}^b$ which we will denote by $\otimes$ instead of $\times$ to avoid confusion with the categorical product. The one element set is the unit of this monoidal structure which is why we denote it by $\mathbf{1}$.

Example 3.1 [**net1**] The standard example of a problem which one encounters if one tries to define the product of two measures one of which is not necessarily finite can be found in [4, p.78]. The source of the problem seems to lie in the fact that while all measures are continuous with respect to countable filtered colimits (cf. [4, Lemma 1.10(a)]) only finite measures are continuous with respect to countable filtered limits ([4, Lemma 1.10(b)]). Since limits are required to produce measurable subsets of the product of two measure spaces (e.g. the diagonal), a pair of measures on the factors can not be canonically extended to a measure on the product.

Remark 3.2 For each (X, A) the diagonal $(X, A) \to (X, A) \otimes (X, A)$ and the projection $(X, A) \to \mathbf{1}$ make (X, A) into a (commutative) comonoid in $\mathcal{K}^b$ with respect to the product $\otimes$. Note however that this structure is not natural with respect to morphisms in $\mathcal{K}$.

Definition 3.3 *[***impplem***] Let $f : (X, A) \to (Y, B)$ be a bounded kernel. An implementation of f is a triple $((\Omega, \mathcal{F}), \mathbf{P}, \xi)$ where $(\Omega, \mathcal{F})$ is a measure space, $\mathbf{P} : \mathbf{1} \to (\Omega, \mathcal{F})$ is a finite measure on $(\Omega, \mathcal{F})$ and $\xi : (\Omega, \mathcal{F}) \times X \to (Y, B)$ is a deterministic morphism such that the diagram*

$$\begin{array}{ccc} X & \xrightarrow{\mathbf{P} \otimes Id} & (\Omega, \mathcal{F}) \otimes X \\ \downarrow & & \downarrow \xi \\ (X, A) & \xrightarrow{f} & (Y, B) \end{array}$$

commutes.

Note that in the definition given above we let X denote the object of the category of kernels corresponding to the set X with the σ-algebra of all subsets. The left vertical arrow in our diagram is the deterministic morphism $X \to (X, A)$ which is the identity on the underlying sets.

Remark 3.4 Explain relation to implementations of randomized algorithms.

Lemma 3.5 (Kolmogorov) *[***kol***] Let $f : X \to (Y, B)$ be a bounded kernel. Then there exists a unique measure μ_f on $(Y, B)^X$ such that for any finite set of pairwise distinct points $x_1, \ldots, x_n$ of X and any finite set $V_1, \ldots, V_n$ of elements of B one has*

$$\mu_f(\cap_{i=1}^n A_{(x_i, V_i)}) = \prod_{i=1}^n f(x_i, V_i)$$

Proof: ???

Example 3.6 [**paths1**] Let $X = T$ be an interval of real line. Then Y^T is the space of paths in Y. An elementary measurable subset $A(t, V)$ in $(Y, B)^T$ is the subset of all paths γ such that $\gamma(t) \in V$. More generally $\cap_{i=1}^n A_{(t_i, V_i)}$ in Y^T is the subset of all paths which pass through V_i at time t_i. Lemma 3.5 asserts that any non-deterministic path $\phi : T \to (Y, B)$ defines a measure on $(Y, B)^T$ such that the "size" of $\cap_{i=1}^n A_{(t_i, V_i)}$ relative to this measure is the product of the probabilities (determined by ϕ) that t_i lands in V_i.

Let $ev : (Y,B)^X \otimes X \to (Y,B)$ be the evaluation morphism $(g,x) \mapsto g(x)$. Our choice of the σ-algebra on Y^X implies immediately that it is a measurable map. Consider μ_f as a morphism $\mathbf{1} \to (Y,B)^X$. Then the diagram

$$\begin{array}{ccc} X & \xrightarrow{\mu_f \otimes Id} & (Y,B)^X \otimes X \\ {\scriptstyle Id}\downarrow & & \downarrow{\scriptstyle ev} \\ X & \xrightarrow{f} & (Y,B) \end{array}$$

commutes and provides a canonical implementation of the morphism f. The obvious extension of this construction to bounded kernels $(X,A) \to (Y,B)$ implies the following result.

Lemma 3.7 *[**hasanimpl**] For any bounded kernel $f : (X,A) \to (Y,B)$ the diagram*

$$\begin{array}{ccc} X & \xrightarrow{\mu_f \otimes Id} & (Y,B)^X \otimes X \\ {\scriptstyle Id}\downarrow & & \downarrow{\scriptstyle ev} \\ (X,A) & \xrightarrow{f} & (Y,B) \end{array}$$

where μ_f is the measure of Lemma 3.5, is an implementation of f.

Remark 3.8 Let $f_\alpha : (X_\alpha, A_\alpha) \to (Y,B)$ be a countable family of morphisms in $\mathcal{K}^b$. Our definitions imply that $\coprod f_\alpha$ is a bounded kernel if and only if the functions β_{f_α} are uniformly bounded. This observation shows in particular that $(\coprod X_\alpha, \coprod A_\alpha)$ is not a coproduct of our family in $\mathcal{K}^b$.

Similarly for $f_\alpha : (X,A) \to (Y_\alpha, A_\alpha)$, the family which sends x to the measure $\sum f_\alpha(x,-)$ is not a bounded kernel unless this measure is finite i.e. unless

$$\sum \beta_{f_\alpha} < \infty$$

everywhere on X, which shows that $(\coprod Y_\alpha, \coprod B_\alpha)$ is not a product of our family in $\mathcal{K}^b$.

One can also see (cf. 5.3 below) that sending a family (X_α, A_α) to the coproduct space $(\coprod X_\alpha, \coprod A_\alpha)$ is not a functor from the category of families of objects in $\mathcal{K}$ to $\mathcal{K}$. These properties make the category of bounded kernels to be of limited use. Instead one uses the stochastic category considered in the following section.

4 The stochastic category

A kernel $f : (X,A) \to (Y,B)$ is called stochastic if for any x one has $f(x,Y) = 1$ i.e. if the corresponding measures are probability measures (in probability theory such kernels are also known as Markov kernels). Composition of stochastic kernels is stochastic. The subcategory generated by stochastic kernels is called the *stochastic category*. We denote it by $\mathcal{S}$.

Example 4.1 [**exsc1**]One obtains an important class of stochastic kernels as follows. Consider an (idealized) randomized computer algorithm A which takes as an input a sequence of real numbers $r_1, \ldots, r_m$ and produces as an output a sequence of real numbers $s_1, \ldots, s_n$. Let us assume that our computer has access only to the usual (i.e. equally distributed) random numbers on the interval $I = [0,1]$. Then such an algorithm defines a map

$$\tilde{a} : \mathbf{R}^m \times I^\infty \to \mathbf{R}^n$$

where $\tilde{a}(s_1, \ldots, s_m; \rho_1, \ldots)$ is the result our algorithm will produce for the input $r_1, \ldots, r_m$ if its i-th request for a random number gives ρ_i. Consider the usual Lebesgue measure λ on I^∞. Then sending every $(r_1, \ldots, r_m)$ to the push-out of λ with respect to

$$\tilde{a}_{|(r_1,\ldots,r_m)\times I^\infty} : I^\infty \to \mathbf{R}^n$$

we get a stochastic kernel $a : \mathbf{R}^m \to \mathbf{R}^n$ which we call the kernel corresponding to A. This kernel takes (r, U) where $r \in \mathbf{R}^m$ and $U \subset \mathbf{R}^n$ to the probability that our algorithm will produce a result lying in U when given $r = (r_1, \ldots, r_m)$ as an input.

If A and B are two randomized algorithms such that the output of A can be used as an input for B we map consider the composed algorithm $B \circ A$. It is easy to see that the stochastic kernel corresponding to $B \circ A$ is the composition $b \circ a$ of the stochastic kernels corresponding to A and B. It is also easy to see that the stochastic kernel corresponding to an algorithm is a deterministic morphism if and only if our algorithm is essentially deterministic i.e. while it may request random numbers at some point the output does not depend on which random number it gets.

Note that for a non-empty (X, A) there are no stochastic kernels from (X, A) to $\emptyset$. Therefore, while $\emptyset$ is an initial object of the stochastic category it is not a finial object. On the other hand for any (X, A) there is exactly one stochastic kernel from (X, A) to $\mathbf{1}$. Therefore, $\mathbf{1}$ is the final object of the stochastic category but not of the category of kernels.

For (X, A) and (X', A') the coproduct $(X \coprod X', A \coprod A')$ is easily seen too be the coproduct of (X, A) and (X', A') in the stochastic category. However it is not the product of (X, A) annd (X', A') in the stochastic category since the sum of two probability measures is not a probability measure.

For any measurable map of measure spaces $(X, A) \to (Y, B)$ the corresponding morphism in $\mathcal{K}$ is stochastic. Therefore the functor from measurable spaces to the category of kernels factors through the stochastic category.

Our description of morphisms from infinite coproducts given above implies the following result.

Lemma 4.2 *[l3] Let (X_α, A_α) be a family of measure spaces. Then $(\coprod X_\alpha, A_\alpha^{\cup})$ is a coproduct of this family in the stochastic category.*

Proof: ???

In view of Lemma 4.2 we will write $\coprod(X_\alpha, A_\alpha)$ instead of $(\coprod X_\alpha, A_\alpha^{\cup})$.

Note also that the finite group quotients of Lemma 2.14 remain quotients in the stochastic category.

The tensor product of two stochastic kernels is a stochastic kernel and therefore the symmetric monoidal structure defined above for the category of bounded kernels gives a similar structure on $\mathcal{S}$.

Example 4.3 *[markov2]*Let G be a set which is finite or countable. We consider G as a measure space with respect to the σ-algebra which contains all subsets of G. Then $Hom_{\mathcal{K}^b}(G, G)$ is the set of matrices $(p_{ij})_{i,j\in G}$ such that $p_{ij} \geq 0$, for any i the sum $p_i = \sum_j p_{ij}$ is finite and the set of numbers p_i is bounded. The set $Hom_{\mathcal{S}}(G, G)$ is the set of stochastic matrices with rows and columns numbered by elements of G. The composition of kernels corresponds in this description to multiplication of matrices. If P is an element of this set and $f : G \to \mathbf{1}$ a morphism in $\mathcal{K}$ (corresponding to a random variable by 2.7) then the sequence of random variables $f_n = f \circ G^n$ is the Markov chain generated by the stochastic matrix P.

For any (X, A) let

$$[\mathbf{tr1}] tr_n = \sum_{i=1}^{n} pr_i : (X, A)^{\otimes n} \to (X, A) \tag{5}$$

be the kernel which sends a point $(x_1, \ldots, x_n)$ to the measure $\sum_{i=0}^{n} \delta_{x_i}$. For $n = 0$ we take tr_0 to be the zero kernel. The following lemma gives an important property of stochastic kernels.

Lemma 4.4 *[**comm**] For any stochastic kernel $f : (X, A) \to (Y, B)$ and any $n \geq 0$ the diagram*

$$\begin{array}{ccc}
(X, A)^{\otimes n} & \xrightarrow{f^{\otimes n}} & (Y, B)^{\otimes n} \\
{\scriptstyle tr_n}\downarrow & & \downarrow{\scriptstyle tr_n} \\
(X, A) & \xrightarrow{f} & (Y, B)
\end{array}$$

commutes.

Proof: In view of the definition of tr_n it is sufficient to verify that $pr_i \circ f^{\otimes n} = f \circ pr_i$ for all i. More generally it is sufficient to see that for a kernel $f : X \to Y$ and a stochastic kernel $f' : X' \to Y'$ one has $pr_Y \circ (f \otimes f') = f \circ pr_X$ i.e. that the square

$$\begin{array}{ccc}
X \otimes X' & \xrightarrow{f \otimes f'} & Y \otimes Y' \\
{\scriptstyle pr_X}\downarrow & & \downarrow{\scriptstyle pr_Y} \\
X & \xrightarrow{f} & Y
\end{array}$$

commutes. Let e be the canonical stochastic kernel from an object to the point. We have

$$pr_Y \circ (f \otimes f') = (Id_Y \otimes e) \circ (f \otimes f') = f \otimes (e \circ f') = f \otimes e = f \circ pr_X$$

where the third equality holds since $e \circ f' = e$ eactly means that f' is stochastic.

5 Branching morphisms and branching category

For a measure space (X, A) let $S^n(X, A) = (X, A)^n / \Sigma_n$ be the n-th symmetric power of (X, A). For $n = 0$ we set $S^0(X, A) := \mathbf{1}$ for all (X, A) including the empty set. We further set

$$S^\bullet(X, A) = \coprod_{n \geq 0} S^n(X, A)$$

Example 5.1 [**ex6**]We obviously have:

$$S^\bullet(\emptyset) = \mathbf{1}$$

and

$$S^\bullet(\mathbf{1}) = \mathbf{N}$$

Lemma 2.14 shows that for each n, $S^n(-)$ is a functor from the category of bounded kernels to itself. Since $S^\bullet(X, A)$ is the coproduct of $S^n(X, A)$ in $\mathcal{K}$ we conclude that $S^\bullet(-)$ is a functor from the category of bounded kernels to the category of all kernels. Finally, since coproduct of stochastic kernels is stochastic we conclude that both the individual symmetric powers $S^n(X, A)$ and the total symmetric power $S^\bullet(X, A)$ are functors from the stochastic category to itself.

Remark 5.2 For a sufficiently nice space (X, A) the space $S^\bullet(X, A)$ is isomorphic to the space of integer-valued measures $M((X, A), \mathbf{Z}_+)$ on (X, A). This interpretation of the total symmetric power appears in some probabilistic texts on branching processes (e.g. [**?**]). The theory of measure valued branching processes studies the analogs of branching processes with the integer-valued measures replaced by more general measures.

Remark 5.3 *[***ex7***]*One can easily see that the total symmetric power $S^\bullet$ is not a functor from $\mathcal{K}^b$ to $\mathcal{K}^b$. Indeed consider a kernel $a : \mathbf{1} \to \mathbf{1}$ where $a > 1$ (see (4)). Then $S^n(a) = a^n$ and $S^\bullet(a)$ is not bounded since the volumes of corresponding measures on $\mathbf{N}$ are $a, a^2, \ldots$ which is an unbounded function on $\mathbf{N}$.

Definition 5.4 *[***d2***] A branching morphism ϕ from (X, A) to (Y, B) is a morphism in $\mathcal{S}$ of the form $(X, A) \to S^\bullet(Y, B)$.*

The functor $S^\bullet(-)$ is an extension to $\mathcal{S}$ of a functor with the same notation and meaning on the category of measure spaces and measurable maps to itself. In particular the obvious monad structure

$$S^\bullet \circ S^\bullet \to S^\bullet$$

$$Id \to S^\bullet$$

of the total symmetric power functor on sets defines a monad structure on $S^\bullet$ on $\mathcal{S}$. We define the *branching category* B as the category of free algebras over $S^\bullet$. The objects of B are again measure spaces (X, A) and morphisms from (X, A) to (Y, B) are the branching morphisms of Definition 5.4.

Remark 5.5 [**notfree**] In view of Lemma 4.2 algebras over $S^\bullet$ are exactly commutative monoids in $\mathcal{S}$ with respect to $\otimes$.

We will write $\phi : [X, A] \to [Y, B]$ for branching morphisms to distinguish them from morphisms in $\mathcal{K}$ and $\mathcal{S}$. Let us describe the composition of branching morphisms more explicitly. Observe first that there is a measurable map of measure spaces

$$m : S^\bullet(Y, B) \times S^\bullet(Y, B) \to S^\bullet(Y, B)$$

which makes $S^\bullet(Y, B)$ into a commutative monoid. In view of Lemma 2.14 and the definition of the symmetric product it shows that any kernel ϕ from (X, A) to $S^\bullet(Y, B)$ in $\mathcal{K}^b$ defines a family of kernels of the form

$$\phi_n : S^n(X, A) \to S^\bullet(Y, B)$$

(where we set ϕ_0 to be identically 1). If the original kernel is stochastic so are the kernels ϕ_n and therefore by Lemma 4.2 they define a kernel

$$\phi_* = \coprod \phi_n : S^\bullet(X, A) \to S^\bullet(Y, B)$$

We can now define the composition of two branching morphisms by the rule:

$$\psi \circ_B \phi := \psi \circ \phi_*$$

Forgetting the $S^\bullet$ algebra structure defines a functor

$$F : B \to \mathcal{S}$$

which takes (X, A) to $S^\bullet(X, A)$ and ϕ to the kernel ϕ_* defined above.

Example 5.6 [**ex8**]Consider morphisms in the branching category of the form $\phi : [\mathbf{1}] \to [\mathbf{1}]$. Since $S^\bullet(\mathbf{1}) = \mathbf{N}$ we may identify this set with the set of probability measures on $\mathbf{N}$. For any ϕ let $p_\phi = \sum p_i t^i$ be the generating function of this measure. This construction identifies $Hom_B([\mathbf{1}], [\mathbf{1}])$ with formal power series $\sum p_i t^i$ satisfying $p_i \geq 0$ and $\sum p_i = 1$. If ϕ, ψ two endomorphisms of $[\mathbf{1}]$ in B then one has

$$[\mathbf{compseries}] p_{\phi\circ\psi} = p_\psi(p_\phi(t)) \tag{6}$$

i.e. in this description the composition of morphisms corresponds to the composition of power series in the reverse order.

Example 5.7 [**ex10**]The previous example has an immediate generalization to branching morphisms of the form $\phi : [\mathbf{n}] \to [\mathbf{n}]$ where $\mathbf{n} := \coprod_{i=1}^n \mathbf{1}$ is the set of n elements considered as a measure space with respect to the maximal σ-algebra. Such morphism is a collection of n probability measures on $\mathbf{N}^n$. If we describe these measures through their generating functions we may identify $Hom_B([\mathbf{n}], [\mathbf{n}])$ with the set of n-tuples $(f_1, \ldots, f_n)$ where each f_i is a formal power series in n-variables with non-negative coefficients satisfying the condition $f_i(1, \ldots, 1) = 1$. The composition of morphisms corresponds to the substitution composition for such n-tuples.

The kernel (5) is clearly invariant under the action of the symmetric group and by Lemma 4.2 it defines a bounded kernel

$$tr_n : S^n(X, A) \to (X, A)$$

which sends the point $x_1, \ldots, x_n$ to the sum of δ-measures $\delta_{x_1} + \ldots + \delta_{x_n}$ (for $n = 0$ our kernel is 0) and which we continue to denote by tr_n. The coproduct of tr_n's is a kernel $tr_* : S^\bullet(X, A) \to (X, A)$. For a stochastic kernel $(X, A) \to S^\bullet(Y, B)$ (i.e. for a branching morphism $\phi : [X, A] \to [Y, B]$) define a kernel

$$tr(\phi) : (X, A) \to (Y, B)$$

as the composition $tr_* \circ \phi$.

Proposition 5.8 [**comm2**] *For any ϕ as above the diagram*

$$\begin{array}{ccc} S^\bullet(X, A) & \xrightarrow{\phi_*} & S^\bullet(Y, B) \\ {\scriptstyle tr_*}\downarrow & & {\scriptstyle tr_*}\downarrow \\ (X, A) & \xrightarrow{tr(\phi)} & (Y, B) \end{array}$$

commutes.

Proof: By definition of ϕ_* it is sufficient to verify that for any n the diagram

$$\begin{array}{ccccc} (X, A)^{\otimes n} & \xrightarrow{\phi^{\otimes n}} & S^\bullet(Y, B)^{\otimes n} & \xrightarrow{m} & S^\bullet(Y, B) \\ {\scriptstyle tr_n}\downarrow & & {\scriptstyle tr_n}\downarrow & & \downarrow{\scriptstyle tr_*} \\ (X, A) & \xrightarrow{\phi} & S^\bullet(Y, B) & \xrightarrow{tr_*} & (Y, B) \end{array}$$

commutes. The right hand side square consists of kernels which take a point to the sum of finitely many points and it is easy to verify its commutativity explicitly. The left hand side square commutes by Lemma 4.4.

Corollary 5.9 *[**main1**] For a pair of branching morphisms $\phi : [X, A] \to [Y, B]$, $\psi : [Y, B] \to [Z, C]$ one has*

$$tr(\psi \circ \phi) = tr(\psi) \circ tr(\phi)$$

Proof: This follows immediately from the explicit description of the composition of branching morphisms given above and Lemma 5.8.

Example 5.10 [**ex11**]Consider a branching morphism $\phi : [\mathbf{1}] \to [\mathbf{1}]$ which we describe through the corresponding probability generating function $p_\phi = \sum p_i t^i$ as in Example 5.6. Then $tr(\phi)$ is a kernel $\mathbf{1} \to \mathbf{1}$ i.e. a non-negative number. One can easily see that

$$tr(\phi) = \sum i p_i = p'_\phi(1)$$

where p'_ϕ is the formal derivative of p_ϕ with respect to t. In other words, $tr(\phi)$ is in this case the expectation value of ϕ. For two morphisms ϕ, ψ of this form Corollary 5.9 asserts that

$$tr(\psi \circ \phi) = tr(\psi) tr(\phi).$$

In view of (6) this follows from the equation

$$(p_\phi \circ p_\psi)'(1) = p'_\psi(1) p'_\phi(p_\psi(1)) = p'_\psi(1) p'_\phi(1)$$

where the last equation holds since the $p_\psi(1) = 1$ because ψ is a stochastic kernel.

Example 5.11 [**ex12**]Consider now branching morphisms $[\mathbf{n}] \to [\mathbf{n}]$ as in Example 5.7. For a morphism ϕ of this form $tr(\phi)$ is a kernel $\mathbf{n} \to \mathbf{n}$ i.e. an $n \times n$-matrix (a_{ij}) with entries from $[0, \infty]$. If we represent ϕ a sequence of power series $(f_1, \ldots, f_n)$ in variables $t_1, \ldots, t_n$ then one gets

$$a_{ij} = \frac{\partial f_i}{\partial t_j}(1)$$

If $\psi = (g_1, \ldots, g_n)$ is another such morphism then the statement of Corollary 5.9 is again equivalent to the formula for the differential of a composition combined with the fact that $g_i(1) = 1$ since ψ is stochastic.

3 A categorical view of Markov processes

1 Processes and Kolmogorov equations

2 Trajectory structures

4 Tonus spaces

1 Tonus spaces

Definition 1.1 [**conus**] *A* conus structure *on a set C is an abelian semi-group structure (with unit 0) together with a map*

$$m : \mathbf{R}_{\geq 0} \times C \to C$$

which makes C into a module over $\mathbf{R}_{\geq 0}$ i.e. such that

$$[\mathbf{eqpo1}] m(r, x + y) = m(r, x) + m(r, y) \tag{7}$$

$$[\mathbf{eqpo3}] m(r+s,x) = m(r,x) + m(s,x) \tag{8}$$

$$[\mathbf{eqpo4}] m(rs,x) = m(r,m(s,x)) \tag{9}$$

$$[\mathbf{eqpo6}] m(1,x) = x \tag{10}$$

$$[\mathbf{eqpo5}] m(0,x) = 0 \tag{11}$$

When no confusion is possible we write rx instead of $m(r,x)$. A set with a conus structure is called a conus space*.*

Definition 1.2 *[**dpo1**] A* tonus structure *on a set C is a topology together with a conus structure such that the addition and the multiplication by scalars are continuous.*

Definition 1.3 *[**dpo2**] Let C_1, C_2 be two conus (resp. tonus) spaces. A morphism $f : C_1 \to C_2$ is a map (resp. a continuous map) which commutes with addition and multiplication by scalars.*

We let T denote the category of tonus spaces.

Proposition 1.4 *[**ppo1**] The category T has all limits. The final object of T is the one point space. For any diagram $\mathcal{D}$ of tonus spaces the underlying topological space of $lim(\mathcal{D})$ is the limit of the corresponding diagram of topological spaces and the same is true for the limit of the underlying diagram of conus spaces and abelian semi-groups.*

Proof: Straightforward.

Proposition 1.5 *[**ppo2**] The category T of tonus spaces has colimits. The initial object of T is the one point space.*

Proof: The statement of the proposition follows from Lemmas 1.6-1.8 below and the usual reduction of general colimits to inductive colimits, reflexive coequalizers and finite coproducts.

Lemma 1.6 *[**lpo5**] Let $(C_\alpha, f_{\alpha\beta} : C_\alpha \to C_\beta)$ be an inductive system of tonus spaces. Let C be the colimit of this sequence in the category of sets which we consider with the colimit topology and the obvious operations of addition and multiplication by elements of $\mathbf{R}_{\geq 0}$. Then C is a tonus space and a colimit of our sequence in T.*

Proof: It follows by direct verification using the fact that inductive colimits commute with finite products in the category of topological spaces.

Lemma 1.7 *[**lpo6**] Let C_1, C_2 be tonus spaces, $f, g : C_1 \to C_2$ two morphisms and $s : C_2 \to C_1$ a common section of f and g (i.e. f, g, s form a reflexive coequalizer diagram). Let C be the coequalizer of f and g in the category of sets which we consider with the coequalizer topology and the obvious operations of addition and multiplication by elements of $\mathbf{R}_{\geq 0}$. Then C is a tonus space and a coequalizer of f and g in T.*

Proof: As in the proof of Lemma 1.6 everything follows by direct verification from the fact that reflexive coequalizers commute with finite products.

Lemma 1.8 *[**lpo7**] Let C_1, C_2 be tonus spaces. Let $C = C_1 \times C_2$ and consider C with the topology of the product and the obvious operations of addition and multiplication by elements of $\mathbf{R}_{\geq 0}$. Then C is a tonus spaces which is both the product and the coproduct of C_1 and C_2 in T.*

Proof: The only non-trivial part of the lemma is that C is the coproduct of C_1 and C_2 i.e. that for any tonus space D the map

$$Hom(C,D) = Hom(C_1,D) \times Hom(C_2,D)$$

given by the composition with the embeddings $C_1 \to C$, $C_2 \to C$ is bijective. It is clearly injective and to verify that it is bijective it is enough to prove that a map $f : C_1 \times C_2 \to D$ which is compatible with the algebraic structures and whose restrictions f_1, f_2 to $C_1 \times \{0\}$ and $\{0\} \times C_2$ are continuous is itself continuous. This follows from the fact that $f = m_D \circ (f_1 \times f_2)$ and the continuity of $m_D : D \times D \to D$.

Definition 1.9 *[**grouplike**] A tonus space C is called group-like if the underlying semi-group is a group.*

For the basic definitions related to the topological vector spaces and pre-ordered vector spaces we follow [**?**].

Lemma 1.10 *[**lpo3**] Let V be a group-like tonus space. Then there exists a unique extension of $m : \mathbf{R}_{\geq 0} \times V \to V$ to a continuous map $m : \mathbf{R} \times V \to V$ satisfying the condition*

$$m(r-s,x) = m(r,x) - m(s,x)$$

and with respect to this map V becomes a topological vector space (over $\mathbf{R}$).

Proof: The uniqueness is obvious. It is also obvious that if m as required exists then it makes V into a topological vector space. To prove the existence consider the map $\tilde{m} : \mathbf{NR} \times \mathbf{R}_{\geq 0} \times V \to V$ of the form $\tilde{m}(r,s,x) = m(r,x) - m(s,x)$. The algebraic properties of m imply that it has a decomposition

$$\mathbf{R}_{\geq 0} \times \mathbf{R}_{\geq 0} \times V \to \mathbf{R} \times V \stackrel{m}{\to} V$$

where the first arrow is defined by $(r,s) \mapsto r-s$. Since the first arrow is a strict topological epimorphism and the composition is continuous we conclude that m is continuous.

Lemma 1.11 *[**lpo4**] Let C be a tonus space and let $C \to V_C$ be the universal map from C as an abelian semi-group to an abelian group. Then V has a unique structure of a tonus space such that $C \to V_C$ is a morphism of tonus spaces. With this structure $C \to V_C$ is the universal morphism from C to a group-like tonus space.*

Proof: By (see e.g. []) we may describe V_C as the set of equivalence classes of pairs (x,y), $x,y \in C$ such that $(x_1,y_1) \cong (x_2,y_2)$ if and only if there exists u such that $x_1 + y_2 + u = x_2 + y_1 + u$. As usual we will write $x - y$ for the equivalence class of (x,y). For $r \in \mathbf{R}_{\geq 0}$ set $r(x,y) = (rx, ry)$. In view of 7 this defines a map $\mathbf{R}_{\geq 0} \times V_C \to V_C$ which takes $x - y$ to $rx - ry$ and one verifies easily that it satisfies the conditions 8-11. Let $\pi : C \times C \to V_C$ be the surjection $(x,y) \mapsto x - y$. Consider V_C as topological space with the topology defined by π i.e. such that U is open in V_C if and only if $\pi^{-1}(U)$ is open in $C \times C$. The universal properties of this topology imply immediately that the addition and multiplication by elements from $\mathbf{R}_{\geq 0}$ are continuous for V and we conclude that V has a structure of a tonus space such that $C \to V_C$ is a morphism of tonus spaces. One can see immediately that such a structure is unique.

Definition 1.12 *[**cancellable**] A tonus space C is called pre-group like if the universal map $C \to V_C$ is an injection i.e. if the underlying semi-group is a semi-group with cancellation.*

Definition 1.13 *[**reduced**] A tonus space C is called reduced if it is pre-group like and the topology on C induced by the map $C \to V_C$ coincides with the original topology.*

Definition 1.14 *[**closedts**] A tonus space C is called closed the corresponding universal map $C \to V_C$ is a closed embedding.*

Clearly any closed tonus space is reduced and any reduced is a pre-group like. It is also clear that any group-like tonus space is closed. To produced counter-examples to other implications we will use the following lemma.

Lemma 1.15 *[**need1**] Let $f : C \to V$ be a monomorphism from a tonus space C to a group-like tonus space V and let V_0 be the set of interior points of $f(C)$ in V. Assume that the following two conditions hold:*

1. *the map $C_0 = f^{-1}(V_0) \to V_0$ is a homeomorphism,*
2. *for any $v \in V$ there exist $x, y \in V_0$ such that $v = x - y$.*

Then $V(f) : V_C \to V$ is an isomorphism.

Proof: Clearly $V(f)$ is bijective as a map of sets and continuous. Let us show that it is open. Let V_0 be the set of interior points of $f(C)$ it is open in V and the restriction of f to $C_0 = f^{-1}(V_0)$ is an isomorphism. Consider the diagram:

$$\begin{tikzcd} C_0 \times C_0 \arrow[r, equal] \arrow[d, "p_0"'] & V_0 \times V_0 \arrow[d, "q_0"] \\ V_C \arrow[r, "V(f)"] & V \end{tikzcd}$$

where the vertical arrows map (u, v) to $u - v$ and f_0 is the restriction of f to C_0. Our conditions imply that q_0 is surjective. Since V_0 is open in V and the subtraction map $V \times V \to V$ is open (follows from the fact that it is isomorphic to the projection $V \times V \to V$ to one of the factors) we conclude that q_0 is also open. This immediately implies that $V(f)$ is open.

Example 1.16 [**contr2**]Not all reduced tonus spaces are closed. Indeed let C be the subset in $\mathbf{R}^2$ which consists of points (x, y) such that $x \geq 0$ and $y > 0$ and the point $(0, 0)$. Considered with the induced topology and the obvious addition and multiplication by scalars C is a tonus space. Lemma 1.15 implies immediately that the embedding $C \to \mathbf{R}^2$ coincides with the universal embedding to a group-like tonus space. Therefore C is reduced but not closed.

Example 1.17 [**contr1**] Not any pre-group like tonus space is reduced. Consider the subset C in $\mathbf{R}^2$ which consists of (x, y) such that $x, y \geq 0$. Let further U be the subset of elements of C of the form $(x, 0)$ where $x > 0$. Consider the topology on C which is generated by the usual topology coming from $\mathbf{R}^2$ together with the condition that U is open. One verifies immediately that the addition and multiplication by scalars are continuous in this topology. On the other hand Lemma **??** again implies that the embedding $C \to \mathbf{R}^2$ is the universal one. Since in the topology on C induced by this embedding U is not open we conclude that C is pre-group like but not reduced.

Example 1.18 [**expo1**]Not all tonus spaces are pre-group like. Indeed, consider the set $\{0, 1\}$ with the discrete topology, the abelian semi-group structure given by $0 + 0 = 0$, $0 + 1 = 1$, $1 + 1 = 1$ and m given by $m(r, 0) = 0$, $m(r, 1) = 1$ if $r \neq 0$ and $m(0, 1) = 0$. These structures satisfiy all the conditions of Definition 1.2 but the resulting tonus space C is not pre-group like since $V_C = 0$. We will see below (Lemma 1.20) however that all Hausdorf tonus spaces are pre-group like. Note that the spaces in Examples 1.16 and 1.17 are both Hausdorf. Thus a Hausdorf tonus space need not be reduced or closed.

Sending C to $(V_C, C_{red}))$ where C_{red} is the image of C in V_C considered with the topology induced from V_C we get (by Lemmas 1.10, 1.11) a functor from tonus spaces to pairs (V, C) where V is a topological vector space and C is a cone in V. Clearly this functor is a full embedding on the subcategory of reduced tonus spaces and the pair (V, C) is in the image of this embedding if and only if any element of V can be written as $x - y$ where x, y are in C. Recall that a pre-ordered topological vector space is a pair as above such that C is closed in V. Therefore, we get the following result.

Proposition 1.19 *[**embed1**] The category of closed tonus spaces is equivalent to the full subcategory of the category of pre-ordered topological vector spaces (V, C) such that any element of V is of the form $x - y$ for $x, y \in C$.*

Lemma 1.20 *[**lpo1**] Let C be a Hausdorf tonus space then one has:*

1. *C is pre-group like i.e. for any x, y, u in C such that $x + u = y + u$ one has $x = y$*
2. *$m(r, 0) = 0$*

Proof: Let us denote $m(r, x)$ by rx. Consider the first claim. By 10 and 8 for any positive integer n we have $nx = \sum_{i=1}^{n} x$. From this by easy induction we get that for x, y, u as above one has $nx + u = ny + u$. By 7 and 9 we get that

$$x + (1/n)u = y + (1/n)u$$

Since C is Hausdorf a sequence may have only one limit and from the continuity of addition and multiplication by a number and 11 we get

$$x = x + 0u = \lim_{n\to\infty}(x + (1/n)u) = \lim_{n\to\infty}(y + (1/n)u) = y + 0u = y.$$

To get the second claim note that by 7 we have $r0 + r0 = r0 = r0 + 0$ and we conclude from the first part of the proof that $r0 = 0$.

Lemma 1.21 *[**hus**] Let C be a Hausdorf tonus space C. Then V_C is Hausdorf.*

Proof: Consider the natural map $\pi : C \times C \to V_C$. If C is Hausdorf then by Lemma 1.20 we have $\pi^{-1}(0) = \Delta$ where Δ is the diagonal. Since in a Hausdorf space the diagonal is closed and since π is a topological epimorphism we conclude that $\{0\}$ is closed in V_C. Since V_C is a topological vector space this implies in the standard way that V_C is Hausdorf.

Let C be a conus space and let $f_\alpha : C \to C_\alpha$ be a collection of conus maps to tonus spaces C_α. Let $t(f_\alpha)$ be the weakest topology on C which makes all the maps f_α continuous. It is easy to see that with this topology C is a conus space. We will say that the topology on C is defined by the collection f_α.

Lemma 1.22 *[**isred1**] Let C be a pre-group like conus space and let $f_\alpha : C \to C_\alpha$ be a collection of morphisms to reduced tonus spaces. Then C with the induced topology is a reduced tonus space.*

Proof: Let $C \to V_C$ and $C_\alpha \to V_\alpha$ be the universal morphisms to group-like spaces. By universality we get commutative squares

$$\begin{array}{ccc} C & \xrightarrow{f_\alpha} & C_\alpha \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p_\alpha} \\ V & \xrightarrow{g_\alpha} & V_\alpha \end{array}$$

such that g_α are continuous. Let $x \in U \subset C$ be an open neighborhood of x in C. We have to show that there is an open neighborhood U' of $p(x)$ in V such that $p^{-1}(U') \subset U$. Since the topology on C is defined by (f_α) there exists a finite set $\alpha_1, \ldots, \alpha_n$ and open neighborhoods $W_1, \ldots, W_n$ of $f_{\alpha_i}(x)$ in C_α such that U contains $\cap f_{\alpha_i}^{-1}(W_i)$. Since each C_α is assumed to be reduced we have $W_i = p_{\alpha_i}^{-1}(W_i')$ for some W_i' open in V_α. The commutativity of our squares imply now that

$$\cap p^{-1} g_\alpha^{-1}(W_i') \subset U.$$

Remark 1.23 [**impo**] It is important to note that (in the notations of Lemma 1.22) the universal topology on V defined by the topology on C need not coincide with the topology induced by the maps $g_\alpha : V \to V_\alpha$. For an example see **??**.

In the following lemma we keep the notations of Lemma 1.22.

Lemma 1.24 *[**isclosed**] Let C be a pre-group like conus space and $f_\alpha : C \to C_\alpha$ a collection of maps to closed tonus spaces such that if $x \in V$ is an element satisfying $g_\alpha(x) \in C_\alpha$ for all α then $x \in C$. Then with the topology defined by (f_α), C is a closed tonus space.*

Proof: By Lemma 1.22 C is reduced. It remains to check that the image of C in V is closed. Let $x \in V$ be an element outside of C. Then by our assumption there exists α such that $g_\alpha(x)$ is outside C_α. Since C_α are closed this implies that there is a neighborhood W of $g_\alpha(x)$ which does not intersect C_α. Then $g_\alpha^{-1}(W)$ is a neighborhood of x which does not intersect C.

2 Embedding $\mathcal{K}^{op} \to T$

Let (X, A) be a measure space and $M^+(X, A)$ the set of non-negative measurable functions on (X, A). It has an obvious structure of a conus space. Define the standard topology on $M^+(X, A)$ by the condition that a set Z is closed if and only if for any sequence f_n of elements of Z such that $f_n \uparrow f$ we have $f \in Z$.

3 Embedding $\mathcal{K} \to T$

Let (X, A) be a measurable space and let $M_+(X, A)$ be as above the set of all bounded measures on (X, A). Any (bounded, non-negative) measurable function $f \in M^+(X, A)$ defines a map

$$f_* : M_+(X, A) \to \mathbf{R}_{\geq 0}$$

Define the *standard topology* on $M_+(X, A)$ as the weakest topology which makes all maps of the form f_* continuous.

Lemma 3.1 *[**lem4**] A map u from a topological space T to $M_+(X, A)$ is continuous with respect to the standard topology if and only if for any $f \in M^+(X, A)$ the composition*

$$f_* \circ u : T \to \mathbf{R}_{\geq 0}$$

is continuous.

Lemma 3.2 *[***lem1***] The set $M_+(X,A)$ considered with the standard topology and the addition and multiplication by elements of $\mathbf{R}_{\geq 0}$ defined in the obvious is a closed, Hausdorf tonus space.*

Proof: The continuity of the addition and multiplication by scalars follow from Lemma 3.1. To see that the standard topology is Hausdorf consider two measures μ_1 and μ_2 such that $\mu_1 \neq \mu_2$. Then there is a measurable subset $U \in A$ such that $\mu_1(U) \neq \mu_2(U)$. Let f be the indicator function of U. Then for any μ, $f_*(\mu) = \mu(U)$ and if V_1, V_2 are two non-intersecting neighborhoods of $\mu_1(U)$ and $\mu_2(U)$ respectively then $f_*^{-1}(V_i)$ give us two non-intersecting neighborhoods of μ_1 and μ_2.

To see that $C = M_+(X,A)$ is closed in the corresponding vector space V we need to check that if μ_1, μ_2 are two measures such that $x = \mu_1 - \mu_2$ is not in C then there exists a neighborhood N of x in V such that $N \cap C = \emptyset$. By Lemma 1.11, V is universal and therefore any map of the form f_* extends to a continuous map $f_* : V \to \mathbf{R}$. Since x is not in C there exists a measurable subset $U \in A$ such that $x(U) = \mu_1(U) - \mu_2(U) < 0$. Let W be a neighborhood of $x(U)$ which lies in $(-\infty, 0)$. Taking f to be the indicator function of U we get a neighborhood $f_*^{-1}(W)$ of x which does not intersect C.

Remark 3.3 [**dense**]Unless A is finite the image of $C = M_+(X,A)$ in the corresponding universal group-like tonus space V has no internal points i.e. the complement to C in V is dense.

Lemma 3.4 *[***lem2***] Let $\phi : (X,A) \to (Y,B)$ be a bounded kernel. Then the composition with ϕ defines a map*

$$\phi_* : M_+(X,A) \to M_+(Y,B)$$

which is a morphism of tonus spaces.

Proof: Follows from Lemma 3.1.

Remark 3.5 [**rem1**]Consider the metric on $M_+(X,A)$ given by

$$[\mathbf{eqem1}]\nu(\mu_1, \mu_2) = sup_{U \in A} |\mu_1(U) - \mu_2(U)| \tag{12}$$

Remark 3.6 The proof of Lemma 3.4 implies that if ϕ is a (sub-)stochastic kernel then the corresponding map $M_+(\phi)$ does not increase the distances between measures.

Remark 3.7 [**rem1**]For any point x of (X,A) we have the δ-measure δ_x concentrated in x. Evaluating ϕ_* on δ_x we get a measure $\phi_*(\delta_x)$ on (Y,B) and one verifies easily that it is exactly $\phi(x,-)$. This shows that for any (X,A), (Y,B) the map

$$Hom_{\mathcal{K}}((X,A),(Y,B)) \to Hom_T(M_+(X,A), M_+(Y,B))$$

is a monomorphism. We will see below in Theorem 3.12 that it is in fact a bijection.

Let μ be a measure on (X,A) and let $X = \coprod_{i=1}^n X_i$ be a partition of X into a disjoint union of measurable subsets. For any $\delta > 0$ denote by $U(\mu, \delta, (X_i))$ the set of all measures λ on (X,A) such that for each $i = 1, \ldots, n$ one has

$$|\mu(X_i) - \lambda(X_i)| < \delta.$$

Clearly $U(\mu, \delta, (X_i))$ is an open neighborhood of μ in the standard topology.

Lemma 3.8 *[***lem55***] Subsets of the form $U(\mu, \delta, (X_i))$ form a fundamental system of open neighborhoods of μ in the standard topology.*

Proof: If $X = \coprod_{i=1}^{n} X_i$ and $X = \coprod_{j=1}^{m} Y_j$ are two measurable partitions of X then $X = \coprod (X_i \cap Y_j)$ is also a measurable partition of X. Let $\delta > 0$ be a real number and k be an integer such that $k \geq n$ and $k \geq m$. Let λ be an element of $U(\mu, \delta/k, (X_i \cap Y_j))$. Then

$$|\mu(X_i) - \lambda(X_i)| = |\sum_{j=1}^{m} (\mu(X_i \cap Y_j) - \lambda(X_i \cap Y_j)| \leq \sum_{j=1}^{m} |(\mu(X_i \cap Y_j) - \lambda(X_i \cap Y_j)| \leq (m/k)\delta \leq \delta$$

i.e. $\lambda \in U(\mu, \delta, (X_i))$. Similarly $\lambda \in U(\mu, \delta, (Y_j))$ and we conclude that the intersection of two subsets of the type we consider contains a third subset of the same type.

The standard topology is generated by the maps $f_* : \mu \mapsto \int f d\mu$ for bounded non-negative measurable functions f. In particular for any μ finite intersections of subsets of the form

$$U(\mu, \epsilon, f) = \{\lambda : |\int f d\mu - \int f d\lambda| < \epsilon\}$$

form a fundamental system of open neighborhoods of μ. It remains to show that any neighborhood of the form $U(\mu, \epsilon, f)$ contains a neighborhood of the form $U(\mu, \delta, (X_i))$ i.e. that for any f and any $\epsilon > 0$ there exists a partition $X = \coprod X_i$ and $\delta > 0$ such that for any λ satisfying

$$|\mu(X_i) - \lambda(X_i)| < \delta$$

we have

$$|\int f d\mu - \int f d\lambda| < \epsilon.$$

Without loss of generality we may assume that $f(x) < 1$ for all $x \in X$. Let $n > 0$ be an integer. For $k = 0, \ldots, n-1$ set $I_k = [k/n, (k+1)/n)$. Then

$$[0,1) = \coprod_{k=0}^{n-1} I_k$$

is a measurable partition of the interval $[0,1)$. Let further $X_k = f^{-1}(I_k)$ and let

$$f_n = \sum_{k=0}^{n-1} k/n F_k$$

where F_k is the indicator function of X_k. By construction we have $f(x) \geq f_n(x)$ and $f(x) - f_n(x) < 1/n$ for all $x \in X$. For any λ we have

$$|\int f d\mu - \int f d\lambda| \leq |\int (f - f_n) d\mu - \int (f - f_n) d\lambda| + |\int f_n d\mu - \int f_n d\lambda| \leq$$

$$\leq |\int (f - f_n) d\mu| + |\int (f - f_n) d\lambda| + \sum_{k=0}^{n-1} k/n |\mu(X_k) - \lambda(X_k)| \leq$$

$$\leq \mu(X)/n + \lambda(X)/n + \sum_{k=0}^{n-1} k/n |\mu(X_k) - \lambda(X_k)| \leq$$

We also have:

$$\lambda(X) = \sum_{k=0}^{n-1} \lambda(X_k)| \leq \sum_{k=0}^{n-1} |\mu(X_k) - \lambda(X_k)| + \sum_{k=0}^{n-1} \mu(X_k) = \sum_{k=0}^{n-1} |\mu(X_k) - \lambda(X_k)| + \mu(X)$$

and therefore

$$|\int fd\mu - \int fd\lambda| \leq 2\mu(X)/n + \sum_{k=0}^{n-1}(k+1)/n|\mu(X_k) - \lambda(X_k)| \leq$$

$$\leq 2\mu(X)/n + (1+1/n)\sum_{k=0}^{n-1}|\mu(X_k) - \lambda(X_k)|$$

To find n, δ such that $U(\mu, \delta, (X_k)_{k=0}^{n-1})$ is contained in $U(\mu, \epsilon, f)$ it is sufficient now to choose n such that $2\mu(X)/n < \epsilon$ and then choose δ such that $(n+1)\delta < \epsilon - 2\mu(X)/n$.

Let $M_*(X, A)$ be the universal group-like tonus space associated with $M_+(X, A)$ i.e. the space of signed measures on (X, A) with the topology defined by the canonical map

$$p : M_+(X, A) \times M_+(X, A) \to M_*(X, A)$$

For any $f \in M^+(X, A)$ the map $f_* : M_+(X, A) \to \mathbf{R}_{\geq 0}$ defines a map $M_*(X, A) \to \mathbf{R}$ which we will also denote by f_*.

Lemma 3.9 *[**imp1**] The topology on $M_*(X, A)$ coincides with the topology defined by the linear functionals f_* for $f \in M^+(X, A)$.*

Proof: Let $\mu = \mu_+ - \mu_-$ be an element of $M_*(X, A)$ and U be a subset in $M_*(X, A)$ which contains μ and such that $p^{-1}(U)$ is open in $M_+(X, A) \times M_+(X, A)$. We need to verify that there exists a finite set $f_1, \dots, f_n$ of elements of $M^+(X, A)$ and $\delta > 0$ such that for any $\lambda = \lambda_+ - \lambda_-$ in $M_*(X, A)$ satisfying

$$|\int f_i d\lambda - \int f_i d\mu| < \delta$$

for all $i = 1, \dots, n$, we have $\lambda \in U$. The condition that $p^{-1}(U)$ is open together with Lemma 3.8 implies that there exists $\epsilon > 0$ and a measurable partition $X = \coprod_{i=1}^m X_i$ such that for any pair of measures λ_+, λ_- satisfying

$$|\lambda_+(X_i) - \mu_+(X_i)| < \epsilon|$$

$$|\lambda_-(X_i) - \mu_-(X_i)| < \epsilon|$$

one has $\lambda_+ - \lambda_- \in U$.

Proposition 3.10 *[**tem1**] The map $f \mapsto f_*$ gives a bijection*

$$M^+(X, A) \to Hom_T(M_+(X, A), \mathbf{R}_{\geq 0}).$$

Its inverse takes a map ϕ of tonus spaces to the function f such that for each $x \in X$ one has $f(x) = \phi(\delta_x)$.

Proof: Let $\phi : M_+(X, A) \to \mathbf{R}_{\geq 0}$ be a morphism.

Corollary 3.11 *[**definedby**] Let $f, g : M_+(X, A) \to \mathbf{R}_{\geq 0}$ be two morphisms of tonus spaces which coincide on measures of the form δ_x for all $x \in X$. Then $f = g$.*

Theorem 3.12 *[**t1**] The functor $\mathcal{K} \to T$ sending (X, A) to $M_+(X, A)$ is a full embedding. I.e. For any measurable spaces (X, A), (Y, B) the map*

$$[\mathbf{mm}] Hom_{\mathcal{K}}((X, A), (Y, B)) \to Hom_T(M_+(X, A), M_+(Y, B)) \tag{13}$$

is a bijection. Its inverse takes a map ϕ of tonus spaces to the kernel ψ such that for each $x \in X$ the measure $\phi(x, -)$ is $f(\delta_x)$.

Proof: We already noted in Remark 3.7 that the map (13) is injective. To show that it is surjective consider a morphism $\phi : M_+(X, A) \to M_+(Y, B)$ of tonus spaces. Let U be a measurable subset of Y and let I_U be its indicator function. The composition of ϕ with the morphism $M_+(Y, B) \to \mathbf{R}_{\geq 0}$ defined by I_U is, by Proposition 3.10 a measurable function on (X, A) whose value on $x \in X$ is $\phi(\delta_x)(U)$. Therefore, a map $\psi : X \times B \to \mathbf{R}_{\geq 0}$ of the form $\psi(x, U) = \phi(\delta_x)(U)$ is a kernel. It remains to show that the map $\psi_* : M_+(X, A) \to M_+(Y, B)$ defined by this kernel is ϕ. We know that it coincides with ϕ on delta measures. Since the measurable functions on (Y, B) distinguish elements of $M_+(Y, B)$ it is sufficient to check that the compositions of ϕ and ψ_* with any map $M_+(Y, B) \to \mathbf{R}_{\geq 0}$ coincide. This follows from Corollary 3.11.

4 Radditive functors on $\mathcal{K}$

Recall that a contravariant functor F from a category C with finite coproducts and initial object 0 is called radditive if $F(0) = pt$ and $F(X \coprod Y) = F(X) \times F(Y)$. We let $R(C)$ denote the full subcategory in the category of all contravariant functors formed by radditive functors. For general properties of radditive functors see [], [].

Lemma 4.1 *[**lrf1**] Let C be a category as above and assume that finite coproducts in C coincide with finite products (in particular $pt = 0$). Then $R(C)$ is equivalent to the category of contravariant functors F from C to the category of abelian semi-groups such that $F(X \coprod Y) = F(X) \times F(Y)$.*

Proof: In the case of an additive C (i.e. under the additional assumption that morphisms in C can be subtracted) the statement is proved in []. The same proof works without subtraction.

5 Accessible spaces

6 Accessible enrichment of $\mathcal{K}$

Let (X, A), (Y, B) be measurable spaces. For any bounded measure μ on (X, A) and a bounded measurable function f on (Y, B) consider the map

$$\eta(\mu, f) : Hom_{\mathcal{K}}((X, A), (Y, B)) \to \mathbf{R}_{\geq 0}$$

sending ϕ to

$$f \circ \phi \circ \mu : \mathbf{1} \to (X, A) \to (Y, B) \to \mathbf{1}.$$

Define the *standard topology* on $Hom_{\mathcal{K}}((X, A), (Y, B))$ as the weakest topology with respect to which all maps $\eta(\mu, f)$ are continuous.

Lemma 6.1 *[**lae1**] The set $Hom_{\mathcal{K}}((X, A), (Y, B))$ with the standard topology and the obvious operations of addition and multiplication by scalar is a closed, Hausdorff tonus space.*

Proof: ???

Lemma 6.2 *[**lem0**] The composition of morphisms in $\mathcal{K}$ defines maps of tonus spaces of the form*

$$Hom_{\mathcal{K}}((X,A),(Y,B)) \otimes Hom_{\mathcal{K}}((Y,B),(Z,C)) \rightarrow Hom_{\mathcal{K}}((X,A),(Z,C)).$$

Proof: ???

Remark 6.3 [**nottopen**] Note that the maps of topological spaces

$$Hom_{\mathcal{K}}((X,A),(Y,B)) \times Hom_{\mathcal{K}}((Y,B),(Z,C)) \rightarrow Hom_{\mathcal{K}}((X,A),(Z,C))$$

defined by composition of morphisms need not be continuous if we take the standard topology on the right and the product of the standard topologies on the left.

7 Notes

To the relativistic Brownian motion. A physical formulation of the problem. There is a particle p moving according to the Brownian motion pattern on a physical line L with a marked Borel subset B. There are three observers X, N_1, N_2 all moving inertially relative to each other. Observer X fixes the act of observation of the particle by N_1 and the result of the observation (particle is in point $l_1 \in L$). He further fixes an act of observation of the same particle by N_2 and bets that N_2 observed the particle in B. What is the probability that he won?

The relative velocities of the observers with respect to each other and to the line are known. Observer X has a clock. For simplicity assume that all the observers are moving along the line L.

Here is another version. There is a physical line L with a 'Brownian motion field' F. An experimenter X which is located at point 0 of L and has a clock T creates an apparatus A which moves along L with a constant speed v. At time $s \in T$ the experimenter emits a light signal. When A receives this signal it places a particle p at its current location on L. From this point on the movement of p is controlled by F. At time $t \in T$ the experimenter emits a second light signal. When A receives this signal it emits a light signal along L which when it reaches p reflects back. When A receives the reflected signal it emits a light signal to X who notices the time of its arrival.

References

[1] Patrick Billingsley. *Probability and measure.* John Wiley & Sons, New York-Chichester-Brisbane, 1979. Wiley Series in Probability and Mathematical Statistics.

[2] Michèle Giry. A categorical approach to probability theory. In *Categorical aspects of topology and analysis (Ottawa, Ont., 1980)*, volume 915 of *Lecture Notes in Math.*, pages 68–85. Springer, Berlin, 1982.

[3] S. MacLane. *Categories for the working mathematician*, volume 5 of *Graduate texts in Mathematics.* Springer-Verlag, 1971.

[4] David Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.