An Introduction to Higher Order Measure Theory and Integration with Quasi Borel Spaces

Andrew Ricketts

4th Year Project Report
Computer Science and Mathematics
School of Informatics
University of Edinburgh

2023

Abstract

Quasi-Borel spaces are a recently proposed foundation for higher order measure theory. Measure theory allows for a consistent formalisation of probability theory yet does not support function spaces. Thus its use as a theory for probabilistic programming languages is limited. Quasi-Borel spaces resolve this issue by taking functions as primitive notions instead of measurable subsets. We present an introduction to both measure and quasi-Borel space theory up to integration showing how the integrals over quasi-Borel spaces are equal to Lebesgue integrals and are also morphisms.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Andrew Ricketts*)

Acknowledgements

I would first like to thank my supervisor Ohad Kammar for his continuous help throughout the year whilst I completed the project and always being available to answer questions. I would like to thank Franz Miltz and Thanakrit Anutrakulchai for giving me their feedback on my report.

Table of Contents

Chapter 1

Introduction

1.1 Probabilistic Programming Languages

Probabilistic Programming languages are used to specify and automatically conduct inference on statistical models. Traditional probability theory, based on measure theory, does not lend easily to the design of these languages using categorical principles since the category of measurable spaces is not cartesian closed. This result proved by Aumann limits the use of higher order functions as there is no collection of measurable subsets such that an evaluation function can be measurable [Aumann, 1961]. This means, in general, there is no notion of function space for two measurable spaces and hence higher order measurable functions, which have become ubiquitous throughout software engineering, cannot be measurably evaluated.

Recently, there has been interest in the study of semantics of probabilistic programming languages. This has led to developments that allow for higher order probability theory using different categorical structures. These developments are focussed on discovering new categories based on objects that replace measurable spaces.

For instance, one paper proposes the notion of measurable cones and stable measurable functions to form a cpo-enriched cartesian closed category [Ehrhard et al., 2017]. Similarly, another presents a cartesian closed category based on quasi-Borel spaces [Heunen et al., 2018]. Both these categories are closed meaning that for any two objects in the category the set of morphisms between them is also an object. Thus function spaces can be defined which allows for higher order constructions.

Moreover, there is the development of a monoidal closed category based on Banach spaces and regular maps [Dahlqvist and Kozen, 2019]. This differs from the previous two developments in that it is not Cartesian. This means that their semantics is a model of linear logic. Cartesian closed categories on the other hand, instead of linear logic, have semantics that model the simply typed lambda calculus. We will focus on quasi-Borel space theory in this dissertation since it has the necessary qualities to model higher order probabilistic programming languages which replaces measurable spaces with a simple definition in comparison to Banach spaces and measurable cones [Bacci et al., 2018].

1.2 Measure Theory

Measure theory is the rigorous study and formalisation of sizes of sets allowing us to define probability theory that avoids paradoxes or unintuitive results like the Banach–Tarski paradox, [Wagon, 1985]. The primitive notions of measure theory are measurable subsets which allow for consistent operations that give us expected results. Given a set, thought of as a sample space, we pair it with its measurable subsets, thought of as event spaces, to define measurable spaces.

Measures map measurable subsets to non-negative values. They also must be countably additive, meaning they distribute over countably infinite collections of disjoint measurable subsets. A probability measure is a measure such that the measure of the sample space is 1. We formally model a random experiment using measurable spaces equipped with a probability measure.

Using these foundations, measurable functions are defined and gives rise to integration theory. In the probabilistic setting, integrals, specifically Lebesgue integrals, correspond with expectation and are used to further develop probability theory. In this dissertation we will introduce measure theory up to measurable functions assuming no prior knowledge to then compare with quasi-Borel spaces.

1.3 Quasi-Borel Spaces

Quasi-Borel spaces provide a new axiomatic framework for probability that can extend to function spaces as discussed previously. Proposed by [Heunen et al., 2018], a quasi-Borel space (QBS) is a tuple consisting of a carrier set equipped with a set of random elements. These random elements are functions that map from the reals to the carrier set such that they contain all constants, they are closed under pre-composition with measurable functions and are closed under recombination over Borel partitions of the reals.

In measure theory, we derive measurable functions from measurable sets. For a QBS we derive its measurable subsets from the set of random elements. This reverse definition is what allows for function space constructions as it does not contradict Aumman [Aumann, 1961]; we define the measurable subsets for any QBS and can construct function space QBSs giving us measurable subsets for a function space QBS. These function spaces consist of the functions between QBSs which preserve the random element structure, called QBS morphisms.

We will present QBS theory and compare it to measure theory. This will allow us to define integration on a QBS and show it is equal to the Lebesgue integral.

1.4 Integration Theory

Integrals are vitally important to probability theory as they form the basis of many theorems involving expectation. Since the aim of QBS theory is to model probability theory with function spaces, we need integrals over measurable functions and QBS morphisms to be equal.

For measure theoretic probability theory, the Lebesgue integral is defined using indicator functions of measurable subsets. The Lebesgue integral of these functions is simply the measure of the measurable subset it is indicating. By linearity these functions are extended to simple functions, which are finite linear combinations of disjoint indicator functions. Their Lebesgue integral is a finite linear combination of the integral of each indicator function multiplied by their respective constants. Then, the Lebesgue integral of any non-negative measurable function is the supremum of the integrals of all simple functions less than or equal to the function.

For QBS theory we will define instead an approximation function for any QBS morphism which maps to non-negative real numbers. This approximation function will map a morphism to a sequence of increasingly complex simple functions such that the limit of this sequence is equal to the morphism given. Then the integral is calculated in the same way as the Lebesgue integral of simple functions. An important result is that this integral itself, and thus expectation, is a QBS morphism.

1.5 Overview and Contributions

This dissertation will present measure theory from the foundations in measurable subsets up to measurable functions. Then the theory of QBS will be introduced and the links with measurability will be established. With this theoretical introduction, integration theory will be the focus of the final technical chapter. We will define an integral over a QBS, show that it is a morphism and that it is equal to the Lebesgue integral.

The structure of this dissertation is as follows:

- Chapter 2 develops measure theory. We begin with the definition of measurable subsets and measurable spaces. Subsequently, we define and examine the properties of measures. Following this, we define measurable functions and random variables, examining their properties and presenting important examples.
- Chapter 3 introduces and develops QBS theory. We define QBS structures and detail their axioms. Morphisms between QBSs are then explored leading to the definition of measurable subsets of a QBS. We then examine four different constructions: subspaces, product, co-product and function spaces.
- Chapter 4 concerns integration theory and defines the integral of a QBS morphism. We show it is a QBS morphism and equal to the Lebesgue integral.
- Chapter 5 gives three examples of relevant work in more detail. We review the original paper [Heunen et al., 2018]. Then we explore an extension of QBS theory in [Vákár et al., 2019] to include recursion and finally, we analyse a use case of QBS theory shown in [Ścibior et al., 2017].
- Chapter 6 is an evaluation of our work and explores the possibility of future research.

Chapter 2

Measure Theoretic Probability

Rigorous probability theory is made possible by measurable sets which circumvent the inconsistencies of previous probabilistic frameworks. Measure theory axiomatises measurable sets such that they are closed under countably many basic set operations; union, intersection and complement. This allows sets to have a rigorous expression of "size" as a measure. These measures are defined to give us a consistent way to calculate the probability of events and are essential to integration theory. Measurable functions, the central subject of integration theory, allow us to map between measurable spaces and formalise random variables. In this chapter we will develop the foundations of measure theoretic probability theory up to measurable functions to motivate and explain the next two chapters.

2.1 Measurable Sets

Our framework requires a formal representation of sample spaces and events. Measure theory gives us measurable spaces. These are pairs of sets, which represents the sample space in probability, and collection of sets, which represents events. This collection is a σ-algebra, which is closed under countably many basic set operations.

Definition 2.1.1. *Let S be a set. Then a collection of sets $\Sigma \subseteq \mathcal{P}(S)$ is a σ-algebra when*

- $S \in \Sigma$.
- *Given a subset $A \in \Sigma$ then $A^c \in \Sigma$.*
- *Given $A_n \in \Sigma, (n \in \mathbb{N})$, then $\bigcup_{n\in\mathbb{N}} A_n \in \Sigma$.*

Example: Let $S = \{0,1\}$. Then, $\Sigma_1 = \{\emptyset, \{0\}, \{1\}, \{0,1\}\} = \mathcal{P}(S)$ is a σ-algebra. Notice due to the complement condition, the empty set is also always an element of a σ-algebra of a set. The smallest σ-algebra of any set S is $\{\emptyset, S\}$.

For probability theory, we focus on the Borel σ-algebra, which is the σ-algebra *generated* by all the open sets, a class of subsets of $\mathbb{R}$. A class of sets is a collection of subsets of a set which share a property, like being open or closed or containing the element

$r \in \mathbb{R}$. Let $\mathcal{C}$ be a class of subsets of S. Then $\sigma(\mathcal{C})$, the σ-algebra *generated* by $\mathcal{C}$, is the smallest σ-algebra Σ on S such that $\mathcal{C} \subseteq \Sigma$. The Borel σ-algebra, $\mathcal{B}_{\mathbb{R}}$, is important as it contains many of the regular subsets that make sense in everyday probability.

Example: $[0,1] = ((-\infty,0) \cup (1,\infty))^c$. We have that $(-\infty,0)$ and $(1,\infty)$ are both open sets so they are both elements of $\mathcal{B}_{\mathbb{R}}$. Their union must also be an element of $\mathcal{B}_{\mathbb{R}}$ as it is a countable union of sets in the σ-algebra. Finally, since $\mathcal{B}_{\mathbb{R}}$ is closed under complement the complement of their union must also be an element of $\mathcal{B}_{\mathbb{R}}$.

A simpler way to show this is a step before where $[0,1] = (-\infty,0) \cap (1,\infty)$ by showing that since σ-algebras are closed under complement and countable union they must also be closed under countable intersection.

Proposition 2.1. *Given* $A_n \in \Sigma, (n \in \mathbb{N})$, *then* $\bigcap_{n\in\mathbb{N}} A_n \in \Sigma$.

Proof. Since $A_n \in \Sigma$ then $A_n^c \in \Sigma$. Thus, we have $\bigcup_{n\in\mathbb{N}} A_n^c \in \Sigma$ by the third axiom which gives $(\bigcup_{n\in\mathbb{N}} A_n^c)^c \in \Sigma$ again by the complement axiom. We have that $(\bigcup_{n\in\mathbb{N}} A_n^c)^c = \bigcap_{n\in\mathbb{N}} A_n$. Therefore, $\bigcap_{n\in\mathbb{N}} A_n \in \Sigma$. $\square$

Since all the open sets are included in $\mathcal{B}_{\mathbb{R}}$ the complement of a union of two open sets must also be in $\mathcal{B}_{\mathbb{R}}$ by definition of σ-algebras.

Example: $(-\infty,a] = \bigcap_{n=1}^{\infty}(-\infty, a+\frac{1}{n}), \forall a \in \mathbb{R}$. Since $(-\infty,a]$ can be expressed as a countable intersection of open intervals, by the definition of a σ-algebra and the fact $\mathcal{B}_{\mathbb{R}}$ contains all open sets, $(-\infty,a]$ must be in $\mathcal{B}_{\mathbb{R}}$.

Below is an important property of generated σ-algebras relating to subsets. Consider a collection of sets and its generated σ-algebra. Then a subset of that collection generates a subset of that generated σ-algebra. Additionally, generated σ-algebra of a σ-algebra is the same σ-algebra.

Proposition 2.2. *Given collections of sets* $\mathcal{U}, \mathcal{V} \subset \mathcal{P}(X)$.

(i) If $\mathcal{U} \subseteq \mathcal{V}$ *then* $\sigma(\mathcal{U}) \subseteq \sigma(\mathcal{V})$

(ii) If $\mathcal{U}$ *is a* σ*-algebra,* $\mathcal{U} = \sigma(\mathcal{U})$

Proof.

(i) We have $\mathcal{U} \subseteq \mathcal{V}$. $\sigma(\mathcal{V})$ is the smallest σ-algebra that contains $\mathcal{V}$.

Therefore, $\mathcal{V} \subseteq \sigma(\mathcal{V})$ and for all other σ-algebras Σ such that $\mathcal{V} \subseteq \Sigma$ then $\sigma(\mathcal{V}) \subseteq \Sigma$.

Let $x \in \mathcal{U}$ then we also have $x \in \mathcal{V}$. This implies that $x \in \bigcap\{\mathcal{B}_i | \mathcal{V} \subseteq \mathcal{B}_i, \mathcal{B}_i \text{ is a } \sigma\text{-algebra}\}$. Therefore, $\mathcal{U} \subseteq \sigma(\mathcal{V})$ since. Hence, $\sigma(\mathcal{V})$ is a σ-algebra over X that contains $\mathcal{U}$. Therefore, $\sigma(\mathcal{U}) \subseteq \sigma(\mathcal{V})$.

(ii) Given $\mathcal{U}$ is a collection of sets then, $\mathcal{U} \subseteq \sigma(\mathcal{U})$. Since $\mathcal{U}$ is a σ-algebra containing $\mathcal{U}$ then $\sigma(\mathcal{U}) \subseteq \mathcal{U}$ since $\sigma(\mathcal{U})$ is the smallest σ-algebra containing $\mathcal{U}$. Therefore, $\mathcal{U} = \sigma(\mathcal{U})$.

□

Now we can use σ-algebras in a measurable space to combine the notions of sample space and events.

Definition 2.1.2. *A measurable space is a pair (S, Σ), where S is a set and Σ is a σ-algebra over S. A Σ-measurable subset of S is an element $A \in \Sigma$*

Usually, we can refer to $A \in \Sigma$ as measurable given the context of some measurable space. When a set A is an element of the Borel σ-algebra we say it is Borel.

Example: In an experiment you toss a coin twice. Thus, your sample space is the set of all the possible results in one trial. $S = \{HH, HT, TH, TT\}$. Let $\Sigma = \mathcal{P}(S)$. Then, the event a head and a tail is obtained is $\{HT, TH\}$ which would be an element of Σ.

Example: Choose n random points in $[0,1]$. The sample space in this case is $[0,1]^n$ and $\Sigma = \mathcal{B}_{[0,1]^n} \subset \mathcal{B}_{\mathbb{R}^n}$. In this example we choose the points uniformly, meaning each point has an equal chance of being selected. We do not yet have a process or method to assign probabilities to these subsets. For instance, we do not know the probability that all n points are less than 0.5. The next section gives the measure theoretic approach of assigning probability to the measurable subsets.

2.2 Measures

Given our measurable space we now want to know the probability of an event. We need this assignment to follow probabilistic intuition. For example, we should be assigning non-exclusive events according to the inclusion-exclusion principle: the total probability of two non-exclusive events should be the sum of the individual probabilities of the events minus the probability of the overlapping events. Measures assign a non-negative value to each set in the σ-algebra. Importantly, measures are also countably additive, which allows for standard probabilistic notions to be consistent in our framework.

Definition 2.2.1. *Let Σ be a σ-algebra and $\mu : \Sigma \to \mathbb{W}$, where $\mathbb{W} = [0, \infty]$. Then, μ is a measure when,*

- $\forall A \in \Sigma$, $\mu[A] \geq 0$.
- $\mu[\emptyset] = 0$.
- ***Countably additive****: Given $A_n \in \Sigma, \forall n \in \mathbb{N}$ then $\mu[\bigcup_{n \in \mathbb{N}} A_n] = \sum_{n \in \mathbb{N}} \mu[A_n]$, when $A_i \cap A_j = \emptyset, \forall i, j \in \mathbb{N}$.*

Example: A very simple measure is the counting measure $\#(A) = |A|, \forall A \in S$. No element of Σ can have cardinality less than 0 thus the first condition is satisfied. The empty set has cardinality 0. Finally, we have that $\sum_{n=0}^{k} |A_n| = |\bigcup_{n=0}^{k} A_n|$ for disjoint A_n as required.

In a probabilistic setting, if we have a measurable space (X,Σ) then we want $\mu[X]=1$. Thus, we define a *probability measure* to be a measure where the measure of the whole set is 1. For finite measures, those for which $\mu[A]<\infty,\forall A$, we can easily construct probability measures.

Proposition 2.3. *Let (S,Σ) be a measurable space and $\mu:\Sigma\to[0,\infty)$ be a finite measure. Then, the measure $\mathbb{P}[A]=\frac{\mu[A]}{\mu[S]}$ is a probability measure if $\mu[S]>0$.*

Proof. Let us check the four conditions.

(i) Since μ is non-negative, $\mu[A]\geq 0,\mu[S]>0$, implies $\frac{\mu[A]}{\mu[S]}\geq 0$.

(ii) Since μ is a measure, $\mathbb{P}[\emptyset]=\frac{\mu[\emptyset]}{\mu[S]}=\frac{0}{\mu[S]}=0$.

(iii) If $\{A_n\}_{n\in\mathbb{N}}$ is a collection of disjoint sets then

$$\mathbb{P}[\bigcup_{n\in\mathbb{N}}A_n]=\frac{1}{\mu[S]}\cdot\mu[\bigcup_{n\in\mathbb{N}}A_n]=\frac{1}{\mu[S]}\cdot\sum_{n\in\mathbb{N}}\mu[A_n]=\sum_{n\in\mathbb{N}}\mathbb{P}[A_n].$$

(iv) We have $\mathbb{P}[S]=\frac{\mu[S]}{\mu[S]}=1$.

$\square$

Example: A more concrete example, if we consider a discrete uniform distribution, then let # be the counting measure. Then, $\mathbb{P}(A)=\frac{\#(A)}{\#(S)}$ is again a probability measure by Proposition 2.3. If we have the experiment of tossing a coin twice then the probability of the result including both a head and a tail is,

$$\mathbb{P}[\{HT,TH\}]=\frac{\#\{HT,TH\}}{\#\{HH,HT,TH,TT\}}=\frac{2}{4}=\frac{1}{2}.$$

Measures have intuitive basic properties to make them usable.

Lemma 2.1. *Let (S,Σ,μ) be a measure space. Then,*

- $\mu[A\cup B]\leq\mu[A]+\mu[B],(A,B\in\Sigma)$
- $\mu[A\cup B]=\mu[A]+\mu[B]-\mu[A\cap B],(A,B\in\Sigma)$ *when* $\mu[A\cup B]\leq\infty$

Since the second property implies the first we prove the second property.

Proof. We can see that $A\cup B=A\cup(B\setminus(A\cap B))$, which is a disjoint union. So by definition of measures

$$\begin{aligned}\mu[A\cup B]&=\mu[A\cup(B\setminus(A\cap B))]\\&=\mu[A]+\mu[(B\setminus(A\cap B))]\text{ by }\sigma\text{-additivity}\\&=\mu[A]+\mu[B]-\mu[A\cap B],\end{aligned}$$

since $\mu[X\setminus Y]=\mu[X\cap Y^c]=\mu[X]-\mu[Y]$ when $Y\subseteq X$. $\square$

More general properties hold for sequences of sets in Σ which can be proved by induction from these properties.

2.2.1 Lebesgue Measure

The standard measure for subsets of Euclidean space is the Lebesgue measure. It is the precise concept of length or volume applied to all Lebesgue measurable sets. To construct the measure we apply a powerful theorem in measure theory, Carathéodory's extension theorem. Extension theorems allow us to take functions, in our case measures, from smaller objects to larger objects, in our case algebras to σ-algebras. The full theorem involves rings of subsets of $\mathbb{R}$ but for our purposes we will take a less general version, Theorem 1.7 appearing in [Williams, 1991]. We will only consider algebras and σ-algebras.

Firstly, an *algebra* is a collection of sets which satisfy the first two σ-algebra conditions and also given two elements of an algebra F, G then $F \cup G$ is also in the algebra. Essentially, the difference between algebras and σ-algebras is the last conditions. Algebras are closed under pairwise union and σ-algebras are closed under countably infinite union.

Theorem 2.2.1. ***Carathéodory's Extension Theorem*** *[Williams, 1991], Let S be a set and let Σ_0 be an algebra on S. Let $\Sigma = \sigma(\Sigma_0)$. If $\mu_0 : \Sigma_0 \to [0,\infty]$ is a countably additive map, then there exists a measure μ on (S,Σ) such that $\mu[A] = \mu_0[A], \forall A \in \Sigma_0$. If $\mu_0 < \infty$ then the extension is unique.*

This says that given an algebra and countably additive set function, also known as a pre-measure, we can extend it to the σ-algebra generated by the algebra. Then, in this new measurable space the pre-measure and measure agree on the algebra and the remaining elements of the σ-algebra now have a measure defined on them. This allows the definition of pre-measures on much simpler algebras so that we can generalise notions to σ-algebras by only caring about the sets that are intuitive. The measure is defined as the outer measure given the pre-measure. The *outer measure* μ of a pre-measure μ_0 is defined as,

$$\mu[A] = \inf\{\sum_{n\in\mathbb{N}} \mu_0[A_n] | A \subset \bigcup_{n\in\mathbb{N}} A_n\}.$$

Applying this theorem, let $S = \mathbb{R}^n$ and for $F \subseteq \mathbb{R}^n$, let $F \in \Sigma_0$ if F can be expressed as a finite union of closed boxes

$$F = \prod_{i=1}^{n}[a_i, b_i] \cup \ldots \cup \prod_{i=1}^{n}[a_i, b_i] = \bigcup_{k=0}^{m}\prod_{i=1}^{n}[a_{k,i}, b_{k,i}],$$

where m is finite. Then, Σ_0 is an algebra. Define $\mu_0 : \Sigma_0 \to [0,\infty)$,

$$\mu_0 = \sum_{k=0}^{m}\prod_{i=1}^{n}(b_{k,i} - a_{k,i}).$$

This gives a countably additive map and by Carathéodory's Extension Theorem there exists a unique measure μ from $\sigma(\Sigma_0)$ to $[0,\infty)$. We have $\sigma(\Sigma_0) = \mathcal{B}_{\mathbb{R}^n}$ and therefore $\mu : \mathcal{B}_{\mathbb{R}^n} \to [0,\infty)$ and $\mu[F] = \mu_0[F]$ for all $F \in \Sigma_0$. From this we can see that the Lebesgue measure generalises volume since the product of the length of sides of each box gives the volume and we simply sum over all boxes. Thus we built a measure that makes intuitive sense on sets we understand but the complicated sets resulting from the countably infinite set operations on the closed boxes are not considered. Their cases handled by the extension theorem and outer measure.

2.3 Measurable Functions

Random variables are an important topic in probability theory as they allow us to formalise values which are random. In measure theory, random variables are measurable functions from probability spaces to a measurable space, usually $(\mathbb{R}, \mathcal{B}_\mathbb{R})$.

We will start with an example. Given two dice, we model rolling both as the probability space $(S, \Sigma, \mathbb{P})$ where S is the cartesian product of $\{1,2,3,4,5,6\}$ with itself; $S = \{1,2,3,4,5,6\}^2$. For instance, an element of S could be $(1,1)$, representing when both dice come up as 1. Then $\Sigma = \mathcal{P}(S)$, thus Σ contains all subsets of S as events. For example, the set $\{(i,j)|i+j \leq 3\} = \{(1,1),(2,1),(1,2)\} \in \Sigma$ represents the event where the sum of the dice is at most 3. $\mathbb{P}$ is the discrete uniform distribution; the probability of rolling any one event like (a,b) is $\frac{1}{36}, a,b \in \{1,2,3,4,5,6\}$.

Then, let $f : S \to \mathbb{R}$ where $f(s_1,s_2) = \frac{(s_1+s_2)}{2}$ be equal to the average of the roll. We now give the definition of a measurable function.

Definition 2.3.1. *Let* (S,Σ)*,* $(\Omega, \mathcal{F})$ *be a measurable spaces. Then a function* $f : S \to \Omega$ *is* $\langle \Sigma, \mathcal{F} \rangle$*-measurable when* $f^{-1}[A] \in \Sigma, \forall A \in \mathcal{F}$*.*

Thus, to be measurable a function must preserve the structure of the measurable spaces; the pre-image of a measurable set in one space must be measurable in the other.

Going back to our example, with $f(s_1,s_2) = \frac{s_1+s_2}{2}$. The function$f$ is measurable, specifically $\langle \mathcal{P}(S), \mathcal{B}_\mathbb{R} \rangle$-measurable. We know it is measurable since the powerset contains all subsets so $f^{-1}[A] \in \mathcal{P}(S)$. If we have that the $\Sigma = \{\emptyset, S\}$ then f would not be measurable since $f^{-1}(0.5) = \{(1,1)\} \notin \{\emptyset, S\}$. We call f a *random variable* since it maps from a probability space to another measurable space.

An alternate definition of measurability for real-valued functions, those which map to $\mathbb{R}$, can be useful when assessing measurability of a function.

Proposition 2.4. *Let* (S,Σ) *be a measurable space. Then, a function* $f : S \to \mathbb{R}$ *is* $\langle \Sigma, \mathcal{B}_\mathbb{R} \rangle$*-measurable if* $\forall a \in \mathbb{R}, f^{-1}[(-\infty, a]] = \{s | f(s) \leq a\} \in \Sigma$

To see why they are equivalent, note that a function which is measurable on the class of sets which generates a σ-algebra is measurable on the whole σ-algebra shown in 2.5. In addition, the collection of sets $\pi(\mathbb{R}) = \{(-\infty, a] | a \in \mathbb{R}\}$ generates $\mathcal{B}_\mathbb{R}$, shown in 2.6.

Proposition 2.5. *Given two measurable spaces* $(X,\Sigma),(Y,\Gamma)$*. Let* $\Gamma = \sigma[\mathcal{A}]$ *for a class of subsets of* $\mathcal{A}$ *of* Y*. Then, a function* $f : X \to Y$ *which is measurable for all* $A \in \mathcal{A}$ *is* $\langle \Sigma, \Gamma \rangle$*-measurable.*

Proof. Let $\mathcal{F} = \{A \in \Gamma | f^{-1}(A) \in \Sigma\} \subset \Gamma$. Thus $\mathcal{F}$ are the sets for which f is measurable. We will now show $\mathcal{F}$ is a σ-algebra and this leads to the fact that $\mathcal{F} = \Gamma$.

Clearly, $\mathcal{F}$ contains the empty set since $f^{-1}[\emptyset] = \emptyset \in \Sigma$ since Σ is a σ-algebra.

If $A \in \mathcal{F}$ then $f^{-1}[A] \in \Sigma$. Then by the complement axiom we have $(f^{-1}[A])^c \in \Sigma$ gives $f^{-1}[A^c] \in \Sigma$. Therefore, $A^c \in \mathcal{F}$.

If $A_1, A_2, \ldots \in \mathcal{F}$ then $f^{-1}[A_i] \in \Sigma$ implies that $\bigcup_{i \in \mathbb{N}} f^{-1}[A_i] \in \Sigma$. Thus, $f^{-1}[\bigcup_{i \in \mathbb{N}}[A_i]] \in \Sigma$. Therefore, $\bigcup_{i \in \mathbb{N}}[A_i] \in \mathcal{F}$.

This shows that $\mathcal{F}$ is a σ-algebra. Then, $\mathcal{A} \subset \mathcal{F}$ since $\Gamma = \sigma(\mathcal{A})$. Thus $\Gamma = \sigma[\mathcal{A}]$ is a subset of $\sigma[\mathcal{F}] = \mathcal{F}$ by Proposition 2.2. Thus we have $\Gamma \subset \mathcal{F}$ and $\mathcal{F} \subset \Gamma$ and therefore $\Gamma = \mathcal{F}$. Then, f is $\langle \Sigma, \Gamma \rangle$-measurable. ☐

Proposition 2.6. *The Borel σ-algebra is generated by $\pi(\mathbb{R})$, $\mathcal{B}_{\mathbb{R}} = \sigma(\pi(\mathbb{R}))$.*

Proof. First we show that $\sigma(\pi(\mathbb{R})) \subseteq \mathcal{B}_{\mathbb{R}}$. We have that $(-\infty, a] = \bigcap_{n \in \mathbb{N}} (-\infty, a + \frac{1}{n})$ which is a countable intersection of open sets thus in the collection of open sets. Therefore, if $O = \{A \subset \mathbb{R} | A \text{ is open}\}$ then $\pi(\mathbb{R}) \subset O$ and $\mathcal{B}_{\mathbb{R}} = \sigma(O)$. Thus by Proposition 2.2, $\sigma(\pi(\mathbb{R})) \subseteq \sigma(O) = \mathcal{B}_{\mathbb{R}}$

Now we can show that all open sets are in $\sigma(\pi(\mathbb{R}))$. That is to say $O \subset \sigma(\pi(\mathbb{R}))$ since then we can say that $\mathcal{B}_{\mathbb{R}} = \sigma(O) \subset \sigma(\sigma(\pi(\mathbb{R}))) = \sigma(\pi(\mathbb{R}))$ by Proposition 2.2.

Note every element of O is a countable union of open intervals. So we can show that $(a, b) \in \sigma(\pi(\mathbb{R}))$ and it will follow that all the open sets are in $\sigma(\pi(\mathbb{R}))$. We have that $(a, b) = \bigcup_{n \in \mathbb{N}} (a, b - \frac{\varepsilon}{n}]$ for $\varepsilon = \frac{b-a}{2}$. We also have that $(a, u] = (-\infty, u] \cap (-\infty, a]^c \in \sigma(\pi(\mathbb{R}))$ and therefore $(a, b) \in \sigma(\pi(\mathbb{R}))$, as required. ☐

Therefore, by Proposition 2.5 and 2.6 we have that Proposition 2.4 is true; measurability of a real valued function is determined by measurability on $(-\infty, a], \forall a \in \mathbb{R}$.

All continuous functions $f : \mathbb{R}^p \to \mathbb{R}^k$ are measurable. This result gives us a very wide range of functions which are imaginable and easily identifiable. However, it is not true that all measurable functions are continuous. For instance, let $f(x) = \mathbb{1}_{\mathbb{Q}}$. The indicator functions of the rationals is discontinuous everywhere but since the rationals are a measurable set the inverse of 1 is measurable and the inverse of 0 is the complement, the irrationals, so must also be measurable.

Proposition 2.7. *For any $k, p \in \mathbb{N}$, if $f : \mathbb{R}^p \to \mathbb{R}^k$ is continuous then f is $\langle \mathcal{B}_{\mathbb{R}^p}, \mathcal{B}_{\mathbb{R}^k} \rangle$-measurable.*

Proof. Let $\mathcal{A} = \{A : A \text{ is open in } \mathbb{R}^k\}$. Then, since f is continuous the pre-image preserves the open property of sets. Thus, $f^{-1}[A]$ is open and hence $f^{-1}[A] \in \mathcal{B}_{\mathbb{R}^p}, \forall A \in \mathcal{A}$. Since $\mathcal{B}_{\mathbb{R}^p} = \sigma(\mathcal{A})$ by Proposition 2.5 f is $\langle \mathcal{B}_{\mathbb{R}^p}, \mathcal{B}_{\mathbb{R}^k} \rangle$-measurable, as required. ☐

Sums and products of measurable functions are also measurable. Importantly, the limits of measurable functions are measurable. This is an important fact used in integration theory as the Lebesgue integral is defined using a limiting operation of measurable functions. Furthermore, some of the most important theorems in measure theory concern themselves with limits.

Proposition 2.8. *Let $(f_n : n \in \mathbb{N})$ be $\langle \Sigma, \mathcal{B}_{\mathbb{R}} \rangle$-measurable functions (shortened to measurable). Then,*

(i) $f_i + f_j$ *is measurable*

(ii) $f_i \cdot g_j$ *is measurable*

(iii) $\inf(f_n : n \in \mathbb{N})$ *is measurable*

(iv) $\liminf(f_n : n \in \mathbb{N})$ *is measurable*

(v) $\limsup(f_n : n \in \mathbb{N})$ *is measurable*

Furthermore, measurable functions compose when their respective σ-algebras align.

Proposition 2.9. *Let $(X,\Sigma),(Y,\Gamma)$ and (Z,Δ) be measureable spaces. If $f : X \to Y$ and $g : Y \to Z$ are $\langle \Sigma, \Gamma \rangle$-measurable and $\langle \Gamma, \Delta \rangle$-measurable functions then $f \circ g$ is $\langle \Sigma, \Delta \rangle$-measurable.*

2.4 Laws

With our random variables we want to know the probability of an event. For instance, let X be the average of two dice rolls. What the probability that the average is 2, $\mathbb{P}(X = 2)$? With measurable functions we can use the measure on the original probability space and push it onto the space which has our answer. We call the measure the law of a random variable or the push-forward measure.

Definition 2.4.1. *Let $(S,\Sigma,\mathbb{P})$ be a probability space and $f : S \to \mathbb{R}$ be a random variable. Then, $\mathcal{L}_f : \mathcal{B}_R \to [0,1]$,*

$$\mathcal{L}_f = \mathbb{P} \circ f^{-1}$$

is the law f.

The law of a f is the push-forward measure of $\mathbb{P}$ along f. We can show that the law of a function is a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Proposition 2.10. *$\mathcal{L}_f$ is a probability measure.*

Proof. (i) Firstly, the empty set has measure 0 under $\mathcal{L}_f$ since $\mathcal{L}_f[\emptyset] = \mathbb{P}[f^{-1}[\emptyset]] = \mathbb{P}[\emptyset] = 0$ and by definition of $\mathbb{P}, \mathcal{L}_f$ ranges between [0,1].

(ii) Next we have that the measure of $\mathbb{R}$ is 1 as $\mathcal{L}_f[\mathbb{R}] = \mathbb{P}[f^{-1}[\mathbb{R}]] = \mathbb{P}[S] = 1$.

(iii) Finally, the pre-image of f preserves the countable additivity property of $\mathbb{P}$.

$$\mathcal{L}_f[\bigcup_{i \in \mathbb{N}} A_i] = \mathbb{P}[f^{-1}[\bigcup_{i \in \mathbb{N}} A_i]] = \mathbb{P}[\bigcup_{i \in \mathbb{N}} f^{-1}[A_i]] = \sum_{i \in \mathbb{N}} \mathbb{P}[f^{-1}[A_i]] = \sum_{i \in \mathbb{N}} \mathcal{L}_f[A_i].$$

□

Going back to our example, the probability that the average roll is 2 is given by $\mathcal{L}_f[2] = \mathbb{P} \circ f^{-1}[2] = \mathbb{P}[\{(1,3),(2,2),(3,1)\}] = \frac{1}{12}$.

Chapter 3

Quasi Borel Spaces

We now introduce a new framework to axiomatise probability theory which does not take measurable sets as a primitive notion but instead takes a set of random elements. The set of random elements are the functions we want to be measurable and from this we derive the collection of measurable subsets. This is motivated by a result by Aumann, which states that given two measurable spaces X, Y and their function space of Borel measurable functions Y^X, there is not in general a σ-algebra such that $\phi : Y^X \times X \to Y$, $\phi(f,x) = f(x)$ is measurable [Aumann, 1961]. Even for some of the simplest examples this fails, for instance if we have a measurable space which has the powerset as the σ-algebra we still do not have measurable evaluation.

This is the motivation behind QBS theory; to develop probability theory from objects that support the construction of function spaces. In this chapter we develop the theory from the foundation, as in the previous chapter, giving an overview which will allow us to define integration in the next chapter.

3.1 Definition

We define quasi-Borel Spaces (QBS) and Figure 3.1 — 3.3 shows the underlying principle of the axioms.

Definition 3.1.1. *A QBS $X = (\underline{X}, M_X)$ consists of a set $\underline{X}$ and a subset M_X of functions $\mathbb{R} \to X$ such that:*

- *Constants: M_X contains all constant functions, $\forall x \in X, \underline{x}(r) := x \in M_X$.*
- *Pre-composition: if $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable and $\alpha \in M_X$ then $\alpha \circ f \in M_X$.*
- *Recombination: for every partition of $\mathbb{R} = \bigcup_{i \in \mathbb{N}} S_i$ by Borel S_i and a countable sequence of functions $\{\alpha_i \in M_X\}_{i \in \mathbb{N}}$. Let $\beta(x) = \alpha_i(x)$ when $x \in S_i$, then $\beta \in M_X$.*

Remark. *For a QBS X, we refer to its carrier set as $\underline{X}$ and its set of random elements as M_X.*

Figure 3.1 shows the first axiom, all constant functions must be in the set of random elements. It assigns all the probability mass onto one point. This provides a deterministic

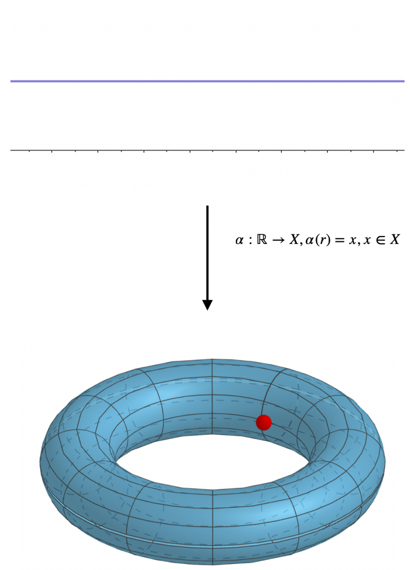

Figure 3.1: Diagram of QBS constant function axiom showing how all constant functions from $\mathbb{R}$ to the space are always random elements.

formulation in the definition of QBS. In the diagram, we use the random element to push the probability over the reals, in this case a uniform distribution, onto the space which results in a singular point.

Figure 3.2 shows the second axiom, closure under pre-composition with a Borel measurable function. Given any measurable function, in this case a simple function of the form $a\mathbb{1}_A + b\mathbb{1}_B$ where $A, B \in \mathcal{B}_\mathbb{R}$, and another random element that maps into the space the resulting composition is also a random element mapping onto the space. In the diagram we then use the composed random element to push the probability distribution over the reals to the space only if those reals values are in the sets A and B.

Figure 3.3 shows the third axiom. We split $\mathbb{R}$ into a countable union of disjoint Borel sets each with a random element assigned to it. Then, the function which maps $r \in \mathbb{R}$, which can only be in one Borel set, using the random element assigned to that Borel set is a random element. In the figure the Borel partition si $\mathbb{R} = [-\infty, 0) \cup [0, \infty]$. For example, if $r = 1$ the recombination would map the probability into the green segment of the space.

Example: Let $\mathrm{Meas}(\mathbb{R}, \mathbb{R})$ be the set of all Borel measurable functions on $\mathbb{R}$. These are functions that are $\langle \mathcal{B}_\mathbb{R}, \mathcal{B}_\mathbb{R} \rangle$-measurable. This set of functions satisfies the QBS axioms. Firstly, all constant functions are measurable. Secondly, composition of measurable functions are measurable as seen in Proposition 2.7. Finally, to show recombination of measurable functions of Borel partitions is measurable let $\mathbb{R} = \bigcup_{i \in \mathbb{N}} S_i$ for disjoint Borel sets and $\beta(x) = a_i(x)$ for $x \in S_i$ where $a_i \in \mathrm{Meas}(\mathbb{R}, \mathbb{R})$. Then,

$$\beta^{-1}((-\infty, c)) = \{\omega | \beta(\omega) < c\} = \bigcup_{i \in \mathbb{N}} \{\omega | a_i(\omega) < c, a_i(\omega) \in S_i\} = \bigcup_{i \in \mathbb{N}} (a_i^{-1}(-\infty, c) \cap S_i)).$$

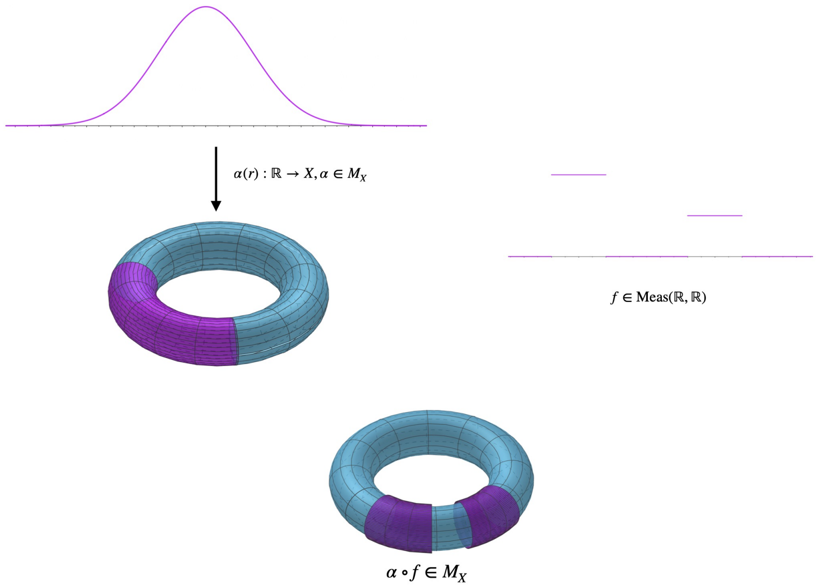

Figure 3.2: Diagram of QBS closure under pre-composition with measurable functions axiom showing how pre-composing any random element with any $f \in \mathrm{Meas}(\mathbb{R}, \mathbb{R})$ is also a random element.

Each set in the union is Borel since a_i is measurable and S_i is Borel, thus their countable union is Borel. Hence $\beta^{-1}((-\infty, a)) = \{\omega | \beta(\omega) < a\}$ is Borel which implies $\beta \in \mathrm{Meas}(\mathbb{R}, \mathbb{R})$. Therefore we have that $(\underline{\mathbb{R}}, \mathrm{Meas}(\mathbb{R}, \mathbb{R}))$ is a QBS.

Two other important examples of QBSs are the discrete and indiscrete QBS structure over any carrier set X.

Example: Let $\underline{X}$ be a set. Let $\mathrm{Set}(\mathbb{R}, \underline{X})$ be the set of all functions from $\underline{\mathbb{R}} \to \underline{X}$. Then, $(\underline{X}, \mathrm{Set}(\underline{\mathbb{R}}, \underline{X}))$ is the indiscrete QBS. The set of all functions contains all the constant functions. Any function precomposed with any measurable function is simply another function so must also be contained in the set. Finally, recombining functions over Borel partitions of $\mathbb{R}$ is again just another function thus must also be contained in the set. Therefore, it is a QBS. It is referred to as indiscrete since we take all the functions as the structure with no selection criteria.

Definition 3.1.2. *Let $\underline{X}$ be a set. A function $\alpha : \underline{\mathbb{R}} \to \underline{X}$ is a σ-simple function if:*

- *The image $\alpha[\underline{\mathbb{R}}] = \{\alpha(r) | r(x) \in \underline{\mathbb{R}}\}$ is countable*
- *$\forall x \in \alpha[\underline{\mathbb{R}}]$, the pre-image $\alpha^{-1}(x) \in \mathcal{B}_{\underline{\mathbb{R}}}$*

The σ-simple functions are the smallest set of functions that a QBS can have since they can be derived directly from the axioms of a QBS. This is why the construction below is called discrete, it is the most selective set of functions that satisfy the axioms.

Proposition 3.1. ***(Discrete QBS)*** *$(\underline{X}, \{\alpha : \underline{\mathbb{R}} \to \underline{X} | \alpha \text{ is } \sigma \text{ simple}\})$ is a QBS.*

Proof. All constants are σ-simple, since the image of $\underline{\mathbb{R}}$ would be the constant and the

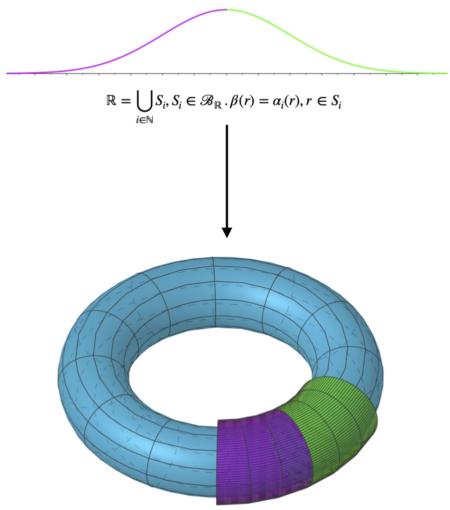

Figure 3.3: Diagram of recombination of random elements.

pre-image of x would always be $\mathbb{R}$.

Precomposing a σ-simple function α with a measurable function f is also σ-simple since $\alpha \circ f[\underline{\mathbb{R}}]$ must be countable since $\alpha[\underline{\mathbb{R}}]$ is countable and $f[\underline{\mathbb{R}}] \subseteq \underline{\mathbb{R}} \implies \alpha[f[\underline{\mathbb{R}}]] \subseteq \alpha[\underline{\mathbb{R}}]$. Since f is measurable $\forall x \in \alpha \circ f[\underline{\mathbb{R}}], f^{-1} \circ \alpha^{-1}(x) = f^{-1}[B] \in \mathcal{B}_{\underline{\mathbb{R}}}, \forall B \in \mathcal{B}_{\underline{\mathbb{R}}}$.

Finally, let β be a recombination of σ-simple functions then $\beta[\underline{\mathbb{R}}] = \{\beta(r) | r \in \underline{\mathbb{R}}\} = \bigcup_{n\in\mathbb{N}} \{\alpha_n(r) | r \in S_n\}$. Each α_n is σ-simple so $\alpha_n[\underline{\mathbb{R}}]$ must be countable therefore $\beta[\underline{\mathbb{R}}]$ is a countable union of countable sets, hence countable. Given $x \in \beta[\underline{\mathbb{R}}]$ we have $\beta^{-1}(x) = \bigcup_{n\in\mathbb{N}} \alpha_n^{-1}(x) \cap S_n$ which is a countable union of intersections between Borel sets since $S_n \in \mathcal{B}_{\underline{\mathbb{R}}}$ and α is σ-simple. Therefore, $\beta^{-1}(x) \in \mathcal{B}_{\underline{\mathbb{R}}}$. $\square$

3.2 Morphisms

An analogue of measurable functions, QBS morphisms allow maps between spaces which preserves the structure of these spaces. Since we have to concern ourselves with sets of random elements we do not follow measurable functions definition using pre-image geometric properties. Instead, we use an algebraic condition to say that a function is a QBS morphism if the function composed with any random element in the domain QBS is a random element of the image QBS.

Definition 3.2.1. *Let $(\underline{X}, M_X), (\underline{Y}, M_Y)$ be QBSs. Then, $f : \underline{X} \to \underline{Y}$ is a QBS morphism if $\forall \alpha \in M_X, f \circ \alpha \in M_Y$. We then say $f : X \to Y$.*

For the space $\mathbb{R} = (\underline{\mathbb{R}}, \text{Meas}(\mathbb{R}, \mathbb{R}))$, the only morphisms between $\mathbb{R}$ and $\mathbb{R}$ are the Borel measurable functions. Moreover, a function $f : \underline{X} \to \underline{\mathbb{R}}$ is a morphism only if $\forall \alpha \in M_X, f \circ \alpha$ is measurable for any QBS (X, M_X) by definition since the random elements of $\mathbb{R}$ are the measurable functions.

Proposition 3.2. *A function $f : \underline{\mathbb{R}} \to \underline{\mathbb{R}}$ is Borel measurable if and only if it is a QBS morphism $f : \mathbb{R} \to \mathbb{R}$.*

Proof. Let $f : \underline{\mathbb{R}} \to \underline{\mathbb{R}}$ be Borel measurable. Then $\forall\alpha \in \text{Meas}(\mathbb{R},\mathbb{R})$, $f \circ \alpha \in Meas(\mathbb{R},\mathbb{R})$ and by definiton f is a QBS morphism $\mathbb{R} \to \mathbb{R}$.

Let f be a QBS morphism. Then $\forall\alpha \in \text{Meas}(\mathbb{R},\mathbb{R})$, $f \circ \alpha \in \text{Meas}(\mathbb{R},\mathbb{R})$. This implies that $(f \circ \alpha)^{-1}(A) \in \mathcal{B}_{\mathbb{R}}, \forall A \in \mathcal{B}_{\mathbb{R}}$ which gives $\alpha^{-1}(f^{-1}(A)) \in \mathcal{B}_{\mathbb{R}}, \forall A \in \mathcal{B}$.

Since it is a morphism for all $\alpha \in \text{Meas}(\mathbb{R},\mathbb{R})$ let $\alpha(x) = x$ then $\alpha^{-1}(\mathcal{B}_{\mathbb{R}}) = \mathcal{B}_{\mathbb{R}}$. This gives $f^{-1}(A) \in \mathcal{B}_{\mathbb{R}}$ and thus f is Borel measurable as required. □

Furthermore, for the indiscrete and discrete spaces we find two other interesting results. Given any QBS A and a set $\underline{X}$ with X the indiscrete space over $\underline{X}$ then all functions of the form $f : \underline{A} \to \underline{X}$ are QBS morphisms of the form $f : A \to X$. This follows from the fact that the indiscrete QBS structure has all valid functions of the form $\alpha : \mathbb{R} \to X$ and therefore $f \circ \alpha$ will also be a valid function. Thus, it is very simple to map into an indiscrete QBS with any function.

In contrast, for discrete spaces the opposite is true, it is very easy to map out of a discrete space. Given any QBS A and a set $\underline{Y}$ with Y the discrete space over $\underline{Y}$ then all functions of the form $f : \underline{Y} \to \underline{A}$ are QBS morphisms of the form $f : Y \to A$. Since all σ-simple functions are apart of every set of random elements of a QBS and composing a σ-simple $\alpha : \mathbb{R} \to Y$ with a function $f : Y \to A$ will give a σ-simple function $f \circ \alpha : \mathbb{R} \to A$ which must be an element of M_A.

As with measurable functions, morphisms compose. All constant functions are morphisms as is the identity function. Finally, all σ-simple functions $\alpha : \mathbb{R} \to A$ for QBS A are morphisms.

Proposition 3.3. *Let A,B,C be QBSs. Then,*

(i) $\forall b \in \underline{B}$, $f : A \to B, f(a) = b$ *is a QBS morphism.*

(ii) $id : A \to A, id(a) = a$ *is a QBS morphism.*

(iii) If $f : B \to C$ *and* $g : A \to B$ *are QBS morphisms then* $f \circ g : A \to C$ *is a QBS morphism.*

(iv) Every σ*-simple function* $\alpha : \mathbb{R} \to A$ *is a QBS morphism.*

Proof. (i) Since all constant functions are in all sets of random elements, $f \circ \alpha = f \in M_B$.

(ii) For $\alpha \in M_A$, $id \circ \alpha = \alpha \in M_A$.

(iii) Since f,g are QBS morphisms, $\forall\beta \in M_A, g \circ \beta \in M_B$ and $\forall\alpha \in M_B, f \circ \alpha \in M_C$ then $f \circ g \circ \beta = f \circ \alpha \in M_C$

(iv) Given a measurable function $f : \mathbb{R} \to \mathbb{R}$ then $\alpha \circ f$ is also σ-simple and hence $\alpha \circ f \in M_A$, since all sets of random elements in any QBS contain all σ-simple functions. To show $\alpha \circ f$ is σ-simple we can show the image of $\underline{\mathbb{R}}$ countable and $\forall x \in \alpha \circ f[\underline{\mathbb{R}}], f^{-1} \circ \alpha^{-1}[x] \in \mathcal{B}_{\underline{\mathbb{R}}}$.

To show it is countable $f[\underline{\mathbb{R}}] \subseteq \underline{\mathbb{R}}$ and $\alpha[f[\underline{\mathbb{R}}]] \subseteq \alpha[\underline{\mathbb{R}}]$. Since α is σ-simple so $\alpha[\underline{\mathbb{R}}]$ is countable and so is its subset $\alpha[f[\underline{\mathbb{R}}]]$.

To show the inverse of any point in the image is a Borel set let $x \in \alpha \circ f$, $f^{-1} \circ \alpha^{-1}(x) = f^{-1}[B]$ for $B \in \mathcal{B}_{\underline{\mathbb{R}}}$ since α is σ-simple. Then, since f is measurable $f^{-1}[B] \in \mathcal{B}_{\underline{\mathbb{R}}}$.

$\square$

With measurable functions, we are now able to define measures. Unlike measure theoretic probability, a measure on a QBS is a tuple which has the idea of push-forward built into it. Since QBSs have random elements from $\mathbb{R}$ to the carrier set, these provide functions which we can push-forward the measure from the space onto $\mathbb{R}$.

Definition 3.2.2. *Let X,Y be QBSs and $f : X \to Y$ be a morphism. Then, a (probability) measure on X is a tuple (α, μ) for $\alpha \in M_X$ and μ a (probability) measure on $\mathbb{R}$. Then, $f \circ \alpha \in M_Y$ and thus $(f \circ \alpha, \mu)$ is a (probability) measure on Y. We say two measures, (α, μ) (β, ν), are equivalent if given a Borel set $A, \mu[\alpha^{-1}[A]] = \nu[\beta^{-1}[A]]$.*

3.3 Borel Subsets

As seen in Chapter 1, measure theory is built on measurable sets. Probability theory still requires sets of events that allow for consistent computations of likelihood. Thus, to achieve a notion of probability theory we need σ-algebras that are somehow related to a QBS. This is done by choosing exactly the sets that make all the random elements measurable. This is a key step which will allow for function spaces. Aumann showed that in general measurable subsets of function spaces do not exist [Aumann, 1961]. However, if we start with a set of functions and define a σ-algebra that makes them measurable then we can construct function spaces.

Definition 3.3.1. *A measurable or Borel subset in a QBS A is a subset $U \subseteq \underline{A}$ such that $\forall \alpha \in M_A, \alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}}$. We denote $\mathcal{B}_A$ as the set of all measurable subsets of A.*

Importantly, the set $\mathcal{B}_A$ is a σ-algebra.

Proof. We use the fact that $\mathcal{B}_{\underline{\mathbb{R}}}$ is a σ-algebra.

Given $U = \emptyset$, $\alpha^{-1}[\emptyset] = \emptyset \in \mathcal{B}_{\mathbb{R}}$.

If $U \in B_A$, then $\alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}}$ which implies $\alpha^{-1}[U]^c = \alpha^{-1}[U^c] \in \mathcal{B}_{\mathbb{R}}$ giving $U^c \in \mathcal{B}_A$.

Finally, $A_n \in \mathcal{B}_A : (n \in \mathbb{N})$, then $\alpha^{-1}[A_n] \in \mathcal{B}_{\mathbb{R}} : (n \in \mathbb{N})$ and by countable unions $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}_{\mathbb{R}}$ which implies $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}_A$. Thus, $\mathcal{B}_A$ is a σ-algebra. $\square$

Now every random element of A is $\langle \mathcal{B}_{\mathbb{R}}, \mathcal{B}_A \rangle$-measurable. We then call the resulting measurable space $\overset{\ulcorner \text{Meas} \urcorner}{A} = (A, \mathcal{B}_A)$ the free measurable space over A.

Example Consider the QBS $\mathbb{R}$. Then, $\mathcal{B}_{\mathbb{R}}$ derived from the QBS is the same as the Borel σ-algebra, $\mathcal{B}_{\underline{\mathbb{R}}}$. This is intuitive to see. The Borel sets of $\mathbb{R}$ are the sets for which all the random elements of $\mathbb{R}$, which are all the Borel measurable functions $\underline{\mathbb{R}}$ to $\underline{\mathbb{R}}$, are measurable. In other words, they are measurable for the Borel σ-algebra.

Example Let $\mathbb{2}$ be the discrete construction over $\{0,1\}$. Then, the Borel subsets are all the elements of the powerset of $\{0,1\}$ since the pre-image of σ-simple functions always maps into $\mathcal{B}_{\mathbb{R}}$ for all the elements in their image. Conversely, consider the indiscrete construction over $\{0,1\}$. Then, the only sets which are Borel measurable under the pre-image of all functions are the empty set and $\mathbb{R}$ so $\mathcal{B}_{\mathbb{2}} = \{\emptyset, \mathbb{R}\}$. This is main reason these constructions are called discrete and indiscrete; the Borel sets correspond to the discrete and indiscrete topologies.

We can now draw a connection between QBS morphisms and measurable functions.

Proposition 3.4. *([Heunen et al., 2018, Proposition 15]) Let (Y, Σ_Y) be a measurable space.*

(i) If (X, M_X) is a QBS then $f : X \to Y$ is a measurable function $(X, \Sigma_{M_X}) \to (Y, \Sigma_Y)$ if and only if it is a QBS morphism $(X, M_X) \to (Y, M_{\Sigma_Y})$.

(ii) If (X, Σ_X) is a standard Borel space, a function $f : X \to Y$ is a morphism $(X, M_{\Sigma_X}) \to (Y, M_{\Sigma_Y})$ if and only if it is a measurable function $(X, \Sigma_X) \to (Y, \Sigma_Y)$.

where $M_{\Sigma_A} = \text{Meas}(\mathbb{R}, A)$ given a QBS A.

This shows we can treat QBS morphisms that map into QBS of the form $(A, \text{Meas}(\mathbb{R}, A))$ as measurable functions when (A, Σ_A) is a measurable space. This allows us to switch between QBS and measurability with respect to certain measurable subsets.

A standard Borel space is a topological space that is either isomorphic to $\mathbb{R}$ or countable, discrete and non-empty. We will consider two important examples, $(\mathbb{2}, \mathcal{P}(\mathbb{2}))$ and $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Thus, a function from $\mathbb{R} \to \mathbb{2}$ is a QBS morphism between $\mathbb{R}$ and $(\mathbb{2}, \text{Meas}(\mathbb{R}, \mathbb{2}))$ if and only if it is $\langle \mathcal{B}_{\mathbb{R}}, \mathcal{P}(\mathbb{2}) \rangle$-measurable. For standard Borel spaces we have that $\Sigma_X = \Sigma_{M_{\Sigma_X}}$.

3.4 Borel Subspaces

Probability theory requires that we deal with subsets. For instance, probabilities can only be non-negative and sometimes we consider smaller examples of spaces for our experiments. Thus in a QBS we can start with a very large, broad QBS like $\mathbb{R}$ and given a subset define a new QBS with the subset as the carrier set and the random elements are those in $\text{Meas}(\mathbb{R}, \mathbb{R})$ which map into the subset. Formally, given a QBS A and a subset $\underline{X}$ of $\underline{A}$ we say the *subspace of A induced by $\underline{X}$* is $X = (\underline{X}, M_X)$ where $M_X = \{\alpha : \mathbb{R} \to X | \alpha \in M_A\}$

Example If we take $\mathbb{R}$ and we want to concern ourselves with $\underline{\mathbb{W}} = [0, \infty]$ we can take the subspace induced by $[0,\infty]$. This would have the set of all non-negative valued measurable functions as the set of random elements, $\mathrm{Meas}(\mathbb{R},\mathbb{W})$.

The function which maps from a subspace back into the original space is a QBS morphism. That is to say

$$\hookrightarrow: X \to A, \hookrightarrow x = x$$

is a QBS morphism since all random elements in the subspace are already random elements in the original space but with the subset as the co-domain. Thus, given $\alpha \in M_X$ we have that $\hookrightarrow \circ\alpha = \alpha \in M_A$.

3.5 Products and Co-Products

Given two QBSs we can always construct their product space, which is also a QBS. These are important as they lead up to the development of measurable evaluation using QBS which measure theoretic probability theory cannot in general always achieve. The random elements in a product of QBSs are tuples of the random elements in the spaces making up the product.

Definition 3.5.1. *Let $(\underline{X},M_X),(\underline{Y},M_Y)$ be QBSs. Then, $(\underline{X}\times\underline{Y},M_{X\times Y})$ is the product space with $M_{X\times Y} = \{\alpha : \mathbb{R} \to X \times Y | \pi_1 \circ \alpha \in M_X, \pi_2 \circ \alpha \in M_Y\} = \{(f,g) : \mathbb{R} \to X \times Y | f \in M_X, g \in M_Y\}$, where $\pi_1 \circ (\alpha_1,\alpha_2) = \alpha_1$ and $\pi_2 \circ (\alpha_1,\alpha_2) = \alpha_2$.*

To see why the product space satisfies the QBS axioms we first see that all the constant functions are elements of $M_{X\times Y}$. Let $\alpha : \mathbb{R} \to X \times Y, \alpha(r) = (x,y)$, then since the functions $\underline{\mathbf{x}}(r) = x \in M_X, \underline{\mathbf{y}}(r) = y \in M_Y$ their tuple must also be in $M_{X\times Y}$.

If we pre-compose with a measurable f for any tuple $(\alpha_1,\alpha_2) \in M_{X\times y}$ then $(\alpha_1,\alpha_2)\circ f = (\alpha_1 \circ f,\alpha_2 \circ f) \in M_{X\times Y}$ since $\alpha_1 \circ f \in M_X$ and $\alpha_2 \circ f \in M_Y$.

Finally, given $\mathbb{R} = \bigcup_{n\in\mathbb{N}} S_i$ for Borel S_i, then $\beta(r) = \alpha_i(r)$ for $r \in S_i$ then $\beta(r) = (\alpha_{i,1}(r),\alpha_{i,2}(r))$. Then, $\pi_1 \circ \beta(r) = \alpha_{i,1}(r) \in M_X$. Similarly, $\pi_2 \circ \beta(r) = \alpha_{i,2}(r) \in M_Y$. Thus $\beta(r) \in M_{X\times Y}$.

The projection functions $\pi_i : X_1 \times X_2 \to X_i$ are QBS morphisms for all i.

Proposition 3.5. *$\forall i, \pi_i : X_1 \times X_2 \to X_i$ are QBS morphisms.*

Proof. Given $f \in M_{X_1\times X_2}$ we have $\pi_i \circ f = \pi_i \circ (f_1, f_2) = f_i \in M_{X_i}$ by definition. Thus, π_i is a QBS morphism. □

Example: The product space of the reals, $\mathbb{R}\times\mathbb{R}$, has the set of random elements $\{(f,g) : \mathbb{R} \to \mathbb{R}\times\mathbb{R} | f,g \in \mathrm{Meas}(\mathbb{R},\mathbb{R})\}$.

Example: If we take the product space of the reals and the discrete space X of any set $\underline{X}$ we have $\mathbb{R}\times X$ with random elements as tuples of measurable functions from $\mathbb{R}$ to $\mathbb{R}$ and σ-simple functions from $\mathbb{R}$ to X. For example, let $f_1(r) = e^r$ and $f_2(r) = x_i$ if $r \in S_i$ for $\mathbb{R} = S_1 \cup S_2 = [0,1]\cup[0,1]^c$ and constants $x_1,x_2 \in \underline{X}$. Then,

$$\lambda r.\left(e^r, \begin{cases} x_1 & r \in [0,1] \\ x_2 & r \in (-\infty,0)\cup(1,\infty) \end{cases}\right)$$

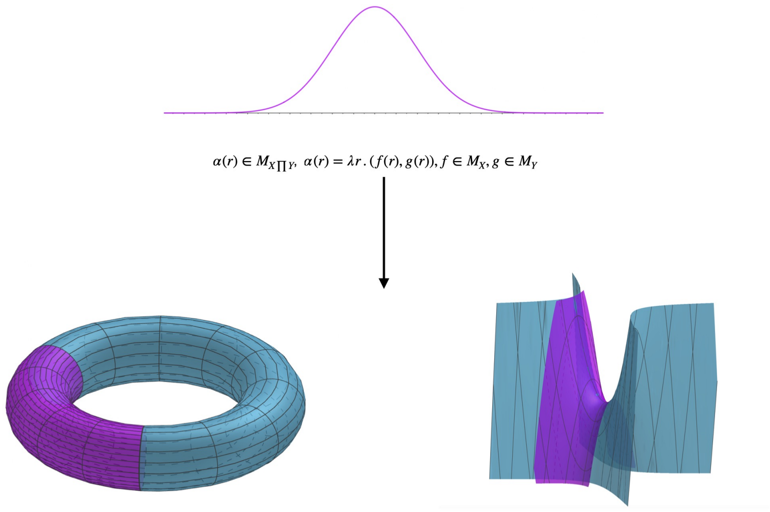

Figure 3.4: Diagram of an example random element in a product QBS.

is a random element.

Co-products can also always be constructed. In contrast to product spaces, the random elements in co-product spaces divide up $\mathbb{R}$ where a division maps to one of the co-products instead of all of $\mathbb{R}$ being mapped to both spaces. For example, if we take the co-product of A, B then a random element in the co-product space would decide for each $r \in \mathbb{R}$ which space we are mapping into.

Definition 3.5.2. *Let* $(\underline{X_1}, M_{X_1}), (\underline{X_2}, M_{X_2})$ *be QBSs. Then,* $(\underline{X_1} \coprod \underline{X_2}, M_{X_1 \coprod X_2})$ *is the co-product space with* $(\underline{X_1} \coprod \underline{X_2}) = (\{1\} \times X_1) \cup (\{2\} \times X_2)$ *and*

$$M_{X_1 \coprod X_2} = \{\lambda r.(f(r), \alpha_{f(r)}(r)) | f \in \text{Meas}(\mathbb{R}, \{1,2\}), (\alpha_i \in M_{X_i})_{i \in \text{Image}(f)}\}.$$

Unpacking this definition, the carrier set is the union of the two sets $\{1\} \times X_1$ and $\{2\} \times X_2$. Meaning, each element of $X_1 \coprod X_2$ is (i, x_i) for $i \in \{1,2\}$ and $x_i \in X_i$. The random elements then are tuples of functions. The first function is a measurable map into the indexing set $\{1,2\}$ and the second function is a random element into the space chosen by the first function. So given $r \in \mathbb{R}$, $f(r)$ decides which space we are mapping into. Then, $\alpha_{f(r)}$ gives us the random element to map into $X_{f(r)}$. Thus, the set of random elements in the co-product space can be very large.

We can check that the coproduct construction satisfies the QBS axioms. Firstly, constant functions of the form $\lambda r.(f(r), \alpha_{f(r)}(r)) = x$ for $x \in X_i$. We have that $f(r) = i$ is measurable since constant functions are measurable. Then, $\alpha_i(r) = x$ is a constant random element in M_{X_i}.

Pre-composing with a measurable function $g : \mathbb{R} \to \mathbb{R}$ gives $\lambda r.(f(g(r)), \alpha_{f(g(r))}(g(r)))$.

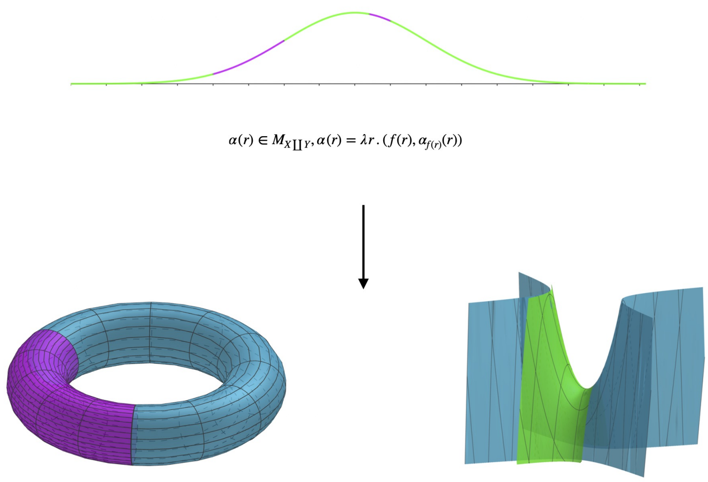

Figure 3.5: Diagram of example random element in a co-product QBS..

Since $f(g(r)) : \mathbb{R} \to \{1,2\}$ is still measurable and α is a random element, we have pre-composition with g is a random element.

Finally, recombination. Let $\mathbb{R} = \bigcup_{n\in\mathbb{N}} S_n$ and $\lambda r.(f(r), \alpha_{f(r)}(r))_n$ be co-product random elements. Then, $\beta(r) = (f(r), \alpha_{f(r)}(r))_n$ when $r \in S_n$ is equal to $\lambda r.(f_n(r), \alpha_{n,f(r)}(r))$ when $r \in S_n$. Since f and α are random elements they are closed under recombination so $f_n(r)$ when $r \in S_n$ and $\alpha_n(r)$ when $r \in S_n$ are both random elements as required.

Example: Take the co-product of $\mathbb{R}$ and the discrete space X for a set $\underline{X}$. Then the set of random elements are

$$\{\lambda r.(f(r), \alpha_{f(r)}(r)) | f \in \mathrm{Meas}(\mathbb{R}, \{1,2\}), \alpha_1 \in \mathrm{Meas}(\mathbb{R}, \mathbb{R}), \alpha_2 \text{ a } \sigma\text{-simple function}\}.$$

For example, let $f(r) = \mathbb{1}_{\mathbb{Q}} + 1$ and $\alpha_1(r) = r^2$ and $\alpha_2(r) = x_2$ for a constant $x_2 \in X$. Thus, at every rational number we map to a constant value in X but everywhere else we are sending r to r^2.

3.6 Function Spaces

In contrast to measure theory, the function space construction always exists. That is to say given any two QBSs we can construct a QBS whose carrier set is the QBS morphisms between the two QBSs. These spaces are particularly useful as they give us structures which provide an evaluation function as a QBS morphism. Let $\mathrm{QBS}(A,B)$ be the set of all QBS morphisms between QBSs A and B.

The definition of a function space QBS requires a set of random elements that map from the reals to the set of QBS morphisms. To achieve this we use the currying operation. To curry a function of the form $X \times Y \to Z$ is to take the function from multiple arguments to one argument. More generally, we can convert evaluating a function of multiple arguments to evaluating multiple functions of one argument. For example, given a function $F(x,y) : X \times Y \to Z$, if we curry F we get $\mathrm{curry}(F) = f : X \to Z^Y$ such that $f(x)(y) = F(x,y)$. Therefore, currying a function returns another function of one argument which maps to a function of one argument. Uncurry is the opposite of curry which takes a function of the form $X \to Z^Y$ and returns a function of two arguments, $X \times Y \to Z$.

Example Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, F(x,y) = x^2 + y^2 - 1$. Then, $\mathrm{curry}(F) = f$ where $f : \mathbb{R} \to \mathbb{R}^{\mathbb{R}}, f(x) = \lambda y.x^2 + y^2 - 1$. For example, $f(4) = \lambda y.16 + y^2 - 1$. Then, we take the function returned by f and plug in a value for y, $f(4)(2) = 16 + (2)^2 - 1 = 16 + 4 - 1 = 19$. Thus, $F(4,2) = f(4)(2)$.

We can now define a function space using this operation.

Definition 3.6.1. *Let A, B be QBSs. Then, B^A is their function space.* $\underline{B^A} = \mathrm{QBS}(A,B) = \{f : A \to B | f \circ \alpha \in M_B, \forall \alpha \in M_A\}$ *and* $M_{B^A} = \{\alpha | \mathrm{uncrurry}(\alpha) \in \mathrm{QBS}(\mathbb{R} \times A, B)\} = \mathrm{curry}[\mathrm{QBS}(\mathbb{R} \times A, B)]$

Thus, the QBS function space of two QBSs A and B has random elements such that when uncurried are QBS morphisms from the product space $\mathbb{R} \times A$ to B. Alternatively, the random elements are the curried QBS morphisms from $\mathbb{R} \times A$ to B. For example given a morphism g from $\mathbb{R} \times A$ to B then by currying g we get a random element from $\mathbb{R}$ to B^A:

$$\mathbb{R} \times A \xrightarrow{g} B \implies \mathbb{R} \xrightarrow{\mathrm{curry}(g)} B^A$$

Example: Examining $\mathbb{R}^{\mathbb{R}}$, this is the QBS with a carrier set as the QBS morphisms from $\mathbb{R}$ to $\mathbb{R}$ which we know by Proposition 3.2 are the Borel measurable functions. Then, the random elements are the functions $\alpha : \mathbb{R} \to \mathbb{R}^{\mathbb{R}}$ such that $\alpha : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are QBS morphisms. These are the curried functions of $\mathrm{Meas}(\mathbb{R} \times \mathbb{R}, \mathbb{R})$.

Example: Take the indiscrete QBS of a set $\underline{A}$. Then $A^{\mathbb{R}}$ is a QBS with the carrier set consisting of QBS morphisms from $\mathbb{R} \to A$ which are all functions from $\mathbb{R}$ to A. Then, the random elements are the curried QBS morphisms from $\mathbb{R} \times \mathbb{R}$ to A which again is every function $\mathbb{R} \times \mathbb{R}$ to A.

To show that the function space is a QBS we again check each condition.

Constants: Let $\alpha : \mathbb{R} \to B^A$ be constant. Then, $\alpha(r) = \lambda a.F(a) \in \mathrm{QBS}(A,B)$. Then,

$$\mathrm{uncurry}(\alpha) = \lambda r.\lambda a.\alpha(r)(a) = \beta(r,a).$$

So we have that $\beta(r,a) = F(a)$. Let $\lambda r.(-, g(r)) \in M_{\mathbb{R} \times A}$ then

$$\lambda r.\beta \circ (-, g)(r) = F(g(r)) \in M_B$$

since $F \in \mathrm{QBS}(A,B)$ and $g \in M_A$.

Pre-composition: To show pre-composing a random element with a measurable function is a random element, we have $f \in \mathrm{QBS}(\mathbb{R},\mathbb{R})$ and $\mathrm{uncurry}(\alpha) \in \mathrm{QBS}(\mathbb{R} \times A, B)$.

Thus,

$$\begin{aligned}\text{uncurry}(\alpha\circ f) &= \lambda r.\lambda a.(\alpha\circ f)(r)(a)\\ &= \lambda r.\lambda a.\alpha(f(r))(a) = \beta(f(r),a).\end{aligned}$$

So we have that $\beta\in\text{QBS}(\mathbb{R}\times A,B)$ and $(f,a)\in\text{QBS}(\mathbb{R},\mathbb{R}\times A)$. Therefore, $\beta\circ(f,a)\in\text{QBS}(\mathbb{R},B)$. Since it is a QBS morphism, we have that $\forall h\in\text{Meas}(\mathbb{R},\mathbb{R}),\beta\circ(f,a)\circ h\in M_B$ thus let h be the identity and hence $\beta\circ(f,a)\in M_B$.

Recombination Let $\mathbb{R}=\bigcup_{i\in\mathbb{N}}S_i$ for $S_i\in\mathcal{B}_{\mathbb{R}}$. For the random elements $F_i\in M_{B^A}$ let uncurry$(F_i)=f_i$ be QBS morphisms from $\mathbb{R}\times A$ to B. We need to show that for $\beta(r)=F_i(r)$ when $r\in S_i$, uncurry(A) is a QBS morphism from $\mathbb{R}\times A$ to B.

Take (ϕ,α) as a random element of $\mathbb{R}\times A$, where ϕ is a Borel measurable function and α is a random element of A. Then we can show uncurry(β) is a QBS morphism by showing uncurry$(\beta)\circ(\phi,\alpha)$ is a random element of B.

Define a new partition of $\mathbb{R}$ such that $\mathbb{R}=\bigcup_{i\in\mathbb{N}}T_i$ where $T_i=\phi^{-1}[S_i]$. Then, let $\beta_i=f_i\circ(\phi,\alpha)$ which is a random element of B since f_i is a QBS morphism. Then by the recombination axiom for QBS B we have that $G(r)=\beta_i(r)$ when $r\in T_i$ is a random element of B. We can now show that $G(r)=\text{uncurry}(\beta)\circ(\phi,\alpha)(r)$.

Fix $r\in\mathbb{R}$ then there is exactly one $i\in\mathbb{N}$ such that $r\in T_i$. Without loss of generality let $r\in T_0$. Then we have

$$G(r)=\beta_0(r)=f_0\circ(\phi,\alpha)(r)=f_0\circ(\phi(r),\alpha(r)).$$

Given we have that $\phi(r)\in S_0$ then

$$\text{uncurry}(\beta)\circ(\phi,\alpha)(r)=F_0\circ(\phi(r))(\alpha(r))=f_0\circ(\phi(r),\alpha(r)),$$

by definition of uncurry. Thus, $G(r)=\text{uncurry}(\beta)\circ(\phi,\alpha)(r)$ so uncurry(β) is a QBS morphism thus β is a random element of B^A.

Using these constructions we can show that the function eval : $B^A\times A\to B$ is a QBS morphism. This is the contrast with measure theory that motivates QBS theory. Aumann showed it was not possible to have spaces that make evaluation measurable but with QBS theory we have spaces which make evaluation a morphism [Aumann, 1961].

Proposition 3.6. *Let* eval : $B^A\times A\to A,\text{eval}(f,a)=f(a)$ *then* eval *is a QBS morphism.*

Proof. Given a random element $F:\mathbb{R}\to B^A$ and a random element $\alpha\in M_A$ we need to show that eval$\circ(F,\alpha)$ is a random element of B.

To prove this we show that eval$\circ(F,\alpha)=f\circ(id,\alpha)$ for uncurry$(F)=f$. Since we have that f is a QBS morphism from $\mathbb{R}\times A$ to B by definition, id is Borel measurable and α is a random element we have that $f\circ(id,\alpha)$ is a random element of B.

Fix $r\in\mathbb{R}$. Then,

$$\text{eval}\circ(F,\alpha)(r)=\text{eval}\circ(F(r),\alpha(r))=F(r)(\alpha(r))=f(r,\alpha(r))$$

by definition of uncurry. Then for $f \circ (id, \alpha)$ we have that

$$f \circ (id, \alpha)(r) = f(r, \alpha(r)).$$

Therefore, $\text{eval} \circ (F, \alpha) = f \circ (id, \alpha)$ and $f \circ (id, \alpha)$ is a random element of B so $\text{eval} \circ (F, \alpha)$ is a random element of B and hence eval is a QBS morphism. □

We can equip the Borel subsets of a QBS A with a set of random elements since each Borel set directly corresponds to an indicator QBS morphism.

Proposition 3.7. *A set U is a Borel subset of $\underline{A}$ if and only if the characteristic function $[- \in U] : A \to \mathbb{2}$ is a QBS morphism to the discrete construction over the 2 element set.*

Proof. Recall if $U \subseteq \underline{A}$ is measurable then $\alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}}, \forall \alpha \in M_A$.

We now need to show that $[- \in U]$ is a QBS morphism which is to show that $[- \in U] \circ \alpha, \forall \alpha \in M_A$ is σ-simple.

First, let $\alpha[\mathbb{R}] = B \subseteq A$ and since α is σ-simple we have that B is countable. Then we have $[- \in U] \circ \alpha[\mathbb{R}] = [B \in U]$ which is either $\{0\}$ or $\{1\}$ and hence countable.

Then, we need to show $\forall x \in \mathbb{2}$,

$$([- \in U] \circ \alpha)^{-1}(x) = \alpha^{-1} \circ [- \in U]^{-1}(x) \in \mathcal{B}_{\mathbb{R}}.$$

If $x = 1$ then

$$\alpha^{-1} \circ [- \in U]^{-1}(1) = \alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}}.$$

If $x = 0$ then

$$\alpha^{-1} \circ [- \in U]^{-1}(0) = \alpha^{-1}[U^c] \in \mathcal{B}_{\mathbb{R}}$$

since $\mathcal{B}_A$ is a σ-algebra. Therefore, $[- \in U]$ is a QBS morphism.

Now we assume the characteristic function is a QBS morphism then $\forall \alpha \in M_A, [- \in U] \circ \alpha$ is a σ-simple function. So, $\alpha^{-1} \circ [- \in U]^{-1}[1] \in \mathcal{B}_{\mathbb{R}} \implies \alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}}$. Thus, U is measurable. □

We now can define the QBS given the Borel sets of a QBS A. These spaces are directly associated with the function space $\mathbb{2}^A$ since we can show a bijection between their random elements.

Definition 3.6.2. *Let X be a QBS. Then let $\mathcal{B}_X = (\underline{\mathcal{B}_X}, M_{\mathcal{B}_X})$ where $\mathcal{U} \in M_{\mathcal{B}_X}$ when $\forall \alpha \in M_X$ the set $\{(r,s) \in \mathbb{R} \times \mathbb{R} | \alpha(s) \in \mathcal{U}(r)\} \in \mathcal{B}_{\mathbb{R} \times \mathbb{R}}$.*

Proposition 3.8. *The map $\mathcal{U} \mapsto \lambda r.[- \in \mathcal{U}(r)]$ is a bijection.*

Proof. We begin by showing that given $\mathcal{U} \in M_{\mathcal{B}_X}$ that $\lambda r.[- \in \mathcal{U}(r)] \in M_{\mathbb{2}^{\mathbb{R}}}$.

Map $\mathcal{U}$ to $\lambda r.[- \in \mathcal{U}(r)]$. Then, uncurry($\lambda r.[- \in \mathcal{U}(r)]$) gives a characteristic function of $\mathcal{U}(r)$ from $\mathbb{R} \times A$ to $\mathbb{2}$. Since $\mathcal{U}(r) \in \underline{\mathcal{B}_A}$ it is measurable and by Proposition 3.7 a characteristic function of a measurable set is a QBS morphism. So we have

$\mathrm{uncurry}(\lambda r.[- \in \mathcal{U}(r)]) \in \mathrm{QBS}(\mathbb{R} \times A, \mathbb{2})$ and by definition of function spaces this implies that $\lambda r.[- \in \mathcal{U}(r)] \in M_{\mathbb{2}^A}$.

Now given $\lambda r.[- \in \mathcal{U}(r)] \in M_{\mathbb{2}^A}$ map it to $\mathcal{U}$. Then, $\mathcal{U} \in M_{\mathcal{B}^A}$ if and only if $\forall \alpha \in M_A$ we have that $\{(r,s) | \alpha(s) \in \mathcal{U}(r)\} \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$. We have that

$$\{(r,s) | \alpha(s) \in \mathcal{U}(r)\} = \{(r,s) | [\alpha(s) \in \mathcal{U}(r)] = 1\}.$$

Let

$$f(r,s) = [\alpha(s) \in \mathcal{U}(r)] = \mathrm{eval}([- \in \mathcal{U}(r)], \alpha(s)),$$

for a fixed $\alpha \in M_A$. Then $f(r,s) \in \mathrm{QBS}(\mathbb{R} \times \mathbb{R}, 2)$ since the characteristic function of $\mathcal{U}(r)$ is a QBS morphism, as is the evaluation map and α. Since $\mathbb{R} \times \mathbb{R}$ and $\mathbb{2}$ are standard Borel space we have $f \in \mathrm{Meas}(\mathbb{R} \times \mathbb{R}, 2)$ by Proposition 3.5. Therefore, we have $\forall \alpha \in M_X$,

$$\{(r,s) | \alpha(s) \in \mathcal{U}(r)\} = f^{-1}[1] \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}.$$

$\square$

Therefore, $\mathcal{B}_A$ is a QBS for any A since $\mathbb{2}^A$ is a QBS, where the random elements of $\mathcal{B}_A$ are mapped from the random elements of $\mathbb{2}^A$ using the bijection from Proposition 3.9.

In summary, we have developed the foundational theory of QBS up to function spaces which allows us to now study integration theory for QBSs.

Chapter 4

Integration

Now that we have the requisite theory we will define an integral over QBSs. Integration theory is the basis of many calculations in probability theory, including expectation which is the cornerstone of many statistics. Measure theoretic probability has the Lebesgue integral which is built up from simple functions. In our approach we will define an approximation function which given any QBS morphism returns a sequence of approximations where each new term increases in complexity. This will allow us to define an integral using the limit of this approximation.

4.1 Spaces

We will be defining our integral with respect to an arbitrary QBS $X = (\underline{X}, M_X)$. The integral is a higher order function mapping from morphisms and probability measures to $[0,1]$. Thus we need a function space of QBS morphisms from X to $[0,\infty]$ and a QBS of probability measures.

Our function space is given by $\mathbb{W}^X$ where $\mathbb{W}$ is the subspace of $\mathbb{R}$ induced by the set $[0,\infty]$. Then $\mathbb{W}^X = (\underline{W^X}, M_{\mathbb{W}^X})$ where $\underline{\mathbb{W}^X}$ is the set of all QBS morphisms from X to $\mathbb{W}$ and

$$M_{\mathbb{W}^X} = \{\alpha \in \text{curry}[\mathbb{W}^{(\mathbb{R}\times X)}] | \text{uncurry}(\alpha) : \mathbb{R} \times X \to \mathbb{W} \text{ is a QBS morphism}\}.$$

We now need to define the space of probability measures. Firstly, we define the measurable space of probability measures.

Definition 4.1.1. *Let (X, Σ_X) be a measurable space. Let $G(X)$ be the set of probability measures on X. Then equip $G(X)$ with $\Sigma_{G(X)}$ as the σ-algebra such that all evaluation maps* $\text{eval}(\mu, B) = \mu[B]$ *are measurable for all $B \in \Sigma_X$.*

Using this we define the space of probability measures of a QBS X, recalling that a probability measure on a QBS X is a tuple of a random element and a probability measure on $\mathbb{R}$.

Let $PX = (\underline{PX}, M_{PX})$ where

$$\underline{PX} = \{(\alpha, \mu) \text{ probability measure on } X\}/\sim .$$

The carrier set of PX is the set of all probability measures up to equivalence. Then let

$$M_{PX} = \{\beta : \mathbb{R} \to PX | \exists \alpha \in M_X, \exists g \in \text{Meas}(\mathbb{R}, G(\mathbb{R})), \forall r \in \mathbb{R}, \beta(r) = [\alpha, g(r)]\},$$

where $G(\mathbb{R})$ is the set of all probability measures on $\mathbb{R}$. The random elements of PX are tuples of functions. The first function is a random element in X. The second function is a measurable map to the probability measures on $\mathbb{R}$. For example, if X is $\mathbb{R}$ then an example of $M_{P\mathbb{R}}$ could be $[x^2, \mathcal{N}]$, which is the case of taking the random element as x^2 and the measurable function to the probability measures is a constant mapping to the normal distribution measure.

4.2 Approximation

We will define an approximation function which takes any QBS morphism in $\underline{\mathbb{W}^X}$ and returns a sequence of Borel sets and weights to represent a simple function. A simple function is one that can be represented as a finite linear combination of indicator functions assigning one value to each disjoint Borel subset over a space. Since the subsets are disjoint we can easily calculate the measure as the sum of the measures of each Borel set.

Definition 4.2.1. *A function f is simple if f can be written as a finite linear combination of indicator functions*

$$f = \sum_{k=1}^{m} a_k \mathbb{I}_{A_k},$$

where $a_k \in \mathbb{W}$ and $A_k \in \mathcal{B}_X$.

Our approximation function will be used to map morphisms to a sequence of simple functions which converge to the morphism. Then given the morphism can be approximated by a simple function we can easily compute the integral as the limit of the integral of each simple function.

Definition 4.2.2. *Let $Approx : \mathbb{W}^X \to (\coprod(\mathcal{B}_X^n, \mathbb{W}^n))^{\mathbb{N}}$,*

$$Approx(f) = \lambda k. \left(\coprod_{i=0}^{k2^k-1} \left(f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k}), \frac{i}{2^k}\right) \right)$$

Let $[\![Approx\ f\ k]\!] : \mathbb{R} \to \mathbb{R}$,

$$[\![Approx\ f\ k]\!](x) = \sum_{i=0}^{k2^k-1} \frac{i}{2^k} \mathbb{I}_{f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k})}(x)$$

The $Approx$ function takes as an input another function, f. It then gives a function in return which maps to different data representations of simple functions, each with an increasing amount of Borel sets. Each Borel set A_k is the pre-image of an increasingly small interval in $\mathbb{W}$ and is assigned the smallest value that f would map any element $x \in A_k$ to. The function $[\![Approx]\!]$ then returns the actual simple function generated by the $Approx$ data representation.

We now claim that the limit of $[\![Approx\ f\ k]\!]$ as $k \to \infty$ is equal to f.

Proposition 4.1. $\lim_{k\to\infty}[\![Approx\ f\ k]\!](x) = f(x)$

Proof. Fix $x \in X$, then x is in exactly one Borel set. Let $x \in f^{-1}[\frac{j}{2^k}, \frac{j+1}{2^k})$. Then,

$$[\![Approx\ f\ k]\!](x) = \frac{j}{2^k},$$

and

$$\frac{j}{2^k} \leq f(x) < \frac{j+1}{2^k} \implies 0 \leq f(x) - [\![Approx\ f\ k]\!](x) < \frac{1}{2^k}.$$

Then as $k \to \infty$, $\frac{1}{2^k} \to 0$ we $f(x) - \lim_{k\to\infty}[\![Approx\ f\ k]\!](x) = 0$ □

Using our approximation function by simple functions of any morphism between X and $\mathbb{W}$ we can define the integral since given a morphism in $\mathrm{QBS}(X,\mathbb{W})$ and a random element $\alpha \in M_X$, their composition is a measurable function. Thus their inverse over intervals of $\underline{\mathbb{W}}$ are elements of the Borel σ-algebra. Furthermore, since we have approximated with simple functions with disjoint Borel sets, probability measures over the union of all the disjoint Borel sets can be approximated as a sum of probability measures over the disjoint Borel sets.

Definition 4.2.3. *Let* $\int : \mathbb{W}^X \times PX \to \mathbb{R}$,

$$\int f d[\alpha,\mu] = \int \lim_{k\to\infty}[\![Approx\ f\ k]\!] \circ \alpha d\mu = \lim_{k\to\infty} \sum_{i=0}^{k2^k-1} \frac{i}{2^k}\mu[\alpha^{-1}[f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k})]]$$

To show this is a QBS morphism we need to show that composition with the random elements of the product space is measurable.

We will need several lemmas to prove this. Firstly, a key part relies on the fact that a function from $\underline{\mathbb{R}}$ to $\underline{X}$ is a random element of QBS X if and only if that function is a QBS morphism from $\mathbb{R}$ to X. We will also show that given a QBS morphism between any two QBSs A and B the inverse of the morphism is a QBS morphism between the Borel sets. We then use these two lemmas to prove that the inverse of the composition of the random elements $\alpha \in M_X$ and the morphisms in $\mathrm{QBS}(X,\mathbb{W})$ are QBS morphisms of a specific form which is a measurable function. Then, we end with an expression which is a composition of measurable functions.

Lemma 4.1. *A function* $\alpha : \underline{\mathbb{R}} \to \underline{X}$ *is a random element if and only if* $\alpha : \mathbb{R} \to X$ *is a QBS morphism*

Proof. If $\alpha \in M_X$ for QBS X then $\forall f \in \mathrm{Meas}(\mathbb{R},\mathbb{R}), \alpha \circ f \in M_X$ by axioms of QBS. Then, since $M_{\mathbb{R}} = \mathrm{Meas}(\mathbb{R},\mathbb{R})$ we have that α is a QBS morphism (since composition with any random element in $\mathbb{R}$ is a random element in M_X).

If α is a QBS morphism then $\forall f \in \mathrm{Meas}(\mathbb{R},\mathbb{R}), \alpha \circ f \in M_X$. Since the identity function $f(x) = \mathrm{id}(x) = x$ is measurable then $\alpha \circ \mathrm{id} \in M_X \implies \alpha \in M_X$. □

Lemma 4.2. *Let $\alpha : A \to B$ be a QBS morphism. Then, $\alpha^{-1} : \mathbb{2}^B \to \mathbb{2}^A$ is a QBS morphism.*

Proof. First we show that $\alpha^{-1}[X]$ is Borel in A for a Borel set $X \in \mathcal{B}_B$. Let X be a Borel set, then recall that $\beta^{-1}[X] \in \mathcal{B}_{\mathbb{R}}, \forall \beta \in M_B$.

Given, α is a QBS morphism we have $\alpha^{-1}[X]$ is Borel if $f^{-1} \circ \alpha^{-1}[X] = (\alpha \circ f)^{-1}[X] \in \mathcal{B}_{\mathbb{R}}, \forall f \in M_A$. Since α is a QBS morphism then $(\alpha \circ f) \in M_B$ and thus since X is Borel in B we have $\alpha^{-1}[X]$ is Borel in A.

A function f is a random element of $M_{\mathbb{2}^A}$ if and only if $\text{uncurry}(f) \in \text{QBS}(\mathbb{R} \times A, \mathbb{2})$. Thus, uncurry is a characteristic function of $U \in \mathcal{B}_{\mathbb{R}\times A}$, $\text{uncurry}(f)(r,a) = [(r,a) \in U]$. This is equivalent to checking if $[a \in U_r]$ where $U_r = \{a|(r,a) \in U\}$. Thus, $f(r) = \lambda a.[a \in U_r]$ for a $U \in \mathcal{B}_{\mathbb{R}\times A}$. By our bijection in Proposition 3.8, we have that $\lambda r.U_r$ are the random elements of $M_{\mathcal{B}_A}$. Similarly for a random element g of $M_{\mathbb{2}^B}$, $g(r) = \lambda b.[b \in V_r]$ for $V \in \mathcal{B}_{\mathbb{R}\times B}$. Then by our bijection $\lambda r.V_r$ are random elements of $M_{\mathcal{B}_B}$.

We have the composition of α^{-1} with a random element of $M_{\mathcal{B}_B}$ is $\lambda r.\alpha^{-1}[V_r] \in \mathcal{B}_A$.

Thus by Proposition 3.7, $\lambda r.[- \in \alpha^{-1}[V_r]] : \underline{\mathbb{R}} \to \underline{\mathbb{2}^A}$. We have that if $a \in \alpha^{-1}[V_r]$ then $\alpha(a) \in V_r$ since $\alpha^{-1}[V_r] = \{a|\alpha(a) \in V_r\}$.

Thus, we can show $[- \in \alpha^{-1}[V_r]] = [- \in V_r] \circ \alpha$ is a random element. To show this, we uncurry giving $[- \in V] \circ (r,\alpha)$. We know that $[- \in V] \in \text{QBS}(\mathbb{R} \times B, \mathbb{2})$ and $\alpha \in \text{QBS}(A,B)$. So $(r,\alpha) \in \text{QBS}(\mathbb{R} \times A, \mathbb{R} \times B)$ since given (f,g) as a random element of $\mathbb{R} \times A$ we have their composition as $(r \circ f, \alpha \circ g)$. Then $\alpha \circ g \in M_B$ and $r \circ f \in \text{Meas}(\mathbb{R},\mathbb{R})$.

Therefore, $[- \in V] \circ (r,\alpha) \in \text{QBS}(\mathbb{R} \times A, \mathbb{2})$ which by definition gives $[- \in V_r] \circ \alpha$ is a random element of $M_{\mathbb{2}^A}$. □

Proposition 4.2. *$\int$ is a QBS morphism*

Proof. We need to show that $\int \circ (\beta_1, \beta_2)$ is measurable, for $\beta_1 \in M_{\mathbb{W}^X}$ and $\beta_2 = [\alpha, g]$ for $\alpha \in M_X, g \in \text{Meas}(\mathbb{R}, G(\mathbb{R}))$. However, we know that limits, sums and scalar multiplication of measurable functions is measurable. So we need to show that

$$F : \mathbb{W}^X \times PX \to [0,1], F(f,[\alpha,\mu]) = \mu[\alpha^{-1}[f^{-1}[a,b)]],$$

is a QBS morphism for all $a,b \in \mathbb{W}$.

We know that $\alpha \in M_X$ is a QBS morphism from $\mathbb{R}$ to X by Lemma 4.1. We also have that $f \in \underline{\mathbb{W}^X}$ and therefore $f \in \text{QBS}(X,\mathbb{W})$. Hence, by Proposition 3.3 (iii) $f \circ \alpha \in \text{QBS}(\mathbb{R},\mathbb{W})$ and by Lemma 4.2 $(f \circ \alpha)^{-1} = \alpha^{-1} \circ f^{-1} \in \text{QBS}(\mathbb{2}^{\mathbb{W}}, \mathbb{2}^{\mathbb{R}})$. Then, given a random element $h \in M_{\mathbb{2}^{\mathbb{W}}}$ we have that $h \in \text{QBS}(\mathbb{R}, \mathbb{2}^{\mathbb{W}})$. Therefore, $\alpha^{-1} \circ f^{-1} \circ h \in \text{QBS}(\mathbb{R}, \mathbb{2}^{\mathbb{R}})$.

We have that a function is in $\text{QBS}(\mathbb{R}, \mathbb{2}^{\mathbb{R}})$ if and only if is a random element of $\mathbb{2}^{\mathbb{R}}$. We know that the random elements of $\mathbb{2}^{\mathbb{R}}$ are the result of currying QBS morphisms in $\text{QBS}(\mathbb{R} \times \mathbb{R}, \mathbb{2})$ which are all the characteristic functions of the Borel sets in $\mathbb{R} \times \mathbb{R}$. Thus, they are of the form $\lambda r.s[(r,s) \in U]$ for $U \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$. This is equivalent to the characteristic function $\lambda r.s[s \in U_r]$ where $U_r = \{x \in \mathbb{R} | (r,x) \in U\}$. Thus, $\lambda r.s.[- \in U_r](s) = [(r,s) \in U]$ so $\mathrm{curry}([- \in U]) = \lambda r.[- \in U_r]$. By our bijective construction we have that $\lambda r.U_r$ are the random elements of $\mathcal{B}_{\mathbb{R}}$.

Thus, we need to show $F \circ (g,h)(r) = g(r)[U_r]$ is measurable. We have that

$$\lambda r.g(r)[U_r] = \mathrm{eval}(g(r),U_r) = \mathrm{eval} \circ (g(r),U_r).$$

We know that by definition g is a measurable function and evaluation is measurable by definition of probability spaces. Thus, we need to show U_r is measurable for the measurable space $(\mathcal{B}_{\mathbb{R}}, \Sigma_{\mathcal{B}_{\mathbb{R}}})$. This is tautological since $\Sigma_{\mathcal{B}_{\mathbb{R}}}$ is the σ-algebra such that the random elements of $\mathcal{B}_{\mathbb{R}}$ are measurable and since all the random elements are $\lambda r.U_r$ we have that $\lambda r.U_r$ is measurable. □

Now that we have our integral is a QBS morphism we want to show that it is equivalent to the Lebesgue integral. We can show this using the the standard machine as in [Williams, 1991]. The standard machine is a step by step method to prove results about measurable functions. We first show that our intended property is true for indicator functions. Then, by linearity we show it is also true for simple functions. Finally, we use the monotone convergence theorem to show it is true for all functions.

Definition 4.2.4. *Given a simple function $f = \sum_{k=1}^{m} a_k \mathbb{I}_{A_k}$, then $\mu_0(f) = \sum_{k=1}^{m} a_k \mu[A_k]$. The set of all simple functions is SF^+.*

Definition 4.2.5. *Given a measurable function $f : X \to \mathbb{W}$, then $\mu(f) = \sup\{\mu_0(h) | h \in SF^+, h \leq f\}$ is the Lebesgue integral.*

Theorem 4.2.1. *Given $f \in \mathbb{W}^X$, then $\int f d[\alpha,\mu] = \mu(f \circ \alpha)$.*

Using the standard machine we start with indicator functions.

Lemma 4.3. *Let $f \circ \alpha = \mathbb{I}_A$ be an indicator function, for $A \in \mathcal{B}_{\mathbb{R}}$. Then $\int f d[\alpha,\mu] = \mu(f \circ \alpha)$.*

Proof. We have $\mu(\mathbb{1}_A) = \mu_0(\mathbb{1}_A) = \mu[A]$. We also have

$$\int f d[\alpha,\mu] = \lim_{k\to\infty} \sum_{i=0}^{k2^k-1} \frac{i}{2^k} \mu[\alpha^{-1}[f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k})]].$$

Case 1: When $i = 2^k$, $\mu[\alpha^{-1}[f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k})]] = \mu[\alpha^{-1}[f^{-1}[1, \frac{2^k+1}{2^k})]]$. As $k \to \infty$, $\frac{2^k+1}{2^k} \to 1$ and so

$$\lim_{k\to\infty} \mu[\alpha^{-1}[f^{-1}[1, \frac{2^k+1}{2^k})]] = \mu[\alpha^{-1}[f^{-1}\{1\}]] = \mu[A].$$

Case 2: When $i = 0$, we have that $\lim_{k\to\infty} \alpha^{-1}[f^{-1}[0, \frac{1}{2^k})] = A^c$ and $0 \cdot \mu[A^c] = 0$.

Case 3: When $i > 0$ and $i \neq 2^k$ we have that $\alpha^{-1}[f^{-1}[\frac{i}{2^k}, \frac{i+1}{2^k})] = \emptyset$ so $\mu[\emptyset] = 0$.

Thus,

$$\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}\mu[\alpha^{-1}[f^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]]=\mu[A]$$

Hence, $\int f d[\alpha,\mu]=\mu[A]$ and therefore $\int f d[\alpha,\mu]=\mu(f\circ\alpha)$. $\square$

The next step is to show it is true for simple functions. The motivation for this is the Lebesgue integral is just an approximation since it is defined as the supremum of all simple functions less than the measurable function in question. Therefore, if we prove it true for simple functions it should follow easily that it is true for all measurable functions.

Lemma 4.4. *Let $f\circ\alpha=\sum_{j=1}^{m}a_j\mathbb{I}_{A_j}$ be a simple function, for $A_j\in\mathcal{B}_{\mathbb{R}}$. Then $\int f d[\alpha,\mu]=\mu(f\circ\alpha)$.*

Proof. We have that $\mu(f\circ\alpha)=\mu(\sum_{k=1}^{m}a_j\mathbb{I}_{A_j})=\sum_{j=1}^{m}a_j\mu(\mathbb{I}_{A_j})$. By Lemma 4.3,

$$\sum_{j=1}^{m}a_j\mu(\mathbb{I}_{A_j})=\sum_{j=1}^{m}a_j\int(\mathbb{I}_{A_j})d\mu.$$

Then,

$$\begin{aligned}a_j\int(\mathbb{I}_{A_j})d\mu&=a_j\cdot\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}\mu[\mathbb{I}_{A_j}^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]\\&=\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{a_j\cdot i}{2^k}\mu[\mathbb{I}_{A_j}^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]\\&=\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}\mu[(a_j\mathbb{I}_{A_j})^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]\\&=\int a_j\cdot\mathbb{I}_{A_j}.\end{aligned}$$

Now we need to show that $\int a\cdot\mathbb{I}_A d\mu+\int b\cdot\mathbb{I}_B d\mu=\int a\cdot\mathbb{I}_A+b\cdot\mathbb{I}_B d\mu$.

$$\begin{aligned}\int a\cdot\mathbb{I}_A d\mu+\int b\cdot\mathbb{I}_B d\mu&=\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}(\mu[(a\mathbb{I}_A)^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]+\mu[(b\mathbb{I}_B)^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})])\\&=\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}\mu[(a\mathbb{I}_A)^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})\cup(b\mathbb{I}_B)^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})]\\&=\lim_{k\to\infty}\sum_{i=0}^{k2^k-1}\frac{i}{2^k}\mu[(a\mathbb{I}_A+b\mathbb{I}_B)^{-1}[\frac{i}{2^k},\frac{i+1}{2^k})],\end{aligned}$$

since A and B are disjoint.

Therefore,

$$\sum_{j=1}^{m} a_j \int (\mathbb{I}_{A_j}) d\mu = \int \sum_{j=1}^{m} a_j \mathbb{I}_{A_j} d\mu$$
$$= \int f \circ \alpha d\mu = \int f d[\alpha, \mu]$$

□

We now use the Monotone Convergence Theorem to show that the integral we defined is the same as the Lebesgue integral for all QBS morphisms. The Monotone Convergence Theorem says that given any sequence of measurable functions that converges to the function f, then the limit of the integrals of the sequence of functions is the integral of the limit of the sequence which is just the function f.

Theorem 4.2.2. ***Monotone Convergence Theorem*** *([Williams, 1991, Theorem 5.3]): If (f_n) is a sequence of non-negative measurable functions such that* $\lim_{n\to\infty} f_n = f$ *then* $\lim_{n\to\infty} \mu(f_n) = f$.

We can now prove theorem 4.2.1

Proof. By the Monotone Convergence Theorem and Proposition 4.1 we have that,

$$\mu(f \circ \alpha) = \mu(\lim_{k\to\infty} [\![Approx\ f\ k]\!] \circ \alpha) = \lim_{k\to\infty} \mu([\![Approx\ f\ k]\!] \circ \alpha).$$

We apply Lemma 4.4 since $[\![Approx\ f\ k]\!] \circ \alpha$ is a simple function.

$$\lim_{k\to\infty} \mu([\![Approx\ f\ k]\!] \circ \alpha) = \lim_{k\to\infty} \int [\![Approx\ f\ k]\!] \circ \alpha d\mu.$$

Then by definition of the integral we have that

$$\lim_{k\to\infty} \int [\![Approx\ f\ k]\!] \circ \alpha d\mu = \lim_{k\to\infty} \lim_{k_1\to\infty} \sum_{i=0}^{k_1 2^{k_1}-1} \frac{i}{2^{k_1}} \mu[\alpha^{-1}[[\![Approx\ f\ k]\!]^{-1}[\frac{i}{2^{k_1}}, \frac{i+1}{2^{k_1}})]].$$

By Proposition 4.1 we have that

$$\lim_{k\to\infty} \lim_{k_1\to\infty} \sum_{i=0}^{k_1 2^{k_1}-1} \frac{i}{2^{k_1}} \mu[\alpha^{-1}[[\![Approx\ f\ k]\!]^{-1}[\frac{i}{2^{k_1}}, \frac{i+1}{2^{k_1}})]]$$
$$= \lim_{k_1\to\infty} \sum_{i=0}^{k_1 2^{k_1}-1} \frac{i}{2^{k_1}} \mu[\alpha^{-1}[f^{-1}[\frac{i}{2^{k_1}}, \frac{i+1}{2^{k_1}})]]$$
$$= \int f d[\alpha, \mu].$$

□

To summarise, we have defined integration on any QBS morphism X to $\mathbb{W}$ for arbitrary X over any probability measure on X. This integral is a QBS morphism and is equal to the Lebesgue integral. Integration theory can be used to immediately define expectation and everything that follows from that like inequalities.

Chapter 5

Related Work

The theory of quasi-Borel spaces is a recent development with the first work published in 2017. Thus most of the recent work has been on the initial development of the theory with focus on the applications to probabilistic programming languages. There have been extensions and additional constructions using QBS theory which provide categories with desirable properties for automatic statistical inference. One major new category are ωQBS which combine QBS structures with a complete partial order (cpo), specially ω-cpo. This allows for the use of recursion types which QBS theory excludes. This focus on domain theory is in pursuit of denotational semantics for probabilistic programming languages. Given earlier sections focus on integration theory the three papers presented in this chapter relate to integration theory as this is the central problem in many Bayesian statistical inference problems. They show the use and extension of QBS theory to development of higher order and recursive statistical probabilistic programming.

5.1 Convenient Category for Higher Order Probability

QBS theory was initially published in [Heunen et al., 2018] where the category of QBS was defined. The work presents the properties of the categorical structure which show its potential of the new theory for applications to probability theory then illustrates, through two important theorems and an example, why QBS theory form a convenient category for higher order probability theory.

After defining the QBS structure, the authors show that QBS can formalise probability theory by replacing the notion of measurable sets with random elements and shows the categorical structure is cartesian closed thus higher order function space construction is supported. This is arguably the most important result as it provides motivation to further investigate the use of the category to probability theory. The author's go on to state an important theorem in probability theory using QBS theory relevant to psuedo-random number generation. Inverse transform sampling is a method to generate random numbers from the the uniform distribution. This method works as we can express the law of a random variable in terms of the uniform distribution. In QBS theory the authors show that this result can be phrased as an equality between quotient spaces

[Heunen et al., 2018, Theorem 26].

De Finetti's theorem is central to Bayesian statistics. Bayesian inference techniques rely on being able to adjust uncertainty based on a sequence of random variables. Thus we cannot rely on an independent identically distributed sequence of random variables for inference as it would imply that $\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1} ... X_1 = x_1) = \mathbb{P}(X_n = x_n)$. De Finetti's theorem uses the weaker condition of exchangeability to circumvent this problem. In QBS theory the authors show this theorem holds for all QBSs. This further shows how the category is suitable for use as a formalisation of probability theory in the interest of use as a basis for probabilistic programming languages which use Bayesian methods for inference.

The authors acknowledge the short comings of the theory, importantly relating to domain theory and recursive types. This motivates the next development, ω-QBSs to extend the theory for recursive types.

5.2 A Domain Theory for Statistical Probabilistic Programmings

Furthering the work on QBS theory, ω-QBS which combine ωcpos for the use of recursive types with the higher order constructions in QBS theory [Vákár et al., 2019]. Using this construction, the authors develop a domain theory for recursive statistical probabilistic programming. Their semantic model allows for commutativity of integrals for higher order types, similar to Fubini's theorem, which further develops models of probability theory for higher order and recursive types. This paper adds to the literature further advancements towards probabilistic programming languages developed using categories that differ from traditional measure theory. Furthermore, the authors relate their model to the semantics of Monte-Carlo. Due to Monte-Carlo methods being a large part of Bayesian statistics, this result about their ω-QBS model further shows the usefulness of QBS like object applied to the development of probabilistic programming languages.

5.3 Denotational Validation of Higher-Order Bayesian Inference

The authors of [Ścibior et al., 2017] present a more concrete use case concerning itself with validation. They use QBS theory to resolve the issues of higher order functions in the probabilistic programming language defined in the paper. The language is equipped with continuous distributions which is novel but only generalises smoothly from the discrete case smoothly due to the combination of quasi-Borel spaces, synthetic measure theory, the defined meta-language for analysing the program and the modular inference blocks which are used to build probabilistic programming languages. These building blocks are used for composing semantic structures that represent probabilistic programming languages.

The authors then go on to validate algorithms used in probabilistic programming languages, namely Sequential Monte Carlo and the Markov Chain Monte Carlo algorithm Metropolis-Hastings. One of the main results was a generalisation of the Metropolis-Hastings-Green theorem to QBS theory. It allows the authors to prove the validity of the Metropolis-Hastings method in their language which is important as it shows that valid higher order probabilistic programming languages can be derived from QBS theory.

Chapter 6

Evaluation

Originally, the aim of the project was to develop statistics from quasi-Borel space theory. This would involve deriving versions of important theorems and bounds from measure theoretic probability for QBS based probability theory. For instance, one potential direction was conditional expectation. However, throughout the academic year, much of the time spent on the project was devoted to first learning both measure theory and quasi-Borel space theory. Due to the significant time investment of learning new areas of mathematics this slowed my progress on developing novel ideas.

6.1 Presentation of Measure and QBS Theory

The majority of this dissertation is a review and introduction to both measure and QBS theory. Measure theory is a rich and expansive field which contains many developments not shown in this dissertation. We only present what is necessary to understand Chapter 4 skipping details which may assist the reader's comprehension of the topic. For instance, we did not include π and λ-systems to construct measures instead starting with σ-algebras. A more foundational and iterative approach to defining measure spaces is to define these systems and use the $\pi-\lambda$-theorem to build σ-algebras. Since π and λ-systems are easier to understand this method may have been more intuitive. Similarly, when defining measures one can start with very small collections of sets and use different extension theorems to build to measures on σ-algebras. We did not take this approach as it would have included details that may derail or confuse the reader. Instead we present everything directly from definition and show enough examples for the reader to use for learning.

Furthermore, QBS theory has other constructions and properties that we did not explore. For instance, subspaces and embeddings have many more details that were not relevant as we only use them to construct the weights QBS $\mathbb{W}$. We opted to omit many of these details as the reader would have found it unnecessary. Additionally, to keep the narrative well structured and precise we detailed our aim of integration theory at the beginning and only presented what was necessary. Presenting the other parts of the theory would have distracted the reader from the objective of presenting integration theory.

6.2 Integration Theory

Once the goal of presenting integration theory was decided I focused on developing the theory using QBS similar to how it is developed in measure theoretic probability theory. Integration theory is built on simple functions and the fact that their measure is easily computable as a weighted sum of measures of disjoint measurable sets. Thus to define integration over QBS I took a similar approach but instead of a sequence of definitions built on simple functions an approximation function was defined which was equal to any morphism in the limit. Then we could immediately state the integral of this function as it had a simple representation. The benefits of this method of building to integration is we do not have to define several different integral types starting form indicator functions. The definition can be stated as the measure of the approximation. However, the method of building integrals from simpler integrals, which may be more intuitive, is lost and the reader must analyse how the approximation function itself behaves before understanding why these are equivalent methods to define the integral.

An important property of the integral is that it is a QBS morphism. This part of the project took a substantial amount of time to accomplish. Firstly, there are many different ways of showing this fact. I started trying to show that when composed with random elements from $\mathbb{W}^X \times PX$ was measurable. This proof would have required purely analytical methods to show. This idea was abandoned due to the fact it did not follow from the theory I built up throughout the year. Instead it would be more intuitive and be easier to understand if it was broken down into a series of QBS theory propositions. Originally, an evaluation function was defined, from the product space of Borel sets and probability measures to $[0,1]$, which could be shown to be measurable by an inductive argument. This would involve showing the sets for which made the evaluation measurable was exactly the Borel σ-algebra. In comparison to the final proof, this method would not of required QBS theory. It would involve only traditional measure theory. Given the project and the previous chapter focuses on QBS theory the final proof is more of a continuation than returning to proofs about measurable sets.

The proof that the QBS integral is the same as the Lebesgue integral uses the standard machine. This method highlights how integration theory is developed. First the property is proved true for indicator functions, then simple functions and from this all measurable functions. The proof is easily readable due to its step by step presentation and adds to the idea that integrals are built from simple functions. However, the proof is not relevant to QBS theory in that it involves no use of QBS theory. Although it is important to show that they are the same given we are aiming for the use of QBS theory as a basis for probability theory and thus the integrals in measure theoretic probability theory should be the same as QBS integrals.

6.3 Future Work

This work does not present any novel ideas and is a summary of the current theory of QBS. Unfortunately, it took longer than expected to reach the required level of theory to pursue original research and therefore here we present possible ideas for future work that could be novel, as suggested in [Kammar, 2022].

Given we have defined integration which extends to expectation the natural next step would be conditional expectation. Conditional expectation is the expected value of a random variable given we know some conditions. Formally, let $(S, \Sigma, \mathbb{P})$ be a probability space, $X : S \to \mathbb{R}$ a random variable and $\mathcal{H} \subseteq \Sigma$. The conditional expectation, $E(X|\mathcal{H})$, is a $\mathcal{H}$-measurable function such that $\int_H E(X|\mathcal{H}) d\mathbb{P} = \int_H X d\mathbb{P}$ for all $H \in \mathcal{H}$. We can see this operation is higher order; $E[X|\mathcal{H}]$ maps from Σ-measurable functions to $\mathcal{H}$-measurable functions.

Conditional expectation is used in a wide variety of statistics. For instance, the existence of conditional expectation can be proved via the Radon-Nikodým theorem [Williams, 1991]. It was shown in [Vákár and Ong, 2018] that there are equivalents of the Radon-Nikodým theorem in QBS theory. Thus it suggests another direction would be to prove that the Radon-Nikodým derivative is a morphism. This would be useful as it would provide measurable derivatives for function spaces.

Bibliography

[Aumann, 1961] Aumann, R. J. (1961). Borel structures for function spaces. *Illinois Journal of Mathematics*, 5(4):614–630.

[Bacci et al., 2018] Bacci, G., Furber, R., Kozen, D., Mardare, R., Panangaden, P., and Scott, D. (2018). Boolean-valued semantics for the stochastic λ-calculus. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 669–678.

[Dahlqvist and Kozen, 2019] Dahlqvist, F. and Kozen, D. (2019). Semantics of higher-order probabilistic programs with conditioning. *arXiv preprint arXiv:1902.11189*.

[Ehrhard et al., 2017] Ehrhard, T., Pagani, M., and Tasson, C. (2017). Measurable cones and stable, measurable functions: a model for probabilistic higher-order programming. *Proceedings of the ACM on Programming Languages*, 2(POPL):1–28.

[Heunen et al., 2018] Heunen, C., Kammar, O., Staton, S., Moss, S., Vákár, M., Ścibior, A., and Yang, H. (2018). The semantic structure of quasi-borel spaces. In *PPS Workshop on Probabilistic Programming Semantics*.

[Kammar, 2022] Kammar, O. (2022). An introduction to statistical modelling semantics with higher-order measure theory. Logic of Probabilistic Programming.

[Ścibior et al., 2017] Ścibior, A., Kammar, O., Vákár, M., Staton, S., Yang, H., Cai, Y., Ostermann, K., Moss, S. K., Heunen, C., and Ghahramani, Z. (2017). Denotational validation of higher-order bayesian inference. *arXiv preprint arXiv:1711.03219*.

[Vákár et al., 2019] Vákár, M., Kammar, O., and Staton, S. (2019). A domain theory for statistical probabilistic programming. *Proceedings of the ACM on Programming Languages*, 3(POPL):1–29.

[Vákár and Ong, 2018] Vákár, M. and Ong, L. (2018). On s-finite measures and kernels.

[Wagon, 1985] Wagon, S. (1985). *The Banach-Tarski Paradox*. Cambridge University Press.

[Williams, 1991] Williams, D. (1991). *Probability with martingales*. Cambridge university press.