A “Pointless” Theory of Probability

by

Qilin Ye

A Dissertation Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(Applied Mathematics, Progressive Degree)

May 2024

Acknowledgements

To be finished.

Table of Contents

Abstract

Traditionally, spaces are modeled as infinite collections of infinitesimal points. In recent decades, however, an alternate view has gained increasing scholarly attention: what if space is *gunky*, consists of regions, and does not admit a smallest, indivisible *atom*, a notion akin to points? In a gunky space, what set of primitives would one adopt in replacement of the Euclidean premises? Are these primitives as expressive as the ones in standard pointy models? And finally, what contradictions would arise in such exotic worlds, and how, if possible, can we resolve them?

In this paper, I provide an overview of previous literature on this area, highlighting some inherent complications that are embedded in a gunky perspective. A particular challenge that arose in the gunky models is the attempt to define a consistent notion of "size," or measure, to regions, and I discuss and evaluate existing solutions to this issue. Building on the notion of measure, I spend the second half of the paper investigating the feasibility of introducing theory of probability to a gunky model. I showed that it is possible to extend a gunky space and equip it with a countably additive probability measure. With this, I further recovered several important concepts in probability theory, including random variables, abstract integration, and several convergence results.

1 Introduction: Why Gunky Spaces?

The traditional understanding of space by mathematicians and philosophers is that space consists of infinitesimal, indivisible parts, called points. This perspective is so well-established that we are barely aware of its existence, taking the assumptions for granted. Indeed, a "**pointy**" model of the world feels so natural: every aspect of elementary math is introduced in this manner. Points are assumed to underpin the structure of most aspects of math. We assume that the state of material objects are described by functions defined on points. Geometry of Euclidean spaces, arithmetic on higher-dimensional spaces, electric fields in physics, among others, are all taught using points. And we say points make up regions and spaces.

Over the past two centuries, with increased focus in recent decades, a growing number of scholars have explored alternative models of space, the "**pointless** models," that omit the fundamental concept of points. Despite a relatively niche direction, scholars early in this field tended to have little collaboration and instead independently derived their own models of pointless geometry. Consequently, the list of scholarly attempts to axiomatize pointless geometry in the early days has been rather extensive, despite its relative lack of popularity compared to its pointy counterpart. The earliest documented exploration of pointless geometry was due to Lobachevsky in 1835, who assumed the primitive notions of *solids* and *contact* between solids. Yet it suffered from obscurity and a lack of rigor, and was discarded by the author himself. The first rigorous treatment of this subject was by Whitehead in the 1920s, where he showed how *solids* and *parthood*, the notion of one solid being contained in another, can be used to define an abstract notion of a point [5]. This approach was called "**gunky**" geometry, a term coined by Lewis [10], and many subsequent work ensued. In a gunky model, every region contains yet smaller regions — the space itself is infinitely divisible, not containing any indivisible points.

Contrasting to pointy spaces in nature, the gunky space also raises several significant motivations for studying it. The first one comes from a pure mathematical perspective: if the continuum of space

consists of points, then assigning the notion of size (or measure) to regions can turn into serious trouble. A classical fact in real analysis shows that there are non-measurable sets such like the Vitali sets, but a more blatant contradiction arises from the Banach-Tarski Paradox. In a pointy space, the unit ball in $\mathbb{R}^3$ (or any higher dimension) can be partitioned into five disjoint sets which, under rigid motions, can be reassembled rigidly into two disjoint unit balls, each congruent to the original one. These partitioned sets, of course, are not Lebesgue measurable. But this fundamentally still leads to problem, since rigid motions preserve the notion of size regardless of the measure we choose, yet through these carefully constructed pathological examples, the post-assembly size is twice the original one. Therefore, the Banach-Tarski Paradox exposes the limitations of point-based measures and encourage the exploration of gunky geometry, where (hopefully) the absence of indivisible points could circumvent these paradoxes.

Another potential motivation for studying a gunky space, as pointed out by Arntzenius, relates to the foundational theories of physics. As Arntzenius [1] notes, in modern physics, in particular non-relativistic quantum mechanics, the representation of a particle's state is by integration probabilistic integrations performed on wave functions which is invariant between functions that differ on a Lebesgue null set. This equivalence suggests a natural inclination towards considering these functions not as distinct entities but as members of the same equivalence class, defined by the relation of differing up to a null set. A gunky model, on the other hand, naturally embraces this idea: the very concept of point-sized differences is systematically eschewed under such settings.

Finally, purely philosophical motivations also justify the pursuit of a gunky space. Russell ([14]) points out that, beyond the actual structure of our universe, metaphysicians are also intrigued by the potential forms of what spaces *could* be, regardless of our current epistemic limitations. Debates are particularly focused on the viability of atomless (i.e. infinitely divisible) gunk as a credible form of matter, challenging models that cannot accommodate such concepts.

This paper is organized as follows. In Section 2, inspired by Tarski's approach ([6], [17]) of recovering standard Euclidean geometrical notions using only *parthood* and *sphere*, I show that it is entirely possible to achieve the same result using properties of isometric transformations on regions in $\mathbb{R}^n$. In Section 3, I provide an overall review of the theory of gunky models, explaining three commonly considered conditions of gunk: *mereological*, *topological*, and *measure-theoretic*. I then introduce a challenge posed by Russell's impossibility result [14] that shows the inherent inconsistency between topological and measure-theoretic gunks. Then, I discuss two recent workaround attempts ([1], [8]) to fix the problem with measures.

By now, we have reached a point where it is reasonable to introduce another layer of abstraction, describing our space via Boolean algebras. Finally, in Section 4, I look into the difficulties in defining a probability space on a Boolean algebra capturing the structure of the spaces used by the previous sections, and offer some partial solutions to them. It is worth pointing out that my approach to "fixing" some of the issues encountered in the Boolean algebra bears traces of Arntzenius [1]'s solution to fixing measure on gunks. In this section, I also attempt to recover several of the important probability theoretic results.

2 A Gunky $\mathbb{R}^n$ under Isometric Transformations

Alfred Tarski was one of the pioneers in exploring a gunky space. Using only the notion of *spheres* (as a special form of solid) and *parthood* (the notion of one solid being contained in another), Tarski [17] proved that it was possible to define all other geometric notion one encounters in Euclidean spaces. In particular, Tarskian spheres are sufficient to help recover the three Euclidean primitives: **point**, **betweenness** (colinearity between three points), and **congruence** (of lines or regions). We will briefly demonstrate the process below using balls in $\mathbb{R}^2$.

(1) B_1 and B_2 are **disjoint** if and only if no ball B_3 is contained in both (satisfying the *parthood* relation to both), written $B_1 \cap B_2 = \emptyset$.

(2) Defining tangency:

- B_1 and B_2 are **externally tangent** if:
  - $B_1 \cap B_2 = \emptyset$, and
  - For any two balls that both contain B_1 and disjoint from B_2, one of them must be contained in the other. In other words, for any B_3, B_4 such that $B_1 \subset B_3$, $B_1 \subset B_4$, and $B_3 \cap B_2 = B_4 \cap B_2 = \emptyset$, either $B_3 \subset B_4$ or $B_4 \subset B_3$.
- B_1 and B_2 are **internally tangent** (assuming $B_1 \subset B_2$, resp. $B_2 \subset B_1$) if:
  - $B_1 \subset B_2$, and
  - For any two balls that both contain B_1 and are contained in B_2, one of them must be contained in the other. In other words, for any B_3, B_4 such that $B_1 \subset B_3 \subset B_2$ and $B_1 \subset B_4 \subset B_2$, either $B_3 \subset B_4$ or $B_4 \subset B_3$.

(3) **Recovering betweenness** (diametrical opposites):

- B_1 and B_2 are **externally diametrical** of a ball B_3 if:

- B_1 and B_3 are externally tangent, and so are B_2 and B_3, and
- If B_1' contains B_1 but is disjoint from B_3, and likewise for B_2, then $B_1' \cap B_2' = \emptyset$.

- B_1 and B_2 are **internally diametrical** of a ball B_3 if:
  - B_1 and B_3 are internally tangent, and so are B_2 and B_3, and
  - If B_1' is disjoint from both B_1 and B_3, and likewise for B_2, then $B_1' \cap B_2' = \emptyset$.

(4) Defining concentric balls: Assuming $B_1 \subset B_2$, they are **concentric** if, for any two balls B_3, B_4 external diametric opposites of B_1 and (internally) tangent to B, they are also at internal diametric opposites of B_2.

(5) **Recovering points**: concentricity defines an equivalence relation, and a **point** is an equivalence class of concentric spheres.

(6) **Recovering congruence**: For another ball B_3, (B_1, B_3) is **equidistant** to (B_2, B_3), written $(B_1, B_3) \equiv (B_2, B_3)$, if there exists a ball B_3' concentric with B_3 such that:

- For all balls B_1' concentric with B_1, either $B_1' \subset B_3'$ or $B_1' \cap B_3' = \emptyset$, and
- Likewise for all balls B_2' concentric with B_2.

In this section, we consider another "pointless" approach of characterizing the Euclidean primitives. Instead of combining it with *parthood*, we consider the effects imposed by rigid motions, or isometric transformations. Since the Euclidean space itself is invariant under any fixed isometric transformation, our derivation of other geometric notions will be purely based on the algebraic structure of regions and isometries, independent from the metric equipped on the space as well as other external qualities. One advantage of such characteristic, as we will see later, is that our derivation is dimension-free and, in fact, capable of instead recovering the dimension of the Euclidean space that we start with.

2.1 Linear Algebraic Properties of Isometries

In this section, we consider the linear algebraic properties isometric transformations on $\mathbb{R}^n$ and identify a very special type of isometry: rotation. Recall that a map $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if for all x, y, $\|x - y\| = \|f(x) - f(y)\|$. Our first claim draws a connection between a specific class of isometries and orthogonal matrices:

Proposition 2.1. *A transformation $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is an origin-preserving (meaning $\psi(0) = 0$) isometry if and only if $\psi(x) = Ax$ for some orthogonal matrix A (meaning AA^T equals the identity).*

Proof. Let ψ be an origin-preserving isometry. Let $\{e_i\}_{i=1}^n$ be the standard Euclidean basis of $\mathbb{R}^n$. Note that an isometry sends triangles to congruent triangles, so in particular law of cosine implies that it also preserves angles and dot products, which we call the "dot product identity." Consequently $\{\psi(e_i)\}_{i=1}^n$ forms an orthonormal basis of $\mathbb{R}^n$ like $\{e_i\}$ does. Our first observation is that ψ is linear. To see this, pick any $u, v, w \in \mathbb{R}^n$.The "dot product identity" shows

$$\psi(u + v) \cdot \psi(w) = (u + v) \cdot w$$

and

$$(\psi(u) + \psi(v)) \cdot \psi(w) = \psi(u) \cdot \psi(w) + \psi(v) \cdot \psi(w) = u \cdot w + v \cdot w = (u + v) \cdot w.$$

Keeping u, v fixed and letting $w = e_i$ shows that $\psi(u+v)$ and $\psi(u)+\psi(v)$ agree coordinate-wise, so they equal. Similarly $\psi(cu) = c\psi(u)$. Therefore ψ is a linear transformation characterized by some matrix A. But then using the dot product identity once again, for any standard basis vectors e_i, e_j,

$$\mathbf{1}[i = j] = (Ae_i) \cdot (Ae_j) = e_j^T A^T A e_i = (A^T A)_{i,j}.$$

This shows A^TA is the identity (and so is AA^T).

The converse is trivial: if $\psi(x) = Ax$ then $(Au) \cdot (Av) = v^TA^TAu = u \cdot v$. Setting $u = v$ we see A maps basis to basis while also preserving length. □

With this result, we are able to fully characterize isometries of $\mathbb{R}^n$ as the composition of an origin-preserving isometry and a translation:

Theorem 2.2. *Every isometry ψ of $\mathbb{R}^n$ can be characterized by an orthogonal matrix $A \in \mathrm{O}_n(\mathbb{R})$ and a "translation" vector b, such that $\psi(x) = Ax + b$.*

Proof. First suppose $\psi(x) = Ax + b$ subject to the constraints above. This implies ψ is an isometry, since orthogonal matrices preserves distance and so does translation.

Conversely, let an isometry ψ be given. We first represent ψ as the composition of an origin-preserving isometry and a translation. To do so, write $\psi = (\psi - \psi(0)) + \psi(0)$. It remains to show that there exists an orthogonal matrix A such that $Ax = \psi(x) - \psi(0)$ for all x. This result now follows from the following lemma. □

It is easy to see that isometries form a group under composition. Among these transformations, **reflections** are of particular significance. These are the transformations that uses a certain **affine subspaces** as mirror, sending points across the mirror along the line perpendicular (orthogonal) to it. Formally a reflection ρ is an isometry of order 2 (i.e. $\rho = \rho^{-1}$ or $\rho^2 = I$, identity): indeed, mirroring twice and one ends up staying at the starting point. To transform u into v via a reflection, the affine mirror must lie in the "center" of the u, v. The canonical example in this case would be a $(n-1)$-dimensional hyperplane containing the "average" $(u+v)/2$ but also remaining orthogonal to the difference, $u - v$. Orthogonality gives rise to the following formula

$$H = \{x \in \mathbb{R}^n : (x-(u+v)/2)\cdot(u-v) = 0\} = \{x \in \mathbb{R}^n : x\cdot(u-v) = (u-v)\cdot(u+v)/2 = (\|u\|-\|v\|)/2\}.$$

From now on, given a reflection ρ, we use $\text{aff}(\rho)$ to denote its corresponding affine subspace. In the next theorem, we show that reflections generate the entire group of isometry.

Theorem 2.3. *Every isometry ψ of $\mathbb{R}^n$ is a composition of at most $n+1$ reflections.*

Proof. We may WLOG assume $\psi(0) = 0$, so that ψ is completely characterized by some $A \in \mathrm{O}_n(\mathbb{R})$, at the cost of at most one reflection. To see this, suppose $\psi(0) \neq 0$. We consider a reflection $\rho^{(0)}$ such that $\text{aff}(\rho^{(0)})$ contains the midpoint $\psi(0)/2$ and is orthogonal to the line crossing 0 and $\psi(0)$. One can also verify that one such example is the reflection across the $(n-1)$-dimensional hyperplane described by $\{u \in \mathbb{R}^n : u \cdot (\psi(0) - 0) = u \cdot \psi(0) = (\|\psi(0)\| - \|0\|)/2 = \psi(0)^2/2\}$.

Now we assume $\psi(0) = 0$ and prove by induction. The case $n = 1$ is clear, since an origin-preserving isometry on $\mathbb{R}$ is either identity or negative identity.

For the inductive step assume the claim holds for $\mathbb{R}^{n-1}$. Consider an arbitrary ψ represented by $A \in \mathrm{O}_n(\mathbb{R})$, and let $\{e_i\}_{i=1}^n$ be the standard basis of $\mathbb{R}^n$. Repeating the argument in the first paragraph, we see that there exists a reflection f_n, possibly identity, such that $f_n(Ae_n) = (f_n \circ \psi)(e_n) = e_n$. Furthermore, since $A \in \mathrm{O}_n(\mathbb{R})$ we know $\|Ae_n\| = \|e_n\|$. This means $\text{aff}(f_n) = \{u \in \mathbb{R}^n : u \cdot (Ae_n - e_n) = (\|Ae_n\| - \|e_n\|)/2 = 0\}$ contains the origin, and therefore $f_n \in \mathrm{O}_n(\mathbb{R})$.

By construction, $f_n \circ \psi$ equals identity on $E = \{0\}^{n-1} \times \mathbb{R}$, the subspace spanned by e_n, whereas its orthogonal complement, $H = \mathbb{R}^{n-1} \times \{0\}$, remains invariant under f_n. Our induction hypothesis states that there exist reflections $f_1, \ldots, f_{n-1}$ whose composition agrees with $f_n \circ \psi$ on H. We can easily define an extension of $f_1 \circ \ldots \circ f_{n-1}$ by setting its n^{th} component to be an identity mapping, i.e., $(f_1 \circ \ldots \circ f_{n-1})(\cdot, x_n) = ((f_1 \circ \ldots \circ f_{n-1})(\cdot), x_n)$. This extended function now coincides with $f_n \circ \psi$ on all of $\mathbb{R}^n$, so $\psi = f_n^{-1} \circ f_1 \circ \ldots \circ f_{n-1}$, and the inductive step is complete.

To sum up, we showed that every origin-preserving isometry of $\mathbb{R}^n$ is the composition of up to n reflections, and more generally, every isometry of $\mathbb{R}^n$ is the composition of up to $n+1$ reflections. $\square$

2.2 Recovering Euclidean Structure via Isometries

Observe that in $\mathbb{R}^n$, any reflection ρ satisfies $\dim \text{aff}(\rho) \leqslant n-1$. For a simple example, in $\mathbb{R}^2$, reflections are characterized by 0- or 1-dimensional affine subspaces, namely, points or lines. And a simple drawing shows that if we have two line reflections across ℓ_1 and ℓ_2, then they commute if and only if ($\ell_1 = \ell_2$ or $\ell_1 \perp \ell_2$). In $\mathbb{R}^n$, the high-dimensional analogy also holds.

Proposition 2.4. *Let ρ_1, ρ_2 be two reflections of $\mathbb{R}^n$, with affine subspaces H_1, H_2. Then they commute if and only if one of the following **commutativity criteria** holds:*

(1) One of the subspaces is contained in the other, or

(2) $H_1 \cap H_2 \neq \emptyset$, and every $u \in H_1 \backslash (H_1 \cap H_2) = H_1 \backslash H_2$ is orthogonal to every $v \in H_2 \backslash H_1$.

For convenience we call these the "reflection commutativity criterion." In either cases, $\text{aff}(\rho_1 \circ \rho_2) = \text{aff}(\rho_2 \circ \rho_1) = H_1 \cap H_2$.

Remark. Despite sounding like an orthogonality criterion, condition (2) cannot be replaced with "H_1 and H_2 are orthogonal." Consider for example two reflections in $\mathbb{R}^3$, one across the xy-plane: $(x, y, z) \mapsto (x, y, -z)$ and one across the yz-plane: $(x, y, z) \mapsto (-x, y, z)$. They clearly commute, but the two planes, when viewed as subspaces, are not orthogonal — they intersect at a line.

This observation suggests a certain structure of composition of commutative reflections. Note that if ρ_1 commutes with ρ_2, then their composition is yet another reflection: $(\rho_1\rho_2)(\rho_1\rho_2)^{-1} = \rho_1\rho_2\rho_2^{-1}\rho_1^{-1} = \rho_1\rho_1^{-1} = I$. This suggests we can define commutative groups of reflection, with identity transformation being the identity.

Theorem 2.5. *In $\mathbb{R}^n$, a commutative group of reflections under composition has at most 2^n elements, and this upper bound can be attained.*

Proof. It's trivial to construct a commutative group of reflections with 2^n elements. For $1 \leqslant i \leqslant n$, define ρ_i to be the mapping that flips the sign of the i^{th} coordinate while keeping others unchanged. It follows that the ρ_i's commute, and that they generate a group of 2^n elements — pick any subset of the n coordinates in $\mathbb{R}^n$ and flip the sign.

Conversely, let G_n be a commutative group of reflections of $\mathbb{R}^n$. For each $1 \leqslant k \leqslant n$, consider $\{\rho \in G_n : \dim \text{aff}(\rho) = k\}$, the subset of reflections whose affine subspace has dimension k. They have to commute pairwise so they have to satisfy the reflection commutativity criterion, and clearly they need to satisfy (2). When their dimensions match, criterion (2) is equivalent to requiring that the normal vectors to these subspaces are orthogonal. In $\mathbb{R}^n$ there exist $\binom{n}{k}$ pairwise orthogonal vectors that can serve as normal vectors to k-dimensional affine subspaces. Therefore $|\{\rho \in G_n : \dim \text{aff}(\rho) = k\}| \leqslant \binom{n}{k}$, and

$$|G_n| = \sum_{k=0}^{n} |\{\rho \in G : \dim \text{aff}(\rho) = k\}| \leqslant \sum_{k=0}^{n} \binom{n}{k} = 2^n. \tag{Δ}$$

Let $|G_n| = 2^n$ be a commutative group of reflections of $\mathbb{R}^n$. There are two "special" elements in this group, namely when $k = n$ and when $k = 0$ as in (Δ). The former corresponds to a reflection whose axis is the entire space — this is the identity transformation. The latter, on the other hand, is the reflection across a 0-dimensional affine subspace, namely a point, if one assumes its existence a priori. We call it the **point reflection** associated with G_n. This point reflection is special in the sense that it "looks the same from every perspective" whereas for every other reflection $\rho \in G_n$, the effect imposed by ρ on a region depends on the location of the region and its relation to the affine subspace associated with ρ.

In other words, the point reflection is the "special" element of G_n displaying a certain invariance property. This should remind one of the notion of conjugation in abstract algebra. Let us quickly recall that if G is a group, then $g, h \in G$ are **conjugates** of each other if there exists a $f \in G$ such that $h = fgf^{-1}$.

It follows by group properties that conjugacy is an equivalence relation, which gives rise to **conjugacy classes**.

Going back to the group of isometries. Let f, g be two isometric transformations. We say x is invariant under g if $x = g(x)$. An immediate result following conjugation is that x is invariant under g if and only if $f(x)$ is invariant under fgf^{-1}:

$$x = g(x) \Rightarrow (fgf^{-1})(f(x)) = (fg)(x) = f(x)$$

and conversely

$$(fgf^{-1})(f(x)) = f(x) \Rightarrow (fgf^{-1})(f(x)) = fg(x) = f(x) \Rightarrow g(x) = x$$

since isometries are injective, directly by definition: $\|f(x) - f(y)\| = 0$ if and only if $\|x - y\| = 0$.

Now we restrict our attention to point reflections — x is invariant under a reflection ρ if and only if x belongs to the affine subspace associated with ρ. Point reflections only have one fixed point, namely the "point" across which the reflection is performed. Therefore, if p is the **point reflection** associated with G_n and f is *any* other isometry (we don't require it to be another reflection), then fpf^{-1} is also a point reflection. Using the characterizations of commutative reflections, we see that these two point reflections p_1, p_2 commute if and only if they are identical — otherwise, $\text{aff}(p_1) \cap \text{aff}(p_2) = \emptyset$, and both condition fails. Put formally:

Definition 2.6. *A **point reflection** p of $\mathbb{R}^n$ is an isometry satisfying the following:*

(1) p is not the identity mapping, and

(2) For any other isometry f, the composition fpf^{-1} either equals p or does not commute with it.

Note that this definition itself does not invoke any notion of points, nor is it dependent on dimensionality, unlike the theorem on maximal commutative group of reflections. Having characterized point reflections, now we may work our way backwards and recover other primitives and notions under standard Euclidean geometry.

First, linearity. Just like how two points define a line, we can fix two point reflections and define a corresponding line reflection. In order to do so, we need to establish a strict partial order on reflections with respect to the dimension and containment of their associated affine subspaces. The idea builds on the following fact: if $H_1 \subset H_2$ are with dimensions $d_1 < d_2$, then there exists an $H_1' \subset H_2$ such that $\dim(H_1') = \dim(H_1)$, and they have orthogonal normal vectors. In order to formally state these, we need to (i) define subspace containment, and (ii) define orthogonal subspaces of the same dimension. Luckily, with conjugation, we have all the ingredients to cook up these notions. We use the symbol $\rho_1 \prec \rho_2$ to represent that ρ_1's subspace is strictly contained in ρ_2's.

Definition 2.7. *Given two commutative reflections ρ_1, ρ_2 of $\mathbb{R}^n$, we say $\boxed{\rho_1 \prec \rho_2}$ if the following holds:*

(1) There exists a translation f leaving ρ_2 invariant under conjugation but not ρ_1, i.e., $\rho_1 \neq f\rho_1 f^{-1}$, but $\rho_2 = f\rho_2 f^{-1}$, and

(2) There does not exist a translation g leaving ρ_1 invariant but not ρ_2.

The conditions impose structural constraints on ρ_1 and ρ_2's subspaces: $\text{aff}(\rho_2)$ has "extra" free dimensions whereas $\text{aff}(\rho_1)$ doesn't. This ensures that if $\rho_1 \prec \rho_2$, then $\text{aff}(\rho_1) \subset \text{aff}(\rho_2)$ with a strictly lower dimension. By definition, $\prec$ is irreflexive and asymmetric. To show transitivity, suppose $\rho_1 \prec \rho_2 \prec \rho_3$, with f_{12} fixing ρ_2 but not ρ_1, and f_{23} fixing ρ_3 but not ρ_2. Then

$$(f_{12}f_{23})\rho_3(f_{12}f_{23})^{-1} = f_{12}(f_{23}\rho_3 f_{23}^{-1})f_{12}^{-1} = f_{12}\rho_3 f_{12}^{-1}.$$

Since f_{12} leaves ρ_2 invariant and $\rho_2 \prec \rho_3$, by definition f_{12} also leaves ρ_3 invariant, so the above equals ρ_3. The other direction is trivial — if a translation leaves ρ_1 invariant under conjugation then by definition it leaves ρ_2 invariant. Repeating this argument once again, we see it must also leave ρ_3 invariant.

While we proved $\prec$ to be a partial order, it is still a much weaker notion than dimensioanlity, since only reflections with overlapping affine subspaces are comparable. Indeed, in $\mathbb{R}^n$ we may compare the dimension of two arbitrary suitably nice regions, but this will suffice in our upcoming definitions.

Definition 2.8. *Given two point reflections p_1 and p_2, the **line reflection** $\ell = \ell(p_1, p_2)$ associated with p_1 and p_2 is the minimal reflection (w.r.t. $\prec$) such that $p_1 \prec \ell$ and $p_2 \prec \ell$. In other words, if $p_1, p_2 \prec \rho$ for some other reflection ρ, then $\ell \prec \rho$.*

It naturally follows that a point reflection p_3 is **colinear** with p_1 and p_2 if $p_3 \prec \ell(p_1, p_2)$. More generally, we can iteratively apply this definition and recover the entire structure of the Euclidean space. Define a point reflection to be of **dimension** 0 and a line reflection of dimension 1. Given two reflections ρ_1, ρ_2 of dimensions $n-1$, we may define a corresponding reflection $\rho = \rho(\rho_1, \rho_2)$ of dimension n to be the minimal reflection w.r.t. $\prec$ satisfying $\rho_1 \prec \rho, \rho_2 \prec \rho$. Finally, given three reflections ρ_1, ρ_2, ρ_3 of the same dimension, we say (ρ_1, ρ_2) is **congruent** to (ρ_2, ρ_3) if there exists a translation τ such that $\tau \circ \rho_1 = \rho_2$ and $\tau \circ \rho_2 = \rho_3$.

To wrap up, we have identified a method that helps us cover the standard Euclidean geometric structure using only the notion of regions and algebraic properties of rigid motions on these regions. We proved that, using commutativity of rotations, it is possible to recover an analogous notion of point using a special type of isometry called point reflections. We then showed that dimensionality and colinearity can be defined by working "backwards" once we have identified point reflections. Finally, congruence is easy with the help of translation, another type of isometry. Throughout this process, we never relied explicitly on the dimension of the ambient space, but were able to recover the dimension through our own characterizations.

3 Incompatibility of Gunk and Measure

After the Whiteheadian perspective was proposed, it has thus become a tradition to model a gunky sapce using Boolean algebra of regular open sets in the Euclidean space. Recall that a **Boolean algebra** [15] is a set $\mathfrak{B}$ equipped with binary operations $\wedge$ (meet), $\vee$ (join), and distinct members $\mathbf{0}$, $\mathbf{1}$ of $\mathfrak{B}$ such that:

(i) $\langle B, \vee, \wedge \rangle$ is a distributive lattice (meaning pairwise meets and joins exist, and the two operations are distributive),

(ii) $x \vee \mathbf{0} = x$ and $x \wedge \mathbf{1} = x$ for all x, and

(iii) $x \vee (-x) = \mathbf{1}$ and $x \wedge (-x) = \mathbf{0}$ for all x.

The following are some immediate consequences of these definitions. We see that Boolean algebra exhibits structures that coincide nicely with many of the spatial relations we are allowed to impose in a pointless model.

(i) We may define a partial order $\leqslant$ on $\mathfrak{B}$, by $x \leqslant y$ if and only if $x \vee y = y$ if and only if $x \wedge y = x$. In the space of regular open sets, this encodes the *parthood* relation. Analogously a strict partial order $<$ can be defined, and $x < y$ means x is a proper part of y.

(ii) The space itself can be considered a region and is represented by $\mathbf{1}$. The $\mathbf{0}$ does not really represent any region, and is included for mathematical convenience. We use $\mathbf{0}$ and $\emptyset$ interchangeably.

(iii) Every region divides the space into two parts: informally, the one belonging to the region itself, and the one that does not. In this mereological sense, there should be nothing else. This correpsonds to the Boolean algebraic complement: joining the two regions gives the top element, i.e., the whole space, and meeting the two regions gives nothing.

Historically, philosophers and mathematicians have explored relevant spatial structures that would be "nice" to have in a model. One of the earliest mereological condition simply demands that the space itself be *atomless*:

Every region has a proper subregion. (**Mereological Gunk**)

A stronger extension of mereological gunk would instead require that not only does every region contain a proper subregion, but the subregion sits well inside it. The approach, following the tradition set by Whitehead [18] and Roeper [12], relies on a primitive called *connectedness*. Every region x would have to contain a subregion that is disjoint from any region disjoint from x:

Every region x contains a subregion that is disjoint from any region disjoint from x. (**Topological Gunk**)

But regions have sizes. While the topological gunk condition provides a statement that demands the qualitative existence of an "interior" part, it lacks a quantitative formulation. To fill in this gap, gunky spaces shouldn't have discrete, quantitatively indivisible chunks. This gives rise to the strongest condition of the three:

With respect to any measure, every region contains a strictly smaller part. (**Measure-Theoretic Gunk**)

All three approaches have been extensively studied. The mereological one was established by Whitehead. The topological condition was discussed in Roeper, and the measure-theoretic gunk was studied by Skyrms [16] and Sikorski [15]. Alternate approaches also exist, such as a nonstandard mereology by Forrest [3] and a metrical appraoch by Gerla [4].

It would be nice to have a model satisfying all three structural constraints on gunky models, but this attempt has been proven impossible by Russell [14], where he claimed that topological gunk and measure-theoretic gunk are inherently incompatible. In the remainder of the section, we will restrict our attention to the Boolean algebra formed by regular open sets in $\mathbb{R}^n$ and reconstruct a famous counterexample involving the Fat Cantor Set. We will then briefly discuss several attempts that address the issues brought up by the Fat Cantor Set.

3.1 Boolean Algebra and the Fat Cantor Set

A set A is **regular open** if $A = \text{Int Cl } A$ (interior of closure). We consider the collection $\mathfrak{B}$ of regular open sets in $\mathbb{R}^n$, equipped with the following operations:

- join: $A \vee B := \text{Int Cl}(A \cup B)$
- meet: $A \wedge B := A \cap B$
- Boolean complement: $-A = \text{Int Cl}(\mathbb{R}^n \backslash A) := \text{Int}(\mathbb{R}^n \backslash A) = \text{Int}(A^c)$
- $\mathbf{0}, \mathbf{1}$ represented by $\emptyset$ and $\mathbb{R}^n$.

Proposition 3.1. *The collection of regular open sets in $\mathbb{R}^n$, with operations defined above, forms a complete Boolean algebra.*

*Proof that $\mathfrak{B}$ **is a Boolean algebra**.* We need to prove that $\mathfrak{B}$ is a distributed, complemented lattice. It is clear that $\langle \mathfrak{B}, \vee, \wedge \rangle$ form a lattice. (ii) is also clear. To prove (iii), it might be convenient to denote $\text{Int Cl}(A)$ by A^{**}, where $A^* := \mathbb{R}^n \backslash \text{Cl } A = (\text{Cl } A)^c$. For any $A \in \text{RO}(\mathbb{R}^n)$,

$$A \vee (-A) = (A \cup (-A))^{**} = (A \cup A^*)^{**}$$

$$= (A^* \cap A^{**})^* = (A^{**} \cap A)^* = \emptyset^* = \mathbf{1} = \mathbb{R}^n$$

and

$$A \wedge (-A) = A \cap \operatorname{Int}(A^c) \subset A \cap A^c = \emptyset.$$

It remains to prove distributivity. First note that

$$A \wedge (B \vee C) = A \wedge (B \cup C)^{**} = A^{**} \cap (B \cup C)^{**}$$

and that

$$(A \wedge B) \cup (A \wedge C) = ((A \wedge B) \cup (A \wedge C))^{**} = ((A \cap B) \cup (A \cap C))^{**} = (A \cap (B \cup C))^{**}.$$

It remains to connect these two chains of equalities — $B \cup C$ is open regular, and

$$X^{**} \cap Y^{**} = X \cap Y = (X \cap Y)^{**} \qquad \text{for any } X, Y \in \mathrm{RO}(\mathbb{R}^n).$$

The other distributive identity can be proven analogously. □

Proof that $\mathfrak{B}$ ***is complete.*** Let I be an index set. Then

$$X = \bigvee_{i \in I} A_i = \operatorname{Int}\operatorname{Cl} \bigcup_{i \in I} A_i \qquad \text{and} \qquad Y = \bigwedge_{i \in I} A_i = \operatorname{Int}\operatorname{Cl} \bigcap_{i \in I} A_i$$

are well-defined in $\mathrm{RO}(\mathbb{R}^n)$.

We show that X is the supremum of $\{A_i\}$. For each $i \in I$,

$$A_i = \operatorname{Int}\operatorname{Cl} A_i \subset \operatorname{Int}\operatorname{Cl} \bigcup_{i \in I} A_i = X.$$

On the other hand if $X' \in \mathrm{RO}(\mathbb{R}^n)$ is another upper bound of $\{A_i\}_{i\in I}$ then $A_i \subset X'$ for all $i \in I$, so $\bigcup A_i \subset X'$, and

$$X = \operatorname{Int}\operatorname{Cl}\bigcup_{i\in I} A_i \subset \operatorname{Int}\operatorname{Cl} X' = X'.$$

That Y is the infimum can be proven analogously. To avoid abuse of symbolic variables we will denote X, Y by $\sup\{A_i\}_{i\in I}$ and $\inf\{A_i\}_{i\in I}$, respectively. □

Unfortunately, a point-less Euclidean geometry does not come without technical challenges. One particular complication is the notion of size or measure:

Proposition 3.2. *Lebesgue measure is not finitely additive on* $\mathfrak{B}$.

Proof. To see this, we appeal to the famous **Smith-Volterra-Cantor set** ([1], [8], [11], [14]), also known as the Fat Cantor Set. Let m denote the Lebesgue measure.

Define $I = [0, 1]$ the unit interval. In the first iteration, remove a subinterval of length $1/4$ from the middle of I, and use I_1 to denote this removed interval: $I_1 = [3/8, 5/8]$. Now $I\backslash I_1$ consists of two disjoint intervals of equal length. In the second iteration, remove two subintervals of length 4^{-2} from each of the two remaining intervals in $I\backslash I_1$, and use I_2 to denote the union of these two removed intervals. Iterative, in the k^{th} iteration we remove intervals of length 4^{-k} from the middle of each of the 2^{k-1} remaining subintervals in $I\backslash\bigcup_{i=1}^{k-1} I_i$ and define I_k accordingly. Since I_k consists of 2^{k-1} disjoint intervals, each with length 4^k, $m(I_k) = 2^{-k}/2$. And clearly $I_i \cap I_j = \emptyset$ for $i \neq j$, so $m(\bigcup_{k\geqslant 1} I_k) = 1/4 + 1/8 + \ldots = 1/2$. We define the Fat Cantor Set to be $\mathcal{C} = I\backslash\bigcup_{k\geqslant 1} I_k$ and it follows that $m(\mathcal{C}) = 1 - 1/2 = 1/2$.

Before visiting pathological examples, it is worth comparing this fat $\mathcal{C}$ against the standard middle-thirds Cantor set, the most notable difference being that $\mathcal{C}$ has a nonzero Lebesgue measure. Despite so, $\mathcal{C}$ is still totally disconencted and is the boundary of $I\backslash\mathcal{C} = \bigcup_{k\geqslant 1} I_k$, just like the standard Cantor set. Given $x \in \mathbb{R}$ and $S \subset \mathbb{R}$, define the point-set distance $d(x, S) = \inf_{s\in S}|x - s|$. It follows that for any $x \in [0, 1] = I$, after the first iteration, $d(x, I\backslash I_1) < 1/2$, and by induction $d(x, I\backslash\bigcup_{i=1}^{k} I_i) < 2^{-k}$. This

shows that given any $x \in [0,1]$ and $\epsilon > 0$, there exists a sufficiently large k such that $(x-\epsilon, x+\epsilon) \cap I_k \neq \emptyset$. In other words, the closure of $\bigcup_{k \geqslant 1} I_k$ is $[0,1]$.

Now we partition $\bigcup_{k \geqslant 1} I_k$ into two parts, a *Big Cantor* defined by $\bigcup_{k \text{ odd}} I_k$ and a *Small Cantor* $= \bigcup_{k \text{ even}} I_k$. (So far we have *Big Cantor*, *Small Cantor*, and $\mathcal{C}$, whose disjoint union is I.) It is immediately clear that these sets are Lebesgue measurable, with $m(\textit{Big Cantor}) = 1/3$ and $m(\textit{Small Cantor}) = 1/6$. While their Lebesgue measures behave nicely under usual set-theoretic operations, things look rather different when we turn into Boolean operations defined above. To see this, observe that *Big Cantor* $\wedge$ *Small Cantor* $= \emptyset$, so $m(\textit{Big Cantor} \wedge \textit{Small Cantor}) = 0$. On the other hand, *Big Cantor* $\vee$ *Small Cantor* $=$ $\operatorname{Int}\operatorname{Cl}(\textit{Big Cantor} \cup \textit{Small Cantor}) = \operatorname{Int}\operatorname{Cl}(\bigcup_{k \geqslant 1} I_k) = \operatorname{Int}([0,1]) = [0,1]$, but

$$1 = m(\textit{Big Cantor} \vee \textit{Small Cantor}) \neq m(\textit{Big Cantor}) + m(\textit{Small Cantor}) = 1/2.$$

Here we constructed examples in $\mathbb{R}$. Higher-dimensional counterparts can be constructed analogously.

□

The problem here is that we constructed two sets A, B such that $A \vee B = \operatorname{Int}\operatorname{Cl}(A \cup B) \supsetneq (A \cup B)$, and this leads to unwanted behaviors. The root of this pathology lies in the fact that both *Big Cantor* and *Small Cantor* (i) have boundaries of positive measure, and (ii) their boundaries are both $\mathcal{C}$. Consequently, "too much" is introduced by the closure operator Cl to the extent where Lebesgue measure breaks. A few modifications have been proposed and explored by previous literature.

3.2 Solution 1: Forcing Null Boundaries

The first one is proposed by Lando and Scott [8], where we simply restrict our attention to sets with boundaries of Lebesgue measure zero, which we denote by $\text{RON}(\mathbb{R}^n) \subset \text{RO}(\mathbb{R}^n)$. (In their paper, they began with regular closed sets instead of regular open sets, but all of the following results hold by taking appropriate

complementations and swapping orders of interior and closure operators.) This set explicitly excludes the pathological *Big/Small Cantor* example above, and indeed preserves finite additivity of Lebesgue measures.

Proposition 3.3. *Lebesgue measure is not finitely additive on* $\mathrm{RON}(\mathbb{R}^n)$ *with operations defined as in* $\mathfrak{B}$.

Proof. Suppose $A, B \in \mathrm{RON}(\mathbb{R}^n)$ and $A \wedge B = A \cap B = \emptyset$. In particular the interiors are disjoint, so

$$\mathrm{Cl}(A) \cap \mathrm{Cl}(B) = (\mathrm{Int}\, A \cup \partial A) \cap (\mathrm{Int}\, B \cup \partial B) = (\partial A \cap \mathrm{Cl}\, B) \cup (\partial B \cap \mathrm{Cl}\, A).$$

Since $m(\partial A) = m(\partial B) = 0$ by assumption, the above set has measure 0. Then

$$\begin{aligned} m(\mathrm{Cl}(A \cup B)) &= m(\mathrm{Cl}(A) \cup \mathrm{Cl}(B)) \\ &= m(\mathrm{Cl}(A)) + m(\mathrm{Cl}(B)) - m(\mathrm{Cl}(A) \cap \mathrm{Cl}(B)) \\ &= m(\mathrm{Cl}(A)) + m(\mathrm{Cl}(B)) = m(A) + m(B). \end{aligned}$$

It remains to notice that $\partial(A \cup B) \subset \partial A \cup \partial B$. Thus $A \cup B$ has null boundary and $m(\mathrm{Cl}(A \cup B)) = m(\mathrm{Int}\, \mathrm{Cl}(A \cup B))$, completing the proof. □

While restricting attention indeed resolves the issue of finite additivity of Lebesgue measure, it still doesn't resolved the issue of countable additivity. Clearly, each I_k as in the construction of $\mathcal{C}$ is in $\mathrm{RO}_0(\mathbb{R}^n)$, but as we already showed,

$$1/2 = \sum_{k \geqslant 1} m(I_k) \neq m(\textstyle\bigvee_{k \geqslant 1} I_k) = m(\mathrm{Int}\, \mathrm{Cl}(\bigcup_{k \geqslant 1} I_k)) = m([0,1]) = 1.$$

This means that the Boolean algebra of regular open sets with null boundaries form an incomplete Boolean algebra, and that its completion is once again $\mathrm{RO}(\mathbb{R}^n)$. Nevertheless, Lando and Scott have shown that this Boolean algebra has many nice properties. $\mathrm{RON}(\mathbb{R}^n)$ sits densely in $\mathrm{RO}(\mathbb{R}^n)$ in the algebraic sense.

It satisfies Roeper (1997)'s axiomatizations of *region-based topology*, a collection of axioms which Roeper proved to be necessary for any collection of pointless regions constructed via an equivalence relation defined on boundary points for pointy regions in a locally compact Hausdorff space, to which the collection of regular open sets belong. And like the previous section, Lando and Scott have also shown that the notion of points, along with the entire pointy topology of $\mathbb{R}^n$, can be identified via certain equivalence relations, and that RON($\mathbb{R}^n$) is not isomorphic to RON($\mathbb{R}^m$) for any $m \neq n$, essentially establishing the notion of dimensionality in a parameter-free approach.

3.3 Solution 2: Equivalence Relations on Sets with Null Difference

Another more classical approach is proposed by Arntznius [1], where he considered **Lebesgue measure algebra**, the algebra of Borel subset of $\mathbb{R}^n$ modulo the ideal of Lebesgue null sets. In his own words, he blurs the differences in regions which "do not correspond to differences in actual physical space." If we define an equivalence relation $A \sim B$ if the symmetric difference $A\Delta B = (A \backslash B) \cup (B \backslash A)$ has measure zero, then naturally all regions in the same equivalence class have the same Lebesgue measure. Therefore, it is well-defined to let $\mu_0([A]) = m(A)$ where $[A]$ is the equivalence class containing A. Since we are dealing with Lebesgue measure algebra, we may relax our assumption on regular open sets and instead directly work on equivalence classes of Borel sets. To this end we redefine the notions of join, meet, and complement on Lebesgue measure algebra canonically:

$$[A] \wedge [B] := [A \cap B] \qquad [A] \vee [B] := [A \cup B] \qquad -[A] := [A^c].$$

Since a countable union of null sets is still null, intuitively,

Proposition 3.4. *μ_0 is a countably additive measure on the Lebesgue measure algebra.*

Proof. Indeed, given pairwise disjoint equivalence classes $\{[A_k]\}_{k\geqslant 1}$ (in the sense that $[A_i] \wedge [A_j] = [\emptyset]$), we define $B_1 = A_1$ and $B_k = A_k \backslash \bigcup_{i=1}^{k-1} A_i$ inductively. It follows that the B_k's are pairwise disjoint, and

$$A_k \Delta B_k = A_k \backslash B_k = A_k \cap \bigcup_{i=1}^{k-1} A_i = \bigcup_{i=1}^{k-1} A_k \cap A_i$$

which is a null set because $A_k \cap A_i$ has measure zero for all $i < k$. Therefore $[A_k] = [B_k]$ for each k, and

$$\begin{aligned}
\mu_0(\bigvee_{k\geqslant 1} [A_k]) = \mu_0(\bigvee_{k\geqslant 1} [B_k]) = \mu_0([\bigcup_{k\geqslant 1} B_k]) & \qquad \text{(Completeness of } \vee\text{)} \\
= m(\bigcup_{k\geqslant 1} B_k) = \sum_{k\geqslant 1} m(B_k) = \sum_{k\geqslant 1} \mu_0([B_k]) & \qquad \text{(Since } \mu_0([A]) = m(A)\text{)} \\
= \sum_{k\geqslant 1} \mu_0([A_k]). & \qquad \text{(Since } [A_k] = [B_k]\text{)} \qquad \square
\end{aligned}$$

This approach clearly solves the predicament of *Big Cantor* and *Small Cantor* as no closure/interior operator is induced during the process of $\vee$ and $\wedge$ with respect to Lebesgue measure algebra. In the measure theoretic sense, this is a satisfying result since we can recover all Borel measurable pointy functions up to a null set of difference.

Arntzenius's approach meets the conditions for Roeper's axiomatizations, but in order to evaluate its compatibility, a topology must first be given to the Lebesgue measure algebra. Arntzenius's construction turns out to satisfy only nine out of the ten axioms by Roeper, failing the last one, which roughly says "if A sits 'strictly inside' B, then there exists a C such that A sits 'strictly inside C, and C sits 'strictly inside' B." Note that this is essentially the condition for Topological Gunk.

Formally, for two Borel sets A, B, Arntzenius defined them to be **connected** if there exists a point p such that any open set containing p intersects both A and B in a non-null region. A region A sit "strictly inside" B if A is not connected with the complement of B.

Proof. To construct a counterexample that breaks Roeper's last axiom, let us once again consider the Fat Cantor Set $\mathcal{C}$. Our goal is to show that any non-null subset of $\mathcal{C}$ is connected to $\mathcal{C}^c$. Recall the definition of $d(x, S)$ at the beginning of the introduction of Fat Cantor Set. Let x be any point in any non-null subset of $\mathcal{C}$ and let $\epsilon > 0$ be given. It follows that there exists a sufficiently large k such that $d(x, \bigcup_{i=1}^{k} I_i) < \epsilon/2$. This means $(x - \epsilon/2, x + \epsilon/2) \cap \bigcup_{i=1}^{k} I_i$ is nonempty. Since $\bigcup_{i=1}^{k} I_i$ is a union of intervals, all with lengths $\geqslant 4^{-k} > 0$, it follows that $(x - \epsilon, x + \epsilon) \cap \bigcup_{i=1}^{k} I_i$ contains at least an interval of positive length. This proves that any arbitrary non-null subset of $\mathcal{C}$ is connected to $\mathcal{C}^c$, so $\mathcal{C}$ violates Roeper's last axiom, and therefore Arntzenius's model does not satisfy Topological Gunk. □

4 An Exploration of Gunky Probability Spaces

If the structure of a gunky space can be used to recapture the essential structures of a pointy space, how about modeling events defined on gunky spaces? In this section, we attempt to recover methods of abstract integration of real-valued random variables from a point-less perspective. Certainly, without appealing to points, some familiar notions in pointy probability theory are now unaccessible. For example, we cannot pursue pointwise limits. And we need new modes of convergence that is compatible with the spatial structure of gunky spaces.

In this section, we begin by exploring the possibilities of extending a Boolean algebra so that it becomes possible to define a countably additive probability measure. With a "nice" probability space as such defined, we then explore the aspects of defining random variable, as well as integration and convergence on them. Like many standard treatments of probability theory, we will begin with the easiest, most well-behaved types of random variables, and work our ways upward toward more abstract, general random variables.

4.1 Defining a Complete, Countably Additive Probability Measure

Let us consider a Boolean algebra $\mathfrak{B}$ and first define a finitely additive probability measure on it. The basic notions are analogous to that of a standard pointy probability theory.

Definition 4.1. *Let $\mathfrak{B}$ be Boolean algebra with $\mathbf{0} = \emptyset$ and $\mathbf{1}$, as well as join operator $\vee$, meet $\wedge$, and complement defined. We define a "gunky" **probability law** $\mathbb{P}$ on $\mathfrak{B}$ to be a real-valued function that is:*

(1) Strictly positive: $\mathbb{P}(x) \geqslant 0$ and $\mathbb{P}(x) = 0$ if and only if $x = \mathbf{0}$, so that every region has positive "size";

(2) $\mathbb{P}(x) \leqslant 1$ and $\mathbb{P}(x) = 1$ if and only if $x = \mathbf{1}$; and

(3) $\mathbb{P}(x \vee y) = \mathbb{P}(x) + \mathbb{P}(y)$ if $x \wedge y = \emptyset$.

Without having looked much into the structure of "gunky" regions, we cannot venture to impose assumptions that are two strong. But one thing we would like to have, with respect to measure-theoretic gunk, is that all regions have positive measure. Nevertheless, right away we can derive a list of intuitive results that align with the pointy probability:

- $\mathbb{P}$ is finitely additive.
- Define a partial order $\leqslant$ by non-strict set inclusion and $<$ its strict counterpart. Then if $x \leqslant y$ (resp. $x < y$), $\mathbb{P}(x) \leqslant \mathbb{P}(y)$ (resp. $\mathbb{P}(x) < \mathbb{P}(y)$).
- Boolean rings. We can define two more operations on $\mathfrak{B}$: $+$ and $\cdot$ so that $(\mathfrak{B}, +, \cdot)$ forms a ring. For $x, y \in \mathfrak{B}$, define $x + y$ to be the **symmetric difference** $(x \wedge -y) \vee (-x \wedge y)$, and define $x \cdot y$ to be $x \wedge y$. The zero of this ring corresponds to $\mathbf{0} = \emptyset$ and the multiplicative identity corresponds to $\mathbf{1}$. It is also worth noting that this ring is **idempotent**: $x \cdot x = x$ for all x.
- Difference: define $x - y$ to be the unique element z such that $z \wedge y = 0$ and $z \vee y = x$. It follows that if $x \leqslant y$ then $\mathbb{P}(x) + \mathbb{P}(y - x) = \mathbb{P}(y)$.
- Inclusion-exclusion. $\mathbb{P}(x) + \mathbb{P}(y) = \mathbb{P}(x \vee y) + \mathbb{P}(x \wedge y)$.

  Proof. Observe that $x \vee y = (x \wedge y) \vee (x - x \wedge y) \vee (y - x \wedge y)$, and that

  $$\mathbb{P}(x - x \wedge y) + \mathbb{P}(y - x \wedge y) = \mathbb{P}(x) + \mathbb{P}(y) - 2\mathbb{P}(x \wedge y).$$

  Rearranging the original equation gives the desired result. □

Of course, we would want stronger formulations of the gunky probability measure — right now it is neither countably additive nor guaranteed to be complete. To approach this, we will define a metric space $(\mathfrak{B}, d)$ from $(\mathfrak{B}, \mathbb{P})$ and define the completion $\overline{\mathfrak{B}}$ of $\mathfrak{B}$. Naturally, a distance metric on elements of

$\mathfrak{B}$ measures how "far" elements are between each other, a natural candidate of which is the symmetric difference $+$. Hence we define $d(x, y) := \mathbb{P}(x + y)$.

It is easy to verify that $d(\cdot, \cdot)$ indeed defines a metric. $d(x, y) = 0$ if and only if $\mathbb{P}(x + y) = 0$, if and only if $x+y = \emptyset$, or equivalently $x = y$. Symmetry is clear. For triangle inequality, note that $a+b \leqslant a \vee b$, so

$$\begin{aligned} d(x, y) = \mathbb{P}(x + y) = \mathbb{P}((x + y) + (z + z)) = \mathbb{P}((x + z) + (y + z)) \\ \leqslant \mathbb{P}((x + z) \vee (y + z)) = \mathbb{P}(x + z) + \mathbb{P}(y + z) = d(x, z) + d(y, z). \end{aligned}$$

With a metric defined, we obtain our first mode of convergence, which we call $\mathfrak{B}$**-convergence**.

Definition 4.2. *We say a sequence $\{x_n\}$ $\mathfrak{B}$-**converges** to a $\mathfrak{B}$-**limit** $x \in \mathfrak{B}$ if $d(x_n, x) \to 0$, or equivalently $\mathbb{P}(x_n + x) \to 0$. Similarly, we say $\{x_n\}$ is $\mathfrak{B}$-**Cauchy** if $\lim_{n\to\infty} \sup_{i,j\geqslant n} d(x_i, x_j) \to 0$, and like usual, we say $(\mathfrak{B}, d)$ is **complete** if every $\mathfrak{B}$-Cauchy sequence also $\mathfrak{B}$-converges in the space.*

As one may suspect, not every space defined in this way is complete. Consider again the example of Fat Cantor Set in RO($\mathbb{R}$), where $\mathbb{P}$ is the Lebesgue measure. By construction, $d(\bigcup_{i=1}^{k-1} I_i, \bigcup_{i=1}^{k} I_i) = 2^{-k}$, so $\{\bigcup_{i=1}^{k} I_i\}_{k\geqslant 1}$ forms a Cauchy sequence with respect to this metric space. Yet this sequence has no limit, since otherwise $\bigcup_{k\geqslant 1} I_k = I\backslash\mathcal{C}$ would have been an open regular set, which we have shown is false, since the interior of its closure is just $[0.1]$.

Our next goal, naturally, would be to complete $(\mathfrak{B}, d)$, since it is well known that every metric space can be completed. The process has nothing exotic in it — it is a standard application of Cauchy completion [11].

We consider $\mathfrak{C}$, the collection of $\mathfrak{B}$-Cauchy sequences in $(\mathfrak{B}, d)$. Let us first clarify the algebraic structures defined on $\mathfrak{C}$. Two sequences $\{x_n\}, \{y_n\}$ are considered the same if they agree term-wise. Viewing $\mathfrak{C}$ as a Boolean ring, the operations are defined via:

- Multiplicative identity (one): the constant sequence of $\mathbf{1}$'s.
- Additive identity (zero): the constant sequence of $\mathbf{0} = \emptyset$'s.
- Addition is defined term-wise: $\{x_n\} + \{y_n\} = \{x_n + y_n\}$ (where the latter is addition defined on $(\mathfrak{B}, +, \cdot)$).
- Multiplication is defined term-wise too: $\{x_n\} \cdot \{y_n\} = \{x_n \cdot y_n\}$ ($\cdot$ from $(\mathfrak{B}, +, \cdot)$ too).

It follows that if we define $\vee$ and $\wedge$ on $\mathfrak{C}$ by

$$\{x_n\} \wedge \{y_n\} := \{x_n\} \cdot \{y_n\} \qquad \text{and} \qquad \{x_n\} \vee \{y_n\} := \{x_n\} + \{y_n\} + \{x_n\} \cdot \{y_n\},$$

then

$$\{x_n\} \vee \{y_n\} = \{x_n \vee y_n\} \qquad \{x_n\} \wedge \{y_n\} = \{x_n \wedge y_n\} \qquad \{x_n\}^c = \{\mathbf{1}\} + \{x_n\} = \{x_n^c\}.$$

That is, $\mathfrak{C}$ can be viewed as a Boolean algebra with respect to these operations.

We say two Cauchy sequences $\{x_n\}, \{y_n\}$ are **co-Cauchy** if $d(x_n, y_n) \to 0$. Define $\overline{\mathfrak{B}}$ to be $\mathfrak{C}$ modulo the equivalence relation of being co-Cauchy. Alternatively, this can be characterized by $\mathfrak{C}$ modulo the ideal $\mathfrak{C}_0$ of sequences with $\mathfrak{B}$-limit $\emptyset$. We write $\overline{\mathfrak{B}} = \mathfrak{C}/\mathfrak{C}_0$. (Our approach to fixing the completeness (and as we later show, countable additivity of $\mathbb{P}$) of $\mathfrak{B}$ is similar to how Arntzenius [1] attempted to fix the issue of countable additivity on a gunky $\mathbb{R}^n$.) Finally, we define a new metric $\overline{d}$ on $\overline{\mathfrak{B}}$ by

$$\overline{d}(x, y) = \overline{d}([\{x_n\}], [\{y_n\}]) := \lim_{n \to \infty} d(x_n, y_n).$$

This is a well-defined metric because of how we constructed our equivalence classes: if $\{y_n\}$ and $\{z_n\}$ are co-Cauchy and $\lim d(x_n, y_n) \to 0$, then $\lim(x_n, z_n) \leqslant \lim d(x_n, y_n) + \lim d(y_n, z_n) = \lim d(x_n, y_n) = 0$. It is, therefore, natural to define a probability law $\overline{\mathbb{P}}$ on $\overline{\mathfrak{B}}$ by

$$\overline{\mathbb{P}}([\{x_n\}]) := \lim_{n\to\infty} \mathbb{P}(x_n).$$

Theorem 4.3. *$\overline{\mathfrak{B}}$ is complete with respect to $\overline{d}$, and $\overline{\mathbb{P}}$ is a countably additive probability measure on $\overline{\mathfrak{B}}$.*

Proof that $\overline{\mathfrak{B}}$ is complete. Let $\{[C_k]\}_{k\geqslant 1}$ be a Cauchy sequence on $(\overline{\mathfrak{B}}, \overline{d})$. Note by definition each C_k is a *sequence* of $\mathfrak{B}$. For each k, discard early terms of C_k, leaving C'_k, the collection of sufficiently late terms so that

$$\sup_{x,y\in C'_k} d(x, y) < \frac{1}{k}.$$

Since a subsequence of a Cauchy sequence is co-Cauchy with the original sequence, for each k, C_k and C'_k belong to the same equivalence class, and therefore $\{[C_k]\}_{k\geqslant 1} = \{[C'_k]\}_{k\geqslant 1}$. Define $c_{k,n}$ to be the n^{th} term of the modified sequence C'_k. Now construct a sequence $X = \{x_n\} \subset \mathfrak{B}$ by setting $x_k = c_{k,k}$, the diagonal sequence of $\{C'_k\}$. The proof is done if we show that $X \in \mathfrak{C}$ and that $\{[C_k]\}$ converges to $[X]$ with respect to $(\overline{\mathfrak{B}}, \overline{d})$.

To show X is $\mathfrak{B}$-Cauchy, let $\epsilon > 0$ be given and let N be sufficiently large so that $\sup_{k,j>N} \overline{d}(C_k, C_j) < \epsilon/3$. Then, for $k, j > N$ and any n,

$$\begin{aligned} d(x_k, x_j) = d(c_{k,k}, c_{j,j}) &\leqslant d(c_{k,k}, c_{k,n}) + d(c_{k,n}, c_{j,n}) + d(c_{j,n}, c_{j,j}) \\ &\leqslant 1/k + 1/j + d(c_{k,n}, c_{j,n}). \end{aligned}$$

Assuming $N > 3\epsilon^{-1}, 1/k+1/j < 2\epsilon/3$. On the other hand by assumption $\overline{d}(C_k, C_j) = \lim_n d(c_{k,n}, c_{j,n}) < \epsilon/3$, so for sufficiently large n, $d(c_{k,n}, c_{j,n}) < \epsilon/3$ as well. Since the above triangle inequality holds for arbitrary n, we conclude that $d(x_k, x_j) < \epsilon$, and therefore X is $\mathfrak{B}$-Cauchy.

Finally, to show $\{[C_k]\}$ converges to $[X]$, for any $\epsilon > 0$, we pick N sufficiently large such that $\sup_{k,j>N} d(x_k, x_j) < \epsilon/2$. Then

$$d(c_{k,j}, x_j) \leqslant d(c_{k,j}, c_{k,k}) + d(c_{k,k}, x_j) = d(c_{k,j}, c_{k,k}) + d(x_k, x_j) \leqslant 1/k + \epsilon/2 < \epsilon$$

if $N > 2\epsilon^{-1}$. Taking limit in j, then in k, we see $\overline{d}([C_k], [X]) = 0$, completing the proof that $\overline{\mathfrak{B}}$ is complete. □

Proof that $\overline{\mathbb{P}}$ is a countably additive probability measure. It is clear that $\overline{\mathbb{P}}$ is a probability measure, since it inherits the structure of $\mathbb{P}$. So we will focus on proving its countable additivity.

Let us first show that $\overline{\mathbb{P}}$ is finitely additive. In particular let $[\{x_n\}]$ and $[\{y_n\}]$ be given. Assume they are disjoint in the sense that $[\{x_n\}] \wedge [\{y_n\}] = [\{\emptyset\}]$. Proving additivity in this case is straightforward by definition:

$$\begin{aligned}
\overline{\mathbb{P}}([\{x_n\}] \vee [\{y_n\}]) &= \overline{\mathbb{P}}([\{x_n \vee y_n\}]) = \lim_{n\to\infty} \mathbb{P}(x_n \vee y_n) && \text{(by definition)} \\
&= \lim_{n\to\infty} (\mathbb{P}(x_n) + \mathbb{P}(y_n) - \mathbb{P}(x_n \wedge y_n)) && \text{(Inclusion-exclusion)} \\
&= \lim_{n\to\infty} \mathbb{P}(x_n) + \lim_{n\to\infty} \mathbb{P}(y_n) - \underbrace{\lim_{n\to\infty} \mathbb{P}(x_n \wedge y_n)}_{=0} \\
&= \overline{\mathbb{P}}([\{x_n\}]) + \overline{\mathbb{P}}([\{y_n\}])
\end{aligned}$$

To prove countable additivity, let $\{[C_k]\} \subset \overline{\mathfrak{B}}$ be pairwise disjoint, where each C_k is a $\mathfrak{B}$-Cauchy sequence. We assume that their infinite join $[C] = \bigvee_{k\geqslant 1} [C_k]$ exists in $\overline{\mathfrak{B}}$, and the goal is to show $\overline{\mathbb{P}}([C]) = \sum_{k\geqslant 1} \overline{\mathbb{P}}([C_k])$. It is known in standard pointy measure theory that finite additivity, combined

with bounded continuity from above, implies countable additivity. Here we adopt a similar approach. To this end, let us consider another sequence with $[X_k] \geqslant [X_{k+1}]$ for each k, and that $\bigwedge_{k\geqslant 1}[X_k] = [\{\emptyset\}]$ (with respect to $\overline{d}$).

Monotonicity of $\overline{\mathbb{P}}$ implies that $\lim \overline{\mathbb{P}}([X_k])$ exists and the sequence is in particular Cauchy. Since $\overline{\mathfrak{B}}$ is complete, $[X_k]$ also $\overline{\mathfrak{B}}$-converges to some limit which we call $[\overline{X}]$. Our goal is, of course, to show that $[\overline{X}] = [\{\emptyset\}]$, but this is obvious: for each $k \geqslant 1$,

$$\begin{aligned}
[\overline{X}] \wedge [X_k] &= \lim_{n\to\infty} [X_n] \wedge \lim_{n\to\infty} [X_k] && \text{(treating } \{[X_k]_{n\geqslant 1}\} \text{ as constant sequence in } n) \\
&= \lim_{n\to\infty} ([X_n] \wedge [X_k]) \\
&= \lim_{n\to\infty} [X_n] = [\overline{X}]. && ([X_k] \text{ is monotonically decreasing})
\end{aligned}$$

This means $[\overline{X}] \leqslant [X_k]$ for each k. Since $\bigwedge_{k\geqslant 1}[X_k] = [\{\emptyset\}]$ we conclude that $[\overline{X}] = [\{\emptyset\}]$, and so $\overline{\mathbb{P}}([X_k]) = 0$. This shows continuity of $\overline{\mathbb{P}}$ at the zero of $\overline{\mathfrak{B}}$.

To complete the proof, intuitively we can "translate" and "invert" the monotonicity and limit. Formally, we can define $[X_k] := [C] + \bigvee_{i\leqslant k}[C_i]$ and $[X] = \bigwedge_{k\geqslant 1}[X_k]$. Here, X_k is the "tail join" of $[C_k]$'s that is disjoint from the early part $\bigvee_{i\leqslant k}[C_i]$, and $[X]$ in some informal sense the lim inf of the $[C_k]$'s. If $[X] \neq [\{\emptyset\}]$ then it is disjoint from any of the $[C_i]$'s, so $[X] \wedge [C] = [\{\emptyset\}]$. Since $[X_k] \leqslant [C]$ it follows that $[X] \wedge [X_k] = [\{\emptyset\}]$ as well. But on the other hand $[X]$ is defined to be the infinite meet of the $[X_k]$'s, so $[X] \wedge [X_k] = [X]$ for every k. Therefore $[X] = [\{\emptyset\}]$, and by continuity at zero, $\overline{\mathbb{P}}([X]) = 0$. Using definition,

$$0 = \overline{\mathbb{P}}([X]) = \lim_{k\to\infty} \overline{\mathbb{P}}([X_k]) = \lim_{k\to\infty} \overline{\mathbb{P}}([C] + \bigvee_{i=1}^{k} [C_i]) = \overline{\mathbb{P}}([C]) - \lim_{k\to\infty} \overline{\mathbb{P}}(\bigvee_{i=1}^{k} [C_i])$$

and the proof is complete! □

4.2 Elementary and Generalized Random Variables

Now that we have shown that every Boolean algebra can be extended to a complete one equipped with a countably additive probability, let us abuse notation and assume from now on that $\mathfrak{B}$ is complete and $\mathbb{P}$ countably additive on it, with all previous operations defined, along with $\mathbf{0} = \emptyset$ and $\mathbf{1}$. Our next goal would be to define random variables and integrations (expected values) on them. Without directly working with points, we naturally consider "random variables" characterized by their values on various regions. To this end, we say a collection of elements $\{x_i\}_{i\in I}$, countable or finite, **partitions** $\mathfrak{B}$ if they are pairwise disjoint, i.e., $x_i \wedge x_j = \emptyset$ for $i \neq j$, and $\bigvee_{i\in I} x_i = \mathbf{1}$. It follows by countable additivity that any partition $\{x_i\}_{i\in I}$ satisfies $\sum_{i\in I} \mathbb{P}(x_i) = 1$. The following definitions are straightforward:

Definition 4.4. *Let $(\mathfrak{B}, \mathbb{P})$ be a complete space. We define the following types of **random variables**:*

- *A **simple random variable** (resp. **elementary random variable**) is a real-valued function X defined on a finite (resp. countable) partition $\{x_n\}$ of $\mathfrak{B}$ that maps each x_i to a constant value. We denote the collection of elementary random variables by $\mathfrak{E}$.*
- *A **constant random variable** is a simple random variable defined on $\{\mathbf{1}, \emptyset\}$.*
- *Given $x \in \mathfrak{B}$, a corresponding **indicator random variable** I_x can be defined as the simple random variable on $\{x, x^c\}$ with $I_x(x) = 1$ and $I_x(x^c) = 0$.*

For simplicity, unless otherwise stated, when given X and is corresponding partition $\{x_n\}$, we assume X does not take duplicate values on different regions: if $x_i \neq x_j$ then $X(x_j) \neq X(x_j)$, for otherwise we can always take out x_i, x_j from $\{x_n\}$ and insert $x_i \vee x_j$ into it.

In some sense, we are defining elementary random variables based on their indicator decomposition — we will justify this statement later. For now, we compare two random variables X on $\{x_n\}$ and Y on $\{y_n\}$ by considering the following "finer" partition consisting of all non-zero meets of $x_i \wedge y_j$. Let us write this as $\{z_m\} := \{x_n\} \oplus \{y_n\}$. It is easy to see that $\{z_m\}$ is still a countable partition of $\mathfrak{B}$. Then both X

and Y can be defined with respect to $\{z_m\}$ (though this violates our just-stated no duplicate assumption), and we define their sum to be the elementary random variable

$$X \leqslant Y \text{ if } X(z_i) \leqslant Y(z_i) \text{ for each } z_i \in \{z_m\} = \{x_n\} \oplus \{y_n\}.$$

We may define $\geqslant$ analogously. Elementary arithmetic of X and Y can be defined via $\{z_m\}$ as well, and the idea behind each should be self-explanatory:

- Addition: $X + Y$ is defined to be the elementary random variable on $\{z_m\}$ by $(X + Y)(z_i) := X(z_i) + Y(z_i)$, or equivalently, $(X + Y)(x_i \wedge y_j) := X(x_i) + Y(y_j)$.
- Scalar multiplication: given $c \in \mathbb{R}$, define the scalar multiple cX as $(cX)(x_i) := cX(x_i)$ for each i.
- Multiplication: XY is defined to be the elementary random variable on $\{z_m\}$ by $(XY)(z_i) := X(z_i)Y(z_i)$, or equivalently, $(X + Y)(x_i \wedge y_j) := X(x_i)Y(y_j)$.

Many nice properties of elementary random variables follow from these definitions: symmetry, linearity, distributivity, and so on. We define the "region-wise" minimum, $X \wedge Y$ and $X \vee Y$, by

$$(X \wedge Y)(z_i) = \min(X(z_i), Y(z_i))$$

or equivalently $(X \wedge Y)(x_i \wedge y_j) = \min(X(x_i), Y(y_j))$, and likewise a "region-wise" maximum $X \vee Y$. Next, we define the positive and negative parts of X, denoted X^+, X^-, by

$$X^+ := X \vee 0 \qquad \text{and} \qquad X^- = -(X \wedge 0)$$

where 0 here represents the constant random variable taking value 0. It follows that $X = X^+ - X^-$ and we define the absolute value of X to be $|X| = X^+ \vee X^-$.

So far, we have analyzed simple random variables on $(\mathfrak{B}, \mathbb{P})$ and have shown that they behave much like simple random variables defined in a pointy setting, e.g., a probability space defined on Borel sets in $\mathbb{R}^n$. With finite addition of simple random variables defined, given X defined on $\{x_i\}_{i=1}^n$, we can naturally represent X as a linear combination of indicator random variables $X = \sum_{i=1}^n X(x_i) I_{x_i}$. Clearly, we would want a similar expression for elementary random variables, where the finite sum is replaced by an infinite sum. To do so, we need to justify the limit and before that, define what it means for a sequence of (elementary) random variables to converge.

Definition 4.5. *We say a sequence of elementary random variables $\{X_n\} \subset \mathfrak{E}$ **converges** to X, written $X_n \to X$, if there exists a decreasing sequence $\{Y_n\} \subset \mathfrak{E}$ such that*

$$\bigwedge_{k \geqslant 1} Y_k \text{ exists and equals } 0 \qquad \text{and} \qquad |X_n - X| \leqslant Y_n.$$

*Analogously, we define $\{X_n\} \subset \mathfrak{E}$ to be **Cauchy** if there exists a decreasing sequence $\{Y_n\} \subset \mathfrak{E}$ such that for each n and all $i, j > n$, $|X_i, X_j| \leqslant Y_n$.*

Intuitively, this is a relatively weak sense of convergence — to draw some analogy to the pointy spaces, this notion is somewhat similar to pointwise convergence: $\bigwedge_{k \geqslant 1} Y_k$ is the infinite term-wise minimum of the sequence $\{Y_n\}$, so if that equals 0, then $Y_n \to 0$ pointwise. It is however slightly stronger than pointwise convergence, in the sense that it still imposes some kind of uniform bound. In pointy probability theory a famous example showing almost sure convergence does not imply convergence in expectation is $X_n = n \cdot I_{[0,1/n]}$, where because of the unboundedneess of n, a point mass "escapes" at 0. Here, such things cannot happen because we required $\{Y_n\}$ to decrease a priori.

Returning to our main goal, let $X \in \mathfrak{E}$ be defined on a countable $\{x_n\}$. To show that X can be represented by $\sum_{i=1}^\infty X(x_i) I_{x_i}$, we approximate this infinite sum by finite sums $\sum_{i=1}^n X(x_i) I_{x_i}$. We define the partial joins $y_n = \bigvee_{i=1}^n x_i$ and notice that the partial sums $\sum_{i=1}^n X(x_i) I_{x_i} = I_{y_n} X$. We apply this

convergence to I_{y_n} and show it converges to $\mathbf{1}$, the constant random variable taking value 1 (or equivalently the indicator of $\mathbf{1} \in \mathfrak{B}$):

$$|I_{y_n} - \mathbf{1}| = \bigwedge_{i=1}^{n} x_i^c \qquad \text{and} \qquad \bigwedge_{i=1}^{n} x_i^c = \emptyset$$

since the infinite join does not contain any x_i, but the x_i's make up the entire space. Therefore $I_{y_n} \to \mathbf{1}$. By the same token we see that $I_{y_n} X \to X$, so the proof is complete, and we claim that

$$X = \sum_{k \geqslant 1} X(x_k) I_{x_k}$$

is indeed an indicator representation of X.

Our next question is, how do we define a more general form of random variables using approximations of elementary random variables? In a pointy setting, given $X \geqslant 0$, it is well-known that X can be approximated pointwise by a sequence $\{X_n\}$ of monotone increasing simple random variables by considering $X\mathbf{1}[X \leqslant n]$ and rounding the values of this variable down modulo 2^{-n}, so that the resulting random variable only takes values of multiples of 2^{-n}, with minimum 0 and maximum n.

In our setting, we also consider convergence of elementary random variables, but instead of appealing to dyadic numbers we again consider equivalence classes of convergent elementary random variables based on the "limit," as we once did during the process of Cauchy completing our initial Boolean algebra.

To achieve this, we first define sequence-wise operations of sequences of elementary random variables. Abusing the notations, given $\{X_n\}, \{Y_n\} \subset \mathfrak{E}$, we define addition, multiplication, scalar multiplication, join / pointwise maximum, and meet / pointwise minimum element-wise, i.e., $\{X_n\} + \{Y_n\} := \{X_n + Y_n\}$ and so on. Once again, we consider the quotient space of Cauchy sequences modulo sequences that converge to zero. Notation-wise, we define $\mathfrak{C}(\mathfrak{E})$ to be the space of Cauchy sequences of elementary random

variables, with operations defined component-wise, and $\mathfrak{C}_0(\mathfrak{E})$ the space of sequences of elementary random variables converging to 0 (the zero constant variable). Then we define the space of **general random variables** (or just **random variable**) to be $\mathfrak{X} := \mathfrak{C}(\mathfrak{E})/\mathfrak{C}_0(\mathfrak{E})$. In other words, we identify each general random variable with the collection of all Cauchy sequences whose "limit" agree, and we shall a general random variable by an equivalence class written as $X = [\{X_n\}]$. This way, $\mathfrak{X}$ has many "nice" properties that align with our intuition. Furthermore, when defining $\mathfrak{X}$ we used a notion of convergence akin to pointwise convergence, but Kappos [7] showed that if X_n converges to X, then there exists another sequence $\{Y_n\}$ that converges to X **uniformly**. That is, there exist constant random variables $\{c_n\}$, $c_n \downarrow 0$, such that $|X_n - X| \leqslant c_n$. In the following sections we consider several properties of random variables defined in this manner, and show that they align nicely with our intuition and pointy counterparts.

We first explore the notion of distribution. In a pointy setting it is well-known that a random variable X is characterized by its distribution function $F(x) = \mathbb{P}(X < x)$. Here, we will show that a similar notions holds on both $\mathfrak{E}$ and $\mathfrak{X}$. Firstly, for any elementary random variable $X \in \mathfrak{E}$ defined on $\{x_i\}$and $c \in \mathbb{R}$, we define

$$\overline{D}_X(c) := [X \leqslant c] = \bigvee_{X(x_i)\leqslant c} x_i \qquad \text{and} \qquad \underline{D}_X(c) := [X < c] = \bigvee_{X(x_i)<c} x_i.$$

Both $\overline{D}_X$ and $\underline{D}_X$ are monotone increasing functions, in the sense that if $c_1 < c_2$ then $\bigvee_{X(x_i)\leqslant c_1} x_i$ is contained in $\bigvee_{X(x_i)\leqslant c_2} x_i$. Since $\{x_i\}$ can be labeled arbitrarily, we may WLOG assume that $X(x_i) < X(x_j)$ for $i < j$, so that

$$\bigvee_{X(x_i)\leqslant c} x_i = \bigvee_{i=1}^{k(c)} x_i$$

where $k(c)$ is the largest index, 0 if nonexistent, such that $X(x_i) \leqslant c$, and likewise for strict inequality and $\underline{D}_X$. Clearly $\lim_{c\to-\infty} k(c) = 0$ and $\lim_{c\to\infty} k(c) = \infty$. But then

$$\lim_{c\to-\infty} \overline{D}_X(c) = \lim_{c\to-\infty} \underline{D}_X(c) = \lim_{k(c)\to 0} \bigvee_{i=1}^{k(c)} x_i = \emptyset \qquad \text{and} \qquad \lim_{c\to\infty} \overline{D}_X(c) = \lim_{c\to\infty} \underline{D}_X(c) = \mathbf{1}.$$

What happens if we now consider $\{X_n\} \subset \mathfrak{E}$ and the corresponding $X = [\{X_n\}] \in \mathfrak{X}$? We can extend the definition of $\overline{D}, \underline{D}$ by dsefining

$$\overline{D}_X(c) := \limsup_{n\to\infty} \overline{D}_{X_n}(c) \qquad \text{and} \qquad \underline{D}_X(c) := \liminf_{n\to\infty} \underline{D}_{X_n}(c),$$

where the lim sup, lim inf are defined with nested operations of $\vee$ and $\wedge$. This definition is compatible with the version defined on $\mathfrak{E}$ because for any $X \in \mathfrak{E}$ we can simply consider the constant sequence $\{X, X, \ldots\}$. To show that this definition is well-defined, we consider $\{X_n\} \in \mathfrak{C}(\mathfrak{E})$ Cauchy and $\{Y_n\} \in \mathfrak{C}_0(\mathfrak{E})$, a sequence converging to 0. Since we may replace $\{Y_n\}$ with a sequence uniformly converging to 0, the difference between X_n and $X_n + Y_n$ is uniformly bounded, and so are $\overline{D}_{X_n}$ and $\overline{D}_{X_n+Y_n}$ (and likewise for $\underline{D}$). Letting $\epsilon \downarrow 0$ the claim follows. Therefore,

Theorem 4.6. *Any $X \in \mathfrak{X}$ is uniquely characterized by its **distribution function** $F : \mathbb{R} \to \mathfrak{B}$ defined by $F(c) = [X < c]$. This is a monotone function with limits $F(-\infty) = \lim_{c\to-\infty} F(c) = \emptyset$ and $F(\infty) = \mathbf{1}$. Further, this function is right continuous, i.e., $\lim_{x\downarrow c} F(x) = F(c)$.*

An important result that we can derive via distribution functions is:

Theorem 4.7. *$\mathfrak{X}$ is complete with respect to arbitrary joins and meets taken over a collection of uniformly bounded random variables. In other words, if $\{X_i\}_{i\in\mathfrak{I}} \subset \mathfrak{X}$ is indexed over any arbitrary I, and there exists $M \geqslant 0 \in \mathfrak{X}$ such that $-M \leqslant X_i \leqslant M$ for all i, then both $\bigvee_{i\in I} X_i$ and $\bigwedge_{i\in I} X_i$ exist and are in $\mathfrak{X}$.*

Proof. $\mathfrak{B}$ is complete, so the arbitrary join $\bigwedge_{i\in I} \underline{D}_{X_i}(c)$ exists for each c. Define $\underline{D}_X(c)$ to be the right limits of $\bigwedge_{i\in I} \underline{D}_{X_i}(x)$, i.e.,

$$\underline{D}_X(c) := \lim_{x\downarrow c} \bigwedge_{i\in I} \underline{D}_{X_i}(x).$$

Then $\underline{D}$ satisfies all the criterion for a distribution function. And it follows by completeness of $\mathfrak{B}$ that the random variable corresponding to $\underline{D}_X$ must coincide with $\bigvee_{i\in I} X_i$, so the closure with respect to arbitrary join is proven. The other case $\bigwedge_{i\in I} X_i$ is analogous. □

Note that the assumption of boundedness is necessary. Since if $X \leqslant Y$, $\underline{D}_X \geqslant \underline{D}_Y$, $-M \leqslant X_i \leqslant M$ implies that for all i, $\underline{D}_M \leqslant \underline{D}_{X_i} \leqslant \underline{D}_{-M}$. Without this assumption the claim fails, for we may construct an example where $\underline{D}_X$ becomes uniformly $\mathbf{0}$, breaking $\lim_{c\to\infty} \underline{D}_X(c) = \mathbf{1}$.

4.3 Expected Values and Convergence Theorems

Let us now turn to defining expected values of random variables. As usual, we start with elementary ones.

Definition 4.8. *Let $X \in \mathfrak{E}$ and write X as $\sum_{k\geqslant 1} X(x_k) I_{x_k}$. If $\sum_{k\geqslant 1} |X(x_k)|\mathbb{P}(x_k) < \infty$, then we define the **expected value** of X, written $\mathbb{E}X$, to be $\sum_{k\geqslant 1} X(x_k)\mathbb{P}(x_k)$.*

More generally, given $X \in \mathfrak{X}$, we know that there exists a sequence $\{X_n\} \subset \mathfrak{E}$ converging uniformly to X. We say X possesses an expected value if in addition each X_n possesses an expected value. Finally, we define L^1 to be the space of all random variables X with $\mathbb{E}|X| < \infty$.

Note that not all $X \in \mathfrak{E}$ possess expected values, since not all of them have absolutely convergent indicator representations. It is clear that the collection of elementary variables with well-defined expected values is closed under addition, scalar multiplication, and maximum ($\vee$) and minimum ($\wedge$). As we generalize the notion of expected values to non-elementary random varaibles, the intuition is that uniform

convergence preserves limit of expected values, so the above is well-defined. This is indeed true: uniform convergence, along with triangle inequality, implies

$$|\mathbb{E}X_n - \mathbb{E}X| \leqslant \mathbb{E}|X_n - X| \to 0.$$

The previously mentioned algebraic operators are also well-defined on $\mathfrak{X}$ with respect to expected values, since the definition is a natural extension of the expected value defined on $\mathfrak{E}$. Consequently, L^1 is also closed under addition, scalar multiplication, maximum, and minimum. This definition turns out to be highly compatible with its pointy counterparts, in the sense that many important convergence theorems hold as well. We first prove the

Theorem 4.9 (Dominated Convergence Theorem). *If $X_n \geqslant 0, X_n \in L^1$ is monotone decreasing with $X_n \to 0$, then the expectations converge as well: $\mathbb{E}X_n \downarrow 0$.*

Proof. We will prove this claim via a multi-step procedure: first we show it holds for simple random variables, then elementary random variables, and finally, (general) random variables.

Step 1: simple random variables. Let each X_n be simple, defined on $\{x_{(n,i)}\}_{i=1}^{c(n)}$. Using indicator representation, we write $X_n = \sum_{i=1}^{c(n)} X_n(x_{(n,i)}) I_{x_{(n,i)}}$. Further WLOG assume that for each fixed n, the $x_{(n,i)}$'s are arranged in the decreasing order based on $X_n(x_{(n,i)})$, i.e., $X_n(x_{(n,i)}) \geqslant X_n(x_{(n,i+1)})$. Let $\epsilon > 0$ be given. Our goal is to show that for sufficiently large n, $\mathbb{E}X_n < \epsilon$. The idea behind the proof is that we show as $n \to \infty$, X_n takes value $> \epsilon$ on a sufficiently small region, whose contribution to $\mathbb{E}X_n$ can be controlled, whereas X_n is sufficiently small on the remainder of the space, and consequently its contribution to $\mathbb{E}X_n$ is also controllable.

Let $M = X_1(x_{1,1})$. By monotonicity of both $\{x_{n,1}, x_{n,2}, \ldots\}$ and $\{X_n\}$, we know $M \geqslant \mathbb{E}X_1 \geqslant \mathbb{E}X_n$ for all n. On the other hand, since the values of $X_n(x_{n,u})$ is decreasing, for each n there exists a $d(n)$ such that the first $d(n)$ terms of $X_n(x_{(n,i)})$ is $\geqslant \epsilon/2$ and the remaining $c(n) - d(n)$ terms are $< \epsilon/2$. As

discussed informally before, for each n we consider the partition of space by $\bigvee_{i=1}^{d(n)} x_{n,i}$ and $\bigvee_{i>d(n)} x_{n,i}$. By the convergence assumption $X_n \to 0$, we must have

$$\bigwedge_{k=1}^{\infty} \bigvee_{i=1}^{d(k)} x_{k,i} = \emptyset$$

for no part of X_n can remain above ϵ forever. But we previously showed $\mathbb{P}$'s countable additivity is equivalent to continuity at $\mathbf{0}$, so $\mathbb{P}(\bigvee_{i=1}^{d(k)} x_{k,i}) \to 0$ as $k \to \infty$, which means for sufficiently large k, $\mathbb{P}(\bigvee_{i=1}^{d(k)} x_{k,i}) < \epsilon/2M$. Therefore, for sufficiently large n,

$$\begin{aligned}
\mathbb{E}X_n &= \sum_{i\geqslant 1} X_n(x_{n,i})\mathbb{P}(x_{n,i}) = \sum_{i=1}^{d(n)} \underbrace{X_n(x_{n,i})}_{\leqslant M} \mathbb{P}(x_{n,i}) + \sum_{i>d(n)} \underbrace{X_n(x_{n,i})}_{\leqslant \epsilon/2} \mathbb{P}(x_{n,i}) \\
&\leqslant M \sum_{i=1}^{d(n)} \mathbb{P}(x_{n,i}) + \frac{\epsilon}{2} \sum_{i>d(n)} \mathbb{P}(x_{n,i}) \\
&\leqslant M \cdot \mathbb{P}(\bigvee_{i=1}^{d(n)} x_{n,i}) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \qquad \text{End of Step 1}
\end{aligned}$$

Step 2: elementary random variables. Now let $\{X_n\} \subset \mathfrak{E}$ be a sequence of elementary random variables, each possessing an expected value and converging downward to 0. The idea is to use Step 1 to approximate these X_n's and show that the error term can be carefully controlled.

More formally, since $\mathbb{E}X_n = \sum_{i\geqslant 1} X_n(x_{n,i})\mathbb{P}(x_{n,i}) < \infty$ there exists a $e(n)$ such that $\sum_{i>e(n)} X_n(x_{n,i})\mathbb{P}(x_{n,i}) < \epsilon 2^{-n}$. The first $e(n)$ terms form a simple random variable $X'_n := \sum_{i=1}^{e(n)} X_n(x_{n,i}) I_{x_{(n.i)}}$. Further define $X''_n := \bigwedge_{i=1}^{n} X'_n$ so that X''_n is monotonically decreasing. Since $X_n \to 0$ and $0 \leqslant X''_n \leqslant X'_n \leqslant X_n$, we know $X''_n \to 0$ monotonically as well, and by the previous part $\mathbb{E}X''_n \downarrow 0$.

How about the remainder? To fix the potential issue of X'_n not being monotone we introduced X''_n, but in doing so, controlling $X_n - X''_n$ becomes slightly more involved as well. We inductively prove that

$X_n - X_n''$ can still be controlled. Of course, $X_1' = X_1 - X_1''$. For the inductive step, we use inclusion-exclusion on X_n'' and X_{n+1}:

$$\mathbb{E}(X_{n+1}') + \mathbb{E}(X_n'') = \mathbb{E}(X_{n+1}' \wedge X_n'') + \mathbb{E}(X_{n+1}' \vee X_n'') = \mathbb{E}(X_{n+1}'') + \mathbb{E}(X_{n+1}' \vee X_n'')$$

$$\leqslant \mathbb{E}(X_{n+1}'') + \mathbb{E}(X_{n+1}' \vee X_n') \leqslant \mathbb{E}(X_{n+1}'') + \mathbb{E}X_n.$$

Rearranging gives

$$\mathbb{E}(X_{n+1}') \leqslant \mathbb{E}(X_{n+1}'') + \mathbb{E}X_n - \mathbb{E}X_n' < \mathbb{E}(X_{n+1}'') + \epsilon.$$

This shows $X_n - X_n''$ is not far from $X_n - X_n''$, so $\limsup \mathbb{E}X_n = \limsup(\mathbb{E}X_n'' + \mathbb{E}(X_n - X_n'')) \leqslant \epsilon$. Since ϵ is arbitrary the proof is complete. End of Step 2

Step 3: (general) random variables. Let $\{X_n\} \subset \mathfrak{X}$ be L^1. For each n, let $\{Y_{n,k}\}_{k \geqslant 1}$ be a sequence of elementary random variables with expected values defined that converge uniformly to X_n. Further WLOG assume each $\{Y_{n,k}\}_{k \geqslant 1}$ is decreasing, for we may otherwise consider $\{\bigwedge_{i \geqslant k} Y_{n,k}\}_{k \geqslant 1}$. Consider the vertical sequence $Z_n = \bigwedge_{i=1}^n Y_{i,n}$. On one hand, since $Y_{n,\cdot} \downarrow X_n$ we must have $Z_n \geqslant X_n$. On the other hand, by construction we also know that $Y_{n,k} \geqslant Z_k$ if $n \leqslant k$. Finally, it is clear that Z_n is decreasing since by assumption $Y_{n,k} \geqslant Y_{n,k+1}$. Combining these three identities along with $X_n \to 0$, we see that $\lim Z_n = \lim X_n = 0$. But Z_n is elementary, so by Part 2 $\mathbb{E}Z_n \downarrow 0$. Finally, noting that $Z_n \geqslant X_n$ so $\mathbb{E}Z_n \geqslant \mathbb{E}X_n$, we conclude that $\mathbb{E}X_n \downarrow 0$. □

A direct consequence of DCT is that if $X_n \to X$ is monotonically decreasing and L^1, then $\mathbb{E}X_n \downarrow \mathbb{E}X$ since we can apply the previous proof to $X_n - X$, a sequence of L^1 random variables converging downward to 0. This also gives rise to the:

Theorem 4.10 (Monotone Convergence Theorem). *If $\{X_n\} \subset L^1$ is monotone increasing and $X_n \to X \in L^1$, then $\mathbb{E}X_n \uparrow \mathbb{E}X$.*

Finally, we prove the

Theorem 4.11 (Fatou's Lemma). *Let $\{X_n\} \subset L^1$. If there exists a uniform bound $M \in L^1$ such that $0 \leqslant X_n \leqslant M$ then*

$$\mathbb{E}(\liminf_{n\to\infty} X_n) \leqslant \liminf_{n\to\infty} \mathbb{E}X_n,$$

where $\liminf_{n\to\infty} X_n$ is defined as $\bigvee_{k=1}^{\infty} \bigwedge_{n\geqslant k} X_n$.

Proof. First, $\liminf X_n$ is well-defined because $\bigwedge_{n\geqslant k} X_n$ is bounded for each k, and $\mathfrak{X}$ is closed under arbitrary joins and meets taken over bounded subsets. Further, since $0 \leqslant \liminf X_n \leqslant M$ we see it is also in L^1. On the other hand note that $\bigwedge_{n\geqslant k} X_n$ is strictly increasing to $\liminf_n X_n$ as $n \to \infty$ and is uniformly bounded by M, so by MCT, $\mathbb{E}(\bigwedge_{n\geqslant k} X_n) \uparrow \mathbb{E}(\liminf_n X_n)$. But by definition $\bigwedge_{n\geqslant k} X_n \leqslant X_k$ for each k, so $\mathbb{E}(\bigwedge_{n\geqslant k} X_n) \leqslant \mathbb{E}X_k$ for each k, and we conclude that

$$\mathbb{E}(\liminf_{n\to\infty} X_n) = \lim_{k\to\infty} \mathbb{E}(\bigwedge_{n\geqslant k} X_n) \leqslant \liminf_{n\to\infty} \mathbb{E}X_n. \qquad \square$$

Bibliography

[1] Arntzenius, Frank (2008). Gunk, topology, and measure. In *Oxford Studies in Metaphysics*, 225-247. Vol 4. Oxford University Press, 2008.

[2] Betti, Arianna, and Leob, Iris (2012). On Tarski's foundations of the geometry of solids. *The Bulletin of Symbolic Logic*, 18(2), 230-260.

[3] Forrest, Peper (2003). Nonclassical mereology and its application to sets. *Notre Dame Journal of Formal Logic*, 43, no.2 (2003): 79-94.

[4] Gerla, Giangiacomo (1990). Pointless metric spaces. *The Journal of Symbolic Logic*, vol. 55, no.1, pp. 207-219.

[5] Gerla, Giangiacomo (1995). Pointless geometries. In *Handbook of Incidence Geometry: Buildings and Foundations*, chapter 18, 1015-1053. North-Holland, 1995.

[6] Givant, Steven R., and Mackenzie, Ralph (1986). *Alfred Tarski: Collected Papers*, vol. 1, Birkhäuser, 1986.

[7] Kappos, Demetrios A. (1969). *Probability Algebras and Stochastic Spaces*. Academic Press.

[8] Lando, Tamar, and Scott, Dana (2019). A calculus of regions respecting both measure and topology. *The Journal of Philosophical Logic*, 48, 825-850 (2019).

[9] Levy, Azriel (1979). *Basic Set Theory*. Dover Publications: 2012.

[10] Lewis, David (1991). *Parts of Classes*. Cambridge: Blackwell.

[11] Pugh, Charles C. (2002). *Real Mathematical Analysis*. Springer.

[12] Roeper, Peter (1997). Region-based topology. *Journal of Philosophical Logic*, 26(3), 251-309.

[13] Rudin, Walter (1987). *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill.

[14] Russell, Jeff S. (2008). The structure of gunk: adventures in the ontology of space. In *Oxford studies in metaphysics*, Vol. 4, 248-274. Oxford: Oxford University Press.

[15] Sikorski, Roman (1964). *Boolean Algebras*. Berlin: Springer Verlag.

[16] Skyrms, Brian (1983). Zeno's paradox of measure. In *Physics, Philosophy and Psychoanalysis*, pp. 223-254.

[17] Tarski, Alfred (1929). Les fondements de la géométrie des corps, *Annales de la Socié'e Polonaise de Mathématiques*, 29-34. Reprinted in Givant and Mackenzie (1986).

[18] Whitehead, Alfred N. (1929). *Process and Reality*. New York: Macmillan.