

A Formalisation of Gallagher’s Ergodic Theorem

Oliver Nash   

Imperial College London, UK

Abstract

Gallagher’s ergodic theorem is a result in metric number theory. It states that the approximation of real numbers by rational numbers obeys a striking “all or nothing” behaviour. We discuss a formalisation of this result in the Lean theorem prover. As well as being notable in its own right, the result is a key preliminary, required for Koukoulopoulos and Maynard’s stunning recent proof of the Duffin-Schaeffer conjecture.

2012 ACM Subject Classification Mathematics of computing → Probability and statistics

Keywords and phrases Lean proof assistant, measure theory, metric number theory, ergodicity, Gallagher’s theorem, Duffin-Schaeffer conjecture

Digital Object Identifier 10.4230/LIPIcs.ITP.2023.23

Supplementary Material *Software (Source Code)*: <https://github.com/leanprover-community/mathlib>, archived at `swh:1:dir:a022d44847c12ca961b4bd0c8bca4f9df70cf7df`

Acknowledgements It is a pleasure to thank Andrew Pollington who suggested this project during the conference Lean for the Curious Mathematician, held at Brown University (ICERM) in 2022. I also wish to thank Anatole Dedecker, Heather Macbeth, Patrick Massot, and Junyan Xu, all of whom were of direct assistance. Lastly I especially wish to thank Sébastien Gouëzel for many helpful suggestions and Kevin Buzzard for many useful conversations.

1 Introduction

In addition to recognising extraordinary achievements of young mathematicians, the Fields Medal provides a valuable service to the wider mathematical community: it draws attention to important recent results. In recent years, such attention has had a significant positive impact in the formalisation community. Buzzard, Commelin, and Massot’s formalisation of the definition of a perfectoid space [4] and Commelin, Topaz et al.’s spectacular success with the Liquid Tensor Experiment [6] were both the results of projects which formalised work of 2018 Fields Medalist Peter Scholze. Amongst other things, these projects demonstrate that today’s proof assistants are capable of handling the complicated constructions of contemporary mathematics.

In 2022, James Maynard was awarded a Fields Medal with a citation that highlighted his work on the structure of prime numbers as well as on *Diophantine approximation*. In the long form of the citation we read that:

Maynard has also produced fundamental work in Diophantine approximation, having solved the Duffin-Schaeffer conjecture with Koukoulopoulos.

Shortly after the announcement of Maynard’s award, Andrew Pollington suggested to the author that a formalisation of Koukoulopoulos and Maynard’s proof of the Duffin-Schaeffer conjecture would be a worthy target for formalisation. Recognising that this would be an enormous undertaking, he suggested focusing on various necessary preliminaries. Perhaps the most important of these is Gallagher’s ergodic theorem [8] (see also Koukoulopoulos and Maynard [12] lemma 5.1). The statement is as follows:

23:2 A Formalisation of Gallagher’s Ergodic Theorem

► **Theorem 1** (Gallagher’s theorem). *Let $\delta_1, \delta_2, \dots$ be a sequence of real numbers and let:*

$$W = \{x \in \mathbb{R} \mid \exists q \in \mathbb{Q}, |x - q| < \delta_{\text{denom}(q)} \text{ i.o.}\}.$$

Then W is almost equal to either \emptyset or \mathbb{R} .

This deserves a few remarks:

- The notation $\text{denom}(q)$ denotes the denominator of q (in lowest terms). It is a strictly positive natural number.
- Special attention should be paid to the letters “i.o.” appearing in the definition of W : these abbreviate the phrase “infinitely often”. The notation means that $x \in W$ iff there exists an *infinite sequence* of rationals q_0, q_1, \dots (which may depend on x) with $\text{denom}(q_0) < \text{denom}(q_1) < \dots$ satisfying $|x - q_i| < \delta_{\text{denom}(q_i)}$ for all i .
- The phrase “almost equal” characterises this as a theorem of *metric number theory*: it means that the sets are equal up to a set of Lebesgue measure zero.
- It is striking and not at all obvious that W should exhibit such dichotomous behaviour.

It is the purpose of this article to discuss the author’s formalisation of Gallagher’s theorem. It was carried out using the Lean proof assistant together with its `mathlib` library [14]. More precisely we formalised the following:

■ **Listing 1** Gallagher’s theorem [↗](#)

```
theorem add_well_approximable_ae_empty_or_univ
  (δ : ℕ → ℝ) (hδ : tendsto δ at_top (ℕ 0)) :
  (∀m x, ¬ add_well_approximable ℚ δ x) ∨ ∀m x, add_well_approximable ℚ δ x :=
```

The notation will be explained in the pages to come. With further work, one could drop the hypothesis $h\delta$, which says that $\delta_n \rightarrow 0$ as $n \rightarrow \infty$. In fact this highlights a curious feature of the proof: one makes two totally separate arguments, a measure-theoretic argument which assumes the hypothesis $h\delta$ and a number-theoretic argument assuming its negation. One then invokes the law of excluded middle to deduce the result unconditionally. The measure-theoretic argument assuming $h\delta$ is much harder and is what we have formalised.

The paper is intended for non-experts and the structure is as follows. In section 2 we outline relevant basic concepts so that we can reinterpret Gallagher’s theorem as a result about the limsup of thickenings of finite-order points in the circle. We also make some general remarks about metric number theory. In section 3 we introduce the most important foundational result required. Lebesgue’s density theorem is the workhorse of Gallagher’s proof. In section 4 we discuss a key measure-theoretic lemma due to Cassels which is of some independent interest in its own right. In section 5 we discuss the results about ergodic maps which we needed, emphasising the ergodicity of certain maps of the circle. In section 6 we introduce points of approximately finite order and use this language to give a proof of Gallagher’s theorem. We finish with section 7 where, amongst other things, we discuss further directions this work could be taken. For the most part we do not enter into the details of proofs. The main exception to this is the proof of Gallagher’s theorem itself since we hope our presentation of Gallagher’s ideas may make them more accessible than other accounts intended for specialists (such as [8] Theorem 1 or [11] Theorem 2.7(B)).

In keeping with `mathlib`’s stress on mathematical unity, all work was added directly to `mathlib`’s master branch in a series of 27 pull requests, collectively adding just over 3,500 net new lines of code. This work is thus automatically available to all future `mathlib` users. Throughout this text we also provide permalinks to relevant locations in `mathlib`; each one is indicated with the symbol [↗](#). We also provide a judiciously chosen set of code listings (such as listing 1 above) containing Lean code. Often our intention is to assist the reader who wishes to compare a key informal statement with its formal equivalent.

2 Basic concepts

We outline some basic concepts to fix notation and to assist non-experts.

2.1 Almost equal sets

Given a measurable space with measure μ , when there is no possibility of ambiguity about the measure, we shall use the notation:

$$s =_{a.e.} t, \tag{1}$$

to say two subsets s, t are almost equal with respect to μ . We recall that this is equivalent to the following pair of measure-zero conditions \square :

$$\mu(s \setminus t) = 0 \quad \text{and} \quad \mu(t \setminus s) = 0. \tag{2}$$

2.2 Obeying a condition infinitely often

Gallagher's theorem concerns a set of points obeying a condition "infinitely often". In general, given a sequence of subsets s_0, s_1, \dots of some background type X the expression $\exists \dots i.o.$ is used to define¹:

$$\{x : X \mid \exists n \in \mathbb{N}, x \in s_n \text{ i.o.}\} = \{x : X \mid \text{the set } \{n \in \mathbb{N} \mid x \in s_n\} \text{ is infinite}\}. \tag{3}$$

In fact there is another expression for this set; it is easy to see that \square :

$$\{x : X \mid \exists n \in \mathbb{N}, x \in s_n \text{ i.o.}\} = \limsup s, \tag{4}$$

where:

$$\limsup s = \bigcap_{n \geq 0} \bigcup_{i \geq n} s_i.$$

When formalising a result about the set of points belonging to some family of subsets infinitely often, one can thus phrase it in language of (3) or in the language of \limsup . We opted for the latter. This was preferable because \limsup makes sense for any complete lattice whereas (3) is specific to the lattice of subsets of a type. All API developed was thus more widely applicable.

In the course of the proof it is useful to work with the \limsup bounded by a predicate $p : \mathbb{N} \rightarrow \text{Prop}$. This can be defined:

$$\limsup_p s = \bigcap_{n \geq 0} \bigcup_{p(i), i \geq n} s_i. \tag{5}$$

The actual definition which we added `filter.blimsup` \square is slightly different so that it also applies in a conditionally complete lattice (such as \mathbb{R}) but we provided a lemma showing the equivalence to (5) for complete lattices (such as `set \mathbb{R}`) \square . Using `blimsup`, we can work with the \limsup of the subfamily defined by the predicate p without having to pass to the subtype of the indexing type. For example given two predicates p, q , we can express the useful identity:

$$\limsup_{p \vee q} s = \limsup_p s \cup \limsup_q s, \tag{6}$$

formally as:

¹ This is standard notation appearing throughout the informal literature.

23:4 A Formalisation of Gallagher’s Ergodic Theorem

■ **Listing 2** `lim sup` bounded by the logical or of two predicates [↗](#)

```
@[simp] lemma blimsup_or_eq_sup :
  blimsup u f (λ x, p x ∨ q x) = blimsup u f p ∪ blimsup u f q :=
```

without involving any subtypes. This is a standard design pattern used throughout `mathlib`.

2.3 Thickenings

Using the language introduced above, the set W appearing in the statement of theorem 1 may be defined as:

$$W = \limsup_{n>0} s,$$

where:

$$s_n = \{x \in \mathbb{R} \mid \exists q \in \mathbb{Q}, \text{denom}(q) = n, |x - q| < \delta_n\}.$$

These subsets s_n have a special form: they are *thickenings*. In general, given a metric space X , if $s \subseteq X$ and $\delta \in \mathbb{R}$, the (open) δ -thickening of s is:

$$\text{Th}(\delta, s) = \{x \in X \mid \exists y \in s, d(x, y) < \delta\}.$$

This generalises the concept of an open ball. Fortunately thickenings already existed [↗](#) in `mathlib` thanks to the work of Gouezel on the Gromov-Hausdorff metric [9]. We can thus express W as:

$$W = \limsup_{n>0} \text{Th}(\delta_n, \{q \in \mathbb{Q} \mid \text{denom}(q) = n\}).$$

Using the language of thickenings turned out to be very convenient formally, not just for Gallagher’s theorem but also for example in lemma 5 (discussed below).

2.4 The circle as a normed group

The subset W appearing in theorem 1 trivially satisfies the periodicity condition²:

$$1 + W = W,$$

and thus descends to a subset of the circle $\mathbb{S} = \mathbb{R}/\mathbb{Z}$. This quotient is actually a *normed* group. We recall that a normed (additive) group G is a group carrying a real-valued function:

$$\begin{aligned} G &\rightarrow \mathbb{R} \\ g &\mapsto \|g\| \end{aligned}$$

such that the distance function $d(g, h) = \|g - h\|$ satisfies the metric space axioms [↗](#).

In our case, given $x \in \mathbb{R}$ representing the coset $\hat{x} \in \mathbb{S}$, the norm obeys:

$$\|\hat{x}\| = |x - \text{round}(x)| \tag{7}$$

where `round(x)` is the nearest integer to x . Since the norm gives us a metric, we may speak of thickenings of subsets of \mathbb{S} .

² If q approximates x then $1 + q$ approximates $1 + x$ with the same error and `denom(1 + q) = denom(q)`.

Furthermore, $\mathbb{Q}/\mathbb{Z} \subseteq \mathbb{S}$ is exactly the set of points of finite order in \mathbb{S} . Gallagher's theorem may thus be regarded as establishing a special property enjoyed by the circle \mathbb{S} in the category of normed groups. As we shall see this is a useful point of view since the proof of Gallagher's theorem depends on the fact that certain transformations are ergodic when regarded as maps $\mathbb{S} \rightarrow \mathbb{S}$.

When this work began, `mathlib` already contained a model of the circle as complex numbers of unit length, called `circle` ³. Although this model is naturally equivalent to \mathbb{R}/\mathbb{Z} , the equivalence uses the exponential and logarithm maps which are irrelevant for our work. We thus introduced a second model called `add_circle` defined to be \mathbb{R}/\mathbb{Z} :

■ **Listing 3** The additive circle ³

```
def add_circle {K : Type*} [linear_ordered_add_comm_group K]
  [topological_space K] [order_topology K] (p : K) :=
  K / zmultiples p
```

When $p = 1$ this is exactly \mathbb{R}/\mathbb{Z} but we allow a general value of p to support other applications³.

As the names suggest, `circle` carries an instance of `mathlib`'s `group` class and `add_circle` carries an instance of `add_group`. It is interesting that `mathlib`'s additive-multiplicative design pattern so conveniently allows both models to coexist.

Substantial API for `add_circle` was then developed, notably an instance of the class `normed_add_comm_group` ³ satisfying the identity (7) ³ and a characterisation of the finite-order points using rational numbers ³:

$$\{y \in \mathbb{S} \mid o(y) = n\} = \{[q] \in \mathbb{S} \mid q \in \mathbb{Q}, \text{denom}(q) = n\}, \quad (8)$$

where the notation $o(y) = n$ means that y has order n .

Using all of our new language, the statement of Gallagher's theorem becomes:

► **Theorem 2** (Gallagher's theorem). *Let $\delta_1, \delta_2, \dots$ be a sequence of real numbers and let:*

$$W = \limsup_{n>0} \text{Th}(\delta_n, \{y \in \mathbb{S} \mid o(y) = n\}).$$

Then $W =_{a.e.} \emptyset$ or $W =_{a.e.} \mathbb{S}$.

Using (7) and (8), theorem 2 is trivially equivalent to theorem 1.

2.5 Metric number theory

Metric number theory is the study of arithmetic properties of the real numbers (and related spaces) which hold "almost everywhere" with respect to the Lebesgue measure. The arithmetic property in the case of Gallagher's theorem is approximation by rational numbers.

To illustrate, consider the set \mathbb{I} of real numbers which have infinitely-many quadratically-close rational approximations:

$$\begin{aligned} \mathbb{I} &= \{x \in \mathbb{R} \mid \exists q \in \mathbb{Q}, |x - q| < 1/\text{denom}(q)^2 \text{ i.o.}\} \\ &= \limsup_{n>0} \text{Th}(1/n^2, \{q \in \mathbb{Q} \mid \text{denom}(q) = n\}). \end{aligned}$$

³ For example `mathlib` uses $p = 2\pi$ to define angles ³.

23:6 A Formalisation of Gallagher’s Ergodic Theorem

It has been known at least since the early 19th Century, that \mathbb{I} is just the set of irrational numbers⁴:

$$\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}. \tag{9}$$

Considering this result from the point of view of metric number theory, we notice that since \mathbb{Q} has Lebesgue measure zero, \mathbb{I} is almost equal to \mathbb{R} . Thus the metric number theorist would be content to summarise (9) by saying that

$$\mathbb{I} =_{a.e.} \mathbb{R},$$

without worrying about exactly which numbers \mathbb{I} contains.

The benefit of metric number theorist’s point of view is that a great many questions have answers of this shape. Gallagher’s theorem is an especially-beautiful example of this phenomenon.

3 Doubling measures and Lebesgue’s density theorem

Lebesgue’s density theorem is a foundational result in measure theory, required for the proof of Gallagher’s theorem. Although we only needed to apply it to the circle, the density theorem holds quite generally and so we took some trouble to formalise it subject to quite weak assumptions⁵.

3.1 Doubling measures

A convenient class of measures for which the density theorem holds is the class of doubling measures (more precisely *uniformly locally* doubling measures).

► **Definition 3.** *Let X be a measurable metric space carrying a measure μ . We say μ is a doubling measure if there exists $C \geq 0$ and $\delta > 0$ such that for all $0 < \epsilon \leq \delta$ and $x \in X$:*

$$\mu(B(x, 2\epsilon)) \leq C\mu(B(x, \epsilon)).$$

where $B(x, r)$ denotes the closed ball of radius r about x .

The corresponding formal definition, which the author added to `mathlib` for the purposes of formalising the density theorem, is:

■ **Listing 4** Definition of doubling measures [↗](#)

```
class is_doubling_measure
  {α : Type*} [metric_space α] [measurable_space α] (μ : measure α) :=
  (exists_measure_closed_ball_le_mul [] : ∃ (C : ℝ ≥ 0), ∀f ε in  $\mathcal{N}[> 0]$ , ∀ x,
    μ (closed_ball x (2 * ε)) ≤ C * μ (closed_ball x ε))
```

The parameter δ is not explicitly mentioned in the code above because we use `mathlib`’s standard notation for the concept of a predicate holding eventually along a filter [↗](#).

⁴ Thanks to Michael Geißer and Michael Stoll, `mathlib` knows this fact [↗](#).

⁵ We were lucky that Sébastien Gouëzel had recently added an extremely general theory of Vitali families which made this possible.

For our application, we needed to apply the density theorem to the Haar measure [15] μ on the circle. Of course this turns out to be the familiar arc-length measure and so the volume of a closed ball of radius ϵ is given by μ :

$$\mu(B(x, \epsilon)) = \min(1, 2\epsilon).$$

Taking $C = 2$ we thus see that the Haar measure on the circle is doubling. We registered this fact using a typeclass instance as follows:

■ **Listing 5** The circle's doubling measure μ

```
instance : is_doubling_measure (volume : measure (add_circle T)) :=
```

The unit circle corresponds to taking $T = 1$, but the code allows any $T > 0$. Thanks to this instance, Lean knows that any results proved for doubling measures automatically holds for the Haar measure on the circle.

3.2 The density theorem

The version of the density theorem which we formalised is:

► **Theorem 4.** *Let X be a measurable metric space carrying a measure μ . Suppose that X has second-countable topology and that μ is doubling and locally finite. Let $S \subseteq X$ and $K \in \mathbb{R}$, then for almost all $x \in S$, given any sequence of points w_0, w_1, \dots and distances $\delta_0, \delta_1, \dots$, if:*

- $\delta_j \rightarrow 0$ as $j \rightarrow \infty$ and,
- $x \in B(w_j, K\delta_j)$ for large enough j ,

then:

$$\frac{\mu(S \cap B(w_j, \delta_j))}{\mu(B(w_j, \delta_j))} \rightarrow 1,$$

as $j \rightarrow \infty$.

Even in the special case $K = 1$ and $w_0 = w_1 = \dots = x$, the result is quite powerful⁶. A point x satisfying the property appearing in the theorem statement is known as a point of density 1. Using this language, Lebesgue's density theorem asserts that almost all points of a set have density 1. In particular if $\mu(S) > 0$ then there must exist a point of density 1 μ . As an example, if $X = \mathbb{R}$ and S is the closed interval $[0, 1]$, the set of points of density 1 is the open interval $(0, 1)$.

In fact the formal version which we added to `mathlib` is very slightly more general since it allows w and δ to be maps from any space carrying a filter. After a preparatory `variables` statement:

■ **Listing 6** Variables for the density theorem

```
variables {α : Type*} [metric_space α] [measurable_space α] (μ : measure α)
[is_doubling_measure μ] [second_countable_topology α] [borel_space α]
[is_locally_finite_measure μ]
```

it looks like this:

⁶ Indeed this is probably the most common version one finds in the literature.

■ **Listing 7** Lebesgue’s density theorem for doubling measures [↗](#)

```
lemma is_doubling_measure.ae_tendsto_measure_inter_div (S : set α) (K : ℝ) :
  ∀m x ∂μ.restrict S, ∀ {ι : Type*} {l : filter ι} (w : ι → α) (δ : ι → ℝ)
  (δlim : tendsto δ l (ℕ[>] 0))
  (xmem : ∀f j in 1, x ∈ closed_ball (w j) (K * δ j)), tendsto (λ j,
  μ (S ∩ closed_ball (w j) (δ j)) / μ (closed_ball (w j) (δ j))) 1 (ℕ 1) :=
```

The method of proof is essentially to develop sufficient API for `is_doubling_measure` to show that such measure spaces carry certain natural families of subsets called Vitali families and then to invoke the lemma `vitali_family.ae_tendsto_measure_inter_div` [↗](#) added by Gouëzel as part of an independent project [10].

■ **4** Cassels’s lemma

A key ingredient in the proof of Gallagher’s theorem is the following result due to Cassels.

► **Lemma 5.** *Let X be a measurable metric space carrying a measure μ . Suppose that X has second-countable topology and that μ is doubling and locally finite. Let s_0, s_1, \dots be a sequence of subsets of X and r_0, r_1, \dots be a sequence of real numbers such that $r_n \rightarrow 0$ as $n \rightarrow \infty$. For any $M > 0$ let:*

$$W_M = \limsup \text{Th}(Mr_n, s_n),$$

then:

$$W_M =_{a.e.} W_1,$$

i.e., up to sets of measure zero, W_M does not depend on M .

This essentially appears as lemma 9 in [5] in the special case that:

- (a) X is the open interval $(0, 1)$,
- (b) μ is the Lebesgue measure,
- (c) s_n is a sequence of points rather than a sequence of subsets.

Reusing the `variables` from listing 6, the formal version of lemma 5 which we added to `mathlib` looks like this:

■ **Listing 8** Cassels’s lemma [↗](#)

```
theorem blimsup_thickening_mul_ae_eq
  (p : ℕ → Prop) (s : ℕ → set α) {M : ℝ} (hM : 0 < M)
  (r : ℕ → ℝ) (hr : tendsto r at_top (ℕ 0)) :
  (blimsup (λ i, thickening (M * r i) (s i)) at_top p : set α) =m[μ]
  (blimsup (λ i, thickening (r i) (s i)) at_top p : set α) :=
```

Several remarks are in order:

- The syntax $s =^m[\mu] t$ is `mathlib`’s notation for sets (or functions) s, t being almost equal with respect to a measure μ . It is the formal equivalent of the popular informal notation (1).
- The type ascriptions `: set α` appear because of an unresolved *typeclass diamond* in `mathlib`’s library of lattice theory. The issue is that the type `set α` is definitionally equal to $\alpha \rightarrow \text{Prop}$. Since `Prop` is a complete boolean algebra [↗](#) it follows [↗](#) that $\alpha \rightarrow \text{Prop}$ is a complete boolean algebra. Unfortunately the definition [↗](#) of the complete boolean algebra structure on `set α`, though mathematically equal, is not definitionally equal

to that on $\alpha \rightarrow \text{Prop}$. Strictly speaking, because `set α` is a type synonym, this is a permissible diamond. The point is that ideally all lemmas and statements about `set α` (including results about its lattice structure) should depend only on its API and not in its definition. However in practice such “definitional abuse” does occur and can be convenient so it would be useful to resolve the diamond.⁷

- Listing 8 is stated in terms of `blimsup`, i.e., a `limsup` bounded by a predicate p . As discussed in section (2.2), this allows us to avoid having to deal with subtypes. We will see that this is convenient when applying this lemma in the proof of Gallagher’s theorem.
- The key ingredient in the proof of Cassels’s lemma is Lebesgue’s density theorem 4. In view of (2), Cassels’s lemma requires us to establish a pair of measure-zero conditions. According to whether $M < 1$ or $M > 1$, exactly one of these two conditions is trivial for the two sets appearing in the statement of Cassels’s lemma 5. To prove the non-trivial measure-zero condition, one argues by contradiction by assuming the measure is strictly positive, applying the density theorem to obtain a point of density 1, and showing that this is impossible for a doubling measure. The only non-trivial dependency is Lebesgue’s density theorem.
- Although the modifications required for the generalisation of this lemma from its original form in [5] are straightforward, the generalisation (c) from points to subsets (equivalently from balls to thickenings) is extremely useful formally. In the application of this lemma required for Gallagher’s theorem, s_n is the set of points of order n in the circle. In the informal literature, the version of lemma 5 for sequences of points can be applied because the circle has only finitely-many points of each finite order and so one can enumerate all points of finite order as a single sequence of points. This would be messy formally (for example it requires choosing arbitrary orderings) and in any case is not necessary given the more general result.

5 Ergodic theory

Ergodic theory is the study of measure-preserving maps. Given measure spaces (X, μ_X) and (Y, μ_Y) , a measurable map $f : X \rightarrow Y$ is measure-preserving if:

$$\mu_X(f^{-1}(s)) = \mu_Y(s),$$

for any measurable set $s \subseteq Y$. For example, given any $c \in \mathbb{R}$, taking Lebesgue measure on both domain and codomain, the translation $x \mapsto c + x$ is always measure-preserving whereas the dilation $x \mapsto cx$ is measure-preserving only if $c = \pm 1$. Fortunately `mathlib` already contained an excellent theory of measure-preserving maps.

5.1 Ergodic maps, general theory

Within ergodic theory, special attention is paid to ergodic maps.

► **Definition 6.** *Let (X, μ) be a measure space and $f : X \rightarrow X$ be measure-preserving. We say f is ergodic if for any measurable set $s \subseteq X$:*

$$f^{-1}(s) = s \implies s \text{ is almost equal to } \emptyset \text{ or } X.$$

⁷ The diamond is recorded in `mathlib` issue 16932 [↗](#). In fact it is only the `Inf` and `Sup` fields in the complete boolean algebra structures that differ definitionally so this should be fairly easy to resolve.

23:10 A Formalisation of Gallagher’s Ergodic Theorem

Ergodicity is key concept in the proof of Gallagher’s theorem and so we added the following definitions to `mathlib`:

■ **Listing 9** Definition of pre-ergodic [↗](#) and ergodic maps [↗](#):

```
structure pre_ergodic (μ : measure α . volume_tac) : Prop :=
  (ae_empty_or_univ : ∀ {s}, measurable_set s →
    f-1 s = s → s =m[μ] (∅ : set α) ∨ s =m[μ] univ)

structure ergodic (μ : measure α . volume_tac) extends
  measure_preserving f μ μ, pre_ergodic f μ : Prop
```

The reason for the intermediate definition `pre_ergodic` is to support the definition of quasi-ergodic maps which we also defined, but which do not concern us here.

We then developed some basic API for ergodic maps including the key result:

► **Lemma 7.** *Let X be a measurable space with measure μ such that $\mu(X) < \infty$. Suppose that $f : X \rightarrow X$ is ergodic, $s \subseteq X$ is measurable, and the image $f(s)$ is almost contained in s , then s is almost equal to \emptyset or X .*

This result is elementary but not quite trivial and appears formally as follows:

■ **Listing 10** Sets that are almost invariant by an ergodic map [↗](#):

```
lemma ae_empty_or_univ_of_image_ae_le [is_finite_measure μ]
  (hf : ergodic f μ) (hs : measurable_set s) (hs' : f '' s ≤m[μ] s) :
  s =m[μ] (∅ : set X) ∨ s =m[μ] univ :=
```

This is not the first time that ergodic maps have been formalised in a theorem prover and so we have kept the above account very brief. Indeed the Archive of Formal Proofs for Isabelle/HOL contains an impressive body of results about ergodic theory due to Sébastien Gouëzel with contributions from Manuel Eberl, available at the Ergodic Theory entry [↗](#). This entry contains many results about general ergodic theory that have not yet been added to `mathlib`. On the other hand, we needed to know that certain specific maps on the circle are ergodic and our formalisations of these results do appear to be the first of their kind. We discuss these next.

5.2 Ergodic maps on the circle

In order to prove Gallagher’s theorem, we needed the following result:

► **Theorem 8.** *Given $n \in \mathbb{N}$, the map:*

$$\begin{aligned} \mathbb{S} &\rightarrow \mathbb{S} \\ y &\mapsto ny \end{aligned}$$

is measure-preserving if $n \geq 1$ and is ergodic if $n \geq 2$.

The fact that $y \mapsto ny$ is measure-preserving follows from general uniqueness results for Haar measures. In fact the result holds for any compact, Abelian, divisible topological group. Thanks to `mathlib`’s extensive theory of Haar measure [15], it was easy to add a proof of this [↗](#). We encourage readers who are encountering this fact for the first time to examine figure 1 and appreciate why this result holds for \mathbb{S} despite failing for \mathbb{R} .

The proof that $y \mapsto ny$ is ergodic is harder. We proved it as corollary of the following lemma. We sketch a proof to give a sense of what is involved; it is not essential that the reader follow the details: the main point is that we needed to use Lebesgue’s density theorem.



■ **Figure 1** The map $f : y \mapsto 2y$ is measure-preserving.

► **Lemma 9.** *Let $s \subseteq \mathbb{S}$ be measurable and u_0, u_1, \dots be a sequence of finite-order points in \mathbb{S} such that:*

- $u_i + s$ is almost equal to s for all i ,
- the order $o(u_i) \rightarrow \infty$ as $i \rightarrow \infty$.

Then s is almost equal to \emptyset or X .

Proof. The result is fairly intuitive: s is almost equal to $u_i + s$ iff it is composed of a collection of $o(u_i)$ components, evenly-spaced throughout the circle, up to a set of measure zero. Since this holds for all i and $o(u_i) \rightarrow \infty$, such components must either fill out the circle or be entirely absent, up to a set of measure zero.

The way to turn the above intuitive argument into rigorous proof is to use Lebesgue's density theorem 4. We must show that if s is not almost empty then $\mu(s) = 1$. Lebesgue tells us that if s is not almost empty it must contain some point d of density 1. Using d , we construct the sequence of closed balls B_i centred on d such that $\mu(B_i) = 1/o(u_i)$. Because $u_i + s$ is almost s ,

$$\mu(s \cap B_i) = \mu(s)/o(u_i) = \mu(B_i)\mu(s).$$

However since d has density 1, we know that:

$$\mu(s \cap B_i)/\mu(B_i) \rightarrow 1.$$

These two results force us to conclude that $\mu(s) = 1$. ◀

The formal version is very slightly more general and appears in `mathlib` as follows:

■ **Listing 11** Formal statement of lemma 9 [↗](#)

```
lemma add_circle.ae_empty_or_univ_of_forall_vadd_ae_eq_self
  {s : set $ add_circle T} (hs : null_measurable_set s volume)
  {ι : Type*} {l : filter ι} [l.ne_bot] {u : ι → add_circle T}
  (hu1 : ∀ i, ((u i) +_v s : set _) =m[volume] s)
  (hu2 : tendsto (add_order_of ∘ u) l at_top) :
  s =m[volume] (∅ : set $ add_circle T) ∨ s =m[volume] univ :=
```

Theorem 8 follows from lemma 9 because any set s satisfying $f^{-1}(s) = s$ for $f : y \mapsto ny$ satisfies $u_i + s = s$ for the sequence:

$$u_i = [1/n^i] \in \mathbb{S}.$$

Note that we need $n \geq 2$ in order to have $o(u_i) = n^i \rightarrow \infty$. The formal statement appears in `mathlib` as follows:

23:12 A Formalisation of Gallagher's Ergodic Theorem

■ **Listing 12** Formal statement of theorem 8 [↗](#)

```
lemma add_circle.ergodic_nsmul {n : ℕ} (hn : 1 < n) :
  ergodic (λ (y : add_circle T), n · y) :=
```

In fact we needed the following mild generalisation of theorem 8:

► **Theorem 10.** *Given $n \in \mathbb{N}$ and $x \in \mathbb{S}$, the map:*

$$\begin{aligned} \mathbb{S} &\rightarrow \mathbb{S} \\ y &\mapsto ny + x \end{aligned}$$

is measure-preserving if $n \geq 1$ and is ergodic if $n \geq 2$.

This follows easily from theorem 8 because if we define the measure-preserving equivalence:

$$\begin{aligned} e : \mathbb{S} &\rightarrow \mathbb{S} \\ y &\mapsto \frac{x}{n-1} + y \end{aligned}$$

then a quick calculation reveals:

$$e \circ g \circ e^{-1} = f,$$

where $f : y \mapsto ny$ and $g : y \mapsto ny + x$. As a result, theorem 10 follows from theorem 8 via:

■ **Listing 13** The reduction of theorem 10 to theorem 8 [↗](#)

```
lemma ergodic_conjugate_iff {e : α ≃m β} (h : measure_preserving e μ μ') :
  ergodic (e ∘ f ∘ e.symm) μ' ↔ ergodic f μ :=
```

6 Gallagher's theorem

6.1 Points of approximate order

Recall the definition of the set $W \subseteq \mathbb{S}$ appearing in the statement of theorem 2:

$$W = \limsup_{n>0} \text{Th}(\delta_n, \{y \in \mathbb{S} \mid o(y) = n\}).$$

Key to the proof of theorem 2 is the way in which the sets $\text{Th}(\delta_n, \{y \in \mathbb{S} \mid o(y) = n\})$ interact with the group structure of \mathbb{S} . We thus made the following definition:

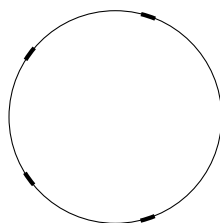
► **Definition 11.** *Let A be a seminormed group, $n \in \mathbb{N}$ (non-zero), and $\delta \in \mathbb{R}$. We shall use the notation:*

$$\mathbb{A}\mathbb{O}(A, n, \delta) = \text{Th}(\delta, \{y \in A \mid o(y) = n\}),$$

for the set of points that have approximate order n , up to a distance δ .

For example, as shown in figure 2, $\mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta)$ is a union of $\varphi(n)$ arcs of diameter 2δ , centred on the points $[m/n]$ with $0 \leq m < n$ and m coprime to n (where φ is Euler's totient function).

The formal counterpart of definition 11 is:



■ **Figure 2** The set of points of approximate order 5 in \mathbb{S} , up to a distance $\delta \approx 0.01$.

■ **Listing 14** Points of approximate order in a normed group \square

```
@[to_additive] def approx_order_of
  (A : Type*) [seminormed_group A] (n : ℕ) (δ : ℝ) : set A :=
  thickening δ {y | order_of y = n}
```

Using this language, the only properties of $\mathbb{A}\mathbb{O}(A, n, \delta)$ that we needed are as follows:

► **Lemma 12.** *Let A be a seminormed commutative group, $\delta \in \mathbb{R}$, $a \in A$, and $m, n \in \mathbb{N}$ (both non-zero). Then⁸:*

- (i) $m \cdot \mathbb{A}\mathbb{O}(A, n, \delta) \subseteq \mathbb{A}\mathbb{O}(A, n, m\delta)$ if m, n are coprime \square ,
- (ii) $m \cdot \mathbb{A}\mathbb{O}(A, nm, \delta) \subseteq \mathbb{A}\mathbb{O}(A, n, m\delta)$ \square ,
- (iii) $a + \mathbb{A}\mathbb{O}(A, n, \delta) \subseteq \mathbb{A}\mathbb{O}(A, o(a)n, \delta)$ if $o(a)$ and n are coprime \square ,
- (iv) $a + \mathbb{A}\mathbb{O}(A, n, \delta) = \mathbb{A}\mathbb{O}(A, n, \delta)$ if $o(a)^2$ divides n \square .

In fact property (iv) holds under the weaker assumption that $r(o(a))o(a)$ divides n where $r(l)$ denotes the radical of a natural number l , but we needed only the version stated in the lemma.

We made one last definition in support of theorem 2:

► **Definition 13.** *Let A be a seminormed group and $\delta_1, \delta_2, \dots$ a sequence of real numbers. We shall use the notation:*

$$\mathbb{W}\mathbb{A}(A, \delta) = \limsup_{n>0} \mathbb{A}\mathbb{O}(A, n, \delta_n),$$

for the set of elements of A that are well-approximable by points of finite order, relative to δ .

Note that $W = \mathbb{W}\mathbb{A}(\mathbb{S}, \delta)$ where W is the set appearing in the statement of theorem 2. The formal counterpart of definition 13 is:

■ **Listing 15** The set of well-approximable elements of a normed group \square

```
@[to_additive] def well_approximable
  (A : Type*) [seminormed_group A] (δ : ℕ → ℝ) : set A :=
  blimsup (λ n, approx_order_of A n (δ n)) at_top (λ n, 0 < n)
```

The additive version of this definition is `add_well_approximable`.

6.2 The main theorem

We are finally in a position to assemble everything and provide a proof of our main result. For the reader's convenience we reproduce the formal statement which appeared above in listing 1:

⁸ If $s \subseteq A$ the notation $m \cdot s$ means $\{my \mid y \in s\}$.

23:14 A Formalisation of Gallagher's Ergodic Theorem

■ Listing 16 Gallagher's theorem [↗](#)

```

theorem add_well_approximable_ae_empty_or_univ
  (δ : ℕ → ℝ) (hδ : tendsto δ at_top (ℵ 0)) :
  (∀m x, ¬ add_well_approximable ℚ δ x) ∨ ∀m x, add_well_approximable ℚ δ x :=

```

The notation $\forall^m x, \dots$ should be read “for almost all $x \dots$ ” and is standard `mathlib` notation [↗](#). Using the lemmas `filter.eventually_eq_empty` [↗](#) and `filter.eventually_eq_univ` [↗](#) the statement in listing 16 is equivalent to:

► **Theorem 14** (Gallagher's theorem with $\delta \rightarrow 0$). *Let $\delta_1, \delta_2, \dots$ be a sequence of real numbers such that $\delta_n \rightarrow 0$ as $n \rightarrow \infty$. Then $\mathbb{W}\mathbb{A}(\mathbb{S}, \delta)$ is almost equal to either \emptyset or \mathbb{S} .*

Proof. For each prime $p \in \mathbb{N}$ we define three sets⁹:

$$\begin{aligned}
 A_p &= \limsup_{n>0, p \nmid n} \mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta_n), \\
 B_p &= \limsup_{n>0, p \parallel n} \mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta_n), \\
 C_p &= \limsup_{n>0, p^2 \mid n} \mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta_n).
 \end{aligned}$$

Let $W = \mathbb{W}\mathbb{A}(\mathbb{S}, \delta)$; bearing in mind (6) it is clear that for any p :

$$W = A_p \cup B_p \cup C_p. \tag{10}$$

We claim that these sets have the following properties:

- (a) A_p is almost invariant under the ergodic map: $y \mapsto py$,
- (b) B_p is almost invariant under the ergodic map: $y \mapsto py + [1/p]$,
- (c) C_p is invariant under the map $y \mapsto y + [1/p]$.

To see why (a) holds, consider:

$$\begin{aligned}
 p \cdot A_p &= p \cdot \limsup_{n>0, p \nmid n} \mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta_n) \\
 &\subseteq \limsup_{n>0, p \nmid n} p \cdot \mathbb{A}\mathbb{O}(\mathbb{S}, n, \delta_n) \\
 &\subseteq \limsup_{n>0, p \nmid n} \mathbb{A}\mathbb{O}(\mathbb{S}, n, p\delta_n) && \text{by lemma 12 part (i)} \\
 &=_{a.e.} A_p && \text{by lemma 5.}
 \end{aligned}$$

A very similar argument shows why (b) holds except using parts (ii), (iii) of lemma 12 instead of part (i).

Claim (c) is actually the most straightforward and holds by direct application of lemma 12 part (iv).

Now if A_p is not almost empty for any prime p , then because it is almost invariant under an ergodic map, lemma 7 tells us that it must be almost equal to \mathbb{S} . Since $A_p \subseteq W$, W must also be almost equal to \mathbb{S} and we have nothing left to prove.

We may thus assume A_p is almost empty for all primes p . By an identical argument, we may also assume B_p is almost empty for all primes p . In view of (10), this means that:

$$W =_{a.e.} C_p \quad \text{for all } p.$$

Thus, by (c), W is almost invariant under the map $y \mapsto y + [1/p]$ for all primes p . The result then follows by applying lemma 9. ◀

⁹ The notation $p \parallel n$ means that p divides n exactly once.

Omitting code comments, the formal version of this ~ 30 line informal proof in `mathlib` requires 101 lines \square .

7 Final words

7.1 Removing the $\delta_n \rightarrow 0$ hypothesis

As mentioned in the introduction, the hypothesis that $\delta_n \rightarrow 0$ in theorem 14 may be removed. A nice follow-up project would be to supply the proof in this case. By replacing δ_n with $\max(\delta_n, 0)$, we may assume $0 \leq \delta_n$ for all n . Given this, if $\delta_n \not\rightarrow 0$, then in fact:

$$\mathbb{W}\mathbb{A}(\mathbb{S}, \delta) = \mathbb{S}.$$

Note that this is a true equality of sets; it is not a measure-theoretic result. The main effort would be to establish some classical bounds on the growth of the divisor-count and totient functions.

In fact Bloom and Mehta have already formalised some of the required bounds as part of their impressive Unit Fractions Project \square formalising Bloom’s breakthrough [2, 3]. Once the relevant results are migrated to `mathlib`, removing the $\delta_n \rightarrow 0$ hypothesis will become even easier.

7.2 The Duffin-Schaeffer conjecture

Given some sequence of real numbers $\delta_1, \delta_2, \dots$, Gallagher’s theorem tells us that $\mathbb{W}\mathbb{A}(\mathbb{S}, \delta)$ is almost equal to either \emptyset or to \mathbb{S} . The obvious question is how to tell which of these two possibilities actually occurs for the sequence in hand. The Duffin-Schaeffer conjecture, now a theorem thanks to Koukoulopoulos and Maynard, provides a very satisfying answer:

$$\mathbb{W}\mathbb{A}(\mathbb{S}, \delta) =_{a.e.} \begin{cases} \emptyset & \text{if } \sum \varphi(n)\delta_n < \infty, \\ \mathbb{S} & \text{if } \sum \varphi(n)\delta_n = \infty. \end{cases}$$

where φ is Euler’s totient function.

That $\mathbb{W}\mathbb{A}(\mathbb{S}, \delta) =_{a.e.} \emptyset$ if $\sum \varphi(n)\delta_n < \infty$ is very easy (it follows from the “easy” direction of the Borel-Cantelli theorem). The converse is extremely hard. It was first stated in 1941 [7] and was one of the most important open problems in metric number theory for almost 80 years.

A formal proof of the converse would be especially satisfying given how elementary the statement of the result is. After Gallagher’s theorem, perhaps the next best target is lemma 5.2 in [12], i.e., theorem 2 in Pollington and Vaughan [13].

7.3 Aistleitner, Borda, and Hauke’s quantitative results

The author is grateful to Christoph Aistleitner who contacted him after reading a preprint of this article. Aistleitner highlighted that Gallagher’s theorem is required for Koukoulopoulos and Maynard’s proof because the combination of Gallagher’s theorem with the results of Pollington and Vaughan in [13] turns the Duffin-Schaeffer conjecture into a purely arithmetic problem. He also highlighted that he, Borda, and Hauke have been able to take Koukoulopoulos and Maynard’s techniques further and obtain *quantitative* results even without using Gallagher’s theorem, see [1].

7.4 Developing against master

It would have been impossible to complete the work discussed here without the extensive theories of algebra, measure theory, topology etc. contained within `mathlib`. As we have said, all of our code was added directly to the `master` branch of `mathlib`; most of it is “library code”, not specific to Gallagher’s theorem.

Although it is harder to develop this way, we believe it is essential in order to permit formalisation of contemporary mathematics. We therefore wish to exhibit this project as further evidence that this workflow can succeed, and we hope to encourage even more people to follow suit.

References

- 1 Christoph Aistleitner, Bence Borda, and Manuel Hauke. On the metric theory of approximations by reduced fractions: a quantitative Koukoulopoulos-Maynard theorem. *Compos. Math.*, 159(2):207–231, 2023. doi:10.1112/S0010437X22007837.
- 2 Thomas F. Bloom. On a density conjecture about unit fractions. (*to appear*), 2021. arXiv:2112.03726.
- 3 Thomas F. Bloom and Bhavik Mehta. The Unit Fractions Project. (*to appear*), 2022. URL: <https://b-mehta.github.io/unit-fractions/>.
- 4 Kevin Buzzard, Johan Commelin, and Patrick Massot. Formalising perfectoid spaces. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, New Orleans, LA, USA, January 20-21, 2020*, pages 299–312, 2020. doi:10.1145/3372885.3373830.
- 5 J. W. S. Cassels. Some metrical theorems in Diophantine approximation. I. *Proc. Cambridge Philos. Soc.*, 46:209–218, 1950. doi:10.1017/s0305004100025676.
- 6 Johan Commelin, Adam Topaz, et al. The Liquid Tensor Experiment. (*to appear*), 2022. see also <https://www.nature.com/articles/d41586-021-01627-2>. URL: <https://github.com/leanprover-community/lean-liquid>.
- 7 R. J. Duffin and A. C. Schaeffer. Khintchine’s problem in metric Diophantine approximation. *Duke Math. J.*, 8:243–255, 1941. URL: <http://projecteuclid.org/euclid.dmj/1077492641>.
- 8 Patrick Gallagher. Approximation by reduced fractions. *J. Math. Soc. Japan*, 13:342–345, 1961. doi:10.2969/jmsj/01340342.
- 9 Sébastien Gouëzel. Formalizing the Gromov-Hausdorff space. *CoRR*, abs/2108.13660, 2021. arXiv:2108.13660.
- 10 Sébastien Gouëzel. A formalization of the change of variables formula for integrals in `mathlib`. In Kevin Buzzard and Temur Kutsia, editors, *Intelligent Computer Mathematics - 15th International Conference, CICM 2022, Tbilisi, Georgia, September 19-23, 2022, Proceedings*, volume 13467 of *Lecture Notes in Computer Science*, pages 3–18. Springer, 2022. doi:10.1007/978-3-031-16681-5_1.
- 11 Glyn Harman. *Metric number theory*, volume 18 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1998.
- 12 Dimitris Koukoulopoulos and James Maynard. On the Duffin-Schaeffer conjecture. *Ann. of Math. (2)*, 192(1):251–307, 2020. doi:10.4007/annals.2020.192.1.5.
- 13 A. D. Pollington and R. C. Vaughan. The k -dimensional Duffin and Schaeffer conjecture. *Mathematika*, 37(2):190–200, 1990. doi:10.1112/S0025579300012900.
- 14 The `mathlib` community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, New Orleans, LA, USA, January 20-21, 2020*, pages 367–381, 2020. doi:10.1145/3372885.3373824.
- 15 Floris van Doorn. Formalized Haar Measure. In Liron Cohen and Cezary Kaliszyk, editors, *12th International Conference on Interactive Theorem Proving (ITP 2021)*, volume 193 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:17, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITP.2021.18.