Probability Theory in the Category of Diffeological Spaces

Wiskerke, Jamie

Bachelor Computer Science & Mathematics

Probability theory in the Category of Diffeological Spaces

Jamie Wiskerke

Henning Basold, Oisín Flynn-Connolly & Rajat Hazra

Bacherlor's Thesis

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl August 22, 2025

Contents

1 Introduction

In some branches of mathematics we would like to do probability theory in a differentiable setting. For example, stochastic differential equations, optimization processes or stochastic gradient descent all require a combination of stochastic and smooth structure. One approach to unifying these fields is through categorical methods, particularly the use of monads, which provide a structured way to handle probability measures in a categorical setting.

The Giry monad, introduced by Michele Giry in 1982 [Gir82], is an example of a probability monad, defined over the category of measurable spaces. It assigns to a space the space of probability measures, equipped with natural transformations that satisfy the monad axioms.

While the Giry monad provides a framework for probability theory on measurable spaces, extending this construction to smooth manifolds or more general differentiable spaces, such as smooth manifolds or extensions thereof, is non-trivial. One problem occurs because probability theory over the category of smooth manifolds often operates outside the category itself. For instance, the space of probability measures on a smooth manifold does not naturally inherit a smooth structure, making it difficult to work within the category of smooth manifolds. This issue highlights the need for a more flexible framework that can accommodate both smoothness and measure-theoretic properties.

Furthermore, not all spaces can be described as a smooth manifold and in particular to describe infinite-dimensional spaces an extension is needed. Some extensions include Banach, Fréchet or Hilbert manifolds. These are all valid contenders for doing probability theory in a differentiable setting, however are quite complicated.

Diffeological spaces are a generalization of smooth manifolds, that can also describe a much broader class of spaces, while retaining the ability to define smooth maps. Unlike the aforementioned extensions, diffeological spaces provide a simpler approach to smoothness, that include infinite-dimensional spaces, singular, quotient and functional spaces. This makes diffeological spaces well-suited for probability theory, where one frequently has to consider spaces of measures, distributions, and other infinite-dimensional objects that carry a natural smooth structure.

The central objective of this thesis, is to adapt the Giry monad to the category of diffeological spaces, **Diff**, by constructing an analogue of the original Giry monad that operates within the smooth setting of diffeological spaces. We start by defining a diffeology on the space of probability measures, ensuring that the evaluation maps are smooth with respect to the piece-wise smooth diffeology on the unit interval. We then verify that the map sending a diffeological space to its space of probability measures is an endofunctor in the category **Diff** and define natural transformations such that they satisfy the monad axioms.

Ultimately, we identify an approach based on the functional diffeology, which supports the construction of a probability monad in a differentiable setting.

The interaction of smooth structures and probability theory is increasingly significant in computer science, especially in areas such as machine learning and deep learning. In particular, diffeological spaces may offer a more flexible framework that allow modelling spaces that appear in shape analysis and Bayesian inference on non-manifold spaces. Furthermore, the categorical approach to probability theory discussed here play a key role in the semantics of probabilistic programming languages, where one can reason about probabilistic composition, such as randomness and side effects.

1.1 Structure of thesis

In section 2 we will discuss some background of previous attempts to combine probability theory with smoothness and explain the motivation of using diffeological spaces. Section 3 provides preliminary definitions and explanation of necessary theory of category theory and diffeological spaces. In section 4 we show how a diffeological space can be constructed as a measurable space and describe how we can construct the space of probability measure on a diffeological space as a diffeological space. Section 5 contains the first attempt of trying to adapt the Giry monad on the category **Diff** and section 6 shows alternative approaches of defining the space of probability measures, in particular the space of probability measures defined using a functional diffeology. Section 7 contains some examples of probability measures on diffeological spaces. Finally, section 8 and 9 concludes the thesis and discusses further areas of research.

2 Background

We will begin by discussing some previous attempts to combine probability theory with some smooth structure. First, we will review how probability measures have been defined on manifolds, particularly in an infinite-dimensional setting. Next we will discuss a coordinate-free formulation of stochastic differential equations (SDEs) via jets, well-aligned with the idea of diffeology.

2.1 Probability theory on manifolds

Probability theory in a smooth setting often requires extending classical measure theory to spaces with some differentiable structure. One challenge is that many constructions in probability theory, such as the space of probability measures, do not naturally inherit a smooth structure. While we can define probability measures on finite-dimensional manifolds using a reference measure, this requires some extra structure.
There are several well-established approaches to define probability measures on manifolds.

Volume forms on Riemaniann manifolds

Definition 2.1 (Riemannian manifold)**.** A *Riemannian manifold* is a smooth manifold M equipped with a *Riemannian metric* g, which assigns to each point $p \in M$ an inner product $g_p : T_pM \times T_pM \to \mathbb{R}$, varying smoothly in p.

One approach of defining probability measures on oriented Riemannian manifolds is to use volume forms normalized to a probability measure [BBI01]. Let M be a manifold with Riemannian metric g. This metric induces a volume form vol_g and if M is compact and oriented, this volume form can be normalized to define a probability measure:

$$\mu_g(A) := \frac{1}{\int_M \mathrm{vol}_g(M)} \int_A \mathrm{vol}_g$$

In particular, every compact, oriented Riemannian manifold carries a canonical probability measure associated to its metric volume [FT22].

Hausdorff measure via charts

Another approach is to extend the theory of Hausdorff measure from Euclidean spaces onto Riemannian manifolds using local charts. Let M be a n-dimensional Riemannian manifold and $\mathcal{H}^n$ be the n-dimensional Hausdorff measure induced by the Riemannian metric. Let $A \subseteq \mathbb{R}^n$ and $V \subseteq M$ open and let $\phi : A \to V$ be a local chart. The pushforward of $\mathcal{H}^n$ using ϕ^{-1} is mutually absolutely continuous to the n-dimensional Lebesgue measure.[Fed14]

Haar measure on Lie groups

For Lie groups, the Haar measure provides a canonical smooth measure. For compact Lie groups, this measure can be normalized to obtain a probability measure [Chi12].

Extension to infinite dimensions

In infinite-dimensional spaces, such as Hilbert spaces, there is no canonical measure. Volume-based constructions fail, which makes defining probability measures more difficult. The paper of Bardelli and Mennucci [EG17] addresses two methods of defining probability measures on manifolds that can be extended to infinite-dimensional manifolds, particularly focusing on Stiefel manifolds modelled on separable Hilbert spaces. These manifolds are important in shape theory and computer vision.

The first approach wraps a Gaussian measure using the exponential map. This technique works well in finite-dimensional spaces, however in infinite dimensions the outcome can be highly sensitive to small changes. A small rotation may map equivalent measures (i.e. mutually absolutely continuous) to singular measures.

The second approach projects a Gaussian measure from a Hilbert space onto a submanifold. In finite dimensions, projections are well-defined almost everywhere, and it is possible to induce a measure on the submanifold. But in infinite dimensions, this projection might not exist of positive probability. However, the authors show that certain infinite-dimensional Stiefel manifold do behave well under such projections.

In infinite dimensions, standard operations such as rotation or projection can heavily change the structure of probability measures. This motivates using alternative frameworks, such as diffeological spaces, that may be better to combine smooth structure with probability theory.

2.2 Intrinsic SDEs via jets

The paper *intrinsic differential equations as jets* by Armstrong and Brigo [AB18] proposes a coordinate-free formulation of Itô stochastic differential equation (SDEs) on manifolds [Itô50] using the concept of *2-jets* of smooth functions. Jets are higher-order generalizations of tangent vectors that encode derivatives of functions at a point.

Earlier attempts to define SDEs on manifolds relied heavily on local coordinate charts and Itô's original calculus, which complicates transformation rules under smooth maps. This has led to attempts to develop coordinate-free approaches of defining SDEs on manifolds, using Stratonovich integrals, which allow SDEs to be represented as flows of vector fields [Elw82]. Further advancements introduced constructions such as second-order tangent vectors, Schwartz morphisms and Itô bundles, which aim to capture second-order effects in a coordinate-free manner [EM89]. However, these approaches are quite complicated.

This has led Armstrong and Brigo to define SDEs intrinsically via 2-jets of curves. By interpreting a SDE as a smooth assignment of a 2-jet at each point on a manifold, it is possible to encode drift (first-order) and diffusion (second-order) data intrinsically, without reference to coordinates.

Definition 2.2 (2-jet)**.** Let M be a manifold and $\gamma_x : \mathbb{R} \to M$ a smooth curve, such that $\gamma_x(0) = x$. The *2-jet* $j_2(\gamma_x)$ at $x \in M$ is the equivalence class of all smooth curves $\tilde{\gamma}_x$ that agree with γ_x up to second order derivative. That is,

$$j_2(\gamma_x) = \{\tilde{\gamma}_x | \tilde{\gamma}_x = x, \tilde{\gamma}_x'(0) = \gamma_x'(0), \tilde{\gamma}_x''(0) = \gamma_x''(0)\}.$$

This equivalence class captures the local behaviour of a stochastic process at $x \in M$, encoding the drift and diffusion of a curve at x.

Definition 2.3 (Intrinsic SDE)**.** An *intrinsic SDE* on a manifold M is defined as a smooth assignment of 2-jets $j_2(\gamma_x)$ at each $x \in M$, written as

$$X_t \rightharpoonup j_2(\gamma_{X_t})(dW_t).$$

Where W_t is a Brownian motion and $j_2(\gamma_{X_t})$ specifies the local behaviour of X_t at time t via the assigned jet.

This definition allows Armstrong and Brigo to define Itô's lemma using the 2-jet to make it coordinate-free.

Theorem 2.4 (Itô's lemma via jets [AB18, Thm 2.3])**.** *If $X_t \rightharpoondown j_2(\gamma_{X_t})(dW_t)$ is a jet-based SDE and $f : M \to \mathbb{R}$ is a smooth function, then the process $f(X_t)$ satisfies the SDE*

$$df(X_t) = df(\gamma'_{X_t}(0))dt + \frac{1}{2}d^2 f(\gamma''_{X_t}(0))dt.$$

The authors introduce a numerical scheme for simulating jet-based SDEs.

Definition 2.5. Let $\mathcal{T}^N = \{0, \delta t, 2\delta t, \ldots, N\delta T = T\}$ be a set of discrete time points for some fixed time T. The 2-jet scheme is defined by

$$X_{t+\epsilon} = \gamma_{X_t}\left(\frac{\epsilon}{\delta t}(W_{t+\delta t} - W_t)\right),$$

with $t \in \mathcal{T}^{N-1}$, $\epsilon \in [0, \delta t]$, $X_0 = x_0$ and γ_{X_t} is the curve associated with the jet at X_t.

Under the assumptions that γ_{X_t} is sufficiently regular, i.e. the first and second order derivatives are Lipschitz and the third order derivative is uniformly bounded. the scheme approximates the jet-based SDE and converges to the classical Itô solution. For γ_x to be sufficiently regular, the first and second order derivatives must be Lipschitz and third derivatives must be uniformly bounded.

Theorem 2.6 (Convergence of 2-jet scheme [AB18, Thm 2.4])**.** *Let X_t be the solution of a classical Itô SDE. The limit of the 2-jet scheme defined in definition 2.5 converges to X_t in mean square, as $\delta t \to 0$.*

This theorem is one of the main results of the paper, which shows that an Itô SDE can be represented by a 2-jet driven by Brownian motion.

Connection to diffeology

In the approach of Armstrong and Brigo, an SDE on a manifold is defined by a smooth assignment of 2-jets at each point. These 2-jets are used to define and approximate stochastic flows in a coordinate-free manner, without the use of charts. This coordinate-free approach aligns well with the idea of diffeological spaces, which are not defined using coordinate charts, but rather with plots. This approach could be used to define SDEs on diffeological spaces.

3 Preliminaries

In this section we will introduce the necessary background on diffeology and category theory, where we discuss one of the main examples of a probability monad, the Giry monad.

3.1 Category theory

Category theory offers an abstract framework for structuring and transferring ideas across various branches of mathematics. Instead of focussing on internal details of mathematical structures (objects), it concentrates on relationships (morphisms) between them. Morphisms can be composed, and this composition is associative and respects identity. In the setting of probability theory, it allows us to compose conditional probabilities in a structure-preserving manner.
As objects in a category are considered atomic, we aim to find universal constructions

by their properties, i.e., how objects relate to all other objects in a category.
A direct application of category theory is functional programming, where types corresponds to objects and functions to morphisms [BW95].

To formulate probability theory over diffeological spaces, we will employ category theory to define a probability monad on the category of diffeological spaces, **Diff**. In this section, we will introduce the basic theory of categories and the notion of monads on a category.

Definition 3.1 (Category)**.** A *category* consists of a collection of objects $X, Y, Z, \dots$ and morphisms between these object $f, g, h, \dots$ such that:

- Every object has an identity morphism $\mathrm{id}_X : X \to X$.
- For any pair of morphisms $f : X \to Y$ and $g : Y \to Z$ there exists a composite morphism $g \circ f : X \to Z$.

These satisfy two axioms:

- *Identity.* For all $f : X \to Y$, we have $\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X$.
- *Associativity.* For any three compositional morphisms $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$, we have $(h \circ g) \circ f = h \circ (g \circ f)$ and denote it simply as $h \circ g \circ f : X \to W$.

Definition 3.2 (Cartesian closed category)**.** A category $\mathcal{C}$ is *Cartesian closed* if:

- For every $X, Y \in \mathcal{C}$, the product $X \times Y$ exists in $\mathcal{C}$.
- For every pair $X, Y \in \mathcal{C}$, there exists an *exponential object* Y^X in $\mathcal{C}$, which can be thought of as the space of maps from X to Y.

3.1.1 Monads

The construction of a monad relies on fundamental categorical concepts such as functors and natural transformations. A functor is a morphism between categories, preserving structure between objects and morphisms. A natural transformation, on the other hand, provides a way to relate two functors.

Monads formalize how we can chain operations that we cannot directly compose. The key idea behind a monad is to describe some extra structure and provide a way to compose such structured values. For example, in functional programming a monad can model computations with side effects, such as randomness.

Definition 3.3 (Functor, [Rie14, Def. 1.3.1])**.** A *functor* F between two categories $\mathcal{C}$ and $\mathcal{D}$ is a map such that

- There exists an object $Fc \in \mathcal{D}$ for each object $c \in \mathcal{C}$.
- For each morphism $f : c \to c' \in \mathcal{C}$ there exists a morphism $Ff : Fc \to Fc' \in \mathcal{D}$, such that the domain is equal to F applied to the domain of f and the codomain of Ff is equal to the codomain of F applied to the codomain of f.

The functor must follow the two functoriality axioms

1. It respects identity morphisms: For each object $c \in \mathcal{C}$, $F(\mathrm{id}_c) = \mathrm{id}_{F(c)}$
2. it respects composition: For any compositional pair f, g in $\mathcal{C}$, $Fg \circ Ff = F(g \circ f)$.

Definition 3.4. An *endofunctor* is a functor whose domain is equal to its codomain.

Definition 3.5 (Natural transformation, [Rie14, Def. 1.4.1])**.** Let $\mathcal{C}, \mathcal{C}'$ be two categories and $F, G : \mathcal{C} \to \mathcal{C}'$ two functors. A *natural transformation*, $\alpha : F \Rightarrow G$, consists of for each object $x \in \mathcal{C}$ a morphism $\alpha_x : Fx \to Gx$ in $\mathcal{C}'$ such that for any morphism $f : x \to y$ in $\mathcal{C}$ the following diagram commutes in $\mathcal{C}'$:

$$\begin{array}{ccc} Fx & \xrightarrow{Ff} & Fy \\ {\scriptstyle \alpha_x}\downarrow & & \downarrow{\scriptstyle \alpha_y} \\ Gx & \xrightarrow[Gf]{} & Gy \end{array}$$

Notation 3.6. Natural transformations are a way to relate two functors on a category, instead of the objects themselves. To distinguish their role from morphisms on objects, we use the notation $\alpha : F \Rightarrow G$.

Definition 3.7 (Monad, [Rie14, Def. 5.1.1])**.** A *monad* on a category $\mathcal{C}$ is a triple (T, η, μ) where:

1. T is and endofunctor $T : \mathcal{C} \to \mathcal{C}$,
2. η, the unit, a natural transformation $\eta : \mathrm{id}_{\mathcal{C}} \Rightarrow T$,
3. μ, the multiplication, a natural transformation $\mu : T^2 \Rightarrow T$,

such that the following diagrams commute:

$$\begin{array}{ccc} T^3 & \overset{T\mu}{\Longrightarrow} & T^2 \\ {\scriptstyle \mu T}\Downarrow & & \Downarrow{\scriptstyle \mu} \\ T^2 & \underset{\mu}{\Longrightarrow} & T \end{array} \qquad\qquad \begin{array}{ccccc} T & \overset{\eta T}{\Longrightarrow} & T^2 & \overset{T\eta}{\Longleftarrow} & T \\ & {\scriptstyle 1_T}\searrow & \Downarrow{\scriptstyle \mu} & \swarrow{\scriptstyle 1_T} & \\ & & T & & \end{array}$$

The presence of this extra structure on morphisms complicates composition. The Kleisli category associated to a monad handles this composition.

Definition 3.8 (Kleisli category associated to a monad)**.** Let $(\mathcal{C}, T, \eta, \mu)$ be a monad on a category $\mathcal{C}$, where $T : \mathcal{C} \to \mathcal{C}$ is a functor, $\eta : \mathrm{id}_{\mathcal{C}} \Rightarrow T$ is the unit, and $\mu : T^2 \Rightarrow T$ is the multiplication natural transformation.

The *Kleisli category* of the monad T, denoted by $\mathcal{C}_T$, is the category defined as follows:

- The objects of $\mathcal{C}_T$ are the same as those of $\mathcal{C}$.
- A morphism $f : X \to Y$ in $\mathcal{C}_T$ is a morphism $f : X \to TY$ in $\mathcal{C}$.
- For morphisms $f : X \to TY$ and $g : Y \to TZ$ in $\mathcal{C}$, the composition $g \circ_T f : X \to TZ$ in $\mathcal{C}_T$ is given by
$$g \circ_T f := \mu_Z \circ Tg \circ f, \quad X \xrightarrow{f} TY \xrightarrow{Tg} TTZ \xrightarrow{\mu_Z} TZ.$$
- The identity morphism $\mathrm{id}_X : X \to TX$ in $\mathcal{C}_T$ is given by the unit $\eta_X : X \to TX$ of the monad.

3.2 Categorical probability theory

A classical technique of applying category theory to probability theory is to define a probability monad on an already existing category, by treating probability distributions, random processes and conditional probabilities as morphisms and functors.

3.2.1 The Giry monad

The Giry monad, introduced by Michele Giry in 1982 [Gir82], is one of the main examples of a probability monad, defined over the category of measurable spaces, **Meas**, and thus one of the main structures used in categorical probability theory. The monad is implicitly used when working with e.g. Markov kernels or probability distributions over probability distributions, used for example in de Finetti's theorem. The Kleisli category associated to the Giry monad gives rise to the category **Stoch** of Markov kernels [Gir82].

Definition and structure

The Giry monad is situated in the category **Meas**, whose objects are measurable space (X, Σ_X) and morphisms are measurable functions.

For a measurable space (X, Σ_X), its functor P assigns the space of probability measures, PX, on X, equipped with the evaluation σ-algebra, the coarsest σ-algebra such that the set of all the evaluation maps

$$\mathrm{ev}_B : PX \to [0,1], \quad \mathrm{ev}_B(\mu) = \mu(B),$$

are measurable for all measurable set $B \in \Sigma_X$.
For measurable $f : X \to Y$, where X and Y are measurable spaces, $Pf : PX \to PY$ is given by the pushforward of measures

$$Pf(\mu) = f_*\mu, \quad f_*\mu(B) = \mu(f^{-1}(B)).$$

The unit $\eta_X : X \to PX$ of the monad maps a point $x \in X$ to its Dirac measure concentrated at x, δ_x. Its multiplication $m_X : PPX \to PX$ flattens a measure of measures by integration:

$$m_X(\nu)(B) = \int_{PX} \mu(B) d\nu(\mu).$$

The Kleisli category

A morphism in the Kleisli category is a measurable map $f : X \to PY$.

The Kleisli category $\mathbf{Meas}_P$ associated with the Giry monad can be seen as the category of Markov kernels. Its objects are again measurable spaces and its morphisms are Markov kernels, measurable maps $f : X \to PY$ representing conditional probability distributions. The composition of two morphisms $f : X \to PY$ and $g : Y \to PZ$ is given by

$$X \xrightarrow{f} PY \xrightarrow{Pg} PPZ \xrightarrow{m_Z} PZ$$

$$(g \circ f)(x)(B) = \int_Y g(y)(B) df(x)(y)$$

3.2.2 The category of Markov Kernels

The category of Markov kernels is often denoted as **Stoch**. The objects of this category are measurable spaces and its morphisms are stochastic or Markov kernels.

Definition 3.9 (Markov kernel)**.** A Markov kernel between two measurable spaces (X, Σ_X) and (Y, Σ_Y) is a map $\kappa : \Sigma_Y \times X \to [0,1]$ such that

1. For all $B \in \Sigma_Y$, the map $x \mapsto \kappa(B|x)$ is Σ_X-measurable
2. For all $x \in X$, the map $B \mapsto \kappa(B|x)$ is a probability measure on Y.

For two morphisms $\kappa : X \to Y$ and $\lambda : Y \to Z$ in **Stoch**, the composition $\lambda \circ \kappa : X \to Z$ is given by the Chapman-Kolmogorov equation:

$$\lambda \circ \kappa(dz|x) := \int_Y \lambda(dz|y)\kappa(dy|x).$$

For any measurable space X, the identity morphism is the Dirac kernel:

$$\mathrm{id}_X(B|x) = \delta(B) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases}$$

3.3 Diffeology

Diffeology is a branch of differential geometry that generalizes the concept of smooth manifolds by using plots to define smooth structures on a set, unlike traditional manifolds, which rely on atlases and local charts. This approach was first introduced by Jean-Marie Souriau in his 1980 paper *Groupes Différentieles* [Sou80] to address limitations in classical differential geometry, in particular when dealing with spaces that are not locally Euclidean or have singularities.

We define a *diffeological space* as a set X equipped with a *diffeology*, a collection of maps, called plots, that are used to characterize smoothness of the space.

Definition 3.10 (Diffeology)**.** Let X be a set. A *diffeological space* is a pair $(X, \mathcal{D})$ with X a set and $\mathcal{D}$ a *diffeology.* A diffeology on X is any set of parametrizations on X satisfying the following three axioms:

1. *Covering.* Every constant map is a plot.
2. *Locality.* Let $p : U \to X$ be a map. If every $r \in U$ has an open neighbourhood $V \subseteq U$ of r such that the restriction $p|_V$ belongs to $\mathcal{D}$, then p itself is a plot.
3. *Smooth compatibility.* Let $p : U \to X$ in be a plot and V a real domain and $F \in C^\infty(V, U)$, then $p \circ F$ is a plot.

Formally, a diffeological space is a pair $(X, \mathcal{D})$, but we shall denote a diffeological space with only a single letter X, where $\mathcal{D}$ is its diffeology.

Definition 3.11 (Smooth map)**.** Let X and Y be two diffeological spaces. A map $f : X \to Y$ is called smooth if for every plot $p : U \to X$, $f \circ p : U \to Y$ is a plot for Y.

$$U \xrightarrow[p]{} X \xrightarrow[f]{} Y$$

Diffeological spaces do not generally come with a topology defined on the space, however every diffeological space has a topology induced by the diffeology, the D-topology, which ensures that every smooth map is also continuous.

Theorem 3.12 (D-topology, [IZ13, Art. 2.8])**.** *Let X be a diffeological space with diffeology $\mathcal{D}$, then there exists a finest topology on X, called the* D-topology*, such that all plots $p \in \mathcal{D}$ are continuous. This means that for every subset $A \subset X$ is open for the D-topology if and only if, for every plot $p \in \mathcal{D}$, $p^{-1}(A) \subset R^n$ is open in the Euclidean topology on the plots domain.*

Theorem 3.13 (Product diffeology, [IZ13, Art. 1.55])**.** *Let $(X, \mathcal{D}_X)$ and $(Y, \mathcal{D}_Y)$ be two diffeological spaces. Then there exists a coarsest diffeology, called the* product diffeology*, $\mathcal{D}_{X\times Y}$ on $X \times Y$, given by*

$$\mathcal{D}_{X\times Y} := \left\{ p = (p_X, p_Y) \,\middle|\, p_X \in \mathcal{D}_X,\, p_Y \in \mathcal{D}_Y \right\},$$

such that $p_X = \pi_X \circ p$ and $p_Y = \pi_Y \circ p$ are the smooth projections to X and Y.

We can define a diffeology on the set of all smooth functions between two diffeological spaces, called the functional diffeology.

Theorem 3.14 (Functional diffeology, [Igl25, Thm. 5.5])**.** *Let X and Y be two diffeological spaces. The set of all smooth maps $C^\infty(X,Y)$ carries a functional diffeology, where a parametrization $p : U \to C^\infty(X,Y)$, $u \mapsto f_u$, is a plot if and only if the associated evaluation map*

$$ev : U \times X \to Y, \quad (u,x) \mapsto p(u)(x)$$

is smooth.

3.3.1 The category of diffeological spaces

The category **Diff** consists of diffeological spaces as object and (diffeologically) smooth maps as morphisms. **Diff** is a Cartesian closed category, making it an ideal contender for extending probability theory to generally smooth spaces.

Proposition 3.15. *The category* **Diff** *is Cartesian closed:*

- *Given $(X,\mathcal{D}_X),(Y,\mathcal{D}_Y)$, their product is $X \times Y$ with the product diffeology.*
- *The exponential object $Y^X = C^\infty(X,Y)$ carries the functional diffeology.*

This property is crucial for applying probability theory in a smooth setting, because many constructions involve function spaces, such as the space of random variables (measurable maps) or transition kernels. In **Diff**, we can treat such function spaces as objects while remaining inside the category, unlike in the category of smooth manifolds, which is not Cartesian closed.

3.3.2 Limitations of manifolds

Classical manifolds require a finite-dimensional, locally Euclidean structure, which excludes many spaces such as infinite-dimensional spaces (e.g., most function spaces) or spaces with singularities (e.g., quotients of manifolds, orbifolds).
A diffeological space has no such restrictions. It is possible to model even infinite-dimensional spaces and spaces with singularities as diffeological spaces.

An example of a highly pathological space that cannot be described by a traditional manifold is the irrational torus, as this is a space with singularities. We can however define a diffeology on the irrational torus.

Example 3.16. Let T_α be the irrational torus given by the quotient $\mathbb{R}/(\mathbb{Z}+\alpha\mathbb{Z})$ for an $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Let $\pi_\alpha : \mathbb{R} \to T_\alpha$ be the canonical projection that maps $x \in \mathbb{R}$ to its equivalence class $[x] \in T_\alpha$.
The set of all parametrizations $p : U \to T_\alpha$ such that for all $u_0 \in U$ there exists an open neighbourhood V of u_0 and a smooth parametrization $q : V \to \mathbb{R}$ such that $\pi_\alpha \circ Q = P|_V$ form a diffeology on T_α.

1. *Covering.* Let $p : U \to T_\alpha$ a parametrization such that for all $u_0 \in U$, $p(u_0) = [x_0]$. Let $q : U \to \mathbb{R}$ a parametrization such that for all $u_0 \in U$, $q(u_0) = x_0$. Then $\pi_\alpha \circ q(u_0) = [x_0] = p(v_0)$ for all $u \in U$.
2. *Locality.* Let $p : U \to T_\alpha$ be a parametrization and $\{U_i\}_{i\in I}$ an open cover of U such that $P|_{U_i}$ is a plot for every $i \in I$. Then for every $u_i \in U_i$, there exists an open neighbourhood $V_i \subseteq U_i$ of u_i and a smooth parametrization $q_i : V_i \to \mathbb{R}$ such that $\pi_\alpha \circ q_i = (p|_{U_i})|_{V_i} = p|_{V_i}$.
3. *Smooth compatibility.* Let $p : U \to T_\alpha$ a plot and $F \in C^\infty(V,U)$. Let $v_0 \in V$, then there exists a neighbourhood $W \subseteq U$ of $f(v_0)$ and smooth $q : W \to \mathbb{R}$ such that $\pi_\alpha \circ q = p|_W$. Let $V' = F^{-1}(W)$ (open as F is continuous), then

$$\pi_\alpha \circ (q \circ F|_{V'}) = (\pi_\alpha \circ q) \circ F|_{V'} = p \circ F|_{V'}.$$

3.3.3 Manifolds as diffeological space

We can describe any traditional manifold as a diffeological space using plots, without mentioning topology or local coordinate charts.

Definition 3.17 (Local smooth map)**.** Let X, Y be diffeological spaces, $A \hookrightarrow X$ and $f : A \to Y$ a map. Then f is called a local smooth map if, for every plot $p : U \to X$, the parametrization $f \circ p : U \to Y$ defined on $p^{-1}(A)$ forms a plot for Y.

Definition 3.18 (Local diffeomorphism)**.** A local diffeomorphism between two diffeological spaces X and Y is an injective local smooth map $f : A \to Y$ with local smooth inverse $f^{-1}(A) : f(A) \to X$.

A manifold is a diffeological space that is locally diffeomorphic to Euclidean space. Diffeological spaces form a category **Diff**, where the objects are diffeological spaces and the morphisms are diffeologically smooth maps, in which the category of manifolds is naturally a subcategory.

4 Probability measures on diffeological spaces

In order to define a probability monad over the category of diffeological spaces, we must first make precise how to view such spaces as measurable and how to equip spaces of probability measures with a suitable diffeology.

We begin by defining a measurable structure on a diffeological space via the D-topology and the associated Borel σ-algebra. Using this measurable structure, we then define a diffeology on the space of probability measures on a diffeological space X, denoted PX. This diffeology ensures that evaluation maps are smooth with respect to a reference diffeology on $[0, 1]$.

Finally, we introduce a subspace of absolutely continuous probability measures $P'X$ with smooth densities relative to a reference measure, when such a measure exists.

Definition 4.1 (Measurable space associated to a diffeological space)**.** Let $(X, \mathcal{D})$ be a diffeological space. The *measurable space associated to* X is the pair $(X, \mathcal{B}(X))$, where $\mathcal{B}(X)$ is the Borel σ-algebra generated by the D-topology on X.

Explicitly, $\mathcal{B}(X)$ is the smallest σ-algebra containing all subsets of X that are open in the D-topology.

4.1 The space of probability measures

We will first consider a useful diffeology using piece-wise smooth maps. The piece-wise smooth diffeology consists of all piece-wise smooth plots.

Definition 4.2 (Piece-wise smooth)**.** Let $U \subseteq \mathbb{R}^n$ and $f : U \to X$ a continuous map. We call f *piece-wise smooth* if there exists finitely many subsets $U_i \subseteq U$ such that $U = \bigcup_{i \in I} U_i$ and $f|_{U_i}$ is smooth for every $i \in I$.

By defining the piece-wise smooth diffeology, as we will obtain more smooth maps, which will be useful when defining the monad unit and multiplication in definition 3.7.

Definition 4.3 (Piece-wise smooth diffeology, [AM23, Def. 7.14])**.** The *piece-wise smooth diffeology* on $[0, 1]$ is defined by the set of all piece-wise smooth parametrizations of $[0, 1]$. We will use the notation $\mathcal{D}_p$ for the piece-wise smooth diffeology on $[0, 1]$.

We use the piece-wise smooth diffeology on $[0,1]$ to include distributions with non-smooth densities or discontinuities, while preserving smoothness.
Let X be a diffeological space. We construct a diffeology on the space of probability measures PX on X as follows:

Definition 4.4 (Evaluation-based diffeology)**.** Let X be a diffeological space and PX be the set of all probability measures on X with respect to the Borel σ-algebra $\mathcal{B}(X)$, generated by the D-topology. For any open subset $U \subseteq \mathbb{R}^n$, a parametrization $p : U \to PX$ is a plot if and only if for every set $B \in \mathcal{B}(X)$ the composition

$$\mathrm{ev}_B \circ p : U \to [0,1], \quad u \mapsto p(u)(B)$$

is a plot for $[0,1]$, where $\mathrm{ev}_B : PX \to [0,1]$ is the evaluation map at B and $[0,1]$ is equipped with the piece-wise smooth diffeology.

Theorem 4.5. *The collection of plots defined above form a diffeology on PX.*

Proof. We will verify the three axioms of a diffeology.

1. *Covering.* Let $p : U \to PX$ be a constant map, i.e. $p(u) = \mu$ for some $\mu \in PX$ and every $u \in U$. Then for any $B \in \mathcal{B}(X)$ the composition with the evaluation map $ev_B \circ p(u) = \mu(B)$ is a constant function, which is piece-wise smooth, thus p is a plot.

2. *Locality.* Let $p : U \to PX$ be a map such that for every $u \in U$, there exists an open neighbourhood $V \subseteq U$ of u, for which $p|_V$ is a plot. Then for any set $B \in \mathcal{B}(X)$ the map $\mathrm{ev}_B \circ p|_V : V \to [0,1]$ is locally piece-wise smooth everywhere on $[0,1]$ by the locality axiom of the diffeology on $[0,1]$, $\mathrm{ev}_B \circ p$ is a plot for $[0,1]$, thus p is a plot for PX.

3. *Smooth compatibility.* Let $p : U \to PX$ a plot and $F \in C^\infty(V,U)$. For every B we have the equality

$$(\mathrm{ev}_B \circ p) \circ F = \mathrm{ev}_B \circ (p \circ F).$$

   Since p is a plot $\mathrm{ev}_B \circ p$ is piece-wise smooth and F is smooth and therefore $(\mathrm{ev}_B \circ p) \circ F$ is piece-wise smooth. Then $\mathrm{ev}_B \circ (p \circ F)$ is piece-wise smooth and thus $p \circ F$ is a plot for PX.

□

5 Adapting the Giry monad

In this section, we will construct a probability monad over the category of diffeological spaces, that is analogous to the Giry monad over the category of measurable spaces [Gir82]. We will first construct a functor P that assigns to each diffeological space X the space of all probability measures, PX. We will define Pf by the pushforward.

Definition 5.1 (Pushforward of a measure)**.** Let (X, Σ_X) and (Y, Σ_Y) be two measurable spaces, $f : X \to Y$ a measurable function and $\mu : \Sigma_X \to \mathbb{R}_+ \cup \{\infty\}$ a measure. The *pushforward* of μ by f is the measure $f_*(\mu) : \Sigma_Y \to \mathbb{R}_+ \cup \{\infty\}$ given by

$$f_*(\mu)(B) = \mu(f^{-1}(B)) \quad \text{for} \quad B \in \Sigma_Y.$$

Lemma 5.2. *Let $f : X \to Y$ be a smooth map between diffeological spaces. If X and Y are equipped with the Borel σ-algebras $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ generated by their respective D-topologies, then f is measurable.*

Proof. Let X and Y be two diffeological spaces equipped with their respective D-topologies and $f : X \to Y$ a smooth map. By Art. 2.9 [IZ13], every smooth map is continuous for the D-topology and thus is f continuous. As we equip X and Y with their respective Borel σ-algebra generated by the D-topology, any continuous map is also measurable, and therefore every smooth f is also measurable. $\square$

Theorem 5.3. *Let $f : X \to Y$ be a smooth map, i.e. a morphism in **Diff**. Then f is also measurable. Define the pushforward as*

$$Pf : PX \to PY, \quad Pf(\mu)(B) = \mu(f^{-1}(B))$$

for $\mu \in PX$ and $B \in \mathcal{B}(Y)$. Then P is a functor in **Diff***.*

Proof. We can construct PX as a diffeological space for any diffeological space X, as seen in the proof of theorem 4.5, meaning for any object X in **Diff**, PX is also in **Diff**. Now, we need to check whether Pf is a morphism in **Diff**. Let $p : U \to PX$ be a plot and $B \in \mathcal{B}(Y)$, then $\text{ev}_B \circ (Pf \circ p) : U \to [0, 1]$ is given by

$$u \mapsto Pf(p(u))(B) = p(u)(f^{-1}(B))$$

and thus $\text{ev}_B \circ (Pf \circ p) = \text{ev}_{f^{-1}(B)} \circ p$. Since p is a plot and $f^{-1}(B) \in \mathcal{B}(X)$, $\text{ev}_{f^{-1}(B)} \circ p$ is a plot for $[0, 1]$, meaning $Pf \circ p$ is a plot for PY and thus Pf is a morphism in **Diff**.
Let X be an object in **Diff** and $\text{id}_X : X \to X$ and $\text{id}_{PX} : PX \to PX$ be identity functions, then $P(\text{id}_X)$ is given by $P(\text{id}_X)(\mu)(B) = \mu(\text{id}_X^{-1}(B)) = \mu(B)$ and id_{PX} is given by $\text{id}_{PX}(\mu)(B) = \mu(B)$, so $P(\text{id}_X) = \text{id}_{PX}$ for any object X.
Let $f : X \to Y$ and $g : Y \to Z$ be two compositional morphisms. Then for any B measurable in X, $P(g \circ f)$ is given by:

$$P(g \circ f)(\mu)(B) = \mu((g \circ f)^{-1}(B)) = \mu(f^{-1}(g^{-1}(B)))$$

and $Pg \circ Pf$ is given by:

$$Pg(Pf(\mu))(B) = Pf(\mu)(g^{-1}(B)) = \mu(f^{-1}(g^{-1}(B))).$$

So $P(g \circ f) = Pg \circ Pf$ and thus P is a functor. $\square$

Now we have proven that P is an endofunctor on the category **Diff**, we will define a unit and multiplication natural transform, that satisfy the monad axioms.

Definition 5.4 (Unit and multiplication)**.** Let X be a diffeological space, We define the unit and multiplication as follows:

1. Unit: The unit $\eta_X : X \to PX$ assigns to each point $x \in X$ the Dirac measure δ_x, where for any measurable $B \subseteq X$:

$$\delta_x(B) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases}$$

2. Multiplication: The multiplication $m_X : PPX \to PX$ is defined by integrating measures over a measure. Let $\nu \in PPX$ and $B \subseteq X$ measurable:

$$m_X(\nu)(B) = \int_{PX} \text{ev}_B d\nu = \int_{PX} \mu(B) d\nu(\mu).$$

The unit and multiplication are defined analogously to the unit and multiplication of the Giry monad. One important difference is that they must be morphisms in the category **Diff**.

5.1 Smoothness of unit and multiplication

When we define the space of probability measures as in definition 4.4, the monad unit and multiplication are not smooth, and thus not morphisms in **Diff**.

Remark 5.5. With P the functor of the monad as defined in definition 4.4 and theorem 5.3, the unit map η_X of the monad is not a morphism in the category **Diff**.

Proof. Consider $\mathbb{R}$ equipped with the standard diffeology. The D-topology coincides with the Euclidean topology. Let $p : \mathbb{R} \to \mathbb{R}$ a plot given by $p(u) = u$. Then $\text{ev}_B \circ \eta_{\mathbb{R}} \circ p(u) = \chi_B(p(u)) = \chi_B(u)$.
Let $B = \mathbb{R} \backslash \mathbb{Q}$, the set of irrationals, which is measurable. Then

$$\text{ev}_B \circ \eta_{\mathbb{R}} \circ p(u) = \begin{cases} 1 & \text{if } u \in B \\ 0 & \text{otherwise} \end{cases}$$

has a discontinuity at every point $u \in \mathbb{R}$, meaning it is not piece-wise smooth and $\eta_{\mathbb{R}}$ is not a morphism. In general, η_X is not a morphism in **Diff**. □

6 Alternative approaches

As in our previous approach the unit and multiplication of the monad are not morphisms, we will discuss some alternative approaches in this section.

6.1 Absolutely continuous measures

One possible approach is to restrict ourselves to absolutely continuous measures on a diffeological space X, with respect to some reference measure on X.

Definition 6.1 (Diffeological space of absolutely continuous measures)**.** Let X be a diffeological space such that there exists a standard measure λ_X on X w.r.t. the Borel σ-algebra generated by the D-topology and PX be the space of all probability measures over X. We define $P'X \subseteq PX$ as the set of all absolutely continuous measures on X w.r.t. a reference measure λ_X, given by

$$P'X := \{\mu \in PX | \exists f : X \to R_{\geqslant 0} \text{ such that } \mu = f\lambda_X \text{ and } f \text{ piece-wise smooth}\}$$

We equip $P'X$ with the subset diffeology from PX, making $P'X$ a diffeological space.

Remark 6.2. The definition of $P'X$ as measures with smooth densities depends on the existence of a reference measure. For general diffeological spaces X, such a reference measure may not exist. Thus, the construction of $P'X$ is not generally available.

With this approach, we encounter the same problem as when attempting to define probability measures on some extension of manifolds to infinite dimensions. There is no canonical infinite-dimensional reference measure, and therefore we exclude many spaces on which we can define the space of absolutely continuous measures $P'X$. As $P'X$ itself is infinite-dimensional, we cannot chain the functor and $P'P'X$ is not defined for any diffeological space X.
In particular, we can only define P' as a functor on a subcategory $\textbf{Diff}_\lambda$ of **Diff**, where the objects are diffeological spaces where a canonical reference measure exists. This would not be an endofunctor, as P' maps objects in $\textbf{Diff}_\lambda$ to objects in **Diff**, but outside $\textbf{Diff}_\lambda$, and we cannot use P' to define a monad.

6.2 The functional diffeology on the space of probability measures

Another strategy of equipping the space of probability measures with a diffeology would be using the functional diffeology on PX. We again interpret the space of probability measures as a diffeological space, where we view $PX \subset C^\infty(X, \mathbb{R})^*$ as a subset of the linear functionals $\phi : C^\infty(X, \mathbb{R}) \to \mathbb{R}$.

We will begin this strategy with the following assumption, which may not hold for arbitrary diffeological spaces.

Conjecture 6.3. *For a diffeological space X, the space of probability measures PX can be identified as a subset of the dual space $C^\infty(X, \mathbb{R})^*$ via,*

$$\mu \mapsto \left(f \mapsto \int_X f d\mu \right)$$

where $f \in C^\infty(X, \mathbb{R})$ and bounded.

Remark 6.4. It is possible that this map is not injective and maps different measures to equivalent linear functionals, making them indistinguishable.

For arbitrary diffeological spaces PX might not be a subset of the linear functionals, so for this section we will assume $PX \subset C^\infty(X, \mathbb{R})^*$.

Definition 6.5 (Functional diffeology on space of probability measures)**.** Let X be a diffeological space and $PX \subset C^\infty(X, \mathbb{R})^*$ the space of all probability measures on X with respect to the Borel σ-algebra, $\mathcal{B}(X)$, generated by the D-topology. For any open subset $U \in \mathbb{R}^n$, a parametrization $p : U \to PX$ is a plot if and only if

$$u \mapsto p(u)(f) := \int_X f dp(u)$$

is smooth for any bounded smooth map $f : X \to \mathbb{R}$.

By the functional diffeology the collection of plots defined above form a diffeology on PX. The condition that f is bounded ensures that all integrals involved are finite and we do not define the trivial diffeology.

Functoriality

We define the effect P has on morphisms in the same way as in definition 5.3, by pushforward of a measure.

Lemma 6.6. *The map Pf, defined by pushforward of a measure μ on a morphism $f : X \to Y$, is smooth.*

Proof. Let $f : X \to Y$ be a morphism in **Diff** and $Pf : PX \to PY$. Let $p : U \to PX$ a plot and $g \in C^\infty(Y, \mathbb{R})$ bounded, then

$$Pf \circ p(u)(g) = f_* p(u)(g) = \int_Y g df_* p(u) = \int_X g \circ f dp(u)$$

As f and g are smooth, $g \circ f \in C^\infty(X, \mathbb{R})$ and p is a plot, then by definition 6.5 $Pf \circ p(u)(g)$ is smooth and thus Pf is smooth. □

As we define Pf as in 5.3, we can prove P is a functor analogously to the proof of 5.3.

Corollary 6.7. *P, as defined in 6.5 forms a functor on the category* **Diff***.*

Unit

We define the monad unit in the same way again. $\eta_X : X \to PX$ maps a point $x \in X$ to its corresponding Dirac measure, δ_x.

Lemma 6.8. *The map $\eta_X : X \to PX$, $x \mapsto \delta_x$ is smooth.*

Proof. Let $p : U \to X$ a plot. Then for any (bounded) $f \in C^\infty(X, \mathbb{R})$, the composition $\eta_X \circ p : U \to PX$ is given by:

$$\eta_X \circ p(u)(f) = \int_X f d\delta_{p(u)} = f(p(u)) = f \circ p(u),$$

which is smooth as composition of smooth maps and thus the monad unit is smooth. □

As $\eta_X : X \to PX$ is a morphism in **Diff**, we can prove $\eta : \mathrm{id}_{\mathbf{Diff}} \Rightarrow P$ is a natural transformation.

Lemma 6.9. *$\eta : id_{\mathbf{Diff}} \Rightarrow P$ is a natural transformation.*

Proof. Let $f : X \to Y$ a morphism. Then for any $x \in X$ and $B \subseteq Y$ measurable,

$$\eta_Y \circ f(x)(B) = \delta_{f(x)}(B)$$
$$Pf \circ \eta_X(x)(B) = Pf(\delta_x)(B) = \delta_x(f^{-1}(B)) = \delta_{f(x)}(B)$$

meaning the following diagram commutes

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \eta_X}\downarrow & & \downarrow{\scriptstyle \eta_Y} \\ PX & \xrightarrow[Pf]{} & PY \end{array}$$

and thus $\eta : \mathrm{id}_{\mathbf{Diff}} \Rightarrow PX$ is a natural transformation. □

Multiplication

Also, the monad multiplication is defined in the same way, by integrating measures over a measure:

$$m_X : PPX \to PX, \quad m_X(\nu)(B) = \int_{PX} \mu(B) d\nu(\mu)$$

Lemma 6.10. *The map $m_X : PPX \to PX$ is smooth.*

Proof. Let $p : U \to PPX$ a plot, then for any bounded $f \in C^\infty(X, \mathbb{R})$, the composition $m_X \circ p(u)(f)$ is given by:

$$m_X(p(u))(f) = \int_{PX} \left(\int_X f d\mu \right) dp(u)(\mu).$$

The inner map $\mu \mapsto \int_X f d\mu$ is smooth by the functional diffeology on PX, and p is a plot for PPX. Then $m_X \circ p(u)(f)$ is a plot for PPX and thus the map m_X is smooth. □

As for the unit we have obtained the multiplication is smooth, and therefore can proceed to prove that $m : PP \Rightarrow P$ forms a natural transformation.

Lemma 6.11. *$m : PP \Rightarrow P$ is a natural transformation.*

Proof. Let $\nu \in PPX$, $\mu \in PX$, $\xi \in PY$ and $B \subseteq X$ measurable. Then

$$\begin{aligned} Pf \circ m_X(\nu)(B) &= m_X(\nu)(f^{-1}(B)) \int_{PX} \mu(f^{-1}(B)) d\nu(\mu) \\ m_Y \circ PPf(\nu)(B) &= \int_{PY} \xi(B) d(PPf(\nu))(\xi) \\ &= \int_{PX} Pf(\mu)(B) d\nu(\mu) = \int_{PX} \mu(f^{-1}(B)) d\nu(\mu) \end{aligned}$$

And the following diagram commutes.

$$\begin{array}{ccc} PPX & \xrightarrow{PPf} & PPY \\ {\scriptstyle m_X}\big\downarrow & & \big\downarrow{\scriptstyle m_Y} \\ PX & \xrightarrow[Pf]{} & PY \end{array}$$

This makes $m : PP \Rightarrow P$ a natural transformation. ☐

Monad structure

With a functor, a unit and a multiplication natural transformation, we can construct a monad on **Diff**. This requires checking if the diagrams in definition 3.7 commute. It should follow in the same way that as in the original paper from Giry that (P, η, m) form a monad on **Diff**, but now the unit and multiplication are smooth maps, instead of only measurable.

Conjecture 6.12. (P, η, m) *forms a monad on* **Diff**.

Proof sketch. Let $\nu \in PX$ and $f \in C^\infty(X, \mathbb{R})$. Then $m_X \circ \eta_{PX}$ is given by

$$m_X \circ \eta_{PX}(\nu) = m_X(\delta_\nu) = \int_{PX} \left(\int_X f d\mu \right) \delta_\nu = \int_X f d\nu.$$

And thus $m_X \circ \eta_{PX}(\nu) = \nu = \mathrm{id}_{PX}$ for all $\nu \in PX$.
Now consider $m_X \circ P\eta_X$. Here $P\eta_X$ is given by the pushforward, i.e. for measurable $B \subseteq X$ and a measure $\nu \in PX$

$$P\eta_X(\nu) = \eta_{X*}\nu.$$

Then $m_X \circ P\eta_X$ is given by

$$\begin{aligned} m_X(P\eta_X(\nu)) &= \int_X f d(m_X(P\eta_X(\nu))) \\ &= \int_{PX} \left(\int_X f d\mu \right) d(P\eta_X(\nu))(\mu) \\ &= \int_X \left(\int_X f d\eta_X \right) d\nu = \int_X f d\nu \end{aligned}$$

And thus $m_X \circ P\eta_X = \mathrm{id}_{PX}$ and the right-most diagram of definition 3.7 commutes.
Let $\xi \in P^3X$. Then $m_X \circ Pm_X : P^3X \to PX$ is given by

$$\begin{aligned} m_X(Pm_X(\xi)) &= \int_{PPX} \left(\int_X f d(m_X(\mu)) \right) d\xi(\mu) \\ &= \int_{PPX} \left(\int_{PX} \left(\int_X f d\mu \right) d\mu(\nu) \right) d\xi(\mu) \end{aligned}$$

And $m_X \circ m_{PX}$ is given by

$$\begin{aligned} m_X(m_{PX}(\xi)) &= \int_{PX} \left(\int_X f d\nu \right) d(m_{PX}(\xi))(\nu) \\ &= \int_{PPX} \left(\int_{PX} \left(\int_X f d\mu \right) d\mu(\nu) \right) d\xi(\mu) \end{aligned}$$

And thus the left-most diagram of definition 3.7 also commutes. □

The proof of this theorem is not exact, as it might be possible that not all integrals converge. Also, since PX might not be a true embedding in the space of linear functionals, it may map different measures to the same linear functional. Therefore, $\int_X f d\nu$ might not be equal to id_{PX}.

7 Examples

In this section, we will discuss two examples. The first example considers the unit circle, which is a space that we can also define as a manifold. For the second example, we will look at the irrational torus, a highly pathological space with singularities.

Example 7.1. Let $S^1 := \{z \in \mathbb{C} : \bar{z}z = 1\}$ be the unit circle, equipped with the *quotient diffeology.* That is, a map $p : U \to S^1$ is a plot if and only if for every $u \in U$ there exists an open neighbourhood V of u and a smooth parametrization $p' : V \to \mathbb{R}$ such that $p|_V : u \mapsto e^{2i\pi p'(u)}$.

We define a probability measure on S^1 by pushing forward a Gaussian measure on $\mathbb{R}$ via the canonical projection $q : \mathbb{R} \to S^1$.

Let $\mu \in \mathbb{R}$ and $\sigma > 0$, and let $\gamma = \mathcal{N}(\mu, \sigma^2)$ be the Gaussian measure on $\mathbb{R}$. The *wrapped normal distribution* on S^1 is defined as the pushforward measure

$$\nu := q_*\gamma,$$

where $q(\theta) = e^{2\pi i\theta}$. The density f of ν with respect to the standard Lebesgue measure on S^1 is given by

$$f(\theta) = \sum_{k\in\mathbb{Z}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\theta - \mu + k)^2}{2\sigma^2} \right), \quad \theta \in [0, 1).$$

Then f is a smooth density function from $S^1 \to \mathbb{R}$, where $\mathbb{R}$ is equipped with the standard diffeology.

Example 7.2. Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ be an irrational number. The *irrational torus* T_α is defined as the quotient space

$$T_\alpha := \mathbb{R}/(\mathbb{Z} + \alpha\mathbb{Z}),$$

and we let $\pi_\alpha : \mathbb{R} \to T_\alpha$ denote the canonical projection sending $x \in \mathbb{R}$ to its equivalence class $[x] \in T_\alpha$.

We equip T_α with the *quotient diffeology*, defined as the pushforward diffeology $\pi_{\alpha*}(\mathcal{D}_\mathbb{R})$. That is, a map $p : U \to T_\alpha$ is a plot if and only if for every $u_0 \in U$, there exists a neighbourhood $V \subset U$ and a smooth map $q : V \to \mathbb{R}$ such that

$$p|_V = \pi_\alpha \circ q.$$

Let γ be the Gaussian probability measure on $\mathbb{R}$. We define a probability measure ν on T_α by pushing forward γ under π_α:

$$\nu := \pi_{\alpha*}\gamma.$$

We can interpret ν via its action on smooth functions. For any $f \in C^\infty(T_\alpha, \mathbb{R})$, we define

$$\phi_\nu(f) = \int_{T_\alpha} f d\nu = \int_\mathbb{R} (f \circ \pi_\alpha) d\gamma.$$

8 Conclusion

In this thesis, we have constructed a probability monad over the category of diffeological spaces, **Diff**, analogous to the Giry monad over the category of measurable spaces. We began by constructing measurable spaces out of diffeological spaces using the Borel σ-algebra generated by the D-topology, and use this to define the space of probability measures PX of a diffeological space X. We then used this to define a functor $P : \mathbf{Diff} \to \mathbf{Diff}$ that maps a diffeological space into its space of probability measures.

We showed what effect the functor P had on morphisms using pushforward of measures along smooth maps and were able to show that for any morphism f in **Diff**, Pf is also a morphism. Furthermore, P respects identity morphisms and composition of morphisms, making P a functor.

In our first attempt, where we used evaluation maps and the piece-wise smooth diffeology on $[0, 1]$ as a reference diffeology, we were able to show that the unit and multiplication of the monad, defined analogously to the unit and multiplication of the Giry monad, are not morphisms in **Diff**, breaking the monad structure. Therefore, we had to find some alternative approach.

We explored two alternative approaches

Restriction to absolutely continuous measures

The first approach was to restrict to absolutely continuous measures on a diffeological space, w.r.t some reference measure. The functor P' maps a diffeological space X to its space of absolutely continuous measures $P'X$. This raised two problems, however. Firstly, there does not exist a canonical reference measure for any diffeological space, so we are not able to define the space of absolutely continuous measures over a diffeological space in general. Secondly, we cannot construct P' as an endofunctor over **Diff**, as the action P' has is not defined for every object and the space $P'X$ does not have a canonical reference measure, making $P'P'X$ undefined.

The functional diffeology on the space of probability measures

The final approach, was to construct a diffeology on the space of probability measures using the functional diffeology. In this approach, we can construct a probability monad over **Diff**, where the unit and multiplication were morphisms in **Diff**.

9 Discussion

The intuitive approach of defining PX using evaluation maps aligns with the construction of the classical Giry monad, however this method fails to make the unit and multiplication smooth.

Defining the space of probability measures using the functional diffeology may offer a more promising solution, as we were able to construct a smooth unit and multiplication. Here, the diffeology is not defined directly in terms of plots for PX, but rather in terms of how the measure acts on smooth test functions. This construction however assumes that every probability measure $\mu \in PX$ can be viewed as a smooth linear functional on the space $C^\infty(X, \mathbb{R})$, that is $PX \subseteq C^\infty(X, \mathbb{R})^*$, which may not hold for all diffeological spaces.

Unlike manifolds and other extensions of smooth manifolds, e.g., Banach or Fréchet manifolds, diffeological spaces do not require coordinate charts. The coordinate-free nature of diffeological spaces align well with the approach of Armstrong and Brigo [AB18] of defining SDEs via jets.

Future work

Several directions of further research are discussed here. These include, but are not limited to:

- **Prove or disprove the conjecture that $PX \subseteq C^{\infty}(X, \mathbb{R})^*$ holds for arbitrary diffeological spaces:** We strongly suspect conjecture 6.3 not to hold for arbitrary diffeological spaces. It may hold for compact spaces. If the inclusion fails for some spaces, we could restrict PX to certain measures, such that they are distinguishable as smooth linear functional. Another possibility is to restrict the function space to functions with compact support, instead of bounded functions.
- **Develop a Kleisli category over Diff, based on the monad defined in theorem 6.12:** With a monad (P, η, m), the next step would be to build the associated Kleisli category. This would involve defining morphisms $X \to PY$ as smooth stochastic maps and understanding how they compose.
- **Explore applications to intrinsic SDEs on diffeological spaces via jets:** Jet spaces and diffeological spaces seem very compatible. Jets may allow defining stochastic differential equations in settings where manifolds are insufficient.

References

[AB18] John Armstrong and Damiano Brigo. Intrinsic stochastic differential equations as jets. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2210):20170559, February 2018.

[AM23] Alireza Ahmadi and Jean-Pierre Magnot. On diffeologies for power sets and measures. *Set-Valued Var. Anal 32, 41*, 2023.

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, 2001.

[BW95] Michael Barr and Charles Wells. *Category Theory for Computing Science*. Number v. 1 in Prentice-Hall international series in computer science. Prentice Hall, 1995.

[Chi12] Gregory S. Chirikjian. *Stochastic Processes on Lie Groups*, pages 361–388. Birkhäuser Boston, Boston, 2012.

[EG17] Bardelli Eleonora and Mennucci Andrea Carlo Giuseppe. Probability measures on infinite-dimensional Stiefel manifolds. *Journal of Geometric Mechanics*, 9(3):291–316, Thu Aug 31 22:00:00 UTC 2017.

[Elw82] Kenneth David Elworthy. *Stochastic Differential Equations on Manifolds*. Lecture note series / London mathematical society. Cambridge University Press, 1982.

[EM89] Michel Emery and Paul André Meyer. *Stochastic Calculus in Manifolds*. Universitext (Berlin. Print). World Publishing Company, 1989.

[Fed14] Herbert Federer. *Geometric Measure Theory*. Classics in Mathematics. Springer Berlin Heidelberg, 2014.

[FT22] Robert Simon Fong and Peter Tino. *Probability Densities on Manifolds*, pages 37–49. Springer International Publishing, Cham, 2022.

[Gir82] Michèle Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, pages 68–85, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg.

[Igl25] Patrick Iglesias-Zemmour. Why Diffeology?, March 2025.

[Itô50] Kiyosi Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Mathematical Journal*, 1:35–47, 1950.

[IZ13] Patrick Iglesias-Zemmour. *Diffeology*. American Mathematical Society, 2013.

[Rie14] Emily Riehl. *Category Theory in Context*. Cambridge University Press, 2014.

[Sou80] Jean-Marie Souriau. Groupes differentiels. In *Differential Geometrical Methods in Mathematical Physics*, pages 91–128, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg.