

Towards a formalisation the Kakeya conjecture in $\mathbb{R}^2$

A thesis presented for the degree of
Master of Science

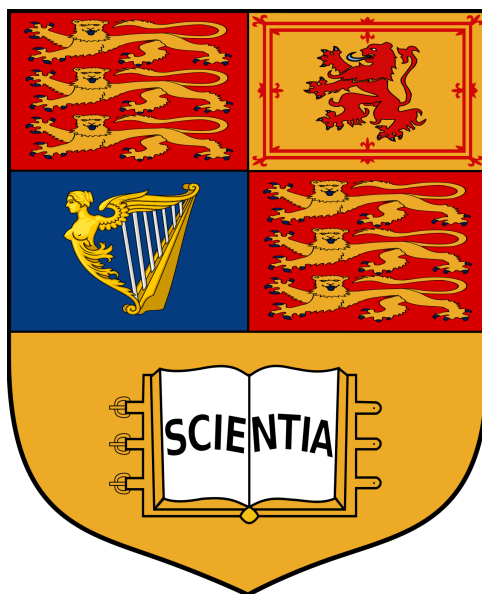
by

Francesco Nishanil Chotuck

CID: 01925507

Supervised by

Dr. Bhavik Mehta



Department of Mathematics
Imperial College London
United Kingdom
September 2025

The work contained in this thesis is my own work unless otherwise stated.

Abstract

A Kakeya set is a subset of $\mathbb{R}^n$ that contains a unit line segment in every direction. A Besicovitch set is a Kakeya set of Lebesgue measure zero. The Kakeya conjecture, one of the central open problems in geometric measure theory, asserts that every Kakeya set has full Hausdorff dimension.

In this thesis we undertake a formalisation in the Lean theorem prover of the existence of Besicovitch sets in $\mathbb{R}^2$, following the construction due to Körner. Our goal is to verify this fundamental result, and also to develop the necessary API in `mathlib` that would allow for the eventual formalisation of the Kakeya conjecture in $\mathbb{R}^2$. The work required us to build a significant amount of infrastructure in analysis, topology, and measure theory in Lean, extending `mathlib` with new results about the Hausdorff distance and Baire category.

As a preliminary step, we provide a complete formalisation of the 1-dimensional Kakeya problem in $\mathbb{R}$, which serves as a simpler case study and illustrates the techniques that are later extended to higher dimensions. The formalisation of the 2-dimensional case remains partial, but the framework established here lays the foundations for further work towards a complete formalisation of the Kakeya conjecture.

*Dedico questa tesi a
Francesco (D'Assisi) Morgese*

Acknowledgements

Hanc quoque dissertationem dicto Sancto Francisco Assisiensi, cuius nomen porto, Italiae patrono caelesti, necnon Episcopo Romano, Summo Pontifici, Italiae Primati atque Civitatis Vaticanae Principi, Papae Francisco, qui mihi inexhaustum afflatus fontem praebet.

First and foremost, I would like to express my deepest gratitude to Bhavik Mehta for many things. I am especially thankful for his generosity in accepting me as a student at a rather late stage in the timeline of this thesis. His patience as I gradually learned the intricacies of Lean, as well as his ability to guide with both rigour and kindness, have been invaluable. He has been not only a dedicated supervisor, mentor, and lecturer, but above all, a true friend—and for this, I remain profoundly grateful.

I owe an immeasurable debt of gratitude to my partner, Hannah Maria Butler, whose constant love, encouragement, and unwavering support have sustained me throughout this journey. Without her presence, understanding, and belief in me, neither this degree nor this thesis would ever have come to fruition.

My heartfelt thanks go also to Yuming Guo, my partner in crime in the pursuit of Lean. Together we navigated its many frustrations, which, though often exasperating, proved in equal measure to be rewarding. The countless hours we spent working side by side, sharing ideas and laughter, and occasionally retreating for a hearty bowl of beef stew, will remain with me as some of the most joyous and memorable moments of this endeavour.

I would also like to thank my dearest friends, in particular Aryan Mangal, Arthur Pierre Louis Aimone, Kishor Sureshkumar, and Jan Olucha Fuentes, whose steadfast support throughout these years has been indispensable. Special thanks are due to Sachchit Prasad, who not only organised the much-needed games nights that provided respite from work, but also offered invaluable guidance in helping me seek professional opportunities.

I am sincerely grateful to the people of Xena, and in particular to Arav Bhattacharyya and Sidharth Hariharan, for their assistance with Lean. My thanks also go to Edison Xie, for the delightful Michelin-star dinners and the wonderful times we shared in the company of friends.

Finally, and most importantly, I extend my deepest thanks to my parents. Their love, sacrifices, and constant encouragement have been the foundation upon which all of this has been built.

Contents

Abstract	ii
Acknowledgements	iv
1 Introduction	1
1.1 The conjecture	1
1.2 Formalising Mathematics	3
1.2.1 Type Theory as a Foundation for Lean	4
1.2.2 Additional Axioms and Classical Reasoning	5
1.3 Structure of the Thesis	5
2 Constructing a Besicovitch set	6
2.1 The Baire Category Theorem	6
2.2 Besicovitch via Baire	7
3 The Keakeya conjecture	11
3.1 The Hausdorff measure and dimension	11
3.2 Hausdorff Tools for the Keakeya Conjecture	14
3.3 The 1-dimensional case	15
3.4 The 2-dimensional case	16
4 Methodology of the Formalisation	17
4.1 Pre-existing Results in <code>mathlib</code>	18
4.2 Defining Besicovitch sets	18
4.3 Körner’s construction	19
4.4 Hausdorff dimension toolkit	28
4.5 Formalising the conjecture	30
5 Conclusion	31
5.1 What was achieved	31
5.2 Broader perspective	32
5.3 Closing remarks	32
References	36

1 Introduction

This thesis attempts the formalisation of Kakeya-type problems in the Lean theorem prover. The Kakeya problem asks how small a subset of Euclidean space can be while still containing a unit line segment in every direction, and the conjecture states that such sets must always have full Hausdorff dimension. In one dimension the statement is immediate, while in two dimensions the behaviour is peculiar: Kakeya sets occupy zero area, yet they are in a precise sense as large as possible, having full Hausdorff dimension.

The motivation for this thesis is twofold. On the one hand, the Kakeya conjecture has long stood as a central problem in analysis. Beyond its geometric charm, the problem connects to some of the most fundamental areas of modern mathematics, including Fourier analysis [11], PDEs [37], and additive combinatorics [28]. Techniques developed in pursuit of Kakeya-type bounds have led to progress on questions as diverse as the behaviour of wave equations [37], restriction estimates [20], and arithmetic problems in finite fields [8]. In this way, Kakeya serves as a kind of testing ground: advances made here often radiate outward, influencing entire branches of mathematics and, indirectly, the many sciences and technologies that rely on them.

On the other hand, the very recent proof by Hong Wang and Joshua Zahl of the 3-dimensional Kakeya conjecture, announced in February 2025, provides a timely backdrop for this work [38]. Their resolution of a case that had resisted progress for decades is a landmark achievement, reaffirming both the depth and the relevance of Kakeya-type questions.

The work proceeds in stages. We begin by introducing and formalising the definitions of Kakeya and Besicovitch sets. Using Körner’s construction, we establish within Lean the existence of Besicovitch sets. We then formalise the Kakeya problem in one dimension, providing a simple but complete case. Finally, we turn to the 2-dimensional conjecture and work towards formalising the tools required for its study.

The contributions of this thesis are the following. It develops a partial formalisation of Körner’s construction of Besicovitch sets and provides a Lean formalisation of the Kakeya problem in one dimension. Along the way, it expands Lean’s community library, `mathlib`, with new results in topology and measure theory, and documents the challenges encountered when translating classical proofs into a formal setting.

In summary, the thesis shows that even the planar Kakeya problem, though long settled, poses significant challenges for mechanisation. Its formalisation highlights both the strengths and the current limitations of Lean in geometric measure theory, while laying the groundwork for future formal developments on higher-dimensional Kakeya conjectures and related problems.

The complete formalisation developed in this thesis is available in a public repository at <https://github.com/FrankieNC/2DKakeyaConjecture>

1.1 The conjecture

The origins of the Kakeya problem trace back to 1917, when Sōichi Kakeya posed a deceptively simple question about the plane: what is the region of the smallest area within which a unit line segment can be rotated through 360° ? Kakeya conjectured that the minimal such region would be a disk of diameter one, reflecting the natural intuition that

the circular symmetry of the disk makes it the most efficient shape [16].

Shortly afterwards, Julius Pál showed that the Kakeya conjecture was incorrect [26]. Pál constructed an explicit example of a smaller region, bounded by a deltoid (see Figure 1), within which a unit line segment can indeed be rotated fully. This “Pál deltoid” has area strictly less than that of the unit disk, disproving Kakeya’s original claim and demonstrating that the disk is not optimal.

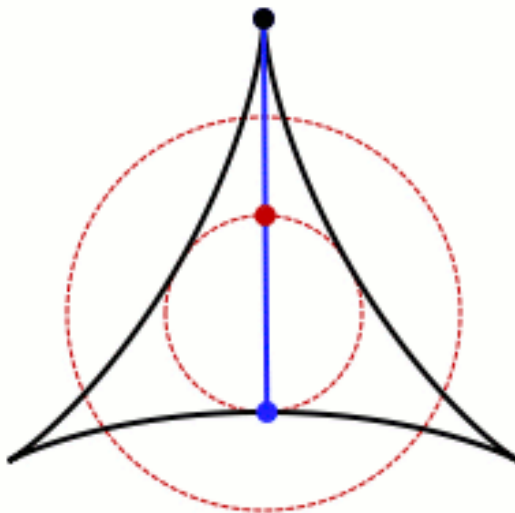


Figure 1: A deltoid-shaped Kakeya set: the black outline represents the boundary, and rotating a unit line segment within the surrounding annulus demonstrates that the set satisfies the Kakeya property. Image source: [Wikimedia Commons \(https://commons.wikimedia.org/wiki/File:Kakeya_needle.gif\)](https://commons.wikimedia.org/wiki/File:Kakeya_needle.gif), licensed under CC BY 2.5.

Sets of this type, first considered by Kakeya in his original question, are now defined as follows.

Definition 1.1 ([9, Section 7.2]). A set, $K \subset \mathbb{R}^n$ is a **Kakeya set** if K contains a unit line segment in every direction.

It was Abram Besicovitch who pushed the Kakeya problem to its most striking conclusion. In 1928, he constructed what are now called *Besicovitch sets*: subsets of the plane of arbitrarily small area—indeed, of Lebesgue measure zero—that nonetheless contain a unit line segment in every direction [3] [4]. This overturned the prevailing intuition of Kakeya, Pál, and their contemporaries, and showed that the problem is far subtler than initially expected.

Besicovitch’s construction sparked a century of subsequent development. During the mid-20th century, variations and refinements of his argument clarified the nature of Besicovitch sets and the methods required to build them. To make this precise, we introduce the formal notion:

Definition 1.2 ([9, Section 7.2]). A set $B \subset \mathbb{R}^n$ is a **Besicovitch set** if it is a Kakeya set of Lebesgue measure zero.

Example 1.3. An explicit illustration of a Besicovitch set in $\mathbb{R}^2$. The set is constructed as a union of line segments arranged to cover every direction while having Lebesgue measure zero.

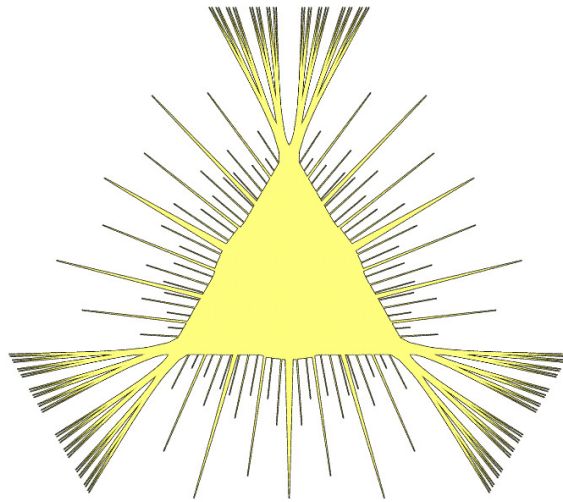


Figure 2: A schematic example of a Besicovitch set in $\mathbb{R}^2$. Image source: [Wikimedia Commons](https://commons.wikimedia.org/wiki/File:KakeyaNeedleSet3.GIF) (<https://commons.wikimedia.org/wiki/File:KakeyaNeedleSet3.GIF>).

In particular, it became clear that while such sets could be negligible from the viewpoint of measure, they could not be “too small” in terms of dimension. The decisive step came in 1971, when Roy Davies proved that every Kakeya set in the plane must have Hausdorff dimension 2 [6]. Thus, although Besicovitch had shown that planar Kakeya sets could have zero area, Davies established that they nonetheless fill the plane in a dimensional sense.

The natural extension of these ideas to higher dimensions leads to the following statement: **Conjecture 1.1** (The Kakeya conjecture). Every Kakeya set in $\mathbb{R}^n$ has full Hausdorff dimension n .

Remark 1.4. The formulation above is the most classical form of the Kakeya conjecture. Over the past century, several equivalent or closely related versions of the conjecture have been studied. These include the statement that Kakeya sets must have full Minkowski dimension, or that the associated maximal function conjecture holds; see for instance the surveys by Wolff [39] and Katz [17].

In $\mathbb{R}$ the conjecture is trivial (see Section 3.3). The case of $\mathbb{R}^2$ is the one of interest in this thesis, since it already displays highly non-trivial behaviour; in the planar setting one can strengthen the statement:

Theorem 1.5 (Kakeya conjecture in $\mathbb{R}^2$). Every Besicovitch set in $\mathbb{R}^2$ has full Hausdorff dimension 2.

This result will serve as the central problem around which this thesis is organised.

1.2 Formalising Mathematics

One of the central themes of this thesis is the idea of *formalising mathematics*: representing definitions, theorems, and proofs in a language precise enough for a computer to check. Unlike the informal arguments found in textbooks or research papers, where intuition and shared conventions often fill in the gaps, a formalised proof requires every logical step to be made explicit. The reward is complete certainty: once verified by a proof assistant, the result can be trusted as correct and reused as a reliable foundation for further work. At the same time, it must be acknowledged that Lean, like any proof assistant, is not

infallible: errors in the underlying kernel, or in the assumptions one chooses to import, would undermine the guarantees. Nevertheless, the kernel is deliberately kept small and has been extensively scrutinised, giving strong confidence in its reliability.

Among the various proof assistants available today—such as Coq, Isabelle, or Agda—this thesis uses the Lean theorem prover [7]. The choice was motivated partly by the author’s own familiarity with Lean, and partly by the fact that Lean has become one of the fastest-growing systems for mathematical formalisation. Its success is driven in large part by its community library, `mathlib` [21], which now contains tens of thousands of theorems covering algebra, analysis, topology, geometry, combinatorics, and beyond. By providing this shared infrastructure, `mathlib` dramatically lowers the entry barrier: new formal developments, such as those undertaken in this thesis, can build directly on existing results rather than starting from scratch.

Throughout this thesis, several contributions were made to `mathlib`; these will be highlighted in the relevant sections, situating this work within the broader collaborative effort to advance the library and the field as a whole.

For readers wishing to learn Lean beyond what is presented here, the community maintains an excellent collection of tutorials and reference material (<https://leanprover-community.github.io/learn>). There is also an interactive version of Lean available online at <http://live.lean-lang.org>, which allows experimentation without installation.

1.2.1 Type Theory as a Foundation for Lean

The logical foundation of Lean is type theory. The development of type-theoretic systems can be traced back to the foundational crises of the early 20th century. Frege’s ambitious programme of reducing all mathematics to logic was famously derailed by Russell’s paradox [31], which exposed inconsistencies in naïve set theory. This prompted the creation of more disciplined frameworks. On one hand, the set-theoretic tradition culminated in Zermelo-Fraenkel set theory with the Axiom of Choice (ZFC), which remains the standard foundation for much of contemporary mathematics. On the other hand, type theory, introduced by Russell and later refined by Quine and others, provided an alternative foundation by avoiding paradoxes through a strict stratification of objects into levels, or *universes*.

The central idea of type theory is that every object comes equipped with a type that prescribes how it can be used. While the analogy with sets is useful, there is a critical distinction: an element may belong to many sets, but in type theory every term has exactly one type. Lean is based on the *Calculus of Inductive Constructions* (CIC) [7], a form of dependent type theory that is expressive enough to describe both computational data structures and mathematical propositions. Dependence here means that types themselves may vary with parameters—for example, one can define a family of vector spaces indexed by their dimension. Together with inductive definitions, products, function spaces, and quotients, CIC provides the expressive power to represent virtually all standard mathematical constructions.

To maintain consistency and avoid paradoxes of self-reference, Lean organises types into a cumulative hierarchy: `Type` : `Type` 1 : `Type` 2 : $\dots$. For most working mathematicians this hierarchy is invisible, but it plays a crucial role in higher-level areas such as category theory and homotopy type theory. Polymorphism adds flexibility: constructions like

Cartesian products, function spaces, and limits are defined once and automatically adapt to any level of the hierarchy.

A cornerstone of the practical utility of type theory is the *Curry-Howard correspondence* [14], which establishes a deep identification between logic and computation. Under this correspondence, propositions are identified with types, and proofs with terms inhabiting those types. A function $a \rightarrow b$ corresponds to a logical implication: given a proof of a , one produces a proof of b . The empty type represents falsehood, while a trivially inhabited type represents truth. This identification ensures that constructing a proof is equivalent to constructing a program of the appropriate type. In Lean, this correspondence is not merely philosophical: it provides the precise mechanism by which the system verifies the correctness of mathematical arguments.

1.2.2 Additional Axioms and Classical Reasoning

While Lean’s underlying foundation is dependent type theory, this alone does not fully capture the reasoning style of modern mathematics. To bridge the gap, Lean incorporates a small but carefully chosen collection of axioms that align its logical framework with ordinary mathematical practice:

- **Propositional extensionality:** if two propositions are logically equivalent, they are identified as equal. This allows us to reason about propositions “up to equivalence,” reflecting the way mathematicians habitually treat statements as interchangeable once their truth conditions coincide.
- **Quotient soundness:** elements related by a quotient relation are regarded as equal. This formalises the intuition that passing to a quotient space collapses distinctions that are immaterial for the mathematics at hand—for instance, rationals defined as equivalence classes of integer pairs.
- **Axiom of choice:** from a non-empty type, one may select an element. This ensures that Lean can replicate many standard results of analysis, algebra, and topology that rely implicitly or explicitly on choice.

Each of these axioms has a precise computational meaning. Propositional and quotient extensionality remove unnecessary bookkeeping, letting formal proofs resemble ordinary ones. The axiom of choice is more delicate: while essential for areas such as functional analysis and measure theory, it introduces non-computable objects, since it asserts the existence of witnesses without providing a construction.

1.3 Structure of the Thesis

The material of this thesis is organised into two main parts.

The first part is expository: it develops the mathematical background required to approach the Keakeya problem in the plane. We introduce the notions of Keakeya and Besicovitch sets, describe the constructions that establish their existence, and explain how one proves the Keakeya conjecture in dimension one and two. The aim here is to give a self-contained account of the mathematics underlying Körner’s construction, preparing the ground for the later formalisation.

The second part concerns the formalisation itself. We describe the implementation of the relevant definitions, constructions, and proofs in the Lean theorem prover, documenting

the new contributions made to the community library `mathlib`, as well as the obstacles encountered in formalising arguments from geometric measure theory. Particular attention is paid to the gap between the informal proofs found in the literature and the fully explicit form required by Lean: the translation process both reveals hidden assumptions and highlights where new lemmas or structural results must be added to the library.

For clarity, we distinguish throughout between *informal* mathematics—the usual “pen-and-paper” exposition—and *formal* mathematics, meaning machine-verified developments in Lean. Results already available in `mathlib` are cited directly rather than reproved, while the new formalised results contributed by this thesis are presented with complete informal proofs alongside their Lean implementations. This both situates the original contributions and illustrates in detail how classical reasoning is systematically re-expressed in a formal environment.

2 Constructing a Besicovitch set

In this section we study the existence of Besicovitch sets.

Besicovitch himself first established the existence of such sets in 1919 and 1928 [3, 4], using delicate geometric arguments to build explicit examples of measure-zero sets containing line segments in all directions. Soon after, Perron [27] and later Besicovitch again [5] refined these methods, while Sawyer [32] developed a different approach based on families of curves. Alternative constructions have since appeared, including probabilistic methods (e.g. Alexander [1]), randomised or game-theoretic approaches (Babichenko et al. [2]), and analytic perspectives such as Körner’s use of Baire category [19] (Many alternative methods have also been developed, see [30], [34], and [15], [18]).

In this thesis we follow Körner’s approach. The choice rests on two considerations. First, it is more suitable for formalisation in Lean than Besicovitch’s original construction, which relies heavily on Euclidean geometry.¹ Körner’s proof instead proceeds through Baire category, topology, and metric space arguments, which align directly with Lean’s existing API. Second, this framework allows the formalisation to contribute new and broadly applicable results in topology and analysis to `mathlib`, strengthening the foundations required for work in geometric measure theory.

2.1 The Baire Category Theorem

The central tool in Körner’s construction is the theory of Baire spaces and the Baire Category Theorem. Before turning to the construction itself, we therefore provide a brief exposition of these results.

Definition 2.1 ([29, Section I.1.]). Let A be a subset of a topological space X .

- A is **dense in X** if $\text{closure}(A) = X$.
- A is **somewhere dense in X** if $\text{interior}(\text{closure}(A)) \neq \emptyset$.
- If A is not somewhere dense in X then it is called **nowhere dense in X** .

Definition 2.2 ([29, Section I.2.]). Let X be a topological space and $Y \subseteq X$.

¹In the Lean/`mathlib` community, “API” refers to the body of lemmas and theorems that make a structure usable in practice; for instance, basic results on open balls and convergence for metric spaces, or on unions and intersections for sets.

- Y is of **first category** (or **meagre**) in X if it can be written as the union of countably many nowhere dense subsets of X .
- Y is of **second category** (or **non-meagre**) in X if it is not of first category.

Meagre sets are best thought of as topologically small. A nowhere dense set contains no open piece, and a meagre set, being a countable union of such sets, is ‘thin everywhere’.

Definition 2.3 ([25, Page 293]). A **Baire space** is a topological space such that every non-empty open subset is of second category.

This definition becomes useful through two classical results that explain when a space is Baire and what this implies in practice.

Theorem 2.4 ([25, Theorem 7.2]). If X is a compact Hausdorff space, or a complete metric space, then X is a Baire space.

Theorem 2.5 (Baire Category Theorem [35, Theorem 3.3.3]). Let X be a complete metric space. Then the intersection of any countable collection of dense open sets in X is dense.

We recall the following terminology, which will be used throughout.

Definition 2.6 ([36, Page 23]). A G_δ set is a subset of a topological space that can be written as a countable intersection of open sets.

What makes the Baire Category Theorem important is that it provides a way to prove existence results indirectly. Since a set of second category cannot be empty, it often suffices to show that the collection of objects with a desired property is of second category in order to guarantee that such objects exist.

2.2 Besicovitch via Baire

Körner’s arguments are formulated in the planar case $\mathbb{R}^2$, though they extend naturally to higher dimensions. In this thesis we restrict attention to $\mathbb{R}^2$. The central idea is to define a suitable collection of sets that already exhibit the structural features of a Besicovitch set, and then apply the theory of Baire to show that this collection is non-empty.

Let $\mathcal{K}$ denote the collection of all non-empty compact subsets of $\mathbb{R}^2$.

To work effectively with this space, we need a way of quantifying how close two compact sets are to one another. This role is played by the *Hausdorff distance*, introduced by Felix Hausdorff in 1918 [12]. Hausdorff’s broader programme was to develop general tools for measuring and comparing sets beyond the reach of classical geometry and analysis, including irregular or ‘pathological’ sets such as Cantor-type constructions.

Definition 2.7 ([19, Section 2]). We define

$$\rho(E, F) = \sup_{e \in E} \inf_{f \in F} \|e - f\| \quad \text{and} \quad d(E, F) = \rho(E, F) + \rho(F, E)$$

for any $E, F \in \mathcal{K}$. We call d the **Hausdorff metric**.

By standard arguments one obtains the following fundamental result.

Theorem 2.8 ([13, Section 28]). The Hausdorff metric d is a complete metric on $\mathcal{K}$.

We study the following set.

Definition 2.9. $\mathcal{P}$ is the collection of all closed subsets P of the rectangle $[-1, 1] \times [0, 1]$ with the following properties:

1. P is the union of line segments joining points of the form $(x_1, 0)$ to points of the form $(x_2, 1)$ with $x_1, x_2 \in [-1, 1]$.
2. If $|v| \leq \frac{1}{2}$ then we can find x_1 and x_2 with $x_2 - x_1 = v$ and $x_1, x_2 \in [-1, 1]$ such that the line segment joining $(x_1, 0)$ to $(x_2, 1)$ lies in P .

We now establish a basic property of $\mathcal{P}$, which will be used to show that this collection forms a complete metric space. The motivation for doing so is that completeness ensures $\mathcal{P}$ is a Baire space, allowing us to apply the theory of Baire in the subsequent argument. **Lemma 2.10** ([19, Lemma 2.2]). $\mathcal{P}$ is a closed subset of $(\mathcal{K}, d)$.

Informal proof. By standard arguments, it suffices to establish that $\mathcal{P}$ is sequentially closed [23, Theorem 6.8]. Thus, let (P_n) be a sequence in $\mathcal{P}$ converging in the Hausdorff metric to some $K \in \mathcal{K}$. Our task is to show that K itself belongs to $\mathcal{P}$, that is, that K inherits properties (1) and (2) from Definition 2.9.

Property (1). Take an arbitrary point $\mathbf{k} \in K$. By the definition of convergence in the Hausdorff metric, there exist points $\mathbf{p}_n \in P_n$ with $\mathbf{p}_n \rightarrow \mathbf{k}$. Since each P_n belongs to $\mathcal{P}$, the point $\mathbf{p}_n$ lies on some line segment of the admissible type, namely

$$\ell_n = \{((1-t)x_{1,n} + tx_{2,n}, t) : t \in [0, 1]\}, \quad x_{1,n}, x_{2,n} \in [-1, 1],$$

which joins $(x_{1,n}, 0)$ to $(x_{2,n}, 1)$.

The pairs $(x_{1,n}, x_{2,n})$ take values in the compact square $[-1, 1]^2$. Passing to a subsequence if necessary, we may therefore assume

$$x_{1,n} \rightarrow x_1, \quad x_{2,n} \rightarrow x_2,$$

for some $x_1, x_2 \in [-1, 1]$. Define ℓ to be the line segment joining $(x_1, 0)$ to $(x_2, 1)$. It is a simple but important observation that $\ell_n \rightarrow \ell$ in the Hausdorff metric: indeed, the parametrisations

$$((1-t)x_{1,n} + tx_{2,n}, t) \quad \text{and} \quad ((1-t)x_1 + tx_2, t)$$

show directly that the vertical cross-sections of ℓ_n converge uniformly to those of ℓ . Thus, every point of ℓ can be approximated by points of ℓ_n and vice versa.

We now claim that $\ell \subset K$. To see this, let $\mathbf{z} \in \ell$. Since $\ell_n \rightarrow \ell$, there exist $\mathbf{z}_n \in \ell_n$ with $\mathbf{z}_n \rightarrow \mathbf{z}$. But each ℓ_n lies in P_n , so $\mathbf{z}_n \in P_n$ for every n . The fact that $d(P_n, K) \rightarrow 0$ implies that $\mathbf{z} \in K$. As $\mathbf{z}$ was arbitrary, $\ell \subset K$.

Moreover, the particular point $\mathbf{k} \in K$ lies on ℓ . Indeed, we already had $\mathbf{p}_n \in \ell_n$ with $\mathbf{p}_n \rightarrow \mathbf{k}$. Choosing for each n the point $\mathbf{q}_n \in \ell$ at the same height as $\mathbf{p}_n$ gives $\|\mathbf{q}_n - \mathbf{p}_n\| \leq d(\ell_n, \ell)$, so $\mathbf{q}_n \rightarrow \mathbf{k}$. Since ℓ is closed, it follows that $\mathbf{k} \in \ell$. In summary: for each $\mathbf{k} \in K$ there exists a line segment ℓ of the admissible form such that $\mathbf{k} \in \ell \subset K$.

Property (2). Fix $v \in [-\frac{1}{2}, \frac{1}{2}]$. For each n , since P_n satisfies property (2), there exist $x_{1,n}, x_{2,n} \in [-1, 1]$ with

$$x_{2,n} - x_{1,n} = v$$

such that the corresponding segment $\ell_n = \ell(x_{1,n}, x_{2,n})$ lies entirely in P_n . Again, by compactness of $[-1, 1]^2$, we may assume (after passing to a subsequence) that

$$x_{1,n} \rightarrow x_1, \quad x_{2,n} \rightarrow x_2.$$

By continuity of subtraction, we have $x_2 - x_1 = \lim_{n \rightarrow \infty} (x_{2,n} - x_{1,n}) = v$. Thus, the limiting segment $\ell = \ell(x_1, x_2)$ has the required slope v .

As before, the convergence $\ell_n \rightarrow \ell$ in the Hausdorff metric, together with $\ell_n \subset P_n$ and $d(P_n, K) \rightarrow 0$, ensures that $\ell \subset K$. Therefore, K contains at least one full line segment of slope v . Since v was arbitrary in $[-\frac{1}{2}, \frac{1}{2}]$, property (2) is satisfied. $\square$

Corollary 2.11. The metric space $(\mathcal{P}, d)$ is complete.

Informal proof. By [24, Theorem 7.1.17], a closed subset of a complete metric space is itself complete. Since $\mathcal{P}$ is closed in the complete metric space $(\mathcal{K}, d)$, it follows immediately that $(\mathcal{P}, d)$ is complete. $\square$

Corollary 2.12. Consequently, $(\mathcal{P}, d)$ is a Baire space.

Informal proof. Since $(\mathcal{P}, d)$ is a complete metric space, it is a standard result that every complete metric space is a Baire space (see [33, Theorem 20.16]). $\square$

We present a technical lemma.

Definition 2.13. If $v \in [0, 1]$ and $\varepsilon > 0$, write $\mathcal{P}(v, \varepsilon)$ for the set of $P \in \mathcal{P}$ with the following property. There exist a finite collection of rectangles $R_1, \dots, R_N$, with sides parallel to the axes, such that, if $y \in [0, 1] \cap [v - \varepsilon, v + \varepsilon]$, then

$$\{x : (x, y) \in P\} \subseteq \left\{ x : (x, y) \in \bigcup_{j=1}^N R_j \right\},$$

and the total length of the finite set of intervals making up $\{x : (x, u) \in \bigcup_{j=1}^N R_j\}$ is strictly less than 100ε .

Lemma 2.14 ([19, Lemma 2.4]). The set $\mathcal{P} \setminus \mathcal{P}(v, \varepsilon)$ is closed and nowhere dense in $(\mathcal{P}, d)$.

Sketch of the informal proof. This theorem is not fully formalised within the present work and is therefore left as a **sorry**. The reasons for this choice are explained in detail in Section 3.3. For the purposes of the present chapter we provide only a sketch of the argument, and refer the reader to the relevant literature for a complete treatment.

Suppose $P \in \mathcal{P}(v, \varepsilon)$, and let $R_1, \dots, R_N$ be a finite family of rectangles establishing the covering and length conditions. The key observation is that these conditions are stable under small perturbations in the Hausdorff metric: if P' is sufficiently close to P , then its horizontal sections in the strip $[v - \varepsilon, v + \varepsilon]$ remain contained in a slight enlargement of the same rectangles. By thickening each R_j slightly, the covering property is preserved while the total horizontal length remains strictly below 100ε . Hence, every P' sufficiently close to P in $(\mathcal{P}, d)$ also belongs to $\mathcal{P}(v, \varepsilon)$, proving openness.

For density, we sketch only the main idea. Given $P \in \mathcal{P}$ and $\delta > 0$, one approximates P by finitely many of its line segments, chosen so that all admissible directions are represented within a small angular tolerance. This produces a nearby set P' with $d(P, P') < \delta$. Within the band $[v - \varepsilon, v + \varepsilon]$, the structure of P' consists of finitely many lines together with finitely many slender triangular regions. These can be covered by a finite collection of axis-parallel rectangles, with total horizontal length bounded by $O(\varepsilon)$. Choosing constants appropriately yields the required 100ε bound, and so $P' \in \mathcal{P}(v, \varepsilon)$. $\square$

Lemma 2.14 establishes the openness and density of the sets $\mathcal{P}(v, \varepsilon)$. These properties are precisely what is needed to deduce the following global category-theoretic result.

Theorem 2.15 ([19, Theorem 2.5]). The set $\mathcal{E}$ of $E \in \mathcal{P}$ such that $\{x : (x, u) \in E\}$ has Lebesgue measure zero is of second category in $(\mathcal{P}, d)$.

Informal proof. For integers $n \geq 1$ and $0 \leq r \leq n$, Lemma 2.14 asserts that $\mathcal{P}(r/n, 1/n)$ is open and dense in $(\mathcal{P}, d)$. Set

$$\mathcal{P}_n := \bigcap_{r=0}^n \mathcal{P}(r/n, 1/n).$$

Being a finite intersection of open dense sets, each $\mathcal{P}_n$ is open and dense, hence of second category.

Fix $P \in \mathcal{P}_n$ and $u \in [0, 1]$. Choose $r \in \{0, \dots, n\}$ with $|u - r/n| \leq 1/n$. By the defining property of $\mathcal{P}(r/n, 1/n)$ there exist finitely many axis-parallel rectangles $R_1, \dots, R_N$ such that

$$\{x : (x, y) \in P\} \subseteq \{x : (x, y) \in \bigcup_{j=1}^N R_j\} \quad \text{for all } y \in [0, 1] \cap [r/n - 1/n, r/n + 1/n],$$

and, for each such y , the finite union of horizontal intervals on the right has total length less than $100/n$. Taking $y = u$ gives

$$\lambda_1(\{x : (x, u) \in P\}) \leq \lambda_1(\{x : (x, u) \in \bigcup_{j=1}^N R_j\}) < \frac{100}{n}. \quad (*)$$

Let

$$\mathcal{P}_* := \bigcap_{n=1}^{\infty} \mathcal{P}_n.$$

Since is a Baire space, the countable intersection of open dense sets is dense and, in particular, of second category. Hence, $\mathcal{P}_*$ is of second category. If $P \in \mathcal{P}_*$, then $(*)$ holds for every $n \in \mathbb{N}$, so $\lambda_1(\{x : (x, u) \in P\}) \leq 100/n$ for all n and therefore

$$\lambda_1(\{x : (x, u) \in P\}) = 0 \quad \text{for all } u \in [0, 1].$$

Thus, $P \in \mathcal{E}$, i.e. $\mathcal{P}_* \subseteq \mathcal{E}$. Since $\mathcal{P}_*$ is of second category, so is $\mathcal{E}$. $\square$

Theorem 2.15 leads directly to the following consequence. By applying Fubini's theorem to pass from information about horizontal slices to the whole set, one obtains the corresponding statement for compact sets of Lebesgue measure zero.

Theorem 2.16 ([19, Theorem 2.3]). The set of $P \in \mathcal{P}$ of Lebesgue measure zero is of second category in $(\mathcal{P}, d)$.

Informal proof. Let λ_1 denote 1-dimensional Lebesgue measure on $\mathbb{R}$ and let $\lambda_2 = \lambda_1 \otimes \lambda_1$ be planar Lebesgue measure on $\mathbb{R}^2$. For $E \subset [-1, 1] \times [0, 1]$ and $u \in [0, 1]$ write

$$E_u := \{x \in [-1, 1] : (x, u) \in E\} \quad (\text{horizontal section at height } u).$$

If $E \in \mathcal{P}$, then E is closed, hence Borel. For fixed u the map $i_u : [-1, 1] \rightarrow [-1, 1] \times [0, 1]$, $x \mapsto (x, u)$, is continuous, so $E_u = i_u^{-1}(E)$ is closed in $[-1, 1]$; in particular, E_u is Lebesgue measurable and

$$\int_{-1}^1 \mathbf{1}_E(x, u) dx = \int_{-1}^1 \mathbf{1}_{E_u}(x) dx = \lambda_1(E_u) \quad \text{for every } u \in [0, 1]. \quad (1)$$

Now apply Fubini's theorem to the non-negative function $\mathbf{1}_E$ on $([-1, 1] \times [0, 1], \lambda_2)$ (the underlying measures are σ -finite):

$$\lambda_2(E) = \int_{[-1, 1] \times [0, 1]} \mathbf{1}_E d\lambda_2 = \int_0^1 \left(\int_{-1}^1 \mathbf{1}_E(x, u) dx \right) du.$$

Invoking (1) yields the slice formula

$$\lambda_2(E) = \int_0^1 \lambda_1(E_u) du. \quad (2)$$

By Theorem 2.15 the class $\mathcal{E} := \{E \in \mathcal{P} : \lambda_1(E_u) = 0 \text{ for all } u \in [0, 1]\}$ is of second category in $(\mathcal{P}, d)$. For each $E \in \mathcal{E}$, (2) gives $\lambda_2(E) = \int_0^1 0 du = 0$; hence $\mathcal{E} \subset \mathcal{Z}$, where $\mathcal{Z} := \{P \in \mathcal{P} : \lambda_2(P) = 0\}$. A superset of a second category set cannot be of first category, so $\mathcal{Z}$ is also of second category. $\square$

To produce a Besicovitch set, start from a member of $\mathcal{P}$ of Lebesgue measure zero. By Theorem 2.16 the subfamily $\{P \in \mathcal{P} : \lambda^2(P) = 0\}$ is of second category in the complete space $(\mathcal{P}, d)$, hence non-empty; fix $P_0 \in \mathcal{P}$ with $\lambda^2(P_0) = 0$. By Definition 2.9(2), for each $v \in [-\frac{1}{2}, \frac{1}{2}]$ there exist $x_1, x_2 \in [-1, 1]$ with $x_2 - x_1 = v$ such that the segment from $(x_1, 0)$ to $(x_2, 1)$ lies in P_0 . Its direction $(v, 1)$ makes angle $\arctan |v| \leq \arctan(1/2) < \pi/6$ with the y -axis and has length $\sqrt{1 + v^2} \geq 1$, so P_0 already contains segments of length at least 1 in every direction inside the cone of aperture $\pi/3$ about the y -axis. Let R_θ be rotation by θ and set

$$B = P_0 \cup R_{\pi/3}P_0 \cup R_{2\pi/3}P_0.$$

The three rotated cones (angles $0, \pi/3, 2\pi/3$) cover all un-oriented directions on the circle; hence B is Kakeya. Rotations preserve Lebesgue measure and the measure of a finite union is the sum of the measures, so $\lambda^2(B) = 0$. Therefore, B is a Besicovitch set.

3 The Kakeya conjecture

In this section we undertake a detailed examination of the Kakeya conjecture in the cases of $\mathbb{R}$ and $\mathbb{R}^2$. Although the 1-dimensional case is elementary, it is not without value: it provides a fully worked example in which the essential definitions can be tested, and it serves as a first benchmark for the process of formalisation in Lean. The 2-dimensional case is of far greater substance.

3.1 The Hausdorff measure and dimension

The Kakeya conjecture is stated in terms of *Hausdorff dimension*, a refinement of the familiar notion of dimension that is designed to handle sets too irregular to be described adequately by integer dimensions. To arrive at the Hausdorff dimension, one first introduces the concept of *Hausdorff measure*. Both ideas were introduced by Hausdorff, in the same work where he also defined the Hausdorff distance [12]. His construction extends Lebesgue's measure theory by permitting coverings with arbitrarily small sets and introducing a dimension parameter, from which the Hausdorff dimension emerges naturally.

We follow the standard exposition in [10, Chapter 2].

Definition 3.1. If U is a non-empty subset of $\mathbb{R}^n$, its **diameter** is defined as

$$|U| := \sup\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x}, \mathbf{y} \in U\},$$

the greatest distance between any two points of U .

Definition 3.2. Let $F \subseteq \mathbb{R}^n$. A countable (or finite) collection of sets $\{U_i\}$ is called a **δ -cover** of F if

$$F \subseteq \bigcup_{i=1}^{\infty} U_i \quad \text{and} \quad 0 < |U_i| \leq \delta \quad \text{for all } i.$$

Definition 3.3 (Pre-Hausdorff measure). For $s \geq 0$ and $\delta > 0$, the **s -dimensional δ -Hausdorff measure** of a set $F \subseteq \mathbb{R}^n$ is defined by

$$\mathcal{H}_\delta^s(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\}.$$

This quantity considers all coverings of F by sets of diameter at most δ and takes the infimum of the corresponding s -weighted sums. As $\delta \rightarrow 0$, these values are monotone increasing, and their limit defines the s -dimensional Hausdorff measure.

Definition 3.4. The **s -dimensional Hausdorff measure** of a set $F \subseteq \mathbb{R}^n$ is defined by

$$\mathcal{H}^s(F) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F).$$

Remark 3.5. For many sets, $\mathcal{H}^s(F)$ is either 0 or ∞ . The critical transition between these regimes is what determines the Hausdorff dimension.

The Hausdorff measure is indeed a measure, satisfying monotonicity, countable sub-additivity, and vanishing on the empty set. In particular:

Theorem 3.6 (Properties of the Hausdorff measure). For any $F, E \subseteq \mathbb{R}^n$,

1. $\mathcal{H}^s(\emptyset) = 0$.
2. If $E \subseteq F$, then $\mathcal{H}^s(E) \leq \mathcal{H}^s(F)$.
3. For any countable family $\{F_i\}$,

$$\mathcal{H}^s\left(\bigcup_{i=1}^{\infty} F_i\right) \leq \sum_{i=1}^{\infty} \mathcal{H}^s(F_i).$$

Remark 3.7. If the sets F_i are disjoint Borel sets, then equality holds in (3).

The Hausdorff measure recovers the familiar notions of length, area, and volume up to a normalising constant. More precisely, if F is a Borel subset of $\mathbb{R}^n$, then

$$\mathcal{H}^n(F) = c_n^{-1} \text{vol}^n(F),$$

where c_n is the volume of the unit ball in $\mathbb{R}^n$. Thus, the Hausdorff measure extends Lebesgue measure, but simultaneously provides a framework capable of handling lower-dimensional or highly irregular sets such as Cantor sets or fractals.

Hausdorff dimension

Returning to the definition of the Hausdorff measure, one observes that for a fixed set $F \subseteq \mathbb{R}^n$ the quantity $\mathcal{H}^s(F)$ is monotone non-increasing in the parameter s . More precisely, if $t > s$ and $\{U_i\}$ is a δ -cover of F , then

$$\sum_i |U_i|^t \leq \sum_i |U_i|^{t-s} |U_i|^s \leq \delta^{t-s} \sum_i |U_i|^s,$$

and taking the infimum over all such covers yields $\mathcal{H}_\delta^t(F) \leq \delta^{t-s} \mathcal{H}_\delta^s(F)$. Passing to the limit as $\delta \rightarrow 0$, it follows that if $\mathcal{H}^s(F) < \infty$, then necessarily $\mathcal{H}^t(F) = 0$ for all $t > s$. Thus, as s varies, the function $s \mapsto \mathcal{H}^s(F)$ exhibits a sharp transition between the value ∞ for small s and the value 0 for large s . The critical threshold is called the *Hausdorff dimension* of F .

Definition 3.8. The **Hausdorff dimension** of a set $F \subseteq \mathbb{R}^n$ is defined by

$$\dim_H(F) = \inf\{s \geq 0 : \mathcal{H}^s(F) = 0\} = \sup\{s \geq 0 : \mathcal{H}^s(F) = \infty\}.$$

Equivalently, one has the dichotomy

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{if } 0 \leq s < \dim_H(F), \\ 0 & \text{if } s > \dim_H(F). \end{cases}$$

At the critical value $s = \dim_H(F)$, the measure $\mathcal{H}^s(F)$ can take any value in $[0, \infty]$. In particular, for a smooth m -dimensional submanifold $F \subseteq \mathbb{R}^n$, one has $\dim_H(F) = m$. Thus, Hausdorff dimension extends the classical notion of dimension to arbitrary sets.

Example 3.9. Let $F \subseteq \mathbb{R}^3$ be the flat unit disc lying in the xy -plane,

$$F = \{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\}.$$

We compute the Hausdorff measures $\mathcal{H}^s(F)$ for different values of s .

When $s = 1$, the measure corresponds to the “length” of the set. Since F contains uncountably many line segments in every direction within the plane, any attempt to cover it by sets of arbitrarily small diameter forces the sum of lengths to diverge. Thus, $\mathcal{H}^1(F) = \infty$.

When $s = 2$, the measure corresponds to “area”. In this case the Hausdorff measure agrees, up to a normalising constant, with planar Lebesgue measure. More precisely,

$$\mathcal{H}^2(F) = \frac{4}{\pi} \text{area}(F).$$

Since the Euclidean area of the unit disc is π , we obtain $\mathcal{H}^2(F) = 4$, a finite, non-zero value.

Finally, when $s = 3$, the measure corresponds to “volume” in $\mathbb{R}^3$. As F is a 2-dimensional surface with no interior in three dimensions, its volume vanishes, giving $\mathcal{H}^3(F) = 0$.

In summary, $\mathcal{H}^s(F)$ is infinite for $s < 2$, zero for $s > 2$, and finite non-zero precisely at $s = 2$. Hence, the critical value is $\dim_H(F) = 2$, in agreement with the classical understanding that smooth 2-dimensional surfaces in $\mathbb{R}^3$ have Hausdorff dimension two.

Theorem 3.10 (Properties of the Hausdorff dimension). Let $E, F \subseteq \mathbb{R}^n$. Then:

1. **Monotonicity.** If $E \subseteq F$, then $\dim_H(E) \leq \dim_H(F)$.
2. **Countable stability.** For a countable sequence $\{F_i\}_{i=1}^\infty$,

$$\dim_H \left(\bigcup_{i=1}^\infty F_i \right) = \sup_i \dim_H(F_i).$$

3. **Countable sets.** If F is countable, then $\dim_H(F) = 0$.
4. **Open sets.** If $F \subseteq \mathbb{R}^n$ is non-empty and open, then $\dim_H(F) = n$.
5. **Smooth sets.** If F is a smooth m -dimensional submanifold of $\mathbb{R}^n$, then $\dim_H(F) = m$.

These properties reinforce the interpretation of the Hausdorff dimension as a genuine extension of classical dimension: it coincides with the expected dimension on regular sets, yet remains meaningful for irregular sets.

3.2 Hausdorff Tools for the Keakeya Conjecture

In preparation for the proof of the Keakeya conjecture in $\mathbb{R}^2$, we first establish a few key properties of the Hausdorff measure and dimension, which will serve as the technical foundation for the arguments to come.

We begin with a regularity property of the Hausdorff measure.

Theorem 3.11 ([9, Theorem 1.6 (a)]). If $E \subseteq \mathbb{R}^n$, then there exists a G_δ set G containing E such that $\mathcal{H}^s(G) = \mathcal{H}^s(E)$.

This result allows one to replace an arbitrary set by a topologically “nice” representative without changing its Hausdorff measure. In particular, it can be used to establish the following fact about the Hausdorff dimension.

Proposition 3.12. Let $E \subseteq \mathbb{R}^n$. Then there exists a G_δ set G containing E such that $\dim_H(G) = \dim_H(E)$.

Informal proof. Let $s := \dim_H(E)$. For each $i \in \mathbb{N}$, apply Theorem 3.11 to the exponent $s + \frac{1}{i} > s$. This yields a G_δ set $G_i \supseteq E$ such that

$$\mathcal{H}^{s+1/i}(G_i) = \mathcal{H}^{s+1/i}(E) = 0.$$

Define

$$G := \bigcap_{i=1}^\infty G_i.$$

Then G is a G_δ set containing E .

Upper bound. Take any $t > s$. Choose $k \in \mathbb{N}$ such that $s + \frac{1}{k} < t$. Since $G \subseteq G_k$, it follows that

$$\mathcal{H}^{s+1/k}(G) \leq \mathcal{H}^{s+1/k}(G_k) = 0.$$

Because the Hausdorff measure is monotone in the dimension parameter (i.e. $\mathcal{H}^u(F) \leq \mathcal{H}^v(F)$ whenever $u > v$), we conclude that

$$\mathcal{H}^t(G) \leq \mathcal{H}^{s+1/k}(G) = 0.$$

Thus, $\mathcal{H}^t(G) = 0$ for all $t > s$, which shows that $\dim_H(G) \leq s$.

Lower bound. Now take $t < s$. By the definition of the Hausdorff dimension, we have $\mathcal{H}^t(E) = \infty$. Since $E \subseteq G$ and Hausdorff measure is monotone in the set variable, it follows that

$$\mathcal{H}^t(G) \geq \mathcal{H}^t(E) = \infty.$$

Hence, $\dim_H(G) \geq s$.

Combining the upper and lower bounds, we obtain $\dim_H(G) = s = \dim_H(E)$. $\square$

Our next step is to establish that orthogonal projections cannot increase the Hausdorff measure. To reach this result, we first develop a general estimate describing how the Hausdorff measure behaves under Hölder-continuous mappings.

Proposition 3.13 ([10, Proposition 2.2]). Let $F \subseteq \mathbb{R}^n$ and $f : F \rightarrow \mathbb{R}^m$ be a mapping such that

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq c \|\mathbf{x} - \mathbf{y}\|^\alpha \quad (\mathbf{x}, \mathbf{y} \in F)$$

for some constants $c > 0$ and $\alpha > 0$. Then for each $s \geq 0$,

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}^s(F).$$

The case $\alpha = 1$ is of particular importance, corresponding to Lipschitz maps. In this situation we immediately obtain the projection inequality.

Lemma 3.14. Let $W \subseteq \mathbb{R}^n$ be a k -dimensional linear subspace, and let $\pi_W : \mathbb{R}^n \rightarrow W$ be the orthogonal projection onto W . For any $E \subseteq \mathbb{R}^n$ and any $s \geq 0$, we have

$$\mathcal{H}^s(\pi_W(E)) \leq \mathcal{H}^s(E).$$

Informal proof. It is straightforward to check that orthogonal projection is 1-Lipschitz, so Proposition 3.13 applies immediately. Indeed, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, write $\mathbf{x} - \mathbf{y} = \mathbf{v} + \mathbf{u}$ with $\mathbf{v} \in W$, $\mathbf{u} \in W^\perp$. Since $\pi_W(\mathbf{x}) - \pi_W(\mathbf{y}) = \mathbf{v}$, the Pythagorean theorem yields

$$\|\pi_W(\mathbf{x}) - \pi_W(\mathbf{y})\| = \|\mathbf{v}\| \leq \|\mathbf{x} - \mathbf{y}\|,$$

establishing the Lipschitz condition. $\square$

3.3 The 1-dimensional case

We begin with the 1-dimensional setting $\mathbb{R}$. Here there is only one relevant direction, since the positive and negative axes are equivalent. Thus, the requirement from Definition 1.1 that a Kakeya set “contain a unit line segment in every direction” reduces to the statement that the set must contain a closed interval of length one. Explicitly, a subset $K \subseteq \mathbb{R}$ is Kakeya if and only if there exists $a \in \mathbb{R}$ such that $[a, a + 1] \subseteq K$.

From this characterisation the Kakeya conjecture in $\mathbb{R}$ is immediate. Since $[a, a + 1]$ is a closed interval, it has positive Lebesgue measure, and in particular

$$\dim_H([a, a + 1]) = 1.$$

Because the Hausdorff dimension is monotone with respect to set inclusion, we obtain

$$\dim_H(K) \geq \dim_H([a, a + 1]) = 1.$$

On the other hand, $K \subseteq \mathbb{R}$ implies $\dim_H(K) \leq 1$. Hence, $\dim_H(K) = 1$.

3.4 The 2-dimensional case

In two dimensions a Kakeya set is easy to visualise, but the structure of a Besicovitch set is far less transparent. Körner's construction shows that such sets can be realised as unions of line segments arranged to satisfy the directional requirement while having Lebesgue measure zero.

Before addressing the proof of the 2-dimensional Kakeya conjecture, we record several technical results that will be required. Their proofs are omitted: although they are not currently available in `mathlib`, they rely on tools from Fourier analysis that lie beyond the scope of this thesis. For details, we refer the reader to the literature.

Theorem 3.15 ([22, Theorem 4.1]). Let $A \subseteq \mathbb{R}^n$ be a Borel set. For $e \in S^{n-1}$, let

$$P_e : \mathbb{R}^n \rightarrow \mathbb{R}, \quad P_e(x) = x \cdot e,$$

denote the orthogonal projection of x onto the line $\{te : t \in \mathbb{R}\}$.

- If $s \leq 1$, then for almost every direction $e \in S^{n-1}$,

$$\dim_H(P_e(A)) = \dim_H(A).$$

- If $s > 1$, then for almost every direction $e \in S^{n-1}$,

$$\mathcal{H}^1(P_e(A)) > 0.$$

Theorem 3.16 ([22, Proposition 6.6]). Let $s \geq m$ and $A \subseteq \mathbb{R}^n$ with $\mathcal{H}^s(A) < \infty$. Then for any m -dimensional linear subspace $V \subseteq \mathbb{R}^n$,

$$\mathcal{H}^{s-m}(A \cap (V^\perp + a)) < \infty \quad \text{for } \mathcal{H}^m\text{-almost all } a \in V.$$

With these preparatory results in hand, we are ready to prove Theorem 1.5.

Informal proof. Let $\Pi : \mathbb{R}^n \rightarrow \mathbb{R}^2$ be the orthogonal projection onto the first two coordinates. If $B \subset \mathbb{R}^n$ is Besicovitch, then $\Pi(B)$ is again Besicovitch and $\dim_H \Pi(B) \leq \dim_H(B)$. Hence, it suffices to prove the theorem in the plane. By Proposition 3.12, there exists a G_δ set $G \supseteq B$ with $\dim_H(G) = \dim_H(B)$, and G is still Besicovitch; replacing B by G we henceforth assume B is G_δ .

For $a \in (0, 1)$, $b \in \mathbb{R}$ and $q \in \mathbb{Q}$ define the half-unit segment

$$I(a, b, q) = \{(q + t, at + b) : 0 \leq t \leq \tfrac{1}{2}\}.$$

Let

$$E_q = \{(a, b) \in (0, 1) \times \mathbb{R} : I(a, b, q) \subset B\}.$$

If $B = \bigcap_{k \geq 1} U_k$ with U_k open and $U_{k+1} \subset U_k$, then

$$E_q = \bigcap_{k \geq 1} \{(a, b) : I(a, b, q) \subset U_k\}.$$

For fixed k the inner set is open: if $I(a_0, b_0, q) \subset U_k$, then $\delta := d(I(a_0, b_0, q), U_k^c) > 0$, and for (a, b) sufficiently close to (a_0, b_0) one has $\sup_{t \in [0, 1/2]} \|(a - a_0)t + (b - b_0)\| < \delta$, whence $I(a, b, q) \subset U_k$. Thus, each E_q is G_δ .

Because B contains a unit segment in every direction, for each $a \in (0, 1)$ there exists a unit segment of slope a contained in B . Taking the first half of that segment, its x -coordinate interval has positive length; choose $q \in \mathbb{Q}$ from that interval and translate vertically to write the half-segment as $I(a, b, q)$. Hence, for the projection $\pi(a, b) = a$,

$$\pi\left(\bigcup_{q \in \mathbb{Q}} E_q\right) = (0, 1).$$

Since π is 1-Lipschitz and $\mathbb{Q}$ is countable, we cannot have $\mathcal{H}^1(E_q) = 0$ for all q (else $\mathcal{H}^1(\bigcup_q E_q) = 0$, forcing $\mathcal{H}^1((0, 1)) = 0$). Thus, there exists $q_0 \in \mathbb{Q}$ with

$$\mathcal{H}^1(E_{q_0}) > 0 \quad \Rightarrow \quad \dim_H(E_{q_0}) \geq 1.$$

For $t \in \mathbb{R}$ consider $p_t : \mathbb{R}^2 \rightarrow \mathbb{R}$, $p_t(a, b) = ta + b$. This is the orthogonal projection onto the line spanned by $(t, 1)$. By the projection theorem [22, Theorem 4.1], for almost every t either

$$\dim_H(p_t(E_{q_0})) = \dim_H(E_{q_0}) \geq 1 \quad \text{if } \dim_H(E_{q_0}) \leq 1,$$

or

$$\mathcal{H}^1(p_t(E_{q_0})) > 0 \quad \text{if } \dim_H(E_{q_0}) > 1.$$

In either case $\dim_H(p_t(E_{q_0})) = 1$ for almost every t (the image lies on a line, so the dimension cannot exceed 1).

For such $t \in [0, 1/2]$ and any $(a, b) \in E_{q_0}$ we have

$$(q_0 + t, p_t(a, b)) = (q_0 + t, at + b) \in I(a, b, q_0) \subset B,$$

hence

$$\{q_0 + t\} \times p_t(E_{q_0}) \subset B \cap \{(x, y) : x = q_0 + t\}.$$

Thus, for a set $T \subset [0, 1/2]$ of positive Lebesgue measure the vertical section

$$B_{q_0+t} := \{y : (q_0 + t, y) \in B\}$$

satisfies $\dim_H(B_{q_0+t}) = 1$.

Assume, towards a contradiction, that $\mathcal{H}^\alpha(B) < \infty$ for some $\alpha \in [1, 2)$. Apply the slicing theorem [22, Proposition 6.6] with $m = 1$ and $V = \{(x, 0) : x \in \mathbb{R}\}$ (the x -axis): for $\mathcal{H}^1$ -almost every x ,

$$\mathcal{H}^{\alpha-1}(B \cap (\{x\} \times \mathbb{R})) < \infty,$$

so $\dim_H(B_x) \leq \alpha - 1 < 1$ for almost every x . This contradicts the existence of the set T of positive measure with $\dim_H(B_{q_0+t}) = 1$. Therefore, $\mathcal{H}^\alpha(B) = \infty$ for all $\alpha \in [1, 2)$, which implies $\dim_H(B) \geq 2$. The reverse inequality $\dim_H(B) \leq 2$ is trivial since $B \subset \mathbb{R}^2$, hence $\dim_H(B) = 2$ in the plane and $\dim_H(B) \geq 2$ in general. $\square$

4 Methodology of the Formalisation

In this section we return to aspects of the formalisation that were mentioned in earlier chapters but not developed in detail. Our aim is to show how the informal arguments presented previously were connected with the resources available in `mathlib`, and to highlight the methodological choices involved in bridging the gap between classical reasoning and its formal counterpart.

The presentation follows the order of the mathematical exposition given above, so that the informal and formal developments can be read in parallel.

4.1 Pre-existing Results in `mathlib`

A central component of the project was identifying which analytic and measure-theoretic results were already available in `mathlib`. Many foundational tools—including the construction of the Hausdorff metric, Hausdorff measure, and Hausdorff dimension, together with their basic monotonicity properties—were already present and could be directly reused. The library also provided general topological results such as the Baire category theorem (e.g. `dense_iInter_of_isOpen`) and the formalisation of first category sets (e.g. `IsMeagre`), as well as the theory of G_δ sets (e.g. `IsGδ`). This existing infrastructure shaped the overall strategy: rather than reproducing established results, the formalisation builds on these foundations and focus on the genuinely new contributions required for the Kakeya problem.

In particular, we are indebted to Sébastien Gouëzel for his work on [Baire’s theorems](#) and defining the [Hausdorff metric](#), and to Yury Kudryashov for his contributions to [\$G_\delta\$ sets](#), the [Hausdorff measure](#), and the [Hausdorff dimension](#). More broadly, this project relied heavily on the collective efforts of the `mathlib` community, whose sustained contributions made this development possible.

4.2 Defining Besicovitch sets

Before proceeding, we adopt a notational convention. We identify $\mathbb{R}^n$ with `Fin n` → $\mathbb{R}$ rather than an iterated product (in particular, $\mathbb{R}^2$ with `Fin n` → $\mathbb{R}$ instead of $\mathbb{R} \times \mathbb{R}$). These types are canonically equivalent in Lean, and we pass between them via the standard equivalence without further comment. We choose the Π -type presentation because the library’s API for `Fin n` → $\mathbb{R}$ (coordinate-wise rewriting, basis vectors, finite-support updates) is richer, making subsequent arguments cleaner and uniform across dimensions.

We encode the Kakeya property by quantifying over unit vectors. A direction in a real normed vector space E is a 1-dimensional subspace; choosing its unit representative v removes the irrelevant scaling. We then ask for a translation of x such that the entire segment $\{x + tv : 0 \leq t \leq 1\}$ lies in the set. This is precisely Definition 1.1 (“contains a unit line segment in every direction”): given the predicate, take the unit vector representing any direction to obtain the required segment. Conversely, if every direction occurs in the classical sense—namely, for each direction there exist two points at distance one whose joining segment lies inside s —then for a unit vector parallel to that segment one may choose the starting point of the segment as x (or the other endpoint if the orientation is reversed).

```
def IsKakeya {E : Type*} [NormedAddCommGroup E] [NormedSpace ℝ E] (s : Set E)
  : Prop :=
  ∀ v : E, ‖v‖ = 1 → ∃ x : E, segment ℝ x (x + v) ⊆ s
```

With `IsKakeya` established, we record a set of basic lemmas—intended for prospective inclusion in `mathlib` rather than later use here—that both sanity-check the notion and supply a lightweight API: monotonicity under inclusion (`IsKakeya.mono`); standard examples (the ambient space and, more generally, any closed ball of radius at least 1; `isKakeya_univ`, `isKakeya_closedBall`); non-emptiness in a non-trivial normed space (`IsKakeya.non-empty`); and an equivalent “subunit” formulation allowing vectors of norm less than or equal to in place of exactly 1 (`isKakeya_iff_subunit`).

Finally, we define the Besicovitch predicate to assert precisely that a set is Kakeya and has Lebesgue measure zero.

```
def IsBesicovitch {n : ℕ} (s : Set (Fin n → ℝ)) : Prop :=
  IsKakeya s ∧ volume s = 0
```

4.3 Körner's construction

The starting point in Körner's argument is the class $\mathcal{P}$ of “Kakeya-type” subsets of the fixed rectangle

$$R := [-1, 1] \times [0, 1] \subset \mathbb{R}^2,$$

built as unions of unit segments that run from the bottom edge $y = 0$ to the top edge $y = 1$. We write

$$\text{seg}_{01}(x_1, x_2) := \{((1-t)x_1 + tx_2, t) : t \in [0, 1]\}.$$

for the segment property (1) of Definition 2.9. Two models of $\mathcal{P}$ are needed: a “plain” set-theoretic version, and a metrised version living inside the space of non-empty compact subsets (so that the Hausdorff metric is available). The former is the most general for stating structural properties; the latter is essential for taking limits and invoking Baire category.

We first define down the ambient rectangle and the elementary segments.

```
def rectangle : Set (Fin 2 → ℝ) := Icc ![-1, 0] ![1, 1]

def segment01 (x1 x2 : ℝ) : Set (Fin 2 → ℝ) := segment ℝ ![x1, 0] ![x2, 1]
```

We record that `rectangle` is closed, bounded, convex and non-empty; this ensures any closed subset $P \subseteq \text{rectangle}$ is automatically compact, and (as noted below) property (ii) will guarantee non-emptiness.

We define $\mathcal{P}$ as follows:

```
def P_collection : Set (Set (Fin 2 → ℝ)) :=
  { P | IsClosed P ∧ P ⊆ rectangle ∧
    (∃ A ⊆ Icc ![-1, -1] ![1, 1], P = (p ∈ A), segment01 (p 0) (p 1)) ∧
    (∀ v : ℝ, |v| ≤ (1/2 : ℝ) →
      ∃ x1 x2 ∈ Icc (-1 : ℝ) 1, x2 - x1 = v ∧ segment01 x1 x2 ⊆ P) }
```

Most of the definition is self-explanatory; we draw note to (1). It encodes Körner's description: “ P is the union of the segments joining $(x_1, 0)$ to $(x_2, 1)$ with $x_1, x_2 \in [-1, 1]$ ”, via an endpoint index set $A \subset [-1, 1]^2$. For $p = (p_0, p_1) \in A$ we read the endpoints as $(p_0, 0)$ and $(p_1, 1)$, and reconstruct

$$P = \bigcup_{p \in A} \text{seg}_{01}(p_0, p_1).$$

Thus, every point of P lies on at least one segment crossing the strip $\{(x, y) : y \in [0, 1]\}$, and conversely every segment indexed by A is contained in P . The choice of A is inessential: the maximal set

$$A := \{(x_1, x_2) \in [-1, 1]^2 : \text{seg}_{01}(x_1, x_2) \subset P\}$$

works, but any smaller index set yielding the same union is equally acceptable.

To run the Baire-category part of the proof one needs a complete metric on $\mathcal{P}$, and the natural choice is the restriction of the Hausdorff metric d on the space of non-empty compact subsets of $\mathbb{R}^2$. For this we duplicate the definition verbatim inside the type of `NonemptyCompacts` (Fin 2 $\rightarrow$ $\mathbb{R}$):

```
def P_collection' : Set (NonemptyCompacts (Fin 2  $\rightarrow$   $\mathbb{R}$ )) :=
  { P | IsClosed (P : Set  $\_$ )  $\wedge$  (P : Set  $\_$ )  $\subseteq$  rectangle  $\wedge$ 
    ( $\exists$  A  $\subseteq$  Icc  $[-1, -1]$   $[1, 1]$ ,
      (P : Set  $\_$ ) = (p  $\in$  A), segment01 (p 0) (p 1))  $\wedge$ 
    ( $\forall$  v :  $\mathbb{R}$ ,  $|v| \leq (1/2 : \mathbb{R}) \rightarrow$ 
       $\exists$  x1 x2  $\in$  Icc  $(-1 : \mathbb{R})$  1, x2 - x1 = v  $\wedge$  segment01 x1 x2  $\subseteq$  P) }
```

There is a tautological correspondence between $\mathcal{P}$ and $\mathcal{P}'$. If $P \in \mathcal{P}$, then P is closed and $P \subseteq R$ with R compact, hence P is compact; by (ii) it is non-empty. Thus, P canonically determines an element of $\mathcal{P}'$. Conversely, if $P \in \mathcal{P}'$, forgetting the compact-non-empty wrapper gives a subset of R satisfying the same axioms, i.e. an element of $\mathcal{P}$. In short, the two definitions are *extensionally*² identical; we use $\mathcal{P}$ when reasoning set-theoretically, and $\mathcal{P}'$ when we need the metric structure (Hausdorff limits, completeness, Baire).

```
theorem P_collection'_image_eq : ( $\uparrow$ ) '' P_collection' = P_collection := by
  ext P
  constructor
  · rintro ⟨Q, hQ, rfl⟩
    exact hQ
  · intro hP
    exact ⟨⟨P, P_collection.isCompact hP⟩, P_collection.non-empty hP⟩, hP, rfl⟩
```

Next, we establish that $\mathcal{P}$ is closed in $(\mathcal{K}, d)$. Among all the components of the formalisation, this proved to be the most technically demanding.

Concretely, the proof requires verifying four separate conditions: first, that the limit set K is itself closed; second, that K is contained within the prescribed rectangle; and finally, that K satisfies the two structural properties that characterise membership in $\mathcal{P}$. Except for the first point, each of these verifications required substantial work and careful argumentation.

To keep the final argument manageable, we isolate each goal (besides the first one) into its own dedicated **theorem**, rather than presenting one monolithic proof. This modularisation was made possible by the tactic `extract_goal`, which we applied after initially formalising the entire argument as a single large lemma—a version so demanding that it required increasing Lean’s heartbeat limit in order to compile. That original proof has since been commented out, but it remains in the codebase for reference.

The closedness of K is immediate from its type: as an element of `NonemptyCompacts`, K is by construction compact, and hence closed.

In order to establish that the limit set K lies inside the rectangle $[-1, 1] \times [0, 1]$, one must verify that this geometric constraint is preserved under convergence in the Hausdorff

²Even though the formulations may differ syntactically, both definitions select exactly the same mathematical object, i.e. their extensions coincide.

metric. Each approximating set P_n is, by definition, contained in the rectangle, but a priori this property need not pass automatically to the limit. Suppose for contradiction that there exists a point $k' \in K$ outside the rectangle. Since the rectangle is closed, the point k' has strictly positive distance from it. Denote this distance by $\delta > 0$. The Hausdorff convergence $P_n \rightarrow K$ guarantees that for sufficiently large n , the sets P_n lie within Hausdorff distance $\delta/2$ of K . Consequently, there exists $y \in P_n$ such that $d(k', y) < \delta/2$. But since each $P_n \subseteq [-1, 1] \times [0, 1]$, the point y must belong to the rectangle. This yields a contradiction, because the very definition of δ asserts that every point in the rectangle lies at least distance δ away from k' . Hence, no such k' exists, and we conclude $K \subseteq [-1, 1] \times [0, 1]$.

This argument is short in prose, but its formalisation in Lean required careful decomposition into smaller lemmas to handle the subtleties of Hausdorff distance and compactness. For instance, the auxiliary result

```
theorem infDist_pos_of_mem_limit_notin_rectangle {k' : Fin 2 → ℝ} (h_notin :
  k' ∉ rectangle) :
  0 < infDist k' rectangle
```

encapsulates the fact that a point outside a closed set has positive distance from it, while the lemma

```
obtain ⟨y, hyP, hy_lt⟩ := exists_dist_lt_of_hausdorffDist_lt hk' hhd (fin_dist
  N)
```

formalises the existence of a point $y \in P_n$ close to k' under small Hausdorff distance.

To establish Property 1, we must show that every point $k \in K$ can be realised as lying on some line segment of the admissible form joining $(x_1, 0)$ to $(x_2, 1)$ with $-1 \leq x_1, x_2 \leq 1$. The challenge is that while each approximant P_n in the defining sequence belongs to the collection $\mathcal{P}$, the limit set K need not retain this property in an obvious way. We therefore need to pass to the limit carefully, using compactness and convergence in the Hausdorff metric.

Fix $k \in K$. Since the P_n converge to K in Hausdorff distance, for each n there exists a point $p_n \in P_n$ arbitrarily close to k . By construction of P_n , each p_n lies on some admissible segment $L_n = \text{seg}_{01}(x_{1,n}, x_{2,n})$, where $\mathbf{x}_n = (x_{1,n}, x_{2,n})$ belongs to the rectangle $[-1, 1]^2$. Thus we obtain a sequence of pairs $(x_{1,n}, x_{2,n})$ in a compact set. By sequential compactness, there exists a subsequence $(x_{1,\varphi(j)}, x_{2,\varphi(j)})$ converging to a limit $(x_1, x_2) \in [-1, 1]^2$. Denote by L the limiting segment $L = \text{seg}_{01}(x_1, x_2)$.

We now appeal to stability of the Hausdorff metric under convergence of endpoints. The lemma

```
theorem tendsto_hausdorffDist_segments_of_tendsto_endpoints
  {ι : Type*} {xn yn : ι → Fin 2 → ℝ} {x y : Fin 2 → ℝ} {l : Filter ι}
  (hx : Tendsto xn l ( x)) (hy : Tendsto yn l ( y)) :
  Tendsto (fun i ↦ hausdorffDist (segment ℝ (xn i) (yn i)) (segment ℝ x y))
  l ( 0) := ...
```

asserts that if the endpoints of the segments converge, then the segments themselves

converge in Hausdorff distance. Consequently, the sequence $L_{\varphi(j)}$ converges to L . Since each $L_{\varphi(j)} \subseteq P_{\varphi(j)}$ and $P_{\varphi(j)} \rightarrow K$ in the Hausdorff metric, one deduces that $L \subseteq K$.

Finally, because $\mathbf{p}_{\varphi(j)} \rightarrow \mathbf{k}$ with each $\mathbf{p}_{\varphi(j)} \in L_{\varphi(j)}$, it follows that $\mathbf{k}$ belongs to the closure of L . But L is closed, being a compact line segment, so in fact $\mathbf{k} \in L$. Thus, we have shown that every point of K lies on some admissible segment with endpoints in $[-1, 1]$.

Formally, in Lean, the argument is decomposed into a chain of carefully designed auxiliary lemmas, each addressing a distinct geometric stability phenomenon.

The lemma

```
theorem exists_segment_in_P_collection' {P_n : ℕ → NonemptyCompacts (Fin 2 → ℝ)}
  {K : NonemptyCompacts (Fin 2 → ℝ)}
  (h_mem : ∀ (n : ℕ), P_n n ∈ P_collection')
  (p_n : (k : Fin 2 → ℝ) → k ∈ K → ℕ → Fin 2 → ℝ)
  (hp_n_mem : ∀ (k : Fin 2 → ℝ) (a : k ∈ K) (n : ℕ), p_n k a n ∈ P_n n)
  (k : Fin 2 → ℝ) (hk : k ∈ ↑K) (n : ℕ) :
  ∃ x ∈ Icc ![-1, -1] ![1, 1], p_n k hk n ∈ segment01 (x 0) (x 1) ∧ segment01
  (x 0) (x 1) ⊆ ↑(P_n n) := ...
```

ensures that every point $\mathbf{p}_n \in P_n$ lies on a segment of the admissible form, contained within the rectangle. This guarantees that the defining structure of the approximants is not lost in the limiting process.

To control the stability of such segments under Hausdorff convergence, we employ

```
lemma exists_point_on_segment01_within_HD
  {a b a' b' : ℝ} {y : Fin 2 → ℝ}
  (hy : y ∈ (segment01 a b)) :
  ∃ t ∈ (segment01 a' b'), dist t y ≤ hausdorffDist (segment01 a b)
  (segment01 a' b') := ...
```

which quantitatively relates points across nearby segments. It asserts that any point on one segment lies uniformly close to some point on another, with the discrepancy bounded by the Hausdorff distance. This lemma is pivotal in showing that points carried along approximating segments converge coherently in the limit.

Finally, endpoint stability was established by Aaron Liu, who formalised bounds on the Hausdorff distance under perturbations of the endpoints.

```
theorem hausdorffDist_segment_left_le_dist
  {E : Type*} {x y z : E}
  [SeminormedAddCommGroup E] [NormedSpace ℝ E] :
  hausdorffDist (segment ℝ x z) (segment ℝ y z) ≤ dist x y := ...
```

This provides the inequality that moving one endpoint by distance δ perturbs the entire segment by at most δ . Together, these results underpin the argument that the limiting set K inherits the segmental structure from its approximants.

By piecing these together, Lean verifies the set-theoretic equality

$$K = \bigcup_{(x_1, x_2) \in A} \text{seg}_{01}(x_1, x_2),$$

where $A \subseteq [-1, 1]^2$ is the collection of endpoint pairs generating segments contained in K . This completes the verification of Property 1.

The second defining property of $\mathcal{P}$ encodes a uniform slope condition: every approximant P_n must contain, for each $|v| \leq \frac{1}{2}$, a line segment whose horizontal displacement is exactly v . The limit set K must therefore inherit the same feature. To make this precise, for each P_n we extract pairs of endpoints $\mathbf{x}_n = (x_{1,n}, x_{2,n}) \in [-1, 1]^2$ with $x_{2,n} - x_{1,n} = v$, such that the segment $\text{seg}_{01}(x_{1,n}, x_{2,n})$ lies entirely inside P_n . At this point, the difficulty is clear: one must control the behaviour of these endpoint pairs under the limiting process and ensure that the horizontal difference condition survives passage to the limit.

To organise this information, we introduce the auxiliary set

$$A := \{p \in [-1, 1]^2 : \text{seg}_{01}(\mathbf{p}_x, \mathbf{p}_y) \subseteq K\}.$$

This set plays a dual role. On one hand, it packages the admissible endpoint data into a closed subset of the square, making it possible to pass compactness arguments from the approximants to the limit in a structured way. On the other hand, it enforces simultaneously the two conditions that matter: the boundedness constraint $(x_1, x_2) \in [-1, 1]^2$ and the geometric constraint $\text{seg}_{01}(x_1, x_2) \subseteq K$. Thus, the problem of showing that K respects the slope condition is reformulated as: for every $|v| \leq \frac{1}{2}$, prove that there exists $p \in A$ with $\mathbf{p}_x - \mathbf{p}_y = v$.

The proof then unfolds as follows. From each P_n , compactness of $[-1, 1]^2$ guarantees a convergent subsequence of endpoints $\mathbf{x}_n \rightarrow \mathbf{x}_\infty = (x_1, x_2)$. Crucially, since the difference $x_{2,n} - x_{1,n} = v$ is exact and independent of n , continuity of subtraction forces $x_2 - x_1 = v$ at the limit. This ensures that the slope condition remains intact. Meanwhile, the Hausdorff convergence of P_n to K guarantees that the limiting segment $\text{seg}_{01}(x_1, x_2)$ lies within K . Therefore, $\mathbf{x}_\infty \in A$ with the required difference, and the set A establishes the property for K . In Lean, this reasoning is formalised in the following:

```

theorem exists_segment_subset_K_of_diff_le_half {P_n : ℕ → NonemptyCompacts
  (Fin 2 → ℝ)} {K : NonemptyCompacts (Fin 2 → ℝ)}
  (h_mem : ∀ (n : ℕ), P_n n ∈ P_collection') (h_lim : ∀ ε > 0, ∃ N, ∀ n ≥ N,
  dist (P_n n) K < ε)
  (h_closed : IsClosed (K : Set (Fin 2 → ℝ)))
  (fin_dist : ∀ (n : ℕ), EMetric.hausdorffEdist ↑(P_n n) (K : Set (Fin 2 → ℝ))
  ≠ ) :
  ∀ v, |v| ≤ 1 / 2 → ∃ x_1 x_2, x_1 ∈ Icc (-1) 1 ∧ x_2 ∈ Icc (-1) 1 ∧ x_2 - x_1 =
  v ∧ segment01 x_1 x_2 ⊆ ↑K :=

```

where the construction of A provides the essential bridge between the combinatorial condition on the approximants and the geometric property of the limit.

The current implementation could certainly be refactored into a shorter and more elegant form. Due to time constraints, we leave it in its present state, but streamlining this part of the code is a natural target for future work.

With this result in place, the formalisation of the fact that $\mathcal{P}_{\text{collection}}$ ’ forms a complete metric space—and hence a Baire space—follows directly from existing lemmas in `mathlib`.

To begin the formalisation of Lemma 2.14, we first need to encode in Lean the geometric objects that appear in its definition. The relevant configurations are planar subsets covered by finitely many axis-aligned rectangles, whose horizontal slices satisfy a uniform thinness condition.

Axis-aligned rectangles are introduced as

```
def axisRect (x1 x2 y1 y2 : ℝ) : Set (Fin 2 → ℝ) :=
  {p | p 0 ∈ Icc x1 x2 ∧ p 1 ∈ Icc y1 y2}
```

and slices at fixed height y are defined by

```
def hSlice (S : Set (Fin 2 → ℝ)) (y : ℝ) : Set ℝ :=
  {x | (![x, y] : Fin 2 → ℝ) ∈ S}
```

together with the vertical “window”

```
def window (v ε : ℝ) : Set ℝ :=
  Icc 0 1 ∩ Icc (v - ε) (v + ε)
```

which restricts attention to a small interval around height v . The thin cover condition from Körner’s statement is then encoded by asserting the existence of a finite family of rectangles that both cover the relevant slices and whose total horizontal measure is bounded by 100ε :

```
def hasThinCover (P : Set (Fin 2 → ℝ)) (v ε : ℝ) : Prop :=
  ∃ (R : Finset (Set (Fin 2 → ℝ))),
    (∀ r ∈ R, ∃ a b c d, r = axisRect a b c d) ∧
    (∀ y, y ∈ window v ε →
      hSlice P y ⊆ hSlice (r ∈ R, (r : Set _)) y) ∧
    (∀ y, y ∈ window v ε → (volume (hSlice (r ∈ R, (r : Set _)) y)).toReal
      < 100 * ε)
```

As a sanity check, we verified that the definition behaves correctly in degenerate cases: for example, a singleton point admits a thin cover given by its trivial rectangle, as formalised in

```
lemma hasThinCover_singleton (v ε : ℝ) (x : Fin 2 → ℝ) (hε : 0 < ε) :
  hasThinCover ({x} : Set (Fin 2 → ℝ)) v ε := by
```

Finally, $\mathcal{P}(v, \varepsilon)$ is represented in Lean as the subset of $\mathcal{P}$ consisting of non-empty compact configurations satisfying the thin cover property:

```
def P_v_eps' (v ε : ℝ) : Set P_collection' :=
  {P | hasThinCover (P : Set _) v ε}
```

We were unable to complete the formalisation of Lemma 2.14. The difficulty was twofold: lack of time, and the fact that the density of $\mathcal{P}(v, \varepsilon)$ in Lean requires handling geometric perturbations of line segments, a step that proved non-trivial. By contrast, the openness direction appears more tractable, and we formalised a number of auxiliary lemmas toward this, such as the behaviour of axis-aligned rectangles under expansion:

```
def axisRectExpand (η a b c d : ℝ) : Set (Fin 2 → ℝ) :=
  axisRect (a - η) (b + η) (c - η) (d + η)
```

```
lemma axisRect_subset_expand {a b c d : ℝ} {η : ℝ} (hη : 0 ≤ η) :
  axisRect a b c d ⊆ axisRectExpand η a b c d := ...
```

```
lemma mem_expand_of_close {a b c d η : ℝ} {p q : Fin 2 → ℝ}
  (hq : q ∈ axisRect a b c d) (hpq : dist p q ≤ η) :
  p ∈ axisRectExpand η a b c d := ...
```

These results were intended as ingredients for the openness proof but remain unfinished, and the corresponding statements in Lean are currently left as **sorry**.

Assuming both openness and density, the conclusion of Lemma 2.14 follows immediately, since `mathlib` already contains the general topological fact that the complement of an open dense set is closed and nowhere dense:

```
theorem P_v_eps_compl_nowhereDense {v ε : ℝ} (hv₀ : 0 ≤ v) (hv₁ : v ≤ 1) (hε :
  0 < ε) :
  IsClosed (P_v_eps' v ε) ∧ IsNowhereDense (P_v_eps' v ε) := by
  simp_rw [isClosed_isNowhereDense_iff_compl, compl_compl]
  exact ⟨isOpen_P_v_eps' hv₀ hv₁ hε, dense_P_v_eps' hv₀ hv₁ hε⟩
```

The central object of Theorem 2.15 is the set $\mathcal{E}$, consisting of those configurations whose horizontal slices have vanishing Lebesgue measure. In Lean this is formalised as

```
def E_set : Set P_collection' := {P | ∀ u ∈ Icc (0 : ℝ) 1, volume (hSlice (P :
  Set (Fin 2 → ℝ)) u) = 0}
```

To approach Theorem 2.15, we rely on the openness and density of the sets $\mathcal{P}(v, \varepsilon)$. These serve as the building blocks for constructing the nested families of sets $\mathcal{P}_n$ and their intersection $\mathcal{P}^*$. For a sequence $\phi : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$, we set

```
def Pn (φ : ℕ → ℝ≥0) (n : ℕ) : Set P_collection' :=
  r ∈ Finset.range n, P_v_eps' ((r : ℝ) * (φ n : ℝ)) (φ n : ℝ)
```

```
def Pstar (φ : ℕ → ℝ≥0) : Set P_collection' := n : ℕ, Pn φ n
```

so that $\mathcal{P}_n$ collects those configurations satisfying the $\mathcal{P}(v, \varepsilon)$ -constraints simultaneously for all $r < n$, and $\mathcal{P}^*$ is their intersection as $n \rightarrow \infty$.

A key point is that we do not fix the canonical sequence $\phi(n) = 1/n$, but rather work with an arbitrary sequence ϕ that satisfies the necessary structural properties (strictly

decreasing, converging to zero, remaining in $(0, 1)$, and respecting the scaling condition). This method—which we dub the “ ϕ -trick”—highlights that only these abstract properties of ϕ are ever used, and not the particular form $1/n$.

To establish Theorem 2.15, it suffices to show that $\mathcal{P}^*$ is of second category in $(\mathcal{P}, d)$. This reduces to proving that $\mathcal{P}^*$ is a dense G_δ set, and then applying the following general result, which the author contributed to `mathlib` (PR #27471):

```
/-- In a non-empty Baire space, any dense ‘Gδ’ set is not meagre. -/
theorem not_isMeagre_of_isGδ_of_dense {s : Set X} (hs : IsGδ s) (hd : Dense s) :
  ¬ IsMeagre s := ...
```

Remark 4.1. In the course of this work the author also contributed additional results to `mathlib`, anticipating their use in the formalisation. These include PR #28363 and PR #27333. Although ultimately unnecessary for the final argument, these contributions nonetheless enrich the library and may find application in related developments.

The remaining arguments rely only on standard principles of topology.

First, we establish that $\mathcal{P}^*$ is a G_δ set. By construction, $\mathcal{P}^*$ is the countable intersection of the sets $P_n(\varphi)$, each of which is itself a finite intersection of sets of the form $\mathcal{P}(v, \varepsilon)$. Since each $\mathcal{P}(v, \varepsilon)$ is open, so too is any finite intersection, and therefore every $P_n(\varphi)$ is open. It follows that

$$\mathcal{P}^* = \bigcap_n P_n(\varphi)$$

is a G_δ set. In Lean this is formalised as the lemma `IsG_Pstar`.

Second, we show that $\mathcal{P}^*$ is dense. Because each $\mathcal{P}(v, \varepsilon)$ is dense, every finite intersection $P_n(\varphi)$ is also dense. Taking the countable intersection then yields that $\mathcal{P}^*$ itself is dense.

In order to formalise Theorem 2.16, we first isolate the relevant collection of sets. We denote by

```
def P_zero_vol : Set P_collection' := {P | volume (P : Set (Fin 2 → ℝ)) = 0}
```

the sub-collection consisting of sets of Lebesgue measure zero. The aim is then to show that the set $\mathcal{E}$ is contained in this collection. This is accomplished by Fubini’s theorem: the hypothesis that every horizontal slice has 1-dimensional measure zero suffices to deduce that the whole 2-dimensional set has vanishing volume.

Formally, this reduction is expressed in the lemma:

```
theorem E_set_subset_PzeroVol : E_set ⊆ P_zero_vol := ...
```

where the proof necessarily passes from the `Fin 2 → ℝ`-model of the plane to the standard product $\mathbb{R} \times \mathbb{R}$, to access the measure-theoretic theorems in `mathlib` concerning product measures and integration using Fubini’s theorem. The measurable equivalence

```
MeasurableEquiv.finTwoArrow : (Fin 2 → ℝ) ≃ ℝ × ℝ
```

provides the bridge, after which the argument is routine: each horizontal section is null by assumption, so Fubini yields zero volume for the entire set. The resulting proof is

somewhat verbose, and while parts (such as the handling of sections outside $[0, 1]$) could be refactored and streamlined, the code compiles and establishes the intended statement. With this inclusion in place, we use a contradiction argument to complete Theorem 2.16. In particular, since $\mathcal{P}_*$ is of second category and contained in $\mathcal{E}$, it follows that P_zero_vol is also of second category. We write this in Lean as

```

theorem P_zero_vol_not_meagre
  (h2φ : ∀ (n : ℕ), φ n ∈ Set.Ioo 0 1) (h3φ : Tendsto φ atTop ( 0))
  (hv0 : ∀ n r, r ∈ Finset.range n → 0 ≤ r * φ n) (hv1 : ∀ n r, r ∈
  Finset.range n → r * φ n ≤ 1) :
  ¬ IsMeagre P_zero_vol := by
intro h
exact E_set_not_meagre φ (Pstar_sub_E_set φ h3φ hv0 hv1) h2φ hv0 hv1 (h.mono
E_set_subset_PzeroVol)

```

To bridge the gap from second category to explicit non-emptiness, one requires the general lemma asserting that a second category subset of a complete metric space cannot be empty. Since this statement was not previously available in `mathlib`, the author contributed it in [PR #27347](#). With this tool in place, the existence theorem becomes immediate:

```

theorem exists_P_with_zero_volume
  (h2φ : ∀ (n : ℕ), φ n ∈ Set.Ioo 0 1) (h3φ : Tendsto φ atTop ( 0))
  (hv0 : ∀ n r, r ∈ Finset.range n → 0 ≤ r * φ n) (hv1 : ∀ n r, r ∈
  Finset.range n → r * φ n ≤ 1) :
  P_zero_vol.non-empty :=
non-empty_of_not_isMeagre (P_zero_vol_not_meagre φ h2φ h3φ hv0 hv1)

```

We did not formalise the final step that produces a Besicovitch set by rotating the seed configuration P_0 . The obstacle is not conceptual but API-level. Although `mathlib` contains planar rotations as [linear isometries](#), what is missing for our purposes is a coherent interface on $\mathbb{R}^2 \simeq \text{Fin } 2 \rightarrow \mathbb{R}$ that simultaneously (i) acts on *sets*, (ii) transports the *segment* structure used to define $\mathcal{P}$, (iii) interacts well with the *Hausdorff metric*, and (iv) preserves the relevant *measures*. In particular, the “rotate P_0 through all angles and reassemble” manoeuvre requires a small library of lemmas of the following schematic form:

- Functoriality on sets: for a linear isometry R_θ (rotation by angle θ),

$$\begin{aligned}
R_\theta(\text{seg}_{01}(x, y)) &= \text{seg}_{01}(R_\theta x, R_\theta y), \\
R_\theta(S \cup T) &= R_\theta S \cup R_\theta T, \\
R_\theta\left(\bigcup_i S_i\right) &= \bigcup_i R_\theta S_i.
\end{aligned}$$

- Compatibility with the Hausdorff metric:

$$d(R_\theta S, R_\theta T) = d(S, T), \quad \theta \mapsto R_\theta(S) \text{ is continuous in } d.$$

- Measure preservation: for Lebesgue measure λ^2 and $\mathcal{H}^1$,

$$\lambda^2(R_\theta E) = \lambda^2(E), \quad \mathcal{H}^1(R_\theta(\ell)) = \mathcal{H}^1(\ell) \quad \text{for any line segment } \ell.$$

- Invariance of the class: if $P \in \mathcal{P}$, then $R_\theta P$ satisfies the analogues of (i)-(ii) in Definition 2.9.

Rotations available in `mathlib` are expressed either via multiplication by complex numbers of unit modulus (in the $\mathbb{C}$ -model of $\mathbb{R}^2$) or as $\mathbb{R}$ -linear isometries. Our development, however, works natively with `Fin 2` $\rightarrow \mathbb{R}$, and the geometry of $\mathcal{P}$ is encoded by the special family of segments

$$\text{seg}_{01}(x_1, x_2) = \{((1-t)x_1 + tx_2, t) : t \in [0, 1]\},$$

whose vertical coordinate *is the parameter* t . Under a genuine rotation the image of seg_{01} is no longer of this canonical form; consequently, to remain within the $\mathcal{P}$ -framework one must compose the rotation with an affine renormalisation (a shear/translation that straightens and recentres the image segment back into $[-1, 1] \times [0, 1]$). Developing this interface appears feasible and largely self-contained; however, its implementation could not be completed within the timeframe of this thesis. Accordingly, its construction and integration are deferred to future work.

4.4 Hausdorff dimension toolkit

In this section we formalise the material of Section 3.2, following the same order of presentation and indicating at each stage which components were already available in `mathlib` and which required new contributions.

Theorem 3.11 was previously formalised by Bhavik Mehta in a private repository that is being prepared to be added to `mathlib`; we therefore take it as an available dependency and do not reproduce its proof here. The subsequent results are stated and proved under this regularity assumption.

A note on methodology: early in the project we anticipated needing the infimum characterisation of Hausdorff dimension, and formalised it accordingly. This result has since been added to `mathlib` (PR #26876). In the end, the characterisation was not required for the arguments in this section. What did prove essential, however, was the finiteness of Hausdorff dimension in finite-dimensional normed spaces, which has also been incorporated into `mathlib` (PR #26970).

The formal proof of Theorem 3.12 is carried out in Lean exactly along the lines of the informal argument, and is implemented as the theorem `exists_G_of_dimH`.

As a preliminary, note that in any finite-dimensional real normed space E one has

$$\dim_H(A) \leq \dim_H(E) = \dim_{\mathbb{R}} E < \infty.$$

Thus, for each $A \subset E$ we may write $\dim_H(A) = D_A \in [0, \infty)$. Finiteness is used only to ensure that exponents remain real; all subsequent arguments are expressed in terms of the Hausdorff measures $\mathcal{H}^s$.

Set $s := D_A$. To construct a G_δ -superset preserving Hausdorff dimension we follow the standard scheme: for a sequence of exponents decreasing to s , adjoin superset approximants and intersect. The ϕ -trick replaces the rigid choice $s + 1/k$ by an arbitrary null-sequence

$$\varphi : \mathbb{N} \rightarrow (0, 1), \quad \varphi(k) \downarrow 0,$$

and we set $s_k := s + \varphi(k)$. Only the properties $\varphi(k) > 0$ and $\varphi(k) \rightarrow 0$ are required, isolating the “arbitrarily close” step in a form directly usable in Lean.

Suppose we have $G'_k \supset A$ with $\mathcal{H}^{s_k}(G'_k) = \mathcal{H}^{s_k}(A)$ and put $G = \bigcap_k G'_k$. The lower bound $\dim_H(A) \leq \dim_H(G)$ follows from monotonicity. For the upper bound assume, towards contradiction, that $\dim_H(G) > s$. Then for some $t > s$ there exists k with $s < s_k < t$. Since $s_k > \dim_H(A)$, we have $\mathcal{H}^{s_k}(A) = 0$ and hence $\mathcal{H}^{s_k}(G'_k) = 0$. Monotonicity in both set and exponent then gives

$$\mathcal{H}^t(G) \leq \mathcal{H}^t(G'_k) \leq \mathcal{H}^{s_k}(G'_k) = 0,$$

contradicting $\dim_H(G) > t$. Thus, $\dim_H(G) = s$.

The formalisation requires some care with Lean’s treatment of dimensions as extended non-negative reals. In order to avoid repeated coercions between $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{\geq 0} \cup \{\infty\}$, we consistently introduce auxiliary variables of the appropriate type at each stage of the argument. For example, the Hausdorff dimension $\dim_H(A)$ is first shown to be finite, which allows us to lift it to a value of type $\mathbb{R}_{\geq 0} \cup \{\infty\}$. This in turn permits the use of monotonicity properties of Hausdorff measures in the $\mathbb{R}_{\geq 0}$ setting, while keeping exponents and measures compatible with Lean’s libraries.

The behaviour of Hausdorff measure under Hölder and Lipschitz maps is already available in `mathlib`, and we make direct use of these results without modification. In particular:

- `HolderOnWith.hausdorffMeasure_image_le` establishes the Hölder case;
- `LipschitzOnWith.hausdorffMeasure_image_le` covers Lipschitz maps defined on subsets;
- `LipschitzWith.hausdorffMeasure_image_le` handles the global Lipschitz case.

These lemmas are sufficient for all applications in this thesis and will be used repeatedly in what follows.

We do not revisit Lemma 3.14 here; it was formalised previously as part of the Imperial College London “Formalising Mathematics” coursework, and has since been added to `mathlib` together with auxiliary results used in its proof (PR #25759).

As noted earlier, Theorem 3.15 and Theorem 3.16 were not formalised in this project. We do, however, include a formal statement of Theorem 3.16, leaving its proof as a **sorry**. At this stage we are cautious about asserting that our formal statement is exactly equivalent to the textbook version: experience shows that issues of normalisation (of $\mathcal{H}^s$), measurability, and “almost-every direction” quantifiers often surface only during a complete formal development. A finished proof would either confirm the equivalence or force minor adjustments to the statement (e.g. hypotheses or explicit constants). For our purposes the proposition records the intended interface; a full formal proof is deferred to future work alongside the necessary Fourier-analytic framework.

```

theorem hausdorffMeasure_slicing
  (n : ℕ)
  (A : Set (EuclideanSpace ℝ (Fin n))) (hA : MeasurableSet A)
  (s : ℝ) (hs : μH[s] A < ∞)
  (k : ℕ) (hks : (k : ℝ) ≤ s)
  (W : Submodule ℝ (EuclideanSpace ℝ (Fin n))) (hW : Module.finrank ℝ W = k)
  (direction : W) (x : W) :
  ∀ x ∂ (μH[(k : ℝ)]).restrict (W : Set (EuclideanSpace ℝ (Fin n))),
  μH[s - k] (A ∩ (AffineSubspace.mk' x W.orthogonal : Set _)) < := by

```

sorry

4.5 Formalising the conjecture

In this section we turn to the formalisation of theakeya conjecture. While a complete proof remains out of reach, we establish the 1-dimensional case in full.

As noted in Section 3.3, in one dimension theakeya condition is equivalent to the requirement that the set contain a closed interval of unit length. This provides the natural starting point for the formalisation.

```
lemma keakeya_contains_unit_Icc {K : Set ℝ} (hK : Isakeya K) :
  ∃ x₀, Icc x₀ (x₀ + 1) ⊆ K := by
  have := hK 1 (by simp)
  rcases this with ⟨x₀, hseg⟩
  exact ⟨x₀, by simp using hseg⟩
```

The proof is immediate: the hypothesis `hK` guarantees, for the unit length 1, the existence of a segment of that length inside K . Unwinding the definition produces an interval $[x_0, x_0 + 1]$ contained in K .

The next step is to connect such intervals to Hausdorff dimension. In Euclidean space, any non-trivial interval has Hausdorff dimension 1. This fact is already established in `mathlib` as a corollary of a general theorem: any subset of $\mathbb{R}^n$ with non-empty interior has Hausdorff dimension n . Hence, for any $a \in \mathbb{R}$, the interval $[a, a + 1]$ has interior $(a, a + 1)$, which is non-empty. The following lemma makes this explicit:

```
lemma dimH_Icc_length_one (a : ℝ) : dimH (Icc a (a + 1)) = 1 := by
  have h : (interior (Icc a (a + 1))).non-empty := by simp [interior_Icc]
  calc
    dimH (Icc a (a + 1)) = Module.finrank ℝ ℝ :=
    Real.dimH_of_non-empty_interior h
    _ = 1 := by simp
```

Remark 4.2. When preparing a pull request formalising the Hausdorff dimension of the unit interval and segment, Professor Antoine Chambert-Loir suggested a more general and conceptually natural approach (PR #27069). Instead of relying on ad hoc arguments, one can observe that any affine map from a finite-dimensional affine space to another affine space is automatically Lipschitz (and indeed Anti-Lipschitz). It then follows that every convex subset of such a space has Hausdorff dimension equal to the dimension of the affine space it generates. This observation provides a clearer and more general framework than the approach adopted here, and we intend to revise the formalisation accordingly in a future contribution to `mathlib`.

Having established that closed intervals of unit length have dimension one, we can promote this to any set containing such an interval: by monotonicity of Hausdorff dimension, its dimension is at least one.

```
lemma dimH_of_contains_Icc {K : Set ℝ} {x₀} (h : Icc x₀ (x₀ + 1) ⊆ K) :
  1 ≤ dimH K := by
  calc
```

```
1 = dimH (Icc x₀ (x₀ + 1)) := (dimH_Icc_length_one x₀).symm
_ ≤ dimH K := dimH_mono h
```

The corresponding upper bound is equally straightforward: every subset of $\mathbb{R}$ has dimension at most one, because $\mathbb{R}$ itself has dimension one and dimension is monotone under inclusion.

```
lemma dimH_le_one_univ : ∀ (K : Set ℝ), dimH K ≤ 1 := fun K ↦ by
  calc
    dimH K ≤ dimH (univ : Set ℝ) := dimH_mono (subset_univ _)
    _ = Module.finite ℝ ℝ := by simp [dimH_univ]
    _ = 1 := by simp
```

Combining the two inequalities, we obtain the desired result:

```
theorem isKakeya.dimH_eq_one (K : Set ℝ) (hK : IsKakeya K) :
  dimH K = 1 := by
  rcases kakeya_contains_unit_Icc hK with ⟨x₀, hsub⟩
  exact le_antisymm (dimH_le_one_univ K) (dimH_of_contains_Icc hsub)
```

In complete analogy, one can formulate the Kakeya conjecture in $\mathbb{R}^2$. At present, however, the corresponding statement remains unproven in Lean and must be left as a **sorry**.

```
theorem 2DKakeya (B : Set ((Fin 2 → ℝ))) (hB : IsBesicovitch B) :
  dimH B = 2 := sorry
```

5 Conclusion

This thesis investigated whether the geometric and analytic arguments underlying the Kakeya–Besicovitch problem can be made fully explicit within a modern proof assistant. The focus was on formalising Körner’s construction of Besicovitch sets in $\mathbb{R}^2$ and on developing the auxiliary infrastructure in Lean’s community library `mathlib` required to support it. The outcome, while not a complete formalisation of Körner’s proof, establishes that the framework is robust enough to carry substantial parts of the argument through to a machine-checked conclusion.

5.1 What was achieved

We conclude by summarising the principal outcomes of this work, together with the supporting `mathlib` infrastructure developed in the process.

1. *Foundational definitions.* We formalised the Kakeya predicate `IsKakeya` in arbitrary finite-dimensional real normed spaces, established its natural invariances and elementary consequences (monotonicity, non-emptiness, canonical examples), and defined the Besicovitch predicate `IsBesicovitch` as “Kakeya with Lebesgue measure zero”. Equivalent formulations in terms of unit vectors and bounded-norm vectors were proved, aligning the formalisation with the classical definitions.
2. *Hausdorff measure and dimension.* We proved that $\dim_H(A) < \infty$ for every $A \subseteq \mathbb{R}^n$ and formalised the G_δ -regularity of Hausdorff dimension: for any $A \subseteq \mathbb{R}^n$ there

exists a G_δ superset $G \supset A$ with $\dim_H(G) = \dim_H(A)$. The key lemma isolates the upper-bound direction $\dim_H(G) \leq \dim_H(A)$ and replaces the classical $1/k$ sequence with an abstract positive null sequence $\varphi(k)$, yielding a robust reusable component. Projection and slicing tools were incorporated from existing `mathlib` results.

3. *Segments and the Hausdorff metric.* We developed a calculus for line segments: end-point control of Hausdorff distance, convergence under endpoint limits, compactness of the normalised segment `segment01`, and a point-selection lemma. These results package the analytic ingredients of Körner’s construction into reusable lemmas.
4. *Körner’s space $(\mathcal{P}, d)$ and Baire category.* We modelled $\mathcal{P}$ both as subsets of the rectangle $[-1, 1] \times [0, 1]$ and as elements of `NonemptyCompacts`, giving direct access to the Hausdorff metric. We proved that $\mathcal{P}$ is closed, hence complete and Baire. For Lemma 2.14 we formalised the openness argument in detail, and provided a documented sketch for density. From this we obtained a full Lean formalisation of Theorem 2.15, establishing that a residual family of sets in $\mathcal{P}$ have null horizontal fibres, and hence that the measure-zero elements of $\mathcal{P}$ form a set of second category.
5. *Contributions to `mathlib`.* Several results were added or prepared as pull requests, including: Hausdorff control for segments; convergence of segments; existence of G_δ supersets of prescribed Hausdorff dimension; utilities for slices and axis-aligned rectangles; and the non-meagreness criterion for dense G_δ sets.

5.2 Broader perspective

Takeya problems remain central to harmonic analysis and geometric measure theory, as highlighted by the recent resolution of the three-dimensional Takeya conjecture by Wang and Zahl. Formalisation adds a complementary perspective: it separates analytic content from organisational structure, forces precise compactness and stability statements, and produces general-purpose lemmas of independent value. The present work demonstrates that Lean and `mathlib` can now support serious projects in geometric measure theory, provided that development continues at the interface between geometry and analysis.

5.3 Closing remarks

This thesis does not resolve the conjecture but develops the formal framework necessary to approach it. The main outcomes are: a formal definition of Takeya and Besicovitch sets; a toolkit for Hausdorff measure and dimension; a calculus for line segments in the Hausdorff metric; and a Baire-category framework for Körner’s space. The remaining steps toward a full formalisation of Körner’s construction are technical rather than conceptual. Completing them would finalise the planar Besicovitch existence argument in Lean and provide a foundation for fully verified treatments of geometric ideas central to harmonic analysis.

This project begun later than others in the course; with additional time, it is likely that a full formalisation of Körner’s construction could have been achieved. Nonetheless, the work carried out here yielded substantial contributions, both conceptual and technical. Future efforts will build on this foundation, with the aim of completing the formalisation and incorporating the results into `mathlib`.

Mathlib References

Below we collect the specific `mathlib` definitions, lemmas, and files cited in the main text. For accessibility, each entry includes the name of the lemma or definition (when applicable), the file path in the repository, and a permanent GitHub link pinned to commit

206c7bb9433658c879d7b34d590f729c6fe69dfe,

which fixes the state of the library at the time of writing. This ensures reproducibility even as `mathlib` evolves.

`dense_iInter_of_isOpen` `Mathlib/Topology/Baire/Lemmas.lean`, lines 61–65. <http://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/Baire/Lemmas.lean#L61-L65>

`IsMeagre` `Mathlib/Topology/GDelta/Basic.lean`, lines 213–214. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/GDelta/Basic.lean#L213-L214>

`IsG $\delta$` `Mathlib/Topology/GDelta/Basic.lean`, lines 56–58. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/GDelta/Basic.lean#L56-L58>

Baire lemmas (file) `Mathlib/Topology/Baire/Lemmas.lean`. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/Baire/Lemmas.lean>

Hausdorff metric (file) `Mathlib/Topology/MetricSpace/HausdorffDistance.lean`. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/MetricSpace/HausdorffDistance.lean>

G δ sets (file) `Mathlib/Topology/GDelta/Basic.lean`. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/GDelta/Basic.lean>

Hausdorff measure (file) `Mathlib/MeasureTheory/Measure/Hausdorff.lean`. <http://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/MeasureTheory/Measure/Hausdorff.lean>

Hausdorff dimension (file)

`Mathlib/Topology/MetricSpace/HausdorffDimension.lean`. <https://github.com/leanprover-community/mathlib4/blob/206c7bb9433658c879d7b34d590f729c6fe69dfe/Mathlib/Topology/MetricSpace/HausdorffDimension.lean>

Second category criterion <https://github.com/leanprover-community/mathlib4/pull/27471>

PR 28363 <https://github.com/leanprover-community/mathlib4/pull/28363>

PR 27333 <https://github.com/leanprover-community/mathlib4/pull/27333>

PR 27347 <https://github.com/leanprover-community/mathlib4/pull/27347>

Linear isometries (file) `Mathlib/Analysis/Complex/Isometry.lean` <https://github.com/leanprover-community/mathlib4/blob/632d3ae291feaf46fa8ea64ba6c65e4068a4ce99/Mathlib/Analysis/Complex/Isometry.lean#L39-L47>

PR 26876 <https://github.com/leanprover-community/mathlib4/pull/26876>

PR 26970 <https://github.com/leanprover-community/mathlib4/pull/26970>

`HolderOnWith.hausdorffMeasure_image_le`

Mathlib/MeasureTheory/Measure/Hausdorff.lean, lines 670–707. <https://github.com/leanprover-community/mathlib4/blob/afb3068a5e9f765af7936d0952ff9e7a64d167b4/Mathlib/MeasureTheory/Measure/Hausdorff.lean#L670-L707>

`LipschitzOnWith.hausdorffMeasure_image_le`

Mathlib/MeasureTheory/Measure/Hausdorff.lean, lines 717–720. <https://github.com/leanprover-community/mathlib4/blob/afb3068a5e9f765af7936d0952ff9e7a64d167b4/Mathlib/MeasureTheory/Measure/Hausdorff.lean#L717-L720>

`LipschitzWith.hausdorffMeasure_image_le`

Mathlib/MeasureTheory/Measure/Hausdorff.lean, lines 728–732. <https://github.com/leanprover-community/mathlib4/blob/afb3068a5e9f765af7936d0952ff9e7a64d167b4/Mathlib/MeasureTheory/Measure/Hausdorff.lean#L728-L732>

PR 27069 <https://github.com/leanprover-community/mathlib4/pull/27069>

References

- [1] Ralph Alexander. Random compact sets related to the kakeya problem. *Proceedings of the American Mathematical Society*, 53(2):415–419, 1975.
- [2] Yakov Babichenko, Yuval Peres, Ron Peretz, Perla Sousi, and Peter Winkler. Hunter, cauchy rabbit, and optimal kakeya sets. *Transactions of the American Mathematical Society*, 366, 07 2012.
- [3] A. S. Besicovitch. Sur deux questions d’intégrabilité des fonctions. *J. Soc. Phys.-Math., Perm*, 2:105–123, 1919.
- [4] A. S. Besicovitch. On kakeya’s problem and a similar one. *Mathematische Zeitschrift*, 27:312–320, 1928.
- [5] A. S. Besicovitch. On fundamental geometric properties of plane line-sets. *Journal of the London Mathematical Society*, s1-39(1):441–448, 1964.
- [6] Roy O. Davies. Some remarks on the kakeya problem. *Mathematical Proceedings of the Cambridge Philosophical Society*, 69:417 – 421, 1971.
- [7] Leonardo de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. The lean theorem prover (system description). In Amy P. Felty and Aart Middeldorp, editors, *Automated Deduction - CADE-25*, pages 378–388, Cham, 2015. Springer International Publishing.
- [8] Zeev Dvir. On the size of kakeya sets in finite fields. *Journal of the American Mathematical Society*, 22(4):1093–1097, 2009.
- [9] K. J. Falconer. *The geometry of fractal sets*. Cambridge University Press, USA, 1985.
- [10] K. J. Falconer. *Fractal geometry : mathematical foundations and applications*. Wiley, Chichester, 1990.
- [11] Jonathan M. Fraser. Fourier analytic properties of kakeya sets in finite fields, 2025.
- [12] F. Hausdorff. Dimension und äußeres maß. *Mathematische Annalen*, 79:157–179, 1919.
- [13] F. Hausdorff. *Set Theory: Translated from the German by John R. Aumann*. Chelsea Publishing Company, 1962.
- [14] William Alvin Howard. The formulae-as-types notion of construction. In Haskell Curry, Hindley B., Seldin J. Roger, and P. Jonathan, editors, *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*. Academic Press, 1980.
- [15] J.-P. Kahane. Trois notes sur les ensembles parfaits linéaires. *L’Enseignement Mathématique*, 15(1):185, 1969.
- [16] Sōichi Kakeya. Some problems on maximum and minimum regarding ovals. *Tohoku Science Reports*, 6:71–88, 1917.
- [17] Nets Katz and Terence Tao. Recent progress on the kakeya conjecture, 2000.
- [18] U. Keich. On lp bounds for kakeya maximal functions and the minkowski dimension in $\mathbb{R}^2$. *Bulletin of the London Mathematical Society*, 31(2):213–221, 1999.
- [19] T. W. Körner. Besicovitch via baire. *Studia Mathematica*, 158(1):65–78, 2003.
- [20] Mark Lewko. Finite field restriction estimates based on kakeya maximal operator

- estimates. *Journal of the European Mathematical Society*, 2014.
- [21] The mathlib Community. The lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020*, pages 367–381, New York, NY, USA, 2020. Association for Computing Machinery.
 - [22] Pertti Mattila. *Fourier Analysis and Hausdorff Dimension*. Number 150 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, United Kingdom, January 2015.
 - [23] B. Mendelson. *Introduction to Topology: Third Edition*. Dover Books on Mathematics. Dover Publications, 2012.
 - [24] M.H. Mortad. *Introductory Topology: Exercises and Solutions*. G - Reference, Information and Interdisciplinary Subjects Series. World Scientific, 2014.
 - [25] J.R. Munkres. *Topology*. Prentice Hall, 2000.
 - [26] Julius Pál. Ein minimumproblem für ovale. *Mathematische Annalen*, 83:311–319, 1921.
 - [27] Oskar Perron. Über einen satz von besicovitsch. *Mathematische Zeitschrift*, 28(1):383–386, 1928.
 - [28] Cosmin Pohoata and Dmitrii Zakharov. Generalized arithmetic kakeya, 2024.
 - [29] R. A McCoy R. C. Haworth. *Baire spaces*. Instytut Matematyczny Polskiej Akademii Nauk, 1977.
 - [30] Hans Rademacher. On a theorem of besicovitch. In *Studies in Mathematical Analysis and Related Topics*, pages 294–296. Stanford University Press, Stanford, Calif., 1962.
 - [31] Bertrand Russell. *Principles of Mathematics*. Routledge Classics. Routledge, London, 2010.
 - [32] Eric Sawyer. Families of plane curves having translates in a set of measure zero. *Mathematika*, 34(1):69–76, 1987.
 - [33] Eric Schechter. Chapter 20 - baire theory. In Eric Schechter, editor, *Handbook of Analysis and Its Foundations*, pages 530–546. Academic Press, San Diego, 1997.
 - [34] I. J. Schoenberg. On the besicovitch-perron solution of the kakeya problem. 1988.
 - [35] C.E. Silva. *Invitation to Ergodic Theory*. Student mathematical library. American Mathematical Society, 2008.
 - [36] Elias M. Stein and Rami Shakarchi. *Measure Theory*, pages 1–48. Princeton University Press, 2005.
 - [37] Terence Tao. From rotating needles to stability of waves: Emerging connections between combinatorics, analysis, and pde. *Notices Amer. Math. Soc.*, 48, 09 2000.
 - [38] Hong Wang and Joshua Zahl. Volume estimates for unions of convex sets, and the kakeya set conjecture in three dimensions, 2025.
 - [39] Thomas Wolff. Recent work connected with the kakeya problem. In *Prospects in Mathematics (Princeton, NJ, 1996)*, pages 129–162. American Mathematical Society, Providence, RI, 1999.