IMPERIAL

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

MSCI RESEARCH PROJECT

Viazovska's Magic Function in Dimension 8: A Formalisation in Lean

Author:
Sidharth Hariharan

Supervisor(s):
Bhavik Mehta

Submitted in partial fulfillment of the requirements for the MSci in Mathematics at Imperial College London

June 4, 2025

Abstract

Hi

Acknowledgments

Plagiarism statement

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

Signature: Sidharth Hariharan
Date: June 4, 2025

Contents

Chapter 1

Introduction

On 5 July, 2022, in Helsinki, Finland, the International Mathematical Union announced the names of the four mathematicians who were to be awarded the Fields Medal, the most coveted prize in the world of mathematics: Hugo Duminil-Copin, June Huh, James Maynard and Maryna Viazovska. Duminil-Copin, Huh and Maynard received this most prestigious honour for making several outstanding contributions to their specific fields of expertise—respectively, statistical physics, geometric combinatorics, and analytic number theory. Viazovska, on the other hand, was awarded the Fields Medal for a more focused line of research: the pursuit if the great mysteries of $\mathbb{R}^8$ and $\mathbb{R}^{24}$, chief amongst them the optimality of the E_8 and Leech lattice sphere packings in these spaces. Her solution in dimension 8 is particularly revolutionary: it uses insights from Fourier analysis and the theory of modular forms to construct a special function—the so-called Magic Function—that, in combination with a previous result by Cohn and Elkies, proves that the E_8 lattice packing is the densest possible sphere packing in $\mathbb{R}^8$. Very shortly afterwards, Cohn, Kumar, Miller, Radchenko and Viazovska were able to use similar ideas to prove that the Leech lattice packing is the densest possible sphere packing in $\mathbb{R}^{24}$.

Before Viazovska's remarkable breakthrough, the optimal sphere packing density was only known in dimensions 1, 2 and 3 [1]. Furthermore, Thomas Hales' solution in dimension 3 [2] was lengthy and involved extensive computer-assisted calculations; in contrast, Viazovska's proof in dimension 8 is elegant and concise. Even before Viazovska was awarded the Fields Medal, her work received wide acclaim from eminent mathematicians across the world: Peter Sarnak described it as "stunningly simple, as all great things are," and Akshay Venkatesh remarked that her Magic Function is very likely "part of some richer story" that connects to other areas of mathematics and physics [3]. Viazovska's work is a truly remarkable achievement in modern mathematics, with its elegance coming from the manner in which the many pieces of the puzzle fit perfectly together. One of the goals of this project is to offer a detailed exposition of one of those pieces: the construction of the so-called 'Magic Function' in dimension 8.

1.1 The Sphere Packing Problem

The Sphere Packing problem is a classical optimsation problem in mathematics. The problem can be formulated as follows.

Problem 1.1.1 (The Sphere Packing Problem in Dimensinon n). *Given some $n \in \mathbb{N}$, what is the densest possible non-overlapping arrangement of n-spheres of equal radius in*

$\mathbb{R}^n$?

Despite its rather straightforward formulation, Problem 1.1.1 is notoriously difficult to solve. Indeed, one obvious question that arises when one looks at the problem statement is how one might define the concept of density. It turns out that the definition is slightly unwieldy, though introducing a periodicity assumption on the sphere packing whose density one wishes to find considerably simplifies this problem.

A key challenge in solving the sphere packing problem in dimension n is the fact that proceeding inductively is not always helpful: 'stacking' the optimal n-dimensional sphere packing onto itself is not guaranteed to yield the optimal sphere packing in $n+1$ dimensions. [1]. In fact, this appraoch is known to fail in dimensions as low as 10 [4]. This is not obvious, not least because the approach does, in fact, succeed in the visualisable dimensions of 1, 2 and 3.

The 1-dimensional case is uninteresting. Visually, one can easily see that the densest possible arrangement of disjoint intervals of the form $(-r, r)$ on the real line consists of intervals centred at all points $2rm$ for $m \in \mathbb{Z}$. Indeed, one can fix r to be $\frac{1}{2}$ by rescaling the real line. The optimal packing therefore consists of open intervals of unit length centred at points on the lattice $\mathbb{Z} \subset \mathbb{R}$.

Figure 1.1: The $\mathbb{Z}$ lattice packing in dimension 1.

Rescaling gives us a powerful—if straightforward—simplification of the sphere packing problem where we can fix the radius of the spheres to a convenient value. Indeed, we only mention rescaling explicitly because it needs to be explicitly dealt with when formalising the problem. We will resume this discussion in . For now, we will take for granted the fact that rescaling does not affect the density of a sphere packing, meaning that we can talk about optimal sphere packings without worrying about the radius of the spheres in question. Bearing in mind that the spheres must all have the same radius, as per the statement of Problem 1.1.1, we will henceforth describe sphere packings simply by describing the points at which the spheres are centred.

Add cross-reference

The sphere packing problem in dimension 2, also known as the circle packing problem, turns out to be more interesting. A reasonable strategy for finding the densest packing is to 'stack' the $\mathbb{Z}$ lattice packing from dimension 1 onto itself in some manner, but the question remains as to exactly how this should be done. 'Stacking' it onto itself would involve extending the lattice $\mathbb{Z} \subset \mathbb{R} \subset \mathbb{R}^2$ into a lattice in $\mathbb{R}^2$ by extending the $\mathbb{R}$-basis $\{(1,0)\}$ of $\mathbb{R}$ (viewed as a subspace of $\mathbb{R}^2$) to an $\mathbb{R}$-basis of $\mathbb{R}^2$, and taking its $\mathbb{Z}$-span.

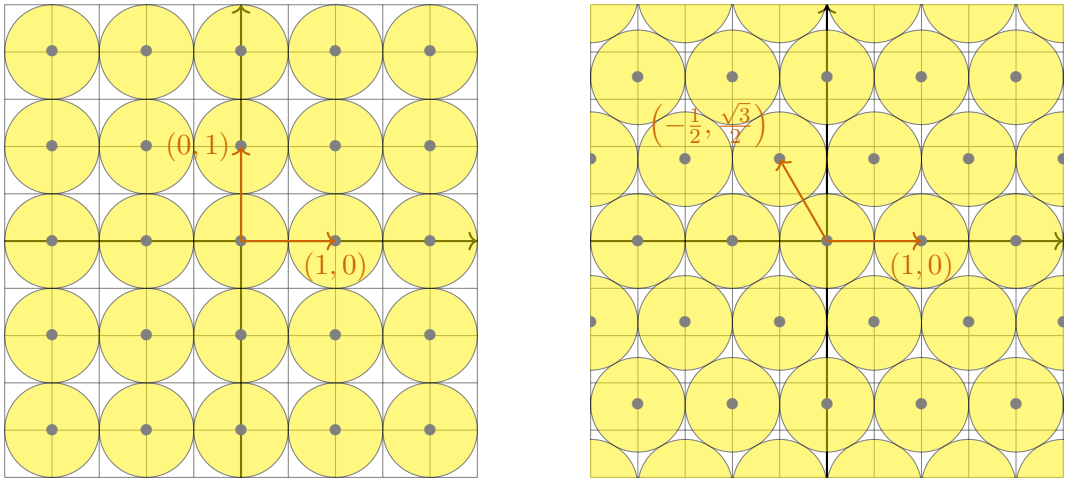
One natural way of doing this is to extend the lattice $\mathbb{Z} \subset \mathbb{R}$ to the lattice $\mathbb{Z}^2 \subset \mathbb{R}^2$ consisting of points with integer coordinates. This corresponds to the natural extension of $\{(1,0)\}$ to the standard $\mathbb{R}$-basis $\{(1,0),(0,1)\}$ of $\mathbb{R}^2$. See Figure 1.2a.

Unfortunately, this packing turns out to be sub-optimal . A better candidate is the A_2 lattice packing, corresponding to the extension of $\{(1,0)\}$ to the A_2 root basis $\left\{(1,0), \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)\right\}$. See Figure 1.2b. This packing is sometimes referred to as the *honeycomb packing* due to the fact that every circle has six neighbours, whose centres form the vertices of a regular hexagon.

Do we want to add the density?

It is well-known that the honeycomb packing is optimal in $\mathbb{R}^2$. What this means is that no circle packing has a density greater than that of the honeycomb packing. The original proof of this fact is attributed to Thue [5], and it is sometimes referred to in the literature as *Thue's Theorem*. Several other mathematicians have since constructed proofs of Thue's Theorem. One approach

based on an idea of Rogers's that does not require particularly sophisticated mathematical tools was outlined by Hales in [6].



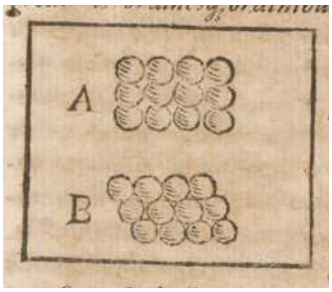
(a) The $\mathbb{Z}^2$ lattice packing. **(b)** The A_2 lattice packing.

Figure 1.2: Circle packings covering the square $\{(x, y) \subset \mathbb{R}^2 \mid -2.5 \leq x, y \leq 2.5\}$.

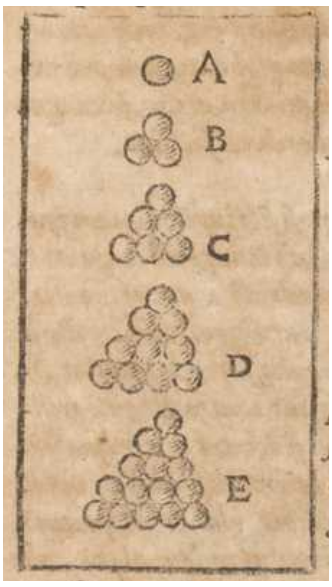
While we do not offer an exposition of Hales's or any other proof of Thue's Theorem, we include a small discussion of the sphere packing problem in two dimensions to offer some intuition as to why the A_2 packing is optimal. For simplicity, we will work under the assumption that the optimal packing in $\mathbb{R}^2$ is some extension of the $\mathbb{Z}$ lattice packing in $\mathbb{R}$, as defined above. We use the strategy of stacking $\mathbb{Z}$ packings on top of each other 'row by row', shifting rows around till their density cannot be further increased.

From Figure 1.2, one can convince oneself with relative ease that the A_2 packing is denser than the $\mathbb{Z}^2$ packing. This makes it denser than any packing that is sparser than the $\mathbb{Z}^2$ packing. In particular, we only need to improve the $\mathbb{Z}^2$ packing to construct the optimal packing. Since vertical shifts only push rows further apart, rendering the packing sparser, it suffices to consider horizontal row shifts. In the $\mathbb{Z}^2$ packing, each circle in a given row is in contact with only one sphere from the row below. A natural improvement is to shift rows about so each circle is in contact with *two* circles from the row below. This results in the A_2 lattice packing. This packing cannot be improved because it is impossible for a circle to be in contact with *three* circles from the row below due to the separation between circles in the one-dimension $\mathbb{Z}$ packing. See Figure 1.3a.

While this is far from a rigorous argument, this approach illustrates why we should not be surprised that the A_2 sphere packing is optimal: the key observation is that the A_2 packing maximises the number of neighbours a circle can have. Admittedly, it is neither obvious why this optimises density nor why the optimal packing involves stacking the $\mathbb{Z}$ packing repeatedly on itself, not least because we have yet to formally define the density of a sphere packing. We merely reassure the reader, at this stage, that the definitions and characteristics of the sphere packing problem in $\mathbb{R}^2$ strongly agree



(a)



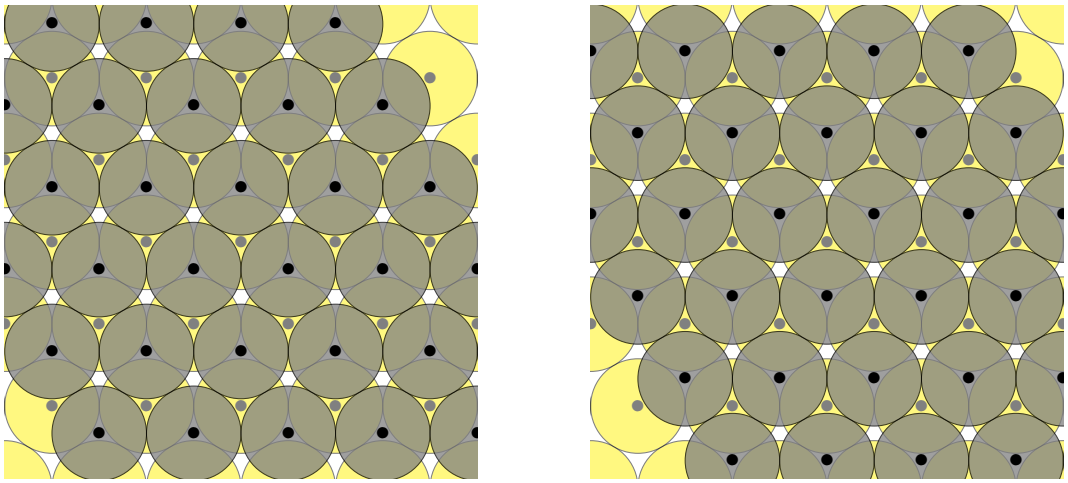
(b)

Figure 1.3: Diagrams from an essay written by Johannes Kepler in Latin in 1611 [7].

with visual intuition. With this, we close our discussion.

In dimension 3, too, it is tempting to replicate this strategy: we can attempt to stack the A_2 packing on top of itself, in layers instead of rows, in such a manner as maximises the number of neighbours a sphere can have. From trial and error, it appears to be the case that a sphere cannot be in contact with more than three neighbours from the layer below. This suggests that the optimal sphere packing in dimension 3 is given by stacking honeycomb arrangements on top of each other with spheres in each layer being nestled in the gaps between three spheres in the layer below.

As it turns out, unlike dimension 2, a characterisation in terms of the number of neighbours in the layer below does not describe a unique packing. In $\mathbb{R}^3$, spheres are so large that it is not possible to stack honeycomb arrangements on top of each other such that *all* gaps between spheres in one layer are occupied by spheres in the next. There is no unique stacking of honeycomb layers such that each sphere has exactly three neighbours in the layer below: in different stackings, the spheres in a layer might fill a different arrangement of gaps between spheres in the layer below, as shown in Figure 1.4. One can construct many different sphere packings in $\mathbb{R}^3$, all of which are as dense as possible, by varying how successive layers are placed. For instance, the sphere packing obtained by successively repeating the arrangement in Figure 1.4a and that obtained by alternating between the arrangements in Figure 1.4a and Figure 1.4b are globally different, despite having the same density and identical layers. The former is referred to as the *face-centred cubic packing* and the latter is referred to as the *hexagonal close-packing*.



(a) **(b)**

Figure 1.4: Two different ways of stacking the honeycomb packing on itself.

The face-centred cubic packing can be visualised by thinking about how spheres can be arranged tetrahedrally. For some n, let $T_n := n(n+1)/2$ be the nth triangular number[1]. Begin by arranging T_n spheres in a triangular formation. On top of this layer, arrange T_{n-1} spheres in a triangular formation, such that each sphere is nestled in the gaps between the spheres in the layer below, as in Figure 1.4a. Continue in this manner till there is only one ball left to be arranged. This leads to an arrangement in which the number of spheres is the nth tetrahedral number[2]. A key characteristic of the face-centred cubic packing is that it consists of such arrangements. In fact, in a 1611 essay whose title has been translated from Latin as *The Six-Cornered Snowflake* [7], it was asserted by Johannes Kepler that spheres cannot be more tightly packed together than they are in a tetrahedral arrangement. This assertion was accompanied

[1] See OEIS A000217
[2] See OEIS A000292

by an illustration: see Figure 1.3b. For over three centuries, this assertion remained unproven, and was referred to as the *Kepler Conjecture*. It was only in 2005 that a paper proving the Kepler Conjecture, written by Thomas Hales, was published [2].

The complexity of the sphere packing problem in dimension 3 is illustrated not only by the time elapsed between Kepler's original assertion and a proof being published but also by the length of Hales's paper. Indeed, in an expository account of his proof published in 2000, five years before the publication of the full paper in the Annals, Hales recounted how a jury of twelve referees, despite having been in deliberation for over a year, had yet to make a "thorough, independent check of the computer code" he had written to perform the elaborate calculations on which "every aspect of [his proof] is based" [6]. In January 2003, at the Joint Math Meetings in Baltimore, USA, Hales announced that he intended to formally verify his proof [8]. The paper authored by Hales and his collaborators on their successful formalisation of his argument was only published in 2017. Therefore, not only did the Kepler Conjecture take close to 400 years to solve, but it took nearly two decades to eliminate any doubt as to the correctness of the solution.

At first glance, this appears to set a dangerous precedent for the sphere packing problem in other dimensions. It well might, for there is much we do not understand about the behaviour of spheres in high dimensions. In the words of Henry Cohn, "each dimension has its own idiosyncracies and charm" [4]. That being said, in the specific cases of dimensions 8 and 24, this turns out to work in our favour.

The solutions in dimensions 8 and 24 are products of the same recipe, which consists primarily of two ingredients. The first is a linear programming bound from a 2003 paper by Henry Cohn and Noam Elkies [9, Theorem 3.1] on all sphere packing densities in $\mathbb{R}^n$. The second is the remarkable insight that the theory of modular forms can be used to obtain tight bounds in dimensions 8 and 24, equal to the densities of the E_8 and Leech lattice packings respectively.

The applicability of the theory of modular forms comes from the formulation of Cohn and Elkies's theorem: if a function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies certain conditions, then *all* sphere packings in $\mathbb{R}^n$ are bounded above by a quantity that depends on f. The trick is therefore not only to find a function satisfying the Cohn-Elkies conditions but to find one for which the Cohn-Elkies bound also corresponds to the density of some sphere packing in $\mathbb{R}^n$. Viazovska's groundbreaking contribution was constructing such a function in dimension 8 using the theory of modular forms. The function is often referred to as the Magic Function, a term we shall adopt in this project because it befits the nature of Viazovska's achievement. A similar approach was used in dimension 24 by Cohn, Kumar, Miller, Radchenko and Viazovska.

As tempting as it is to continue our discussion on the sphere packing problem, this project does consist of two parts: a mathematical examination of the construction of Viazovska's Magic Function in dimension 8 and a formalisation thereof. We will therefore pause this discussion and take a detour into the world of formalisation, which will offer context for the second part of this project.

1.2 The Formalisation Movement

While Hales announced his intent to formally verify his proof of the Kepler Conjecture in 2003, it was not till 2006, after Hales's solution appeared in the *Annals*, that a formal description of Hales's formalisation project was published. Of his motivations, Hales wrote:

> *In response to the lingering doubt about the correctness of the proof, at the beginning*

> *of 2003, I launched the* Flyspeck *project, whose aim is a complete formal verification of the Kepler Conjecture. In truth, my motivations for the project are far more complex than a simple hope of removing residual doubt from the minds of few referees. Indeed, I see formal methods as fundamental to the long-term growth of mathematics.* [10]

Formal theorem proving was not unheard of in 2006. Interactive theorem provers, such as Coq and PRL, have existed since the 1980s. However, it was still a relatively young field, and the amount of mathematics that had been formalised was limited. Hales's project was immensely ambitious, and the fact that it succeeded, despite taking over a decade, is impressive.

There is something prophetic about Hales's "far more complex" motivations for launching the Flyspeck project. The field of formal theorem proving has grown rapidly in the last decade, and interactive theorem provers like Lean are slowly making their way into mainstream mathematics. An excellent example of this is the formal verification of Gowers, Green, Manners and Tao's proof of Marton's Conjecture [11], which was formally verified in Lean in just three weeks. In particular, their proof was formally verified *before* their paper was submitted for publication. The paper appeared in the *Annals* in March 2025.

There are many advantages of formal theorem proving. One advantage is the fact that formally proved theorems are verified by a proof assistant. When code written in proof assistants is compiled, if there are no errors, then the proof can be thought of as being 'correct', in the sense of being consistent with the axioms of the proof assistant.

Finish

1.3 The Scope of this Project

In November 2023, I had the privilege of meeting Maryna Viazovska while pursuing an exchange programme at the Swiss Federal Institute of Technology, Lausanne, where she is based. We began discussing formalising her solution to the sphere packing problem in 8 dimensions, and soon initiated a collaboration with Christopher Birkbeck, Seewoo Lee, and Gareth Ma, with invaluable assistance from Kevin Buzzard, Utensil Song, and Patrick Massot. On 31 May 2024, Viazovska formally announced at the ICMS workshop *Formalisation of Mathematics: Workshop for Women and Mathematicians of Minority Gender* that we would be attempting to formalise her groundbreaking paper.

Viazovska's original paper [12] is divided into five sections. The first section introduces sphere packings and develops basic theory; the second discusses the Cohn-Elkies linear programming bounds [9, Theorem 3.1]; the third offers some background on the theory of modular forms; the fourth constructs two radial, Schwartz Fourier eigenfunctions with double zeroes at almost all points on the E_8 lattice; and finally, the fifth uses these eigenfunctions to construct the "Magic Function", a Schwartz function that satisfies the conditions of Cohn and Elkies's theorem to give an upper bound for all sphere packings in $\mathbb{R}^8$ that is equal to the density of the E_8 packing. The first two sections were formalised collaboratively in July and August 2024, and the third section is actively being worked on by Birkbeck and Lee. This project focuses on formalising the fourth and fifth sections of Viazovska's paper. The code written for this section is primarily my own, and I have credited the contributions of others where appropriate.

The primary objective of this thesis is to offer a mathematical exposition of the fourth and fifth sections of Viazovska's original paper and to provide an account of the formalisation process.

Say where we are with the formalisation before submitting.

Chapter 2

The Sphere Packing Problem in Dimension 8

The purpose of this chapter is to conduct a detailed examination of Viazovska's original paper solving the sphere packing problem in dimension 8 [12] and the formalisation blueprint [13]. While we have already seen the high-level idea in Section 1.1, in this chapter, we will take a closer look at the mathematical details.

We will begin by providing precise mathematical definitions for sphere packings, densities, and the sphere packing constant. We will then discuss the linear programming bound conceived by Cohn and Elkies [9, Theorem 3.1] (or, more precisely, the slight modification thereof that is more directly applicable: see [12, Theorem 2] and [13, Theorem 5.1]). Finally, we will include a small discussion on the theory of modular forms and establish its relevance to the subsequent chapters of this thesis, which will focus on the construction of the 'Magic Function' (denoted as g in [12, Theorem 3]).

In each section, we will not only mathematical details from the sources outlined above but also an overview of the choices made and challenges encountered when formalising these notions. While this chapter primarily concerns the contents of the first three sections of [12], which are not within the scope of this project, there is an undeniable relevance of both the informal and the formal definitions and results. It is therefore necessary to include a detailed treatment thereof before we can construct g and prove it satisfies the desired 'Magic' properties.

2.1 Preliminaries

Before we begin defining things formally, we must include a small disclaimer about the terminology we have been using—and will continue to use—in this project. While Problem 1.1.1 is usually referred to as the *sphere* packing problem, a sphere is not usually thought to have an interior. Typically, in any metric space X with metric d, the *sphere* of radius $r \geq 0$ centred at $x \in X$ is defined to be $\{y \in X \mid d(x, y) = r\}$. In other words, the sphere consists only of a surface. In contrast, the sphere packing problem involves packing *solid balls*. One can see why, in [6], Hales opines that a more proper term for the problem would be the *ball packing problem*. Nevertheless, in this project, we will continue to use the standard terminology, but we include this disclaimer so the reader bears in mind two things: first, that we will often mean 'ball' when we use the word 'sphere', and second, that we work with balls instead of spheres in Lean. We will also mention that it is convenient to require that the balls in question be open, so that the

condition that spheres cannot overlap but merely touch tangentially can be shortened to that of disjointedness. We introduce notation.

Notation. For some $d \in \mathbb{N}$, $x \in \mathbb{R}^d$ and $r > 0$, we denote

$$B_d(x, r) := \left\{ y \in \mathbb{R}^d \;\middle|\; \|x - y\| < r \right\}$$

We organise this section into three subsections. The first defines fundamental notions about sphere packings. The second introduces the properties of two important, and closely related, classes of sphere packings, namely, lattice packings and periodic packings. The third subsection studies the most important sphere packing for our project: the E_8 lattice packing.

2.1.1 Sphere Packing Fundamentals

We begin by defining a sphere packing. As we have stated, we want sphere packings to consist of disjoint spheres of the same radius. Given that lying on the interior of a certain sphere corresponds to being within some distance from its centre, we can capture this notion of disjointedness by imposing a separation condition on the set of centres of the sphere packing.

Definition 2.1.1 (Sphere Packing). Fix $d \in \mathbb{N}$ and $X \subset \mathbb{R}^d$. Assume that there exists a real number $r > 0$, known as the **separation radius**, such that

$$\|x - y\| \geq r$$

for all distinct $x, y \in X$. We define the **sphere packing with centres at** X to be

$$\mathcal{P}(X) := \bigcup_{x \in X} B_d(x, r)$$

Note that the assumption that a separation radius exists is very important.

Non-Example 2.1.2. Let $d = 1$ and $X = \mathbb{R}$. Consider the set

$$\bigcup_{x \in \mathbb{R}} B_1(x, r) = \bigcup_{x \in \mathbb{R}} (x - r, x + r)$$

For any $r > 0$, the above union is all of $\mathbb{R}$. However, it does not make sense to construct a sphere packing whose set of centres is the entirety of $\mathbb{R}$, as this would involve spheres overlapping. It is precisely to avoid such constructions that we impose the condition that r be a separation radius on the set of centres.

Since all the information about a sphere packing is encoded in its set of centres and the corresponding separation radius (which must exist in order for the set of centres to be a valid set of centres for a sphere packing), we decided that a sphere packing would be formalised purely as a set of centres with a valid separation, and that a separate definition would be made to obtain the open balls that constitute the packing. We packaged the data of

- the set of centres

- the separation radius
- the (automatically checked) condition that the separation radius is positive
- the condition that the set of centres is, indeed, separated by this radius

into a `structure` called `SpherePacking`: see [14, `SpherePacking.Basic.SpherePacking`].

Is this an acceptable way of citing the documentation? Should the repo be made public?

We now define finite density, an indicator of how much of a bounded region of space a sphere packing covers.

Definition 2.1.3 (Finite Density). Let $\mathcal{P}$ be a sphere packing. For all $R > 0$, define the **finite density** to be

$$\Delta_{\mathcal{P}}(R) := \frac{\mathrm{Vol}(\mathcal{P} \cap B_d(0, R))}{\mathrm{Vol}(B_d(0, R))}$$

where Vol is the Lebesgue measure on $\mathbb{R}^d$.

Finite density is a somewhat local notion, in that it expresses sphere how closely packed spheres are in a bounded region. The sphere packing problem, on the other hand, examines the notion of closeness on a more global level. While taking the limit of finite densities as the radius of the bounding region approaches infinity might seem like a natural way to define density, it is not obvious that this limit always exists. Therefore, we define density to be the limit superior instead.

Definition 2.1.4 (Density). Let $\mathcal{P}$ be a sphere packing. Define the **density** of $\mathcal{P}$ to be

$$\Delta(\mathcal{P}) := \limsup_{R \to \infty} \Delta_{\mathcal{P}}(R)$$

where $\Delta_{\mathcal{P}}(R)$ is the finite density of $\mathcal{P}$, as defined in Definition 2.1.3.

As one might expect, finite density and density are invariant under scaling.

Proposition 2.1.5. *Let $\mathcal{P}$ be a sphere packing. Fix $\lambda > 0$. Denoting by $\lambda\mathcal{P}$ the sphere packing obtained by scaling the spheres and the set of centres in $\mathcal{P}$ by a factor of λ, we have*

$$\Delta_{\mathcal{P}}(R) = \Delta_{\lambda\mathcal{P}}(\lambda R)$$

for all $R > 0$. Similarly, we have

$$\Delta(\mathcal{P}) = \Delta(\lambda\mathcal{P})$$

The sphere packing problem asks for the sphere packing that achieves the highest possible density. We can be formal about the notion of the highest possible density.

Definition 2.1.6 (Sphere Packing Constant). The **sphere packing constant** in $\mathbb{R}^d$, for

any $d > 0$, is defined to be

$$\Delta_d := \sup\Big(\big\{\Delta_{\mathcal{P}} \;\big|\; \mathcal{P} \text{ is a sphere packing in } \mathbb{R}^d\big\}\Big)$$

Proposition 2.1.5 tells us that it suffices to take the supremum over sphere packings of separation 1, because a sphere packing $\mathcal{P}$ can be scaled down by its separation radius to obtain a sphere packing of separation 1.

Proposition 2.1.7. *For all d, we have*

$$\Delta_d = \sup\Big(\big\{\Delta_{\mathcal{P}} \;\big|\; \mathcal{P} \textit{ is a sphere packing in } \mathbb{R}^d \textit{ with separation } 1\big\}\Big)$$

As intuitive as this result might seem, we do mention it here because it is something we have to deal with explicitly when formalising the theory of sphere packings in Lean. The first instance when we really see it in action is the proof of Theorem 2.2.5.

The objective of the sphere packing problem in any dimension d is to optimise the sphere packing constant Δ_d. As we have seen, this is a highly non-trivial thing to do. We can, however, offer a trivial upper-bound on sphere packing density.

Lemma 2.1.8. *For any sphere packing $\mathcal{P}$ and $R > 0$, we have that $\Delta_{\mathcal{P}}(R) \leq 1$.*

Proof. This is an immediate consequence of the fact that $\mathcal{P} \cap B_d(0, R) \subseteq B_d(0, R)$. □

This immediately gives us the following basic facts.

Corollary 2.1.9. *For any sphere packing $\mathcal{P}$, we have that $\Delta_{\mathcal{P}} \leq 1$.*

Corollary 2.1.10. *For any $d \in \mathbb{N}$, $\Delta_d \leq 1$.*

This is not a very good upper-bound. However, it tells us that the sphere packing constant in any number of dimensions is a finite real number in the interval $(0, 1]$. There is still some work to be done before we can give better bounds on the sphere packing constant. Furthermore, it is unclear whether the sphere packing constant actually is, for a general d, the density of a sphere packing in $\mathbb{R}^d$. Nevertheless, this is a good starting point.

A great deal of basic sphere packing API in Lean was developed in July 2024 for the project to formalise Viazovska's solution in dimension 8. The majority of the code was written by Gareth Ma, who also made significant improvements to the design choices I had made when setting up the project. The definitions and results in this section have all been formalised, and information about the code that has been written can be found in the project documentation [14, `SpherePacking.Basic.SpherePacking`].

In the next subsection, we discuss a special class of sphere packings that have periodicity properties with respect to lattices.

2.1.2 Lattice and Periodic Sphere Packings

We begin by defining lattices and briefly commenting on existing `Mathlib` API on lattices. There are primarily two ways in which lattices are defined in mathematical literature. A lattice in some Euclidean space $\mathbb{R}^n$ is either described as the $\mathbb{Z}$-span of some $\mathbb{R}$-basis of $\mathbb{R}^n$ or as a discrete, co-compact subgroup. One can borrow characteristics from both definitions to construct other equivalent definitions.

The characteristics described in both definitions do exist `Mathlib`. However, given that one of the objectives of creating a unified mathematics library is centralisation, a combination of these definitions is used as the *definition* of a `class` that we call `IsZLattice` and information about its many properties, as well as the $\mathbb{Z}$-span construction, are encoded in theorems. In particular, we have a theorem that tell us that every lattice is a free $\mathbb{Z}$-submodule, meaning it has a $\mathbb{Z}$-basis, and that this $\mathbb{Z}$-basis is actually an $\mathbb{R}$-basis of the ambient space. Furthermore, we have a result that every object constructed in that manner is a lattice. Results about types bearing the `IsZLattice` instance (ie, lattices as they are defined in `Mathlib`) live in the `ZLattice` namespace, whereas results about $\mathbb{Z}$-spans of $\mathbb{R}$-bases live in the `ZSpan` namespace.

Create citation for Mathlib.

We begin by stating the `Mathlib` definition of a lattice.

Definition 2.1.11 (Lattice). A **lattice** in a Euclidean space $\mathbb{R}^n$ is a discrete $\mathbb{Z}$-submodule of $\mathbb{R}^n$ such that its $\mathbb{R}$-span contains every element in $\mathbb{R}^n$.

The `Mathlib` definition is more general, and works for any normed vector space over a normed field. Here, the word 'discrete' means that the lattice is discrete in a topological sense, meaning that the subspace topology on the lattice is precisely the discrete topology.

Definition 2.1.12 (Periodic Sphere Packing). Let $\Lambda \subset \mathbb{R}^d$ be a lattice. We say a sphere packing $\mathcal{P}(X)$ with spheres centred at points in $X \subset \mathbb{R}^d$ is **periodic with respect to** Λ, or **Λ-periodic**,

$$\lambda + X = X$$

ie, for all $\lambda \in \Lambda$ and $x \in X$, we have that $\lambda + x \in X$.

We define Periodic Sphere Packings in Lean as extending the definition of Sphere Packings by creating a `structure` called `PeriodicSpherePacking` that packages the additional data of

- the lattice, viewed as a $\mathbb{Z}$-submodule of the ambient space $\mathbb{R}^d$
- the condition that the set of centres is periodic with respect to this $\mathbb{Z}$-submodule
- the (automatically checked) condition[1] that the $\mathbb{Z}$-submodule is discrete
- the (automatically checked) condition[1] that the discrete $\mathbb{Z}$-submodule is a lattice

The definition is in [14, `SpherePacking.Basic.SpherePacking`].

Lattice packings are a special class of periodic packings.

[1] more precisely, the automatically inferred instance

Definition 2.1.13 (Lattice Packing). Let $\Lambda \subset \mathbb{R}^d$ be a lattice. The Λ **lattice packing** is the sphere packing with centres at points in Λ. Such a sphere packing admits a separation radius because Λ is discrete and is Λ-periodic because Λ is closed under addition.

In Section 2.1.3, we will briefly examine a specific lattice packing, the E_8 lattice packing.

The periodicity property of a periodic sphere packing can be exploited to derive a more convenient formula for its density.

Proposition 2.1.14. *Let $\mathcal{P}(X)$ be a sphere packing with centres at $X \subset \mathbb{R}^d$ and separation r that is periodic with respect to some lattice $\Lambda \subset \mathbb{R}^d$. We have that*

$$\Delta_{\mathcal{P}(X)} = \left|{}^{X}\!/_{\Lambda}\right| \frac{\mathrm{Vol}\left(B_d\left(0, \frac{r}{2}\right)\right)}{\mathrm{Vol}\left({}^{\mathbb{R}^d}\!/_{\Lambda}\right)} \tag{2.1.1}$$

where $\left|{}^{X}\!/_{\Lambda}\right|$ is the number of orbits of the additive Λ-action on X and $\mathrm{Vol}\left({}^{\mathbb{R}^d}\!/_{\Lambda}\right)$ is the volume of the fundamental domain of the Λ-action on $\mathbb{R}^d$.

The proof is beyond the scope of this M4R project, but was formalised in Summer 2024: see `PeriodicSpherePacking.density_eq'` in [14, `SpherePacking.Basic.PeriodicPacking`].

Just as we defined the sphere packing constant for any dimension $d \in \mathbb{N}$, we can define a *periodic* sphere packing constant in any dimension.

Definition 2.1.15 (Periodic Sphere Packing Constant). For all $d \in \mathbb{N}$, define the **periodic sphere packing constant in dimension** d to be

$$\Delta_d^{\text{periodic}} = \sup\left(\left\{\Delta_{\mathcal{P}} \;\middle|\; \mathcal{P} \text{ is a periodic sphere packing in } \mathbb{R}^d\right\}\right)$$

The power of periodic sphere packings is illustrated by a rather surprising fact.

Proposition 2.1.16. *For all $d \in \mathbb{N}$,*

$$\Delta_d = \Delta_d^{\text{periodic}}$$

We do not prove this result here, as it is beyond the scope of this M4R. A proof can be found in [9, Appendix A]. We do, however, mention that Proposition 2.1.16 can be combined with Proposition 2.1.7 to give the following.

Proposition 2.1.17. *For all $d \in \mathbb{N}$,*

$$\Delta_d = \sup\left(\left\{\Delta_{\mathcal{P}} \;\middle|\; \mathcal{P} \textit{ is a periodic sphere packing in } \mathbb{R}^d \textit{ with separation } 1\right\}\right)$$

Proposition 2.1.16 tells us that finding a sphere packing that satisfies the *periodic* sphere packing constant gives us the optimal sphere packing in dimension d. Proposition 2.1.17 allows

us to focus our search even more. We will exploit these two facts in Section 2.2, where we will construct an upper bound for all sphere packing densities in dimension d by constructing an upper-bound for the periodic sphere packing constant in dimension d. When constructing this upper-bound, we will exploit the fact that periodic sphere packings admit a 'nice' density formula (cf. Proposition 2.1.14). The results we have seen about periodic sphere packings will thus greatly simplify our task of finding the optimal sphere packing in dimension 8.

We will end with a discussion on dual lattices, which will come up soon when we discuss the Poisson Summation Formula, an integral tool that not only proves the all-important Cohn-Elkies Linear Programming Bound but also narrows our search for the elusive magic function by pointing us towards Schwartz functions.

The dual lattice of a lattice is essentially its dual space analogue. Viewing a lattice in $\mathbb{R}^d$ as a free $\mathbb{Z}$-submodule of $\mathbb{R}^d$, we can view its dual lattice as its dual $\mathbb{Z}$-submodule of the dual $\mathbb{Z}$-module $\left(\mathbb{R}^d\right)^*$. Indeed, this is how we use the dual lattice of a lattice in Lean. For our purposes, however, we offer a slightly more convenient definition.

Definition 2.1.18 (Dual Lattice). Fix $d > 0$ and let $\Lambda \subset \mathbb{R}^d$ be a lattice. We define the **dual lattice** of Λ to be

$$\Lambda^* := \left\{ y \in \mathbb{R}^d \;\middle|\; \langle x, y \rangle \in \mathbb{Z} \text{ for all } x \in \Lambda \right\}$$

If we identify each $y \in \Lambda^*$ with the map $\langle \text{-}, y \rangle \in \left(\mathbb{R}^d\right)^*$, then we can see that Λ^* is indeed the dual $\mathbb{Z}$-submodule of $\Lambda \subset \mathbb{R}^d$ in the dual space $\left(\mathbb{R}^d\right)^*$. The primary convenience of our definition is that it views Λ^* as a subset of $\mathbb{R}^d$. This will be useful.

We are now ready to discuss a special sphere packing in $\mathbb{R}^8$: the E_8 sphere packing.

2.1.3 The E_8 Lattice Packing

It is quite remarkable that E_8 should show up when discussing sphere packings. At its core, E_8 is an irreducible root system. It shows up in the classification of important classes of objects like irreducible Coxeter groups, crystallographic Coxeter groups, and semi-simple Lie algebras over $\mathbb{C}$. E_8 is not a classical root system but an *exceptional* root system, meaning that the geometric properties of its roots cannot be found in irreducible root systems in all dimensions.

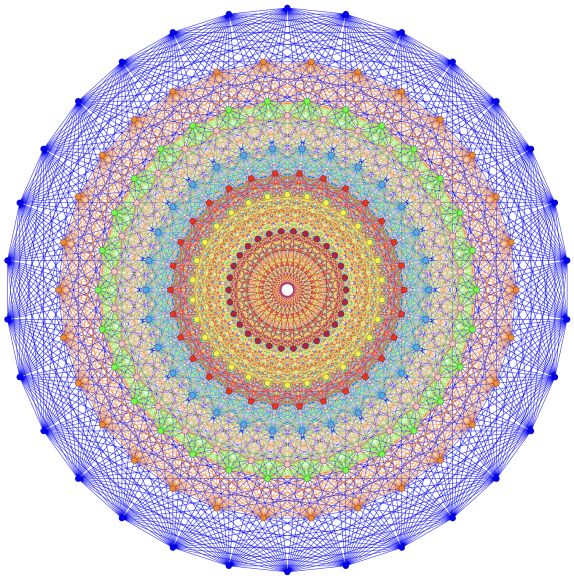


Figure 2.1: The Coxeter projection of the E_8 root system. [15]

The E_8 root system consists of 240 vectors in $\mathbb{R}^8$ that are permuted by a certain finite subgroup of the 8-dimensional orthogonal group. This group is sometimes referred to as the E_8 Coxeter group or as the Weyl group of the E_8 lattice. These roots can be divided into 8 orbits, each of which corresponds to one of the 'layers' of concentric circles in Figure 2.1. The dots in the figure correspond to projections of the roots onto a plane on which a specific type of element of the Coxeter group, known as a

Coxeter element, acts as a rotation. This visualisation offers a convenient—and aesthetically pleasing—means of visualising this collection of 8-dimensional vectors and appreciating some of its symmetry.

There is more than one way of characterising the E_8 lattice, denoted Λ_8. The 'ZSpan' characterisation is that it is the $\mathbb{Z}$-span of the so-called *simple roots* of the E_8 root system, the simple roots being a distinguished basis of $\mathbb{R}^8$ that is contained in the E_8 root system. Another characterisation that follows from a 1938 paper by Louis Mordell is that up to isomorphism, the E_8 lattice is the unique positive-definite, even, unimodular lattice in $\mathbb{R}^8$. We instead give the following explicit definition of the E_8 lattice, which we attempted to reconcile with the 'Zspan' characterisation in Lean.

Definition 2.1.19 (The E_8 Lattice). The E_8 lattice consists of all vectors in $\mathbb{R}^8$ such that either all coordinates are integers or all coordinates are half-integers and the sum of all coordinates is even. That is,

$$\Lambda_8 := \left\{ (x_1, \dots, x_8) \in \mathbb{Z}^8 \cup \left(\mathbb{Z} + \frac{1}{2}\right)^8 \;\middle|\; \sum_{i=1}^{8} x_i \equiv 0 \pmod 2 \right\}$$

For the construction of the magic function, we will primarily rely on two facts.

Theorem 2.1.20 (Properties of the E_8 Lattice). *The following are true of the E_8 lattice.*

1. *For all $x, y \in \Lambda_8$, $\|x - y\| \geq \sqrt{2}$.*
2. *The density of the sphere packing centred at points in Λ_8 with separation $\sqrt{2}$ is*

$$\frac{\pi^4}{384} \approx 0.2536695$$

The sphere packing to which we refer in the second point of the above theorem is precisely the E_8 sphere packing.

Theorem 2.1.21. *The dual lattice of the E_8 lattice is the E_8 lattice itself. That is, $(E_8)^* = E_8$ as free $\mathbb{Z}$-submodules of $\mathbb{R}^8$.*

This follows directly from the definitions of the dual lattice (Definition 2.1.18) and the E_8 lattice (Definition 2.1.19), and we omit the proof. It will be useful later on, as it will give us an insight into the desired properties of the magic function.

Finally, we mentiont that the covolume of the E_8 lattice is 1.

Theorem 2.1.22. $\mathrm{Vol}\left(\mathbb{R}^8 \big/ \Lambda_8\right) = 1.$

The way this can be proven is by showing that the determinant of the matrix consisting of an explicit basis of simple roots of the E_8 root system has determinant 1. We do not go into detail here.

The reason we stated the above properties is because we can use them to simplify the computa-

tions in the Poisson Summation Formula over the E_8 lattice. We will discuss this formula in the next section, but we will not examine the specific case of the E_8 lattice till we get to Chapter 3.

2.2 The Cohn-Elkies Linear Programming Bounds

Arguably, the result that most radically changed the sphere packing game was the linear programming bound constructed by Henry Cohn and Noam Elkies [9, Theorem 3.1]. The bound transforms the sphere packing problem from a geometric one to an analytic one. For all $d \in \mathbb{N}$, it posits the existence of a family of upper-bounds on the sphere packing constant Δ_d, indexed by functions $f : \mathbb{R}^d \to \mathbb{R}$ that satisfy certain conditions.

The power of this result is that it offers a systematic approach to prove that a certain sphere packing $\mathcal{P}$ is optimal in $\mathbb{R}^d$. The optimality condition means that the density of $\mathcal{P}$ is equal to the sphere packing constant Δ_d, which is equivalent to requiring that the density of $\mathcal{P}$ be greater than or equal to the density of any other packing in $\mathbb{R}^d$. The theorem proved by Cohn and Elkies tells us we can accomplish this by

1. identifying a function $f : \mathbb{R}^d \to \mathbb{R}$ that satisfies the conditions of the theorem, giving an upper-bound for Δ_d, and

2. showing that the upper-bound indexed by f is equal to the density of the packing $\mathcal{P}$.

As simple as this sounds, it took close to fourteen years from the publication of Cohn and Elkies's paper for it to be used to concretely construct an optimal sphere packing. The real trick is to construct the right function f to use in the process outlined above. We shall soon see the non-triviality of this task first-hand.

The original result [9, Theorem 3.1] is a bit different from the version that was chosen to be formalised. Firstly, the original result was stated for a very general class of functions known as *admissible functions*. For our purposes, however, it suffices to look at Schwartz functions, which are not only admissible but also have useful properties that we will exploit later. We will remark, however, that at the time when Cohn and Elkies proposed their bound, it was not known that the solution to the sphere packing problem in dimensions 8 and 24 would only involve Schwartz functions. Furthermore, it might be possible that solutions in other dimensions would require the full generality of Cohn and Elkies's original result. Nevertheless, we will restrict our attention to Schwartz functions for the time being, not only because it is sufficient for our purposes but also because the theory of Schwartz functions has been developed quite substantially in `mathlib`.

Another minor difference between the original result and the version we work with is that the original result was stated as an upper-bound on all *centre densities* of sphere packings. The centre density of a sphere packing is merely a rescaling of its density by a factor of $\mathrm{Vol}(B_d(0,1))^{-1}$. Instead of encoding the information of the amount of sphere packing volume per unit volume of the ambient space, the centre density encodes the information of the number of centres of the sphere packing per unit volume of the ambient space. We sidestep these nuances by stating the result in terms of quantities we have defined.

Before we prove the result, we need to state an important dependency: the **Poisson Summation Formula over Lattices**. Again, we only state it for Schwartz functions, but it applies in greater generality. The Formula does not exist in `mathlib`, and there is some debate as to the generality in which it must be stated in the library. It has been stated as part of the sphere packing formalisation effort, but has not been proven.

include code reference

The Poisson Summation Formula essentially relates a sum of a Schwartz function over a lattuce to a sum of its Fourier transform over another lattice. Therefore, before we state the Poisson Summation Formula, we say a few words about Fourier transforms, Fourier inversion, and the formalisation of these definitions in Lean.

2.2.1 Fourier Analysis and the Poisson Summation Formula

The subject of Fourier analysis is deep, and has applications to a number of areas in pure and applied mathematics. There are deeper undertones to the role of the Fourier transform in Viazovska's proof of the optimality of the E_8 lattice packing in $\mathbb{R}^8$: much of the underlying motivation comes from broader Fourier interpolation results that are beyond the scope of this thesis and this formalisation project. In this subsection, we do not do more than define the Fourier transform and its inverse. We will also very briefly discuss the formalisation of these definitions in Lean.

We begin by defining the Fourier transform of a function.

Definition 2.2.1 (Fourier Transform). Fix $m, n \in \mathbb{N}$ and let $f : \mathbb{R}^m \to \mathbb{C}^n$ be an L^1 function. We define the **Fourier transform** of f to be the function

$$\mathcal{F}(f) : \mathbb{R}^m \to \mathbb{C}^n : \xi \mapsto \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i \langle x, \xi \rangle}\, \mathrm{d}x$$

In the `mathlib` definition of the Fourier transform , we have no assumptions on f. This is because the integral of a function that is not integrable is defined to be zero. cite

As is often done in the literature, we adopt the following alternative notation for the Fourier transform.

Notation. Given an L^1 function $f : \mathbb{R}^m \to \mathbb{C}^n$, we denote its Fourier transform by $\widehat{f}$.

We use the $\widehat{\ }$ and $\mathcal{F}$ notations interchangeably.

We can also define the inverse Fourier transform.

Definition 2.2.2 (Inverse Fourier Transform). Fix $m, n \in \mathbb{N}$ and let $f : \mathbb{R}^m \to \mathbb{C}^n$ be an L^1 function. We define the **inverse Fourier transform** of f to be the function

$$\mathcal{F}^{-1}(f) : \mathbb{R}^m \to \mathbb{C}^n : x \mapsto \int_{\mathbb{R}^d} f(\xi)\, e^{2\pi i \langle x, \xi \rangle}\, \mathrm{d}\xi$$

Again, in the `mathlib` definition, we have no assumptions on f.

Confusingly, the inverse Fourier transform is not always the inverse of the Fourier transform—at least, not as per the `mathlib` definition, because we do not assume integrability. Proving "Fourier inversion" formally—that is, that $\mathcal{F}(\mathcal{F}(f)) = f$ for some function f—is therefore highly nontrivial, and requires assumptions on f. We will see that this is true for an important class of functions called Schwartz functions.

That being said, it turns out the following result is true for *all* functions f, if we define the

Fourier and inverse Fourier transforms the way they are defined in `mathlib`.

Lemma 2.2.3. *Let $m, n \in \mathbb{N}$ and let $f : \mathbb{R}^m \to \mathbb{C}^n$ be L^1. For all $x \in \mathbb{R}^m$, we have that*

$$\mathcal{F}^{-1}(f)(x) = \mathcal{F}(f)(-x)$$

The proof is straightforward, and has been formalised in `mathlib`.

We now recall the definition of the dual of a lattice (Definition 2.1.18). The reason why the concept of a dual lattice is relevant is that it allows us to state the **Poisson Summation Formula**, which relates the sum of a function over a lattice to the sum of its Fourier transform over the dual lattice.

Theorem 2.2.4 (Poisson Summation Formula over Lattices). *Let $d > 0$ and let $\Lambda \subset \mathbb{R}^d$ be a lattice. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Schwartz function. For all vectors $v \in \mathbb{R}^d$, we have*

$$\sum_{\ell \in \Lambda} f(\ell + v) = \frac{1}{\mathrm{Vol}\left(\mathbb{R}^d / \Lambda\right)} \sum_{m \in \Lambda^*} \widehat{f}(y) e^{-2\pi i \langle v, m \rangle}$$

Variants of this classical result and its proof can be found in several sources, such as [16, Chapter VII, §6, Proposition 15] and [17, Chapter VII, §7, Equation (VII, 7:5)]. While it has been stated in Lean, it has not been proven yet for lattices other than $\mathbb{Z} \subset \mathbb{R}$ due to a multitude of challenges associated with generalising the argument to higher dimensions.

Armed with this important result, we are ready to state and prove the Cohn-Elkies Linear Programming Bound for Schwartz functions.

2.2.2 The Cohn-Elkies Linear Programming Bound

We now state the most important intermediate result in the proof of the optimality of the E_8 lattice packing in $\mathbb{R}^8$.

Theorem 2.2.5 (Cohn and Elkies, 2003 [9, Theorem 3.1]). *If $f : \mathbb{R}^d \to \mathbb{R}$ is a Schwartz function satisfying the conditions*
(CE1) *f is not identically zero.*
(CE2) *For all $x \in \mathbb{R}^d$, if $\|x\| \geq 1$ then $f(x) \leq 0$.*
(CE3) *For all $x \in \mathbb{R}^d$, $\widehat{f}(x) \geq 0$.*
then we have the following bound on the sphere packing constant Δ_d:

$$\Delta_d \leq \frac{f(0)}{\widehat{f}(0)} \cdot \mathrm{Vol}\left(B_d\left(0, \frac{1}{2}\right)\right)$$

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Schwartz function satisfying the conditions (CE1)-(CE3). By Proposition 2.1.16, it suffices to prove that

$$\Delta_d^{\text{periodic}} \leq \frac{f(0)}{\widehat{f}(0)} \cdot \mathrm{Vol}\left(B_d\left(0, \frac{1}{2}\right)\right)$$

Fix a sphere packing $\mathcal{P}$ that is periodic with respect to some lattice Λ. We need to show that $\Delta_{\mathcal{P}} \leq \frac{f(0)}{\widehat{f}(0)} \cdot \mathrm{Vol}\big(B_d\big(0, \frac{1}{2}\big)\big)/$ By Proposition 2.1.5, we can assume that the separation radius of $\mathcal{P}$ is 1: if it is not, we can scale it appropriately, and this will not affect its density. Denote the set of centres of $\mathcal{P}$ by X. By Proposition 2.1.14, we need to show that

$$\frac{\big|X/\Lambda\big|}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \cdot \mathrm{Vol}\Big(B_d\Big(0, \frac{1}{2}\Big)\Big) \leq \frac{f(0)}{\widehat{f}(0)} \cdot \mathrm{Vol}\Big(B_d\Big(0, \frac{1}{2}\Big)\Big) \tag{2.2.1}$$

It turns out to be easier to show the equivalent inequality

$$\frac{\big|X/\Lambda\big|^2}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \cdot \widehat{f}(0) \leq \big|X/\Lambda\big| \cdot f(0) \tag{2.2.2}$$

Applying (CE2) to our assumption that $\|x - y\| \geq 1$ for all distinct $x, y \in X$, one can show that

$$\big|X/\Lambda\big| \cdot f(0) \geq \sum_{x \in X} \sum_{y \in X/\Lambda} f(x - y) \tag{2.2.3}$$

One can then apply the Poisson Summation Formula (cf. Theorem 2.2.4) and perform some computations to show that

$$\sum_{x \in X} \sum_{y \in X/\Lambda} f(x - y) = \frac{1}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \sum_{m \in \Lambda^*} \widehat{f}(m) \cdot \left| \sum_{x \in X/\Lambda} e^{2\pi i \langle x, m \rangle} \right|^2$$

(CE3) tells us this is a sum of non-negative real numbers. We can therefore bound this sum below by the term corresponding to $m = 0$. Therefore,

$$\begin{aligned} \frac{1}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \sum_{m \in \Lambda^*} \widehat{f}(m) \cdot \left| \sum_{x \in X/\Lambda} e^{2\pi i \langle x, m \rangle} \right|^2 &\geq \frac{1}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \widehat{f}(0) \cdot \left| \sum_{x \in X/\Lambda} e^{2\pi i \langle x, 0 \rangle} \right|^2 \\ &= \frac{\big|X/\Lambda\big|^2}{\mathrm{Vol}\big(\mathbb{R}^d/\Lambda\big)} \cdot \widehat{f}(0) \end{aligned} \tag{2.2.4}$$

Putting these computations together gives us the desired result. □

We have a proof of this result in our repository, but some of the lemmas on which it depends do not have complete (ie, "`sorry`-free") proofs. The details we have skipped are not relevant to the scope of this thesis, but the details we have included will give us an important condition that the magic function should obey. Therefore, not only the statement but also the proof of Theorem 2.2.5 will be important to understand the construction of the magic function.

We are now ready to discuss a key ingredient in the construction of the magic function: the theory of modular forms. We will not explore this in too much detail, as it does not fall within the scope of this thesis. We will not venture too far beyond the fundamentals.

2.3 A Word on Modular Forms

In this section, we give a brief introduction to the theory of modular forms. Birkbeck, Loeffler and others have formalised several results in the theory of modular forms, and a significant portion of their work has been merged into `mathlib`. Definitions and results from this section

that pertain to Viazovska's solution to the sphere packing problem in $\mathbb{R}^8$ that do not feature in `mathlib` are being actively formalised by Birkbeck and Lee, with contributions from Ma.

First, we introduce the following useful notation.

Notation. For the remainder of this paper, denote the Complex upper half-plane by $\mathbb{H}$. That is, define $\mathbb{H} := \{z \in \mathbb{C} \mid 0 < \mathrm{Im}(z)\}$.

This corresponds to the `mathlib` notation for the upper half-plane.

A key motivating idea in the study of modular forms is the study of the action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{H}$ by Möbius transformations via

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z := \frac{az+b}{cz+d}$$

That matrix multiplication corresponds to the composition of Möbius transformations is a well-known fact in Complex Analysis. One can hence show that the above is indeed a group action.

Both the identity $I \in \mathrm{SL}(2,\mathbb{Z})$ and the negative identity $-I \in \mathrm{SL}(2,\mathbb{Z})$ have the same (trivial) action on $\mathbb{H}$. Indeed, the $\mathrm{SL}(2,\mathbb{Z})$ action descends to a faithful action of $\mathrm{PSL}(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z})\big/\{\pm I\}$ on $\mathbb{H}$. Since we are more interested in the *actions* of matrices in $\mathrm{SL}(2,\mathbb{Z})$ and $\mathrm{PSL}(2,\mathbb{Z})$ than we are in their entries, we often do not distinguish between the two groups.

The `mathlib` definition of a modular form is more general than the first definitions of modular forms often seen in literature (see [16, Chapter VII, §2, Definition 4] and [18, Definition 1.1.2]), and instead matches subsequent definitions that generalise these first definitions. Modular forms are usually described as functions that are holomorphic on the upper half-plane that are invariant under the $\mathrm{SL}(2,\mathbb{Z})$-action up to an *automorphy factor* of a certain *weight*. This *weight* is defined as the *weight of the modular form*. However, one is often interested in invariance under not all of $\mathrm{SL}(2,\mathbb{Z})$, but certain *principal congruence subgroups* or subgroups containing such subgroups, known as *congruence subgroups*. Each principal congruence subgroup has a *level*, which is defined to be the *level* of a modular form whose congruence subgroup is principal. Modular forms, as they are defined in `mathlib` and in the blueprint [13], are therefore indexed by two properties: a *congruence subgroup of* $\mathrm{SL}(2,\mathbb{Z})$, which indicates the scope of invariance under the $\mathrm{SL}(2,\mathbb{Z})$-action, and a *weight*, which gives the extent of invariance under the action of elements of the subgroup in question.

To give a complete definition of the weight of a modular form, we need to define automorphy factors and the slash action notation.

Definition 2.3.1 (Automorphy Factors and Slash Actions). Fix $k \in \mathbb{Z}$, $z \in \mathbb{H}$ and $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{Z})$. Define the **automorphy factor of weight** k to be

$$j_k(z,\gamma) := (cz+d)^{-k} \tag{2.3.1}$$

For any function $f : \mathbb{H} \to \mathbb{C}$, with k and γ as above, the **slash operator** maps f to a new function $f \mid_k \gamma : \mathbb{H} \to \mathbb{C}$ given by

$$(f \mid_k \gamma)(z) := j_k(z,\gamma)\, f(\gamma \cdot z) = (cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right) \tag{2.3.2}$$

The action of γ mapping f to $f\mid_k \gamma$ via the weight k slash operator is sometimes referred to as a **slash action**.

It is clear, from the above definition, that $f\mid_0 \gamma = f\circ\gamma$ for al $\gamma \in \mathrm{SL}(2,\mathbb{Z})$. That is, if $f = f\mid_0 \gamma$, then $f = f\circ\gamma$, that is, f is invariant under composition with (the action of) γ. If $f = f\mid_k \gamma$ for some $k \in \mathbb{Z}$ and $\gamma \in \mathrm{SL}(2,\mathbb{Z})$, we can view the weight k as indicating the 'extent of invariance' of f under composition with γ.

To give a complete definition of the congruence subgroup/level of a modular form, we need to define congruence subgroups. The idea is to express the scope of the slash-invariance exhibited by a modular form with respect to the action of $\mathrm{SL}(2,\mathbb{Z})$—that is, the set of elements of $\mathrm{SL}(2,\mathbb{Z})$ under which we have slash-invariance—in the language of modular arithmetic.

Definition 2.3.2 (Congruence Subgroup). Fix $N \in \mathbb{N}$. The **level N principal congruence subgroup** of $\mathrm{SL}(2,\mathbb{Z})$, denoted $\Gamma(N)$, is defined to be the kernel of the surjective group homomorphism from $\mathrm{SL}(2,\mathbb{Z})$ to $\mathrm{SL}\big(2,\mathbb{Z}/N\mathbb{Z}\big)$ that comes from reducing modulo N. That is,

$$\Gamma(N) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod N \right\} \tag{2.3.3}$$

More generally, a subgroup Γ of $\mathrm{SL}(2,\mathbb{Z})$ is called a **congruence subgroup** if $\Gamma(N) \subset \Gamma$ for some $N \in \mathbb{N}$.

We now have enough to define what it means for a holomorphic function to be invariant under the slash action of a congruence subgroup. In the definition of modular forms, however, we include an additional condition that is often referred to as *holomorphicity at $i\infty$*, the purpose of which is to ensure that spaces of modular forms, which turn out to admit $\mathbb{C}$-vector space structures, are, in fact, finite-dimensional [19].

The theory of modular forms is often thought to lie in the very rich intersection of algebra and analysis. Our definitions so far have been largely algebraic, but our next one is analytic. Consider the mapping $q : \mathbb{H} \to \mathbb{C} : z \mapsto e^{2\pi i z}$. This maps $\mathbb{H}$ to the punctured, open unit disc

$$D := \{w \in \mathbb{C} \mid 0 < |q| < 1\}$$

Indeed, for all $z \in \mathbb{H}$, writing $z = x + iy$ for $x, y \in \mathbb{R}$ with $y > 0$, we have

$$|q(z)| = \left|e^{2\pi i(x+iy)}\right| = \left|e^{2\pi i x}\right| \cdot \left|e^{-2\pi y}\right| < 1$$

with $0 \notin q(\mathbb{H})$ but $q(z) \to 0$ as $\mathrm{Im}(z) = y \to \infty$. Now, we know that the holomorphic functions from $D \to \mathbb{C}$ are precisely those that have Laurent expansions of the form

$$\sum_{n=0}^{\infty} c_n w^n$$

for all $w \in D$. If we write $w = q(z)$ for $z \in \mathbb{H}$, the above series turns out to be a *Fourier expansion*. We can hence make the following definition for holomorphicity at $i\infty$.

Definition 2.3.3 (Holomorphicity at $i\infty$). We say a function $f : \mathbb{H} \to \mathbb{C}$ is **holomorphic at** $i\infty$ if f admits a Fourier expansion of the form

$$f(z) = \sum_{n=0}^{\infty} c_n q(z)^n = \sum_{n=0}^{\infty} c_n e^{2\pi i n z}$$

That is, f admits a Fourier expansion with no negative powers of $q(z)$.

The holomorphicity of f at $i\infty$ essentially means that the Fourier expansion of f is a holomorphic $D \to \mathbb{C}$ function in $q(z)$, with the added constraint that $|f(z)|$ remains bounded as $\mathrm{Im}(z) \to \infty$, that is, the corresponding $D \to \mathbb{C}$ function in $q(z)$ extends to a holomorphic function that is defined and bounded at 0. There is a rich theory of functions where $c_0 = 0$, but we will not explore that theory here.[2]

We are now ready to define modular forms. Intuitively, a modular form is a function that satisfies the above definitions in a slash-invariant manner. More precisely, we have the following.

Definition 2.3.4 (Modular Form). Fix $k \in \mathbb{Z}$ and let Γ be a congruence subgroup of $\mathrm{SL}(2, \mathbb{Z})$. We say a function $f : \mathbb{H} \to \mathbb{C}$ is a **modular form of weight** k **with respect to** Γ if f is **invariant** under the slash action of Γ and **holomorphic at** $i\infty$ under the slash action of $\mathrm{SL}(2, \mathbb{Z})$. That is,

1. For all $\gamma \in \Gamma$, $f \mid_k \gamma = f$ (cf. Definition 2.3.1).
2. For all $\gamma \in \mathrm{SL}(2, \mathbb{Z})$, $f \mid_k \gamma$ is holomorphic at $i\infty$ (cf. Definition 2.3.3).

We denote by $M_k(\Gamma)$ the space of modular forms of weight k and congruence subgroup Γ. If $\Gamma = \Gamma(N)$ for some $N \in \mathbb{N}$, we say an element of $M_k(\Gamma)$ has **level** N.

There is an immensely rich theory of modular forms, and for the purposes of practicality, it was decided not to explore this theory in great detail in this project, particularly because the formalisation of the aspects of Viazovska's proof that stem from this theory is being led by Birkbeck, Lee and Ma. We will instead use the remainder of this section to discuss three specific (families of) modular forms and those of their properties that Viazovska uses to construct her magic function.

2.3.1 The Eisenstein Series

The Eisenstein Series are an important family of slash-invariant forms that will prove essential to the construction of the magic function. The Eisenstein Series whose *weight* is an even integer that is at least 4 are modular forms, though we will also need to work with the Eisenstein Series of weight 2, which, despite not being a modular form, is sufficiently well-behaved for our purposes. We will define it separately from those Eisenstein Series that are modular forms.

Let $k \geq 4$ be an even integer. We denote by E_k the weight k Eisenstein Series. There is more than one way to define E_k. In this report, we give the definition that was formalised by Birkbeck for this project. Birkbeck's definition in the project repository is a particular case of the `mathlib` definition, which defines it as a `ModularForm` structure combining the function `eisensteinSeries` with the proofs of the properties that make it a modular form. The `mathlib` definition is more general than the one we study here, and involves imposing congruence con-

[2] Modular forms with this property are known as **cusp forms**. One modular form we will need to construct the magic function is the discriminant form, which will turn out to be a cusp form.

ditions on the subsets of the lattice $\mathbb{Z}^2$ over which the Eisenstein Series are summed. It is not necessary for this project.

Definition 2.3.5 (The Eisenstein Series of Even Weight ≥ 4). For $k \geq 4$ even, define the **weight k Eisenstein Series** to be the function $E_k : \mathbb{H} \to \mathbb{C}$ given by

$$E_k(z) := \frac{1}{2} \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ \gcd(m,n)=1}} \frac{1}{(mz+n)^k} \tag{2.3.4}$$

with the defining summation converging absolutely.

Note that the Eisenstein Series can also be defined as

$$E_k(z) = \frac{1}{2\zeta(k)} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \frac{1}{(mz+n)^k} \tag{2.3.5}$$

with ζ here denoting the Riemann zeta function. It is shown in [18, Equation (4.1), pp. 109-110] that this definition matches the definition formalised by Birkbeck in the project repository and stated informally in Definition 2.3.5.

Blueprint gives zeta def whereas repo gives coprime def. WHOOPSIE!

It is shown in [18, pp. 4-5] that E_k is a weight k, level 1 modular form for even integers $k \geq 4$. That is, E_k is invariant under the weight k slash-action of every element of $\mathrm{SL}(2,\mathbb{Z})$. As important special cases of this, E_k satisfies two important functional equations.

Proposition 2.3.6. *For all even $k \geq 4$ and $z \in \mathbb{H}$, the following both hold:*

$$E_k(z+1) = E_k \tag{2.3.6}$$

$$E_k\left(-\frac{1}{z}\right) = z^k E_k(z) \tag{2.3.7}$$

Proof. Both of these are just slash-invariance properties in disguise. We have

Fix slash formatting

$$E_k(z+1) = \left(E_k \left|_k \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right.\right)(z) = (0z+1)^k E_k(z) = E_k(z)$$

Similarly, we have

$$E_k\left(-\frac{1}{z}\right) = \left(E_k \left|_k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right.\right)(z) = (1z+0)^k E_k(z) = z^k E_k(z)$$

as required. □

The functional equations (2.3.6) and (2.3.7) yield similar results for an important function that will be used in constructing the magic function. We will explore this idea in Chapter 4.

One of the most important properties of the Eisenstein Series—at least, for our purposes—is that their Fourier coefficients[3] grow polynomially. We will be particularly interested in E_4 and E_6, which are defined as above, and their cousin E_2, which we will treat separately. These

[3]The slash-invariant properties of modular forms mean that they have periodicity properties. Computing their Fourier series is hence a natural strategy when attempting to dissect their properties.

functions show up in the definition of Viazovska's magic function, and the polynomial growth property allows us to prove that the magic function is Schwartz.

Our strategy to prove that the Fourier coefficients have polynomial growth will be to compute them explicitly. First, we need to define the arithmetic function $\sigma_k(n)$, which is defined in `mathlib` as `ArithmeticFunction.sigma`.

Definition 2.3.7 (The σ-Function). The **σ-function** $\sigma : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is given by

$$\sigma_k(n) := \sum_{d|n} d^k$$

In `mathlib`, for every natural number `k`, `ArithmeticFunction.sigma k` is defined as an `ArithmeticFunction` $\mathbb{N}$ structure, meaning it is an $\mathbb{N} \to \mathbb{N}$ map that maps 0 to 0.

The reason we defined the σ-function is that the Fourier coefficients of the Eisenstein series are given in terms of σ.

Theorem 2.3.8. *For all even $k \geq 4$ and $z \in \mathbb{H}$, $E_k(z)$ can be expressed as the Fourier series*

$$E_k(z) = 1 + C_k \sum_{n=1}^{\infty} \sigma_{k-1}(n)\, e^{2\pi i n z} \tag{2.3.8}$$

where

$$C_k = \frac{1}{\zeta(k)} \cdot \frac{(-2\pi i)^k}{(k-1)!} \tag{2.3.9}$$

In particular, $C_4 = 240$ and $C_6 = -504$. That is, E_4 and E_6 have the following Fourier expansions:

$$E_4(z) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) e^{2\pi i n z} \tag{2.3.10}$$

$$E_6(z) = 1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) e^{2\pi i n z} \tag{2.3.11}$$

Here, again, the repo disagrees with the blueprint. FIX!

The statement and proof of the general Fourier expansion of E_k for even $k \geq 4$ have been formalised by Birkbeck in the Sphere Packing repository. Substituting $k = 4$ and $k = 6$ in the expression for C_k and evaluating it using software like Wolfram|Alpha gives the desired result.

Now, it is immediate that the Fourier coefficients exhibit polynomial growth: for all $k, n \in \mathbb{N}$, $\sigma_k(n)$ is a sum of at most n numbers that are each at most n^k, meaning $\sigma_k(n) \leq n^{k+1}$.

For the remainder of this subsection, we will focus on a cousin of the weight ≥ 4 Eisenstein Series: the weight 2 Eisenstein Series, denoted E_2. The reason why we treat E_2 separately is that it is not a modular form. Furthermore, it cannot be defined via the summation used in Equation (2.3.4) or Equation (2.3.5): unfortunately, when $k = 2$, these sums do not converge absolutely. That being said, Birkbeck has shown formally that for all $m \in \mathbb{Z}$, $z \in \mathbb{H}$, and $k \geq 2$,

the summation

$$\sum_{n\in\mathbb{Z}} \frac{1}{(mz+n)^k}$$

converges absolutely. He then shows, through several sorry-free lemmas, that

$$\lim_{N\to\infty} \sum_{m=-N}^{N-1} \sum_{n\in\mathbb{Z}} \frac{1}{(mz+n)^k}$$

exists, allowing us to define E_2 in the following manner.

Definition 2.3.9 (E_2). For all $z \in \mathbb{H}$, define

$$E_2(z) := \frac{1}{2\zeta(2)} \lim_{N\to\infty} \sum_{m=-N}^{N-1} \sum_{n\in\mathbb{Z}} \frac{1}{(mz+n)^k}$$

The difference between this definition and (2.3.5) with $k = 2$ is that here, we specify an order of summation for the outer sum, whereas for $k \geq 4$, in both (2.3.5) and (2.3.4), the order is immaterial due to absolute convergence. Interestingly, the Fourier expansion of E_2 agrees with (2.3.8).

Theorem 2.3.10. *For all $z \in \mathbb{H}$, $E_2(z)$ can be expressed as the Fourier series*

$$E_2(z) = 1 - 24\sum_{n=1}^{\infty} \sigma_1(n) e^{2\pi i n z} \tag{2.3.12}$$

Birkbeck gives a formal proof of this over the course of several sorry-free lemmas in SpherePacking.ModularFo
Interestingly, substituting $k = 2$ in (2.3.9) yields precisely -24. Moreover, the same argument we used earlier demonstrates that the Fourier coefficients of E_2 also grow polynomially. We will mention this result again in Chapter 4, where we will prove that the magic function is Schwartz.

We end our discussion on the Eisenstein Series by giving an explicit counterexample to weight 2, level 1 slash-invariance that shows that E_2 is not a weight 2, level 1 modular form.

Lemma 2.3.11. *For all $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{Z})$, we have*

$$E_2 \mid_2 \gamma = (cz+d)^{-2} E_2\left(\frac{az+b}{cz+d}\right) = E_2(z) - \frac{6ic}{\pi(cz+d)}$$

The proof uses results about the discriminant form, which we define in the next subsection. We do not prove the above result here, as it is significantly beyond the scope of this project, but we point the reader to the blueprint [13, Lemma 6.39].

Update blueprint reference before submitting

2.3.2 The Discriminant Form

The discriminant form is a weight 12, level 1 modular form. As was briefly alluded to earlier, it is a cusp form. It is defined in terms of the Eisenstein series E_4 and E_6.

Definition 2.3.12 (The Discriminant Form). The **discriminant form** Δ is defined by

$$\Delta := \frac{E_4^3 - E_6^2}{1728} \tag{2.3.13}$$

The discriminant form has important positivity and non-vanishing properties that we will use repeatedly, either directly or indirectly, in the construction of the magic function. The discriminant form will often show up in denominators, making these properties essential to prove properties like holomorphicity. The key to these properties is the so-called product formula.

Theorem 2.3.13 (Product Formula for Δ). *For all $z \in \mathbb{H}$, $\Delta(z)$ is expressible as the following infinite product:*

$$\Delta(z) = e^{2\pi i z} \prod_{n=1}^{\infty} \left(1 - e^{2\pi i n z}\right)^{24} \tag{2.3.14}$$

A proof can be found in [16, Chapter VII, §4, Theorem 6, p. 95]. Birkbeck has shown formally that the above product converges for all $z \in \mathbb{H}$.

As a remark, we mention that the theory of infinite products is not as well-developed in Lean as the theory of infinite sums. The definition of convergence of infinite products in `mathlib` is designed to yield a strong notion of convergence of infinite sums involving invariance under rearrangements, and is stronger than the notion of pointwise convergence. We do not discuss the details here, but note that the condition is sufficiently strong for our purposes.

We now state the positivity and nonvanishing properties of Δ that we will use when constructing the magic function.

Corollary 2.3.14. *The discriminant form has the following important properties.*

1. *For all $t > 0$, we have $\Delta(it) > 0$. That is, Δ is real and positive on the positive imaginary axis.*
2. *For all $z \in \mathbb{H}$, $\Delta(z) \neq 0$. That is, Δ is nonvanishing on the upper half-plane.*

2.3.3 The Theta Functions

In this subsection, we define and state some basic properties of the Theta functions Θ_2, Θ_3 and Θ_4, the fourth powers of which define the corresponding H-functions. The H-functions will be important ingredients in the construction of the magic function.

Definition 2.3.15 (The Θ- and H-Functions). Define $\Theta_2, \Theta_3, \Theta_4 : \mathbb{H} \to \mathbb{C}$ by

$$\Theta_2(z) = \sum_{n \in \mathbb{Z}} e^{\pi i \left(n + \frac{1}{2}\right)^2 z}$$

$$\Theta_3(z) = \sum_{n \in \mathbb{Z}} e^{\pi i n^2 z}$$

$$\Theta_4(z) = \sum_{n \in \mathbb{Z}} (-1)^n e^{\pi i n^2 z}$$

for all $z \in \mathbb{H}$. Define $H_2, H_3, H_4 : \mathbb{H} \to \mathbb{C}$ by

$$H_2 = \Theta_2^4 \qquad H_3 = \Theta_3^4 \qquad H_4 = \Theta_4^4$$

It can be shown that the H-functions are modular forms of weight 2 and level 2.

Given the manner in which the H-functions are defined, it is tedious to compute their Fourier expansions explicitly. However, the purpose of computing the Fourier expansions of the Eisenstein Series was to determine that their Fourier coefficients grow polynomially. It turns out that in the case of the H-functions, we can do this without explicitly computing their Fourier series.

The Fourier coefficients of H_3 and H_4 grow polynomially because those of Θ_3 and Θ_4 grow polynomially[4]: defining

$$c_3(m) = \begin{cases} 1 & \text{if } m = n^2 \text{ for some } n \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}$$
$$c_4(m) = \begin{cases} (-1)^n & \text{if } m = n^2 \text{ for some } n \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}$$

it is clear that $|c_3(m)|, |c_4(m)| \leq 1$ for all $m \in \mathbb{Z}$. The Fourier expansions of Θ_3 and Θ_4 are then given by

$$\Theta_3(z) = \sum_{m \in \mathbb{Z}} c_3(m)\, e^{i\pi m z}$$
$$\Theta_4(z) = \sum_{m \in \mathbb{Z}} c_4(m)\, e^{i\pi m z}$$

The fact that the Fourier coefficients of H_3 and H_4 also grow polynomially can then be deduced by expressing Θ_3^4 and Θ_4^4 as iterated sums. This is tedious, and we do not do it here.

Unfortunately, due to the fractional term in the exponents of the summands in the definition of Θ_2, it is not possible to use the same technique to show that its Fourier coefficients grow polynomially. Fortunately, we can still prove the result for H_2, because raising Θ_2 to the fourth power gets rid of the fractional exponent. That is,

$$H_2 = \Theta_2^4 = \left(\sum_{n \in \mathbb{Z}} e^{\pi i \left(n+\frac{1}{2}\right)^2 z}\right)^4 = \left(\sum_{n \in \mathbb{Z}} e^{\pi i \left(n^2+n+\frac{1}{4}\right) z}\right)^4$$
$$= \left(\sum_{n \in \mathbb{Z}} e^{\pi i \left(n^2+n\right) z}\, e^{\frac{\pi i z}{4}}\right)^4 = \left(e^{\frac{\pi i z}{4}}\right)^4 \left(\sum_{n \in \mathbb{Z}} e^{\pi i \left(n^2+n\right) z}\right)^4 = e^{\pi i z} \left(\sum_{n \in \mathbb{Z}} e^{\pi i \left(n^2+n\right) z}\right)^4$$

This can be explicitly computed as an iterated sum with coefficients that grow polynomially.

Finally, we mention some important relations that we will take advantage of when proving properties about the magic function. Some are given as slash actions of elements of $\mathrm{SL}(2, \mathbb{Z})$, so we define some notation first.

[4] Proposition 4.2.2 explicitly establishes this fact.

Notation. Denote

$$S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We now state important properties of the H-functions.

Proposition 2.3.16. *The following slash-action relations hold. Furthermore, the H-functions are related to the other modular forms we have explored in the following manner.*

1.

Finally, the H-functions satisfy the Jacobi identity:

$$H_2 + H_4 = H_3$$

We do not prove this proposition here, as it is beyond the scope of this thesis. Proofs of the individual results can be found in [13].

Chapter 3

A Roadmap to Constructing the Magic Function

We mentioned, in the introduction, that the scope of this project is to construct Viazovska's Magic Function in Lean and prove that it satisfies certain specific properties, such as satisfying the hypotheses of the Cohn-Elkies Linear Programming Bound. In this chapter, we will outline the steps we will take to achieve this goal. In particular, we will list all the conditions we need to prove that the Magic Function satisfies. Our approach will be to construct the Magic Function in terms of two intermediary functions. Proving it satisfies the necessary conditions will then be a matter of proving that these intermediary functions satisfy certain properties. We will list these properties as well.

3.1 Radial Schwartz Functions

In the statement of Theorem 2.2.5, we require the function in terms of which we bound the sphere packing constant in dimension d to be Schwartz. However, we have yet to formally define what this means. Intuitively, a Schwartz function is a smooth function whose every derivative decays faster than any inverse power of the norm. Below, we give a more formal definition that is adapted from the definition of the `structure SchwartzMap` in `mathlib`.

cite

Definition 3.1.1 (Schwartz Function). Let E and F be normed $\mathbb{R}$-vector spaces. We say that $f : E \to F$ is **Schwartz** if it is infinitely continuously differentiable and for all $n, k \in \mathbb{N}$, there exists some $C \in \mathbb{R}$ such that for all $x \in E$,

$$\|x\|^k \cdot \left\| f^{(n)}(x) \right\| \leq C \tag{3.1.1}$$

One can show that $\mathbb{R}$-linear combinations of Schwartz functions are Schwartz functions. Then, given any E and F, we can define the Schwartz space $\mathcal{S}(E, F)$.

Definition 3.1.2 (Schwartz Space). Let E and F be normed $\mathbb{R}$-vector spaces. We define the **Schwartz space** $\mathcal{S}(E, F)$ to be the set of all Schwartz functions from E to F, viewed as a vector space over $\mathbb{R}$.

cross reference

At the outset, it might appear that the reason we are interested in Schwartz functions is that this is a requirement of the Poisson Summation Formula, which is used in the proof of the Cohn-Elkies Linear Programming Bound (Theorem 2.2.5). However, this turns out to be a sufficient condition for the Poisson Summation Formula to hold, not a necessary condition. There is a deeper reason why we are interested in Schwartz functions: the Cohn-Elkies Conditions immediately show us that we should also consider the properties of the Fourier transform of the magic function, and Fourier transforms of Schwartz functions turn out to be Schwartz. In fact, we can say something stronger when we view the Fourier transform as an operator on the Schwartz space.

Theorem 3.1.3. *Let V be a finite-dimensional inner-product space over $\mathbb{R}$ and let E be a normed vector space over $\mathbb{C}$. The Fourier transform*

$$\mathcal{F} : \mathcal{S}(V, E) \to \mathcal{S}(V, E) : f \mapsto \widehat{f}$$

is a linear isomorphism of $\mathcal{S}(V, E)$.

A formal proof of this result can be found in `mathlib`.

It turns out that there is another condition we can impose to simplify our hunt for the magic function. The key idea is to find a function satisfying the conditions (CE1)-(CE3). Observe, for $x \in \mathbb{R}^d$, that (CE1) does not depend on x, and (CE2) and (CE3) only depend on $\|x\|$. This allows us to narrow our search to **radial functions**, which we define as follows.

Definition 3.1.4 (Radial Functions). Let E be a normed $\mathbb{R}$-vector space and α an arbitrary set. We say that $f : E \to \alpha$ is **radial** if for all $x, y \in E$, if $\|x\| = \|y\|$, then $f(x) = f(y)$.

Radial Schwartz functions interact with the Fourier Transform in an even nicer way than ordinary Schwartz functions.

Proposition 3.1.5. *Let $f : \mathbb{R}^d \to \mathbb{C}$ be a radial Schwartz function. Then,*

$$\mathcal{F}(\mathcal{F}(f)) = f$$

Proof. Fix $x \in \mathbb{R}^d$. Applying Lemma 2.2.3 to $\mathcal{F}(f)$ and $-x$, we have

$$\mathcal{F}(\mathcal{F}(f))(x) = \mathcal{F}^{-1}(\mathcal{F}(f))(-x) = f(-x) = f(x)$$

where the fact that $\mathcal{F}^{-1}(\mathcal{F}(f)) = f$ follows from the fact that f is Schwartz and the fact that $f(-x) = f(x)$ follows from the fact that f is radial. □

Proposition 3.1.5 indirectly provides us with a mechanism for constructing the magic function that is beautifully outlined in [1]. Observe that if f is a radial Schwartz function, we can write

$$f = \underbrace{\frac{f - \widehat{f}}{2}}_{=:f_-} + \underbrace{\frac{f + \widehat{f}}{2}}_{=:f_+}$$

The functions f_- and f_+ have the properties that

$$\mathcal{F}(f_-) = \frac{1}{2}\left(\mathcal{F}(f) - \mathcal{F}\left(\widehat{f}\right)\right) = \frac{1}{2}\left(\widehat{f} - f\right) = -f_-$$
$$\mathcal{F}(f_+) = \frac{1}{2}\left(\mathcal{F}(f) + \mathcal{F}\left(\widehat{f}\right)\right) = \frac{1}{2}\left(\widehat{f} + f\right) = f_+$$

where we use Proposition 3.1.5 to show that $\widehat{\widehat{f}} = f$. In other words, f_- and f_+ are **eigenfunctions of the Fourier transform** with eigenvalues -1 and $+1$ respectively. In fact, if

$$f = \lambda f_1 + \mu f_2$$

for *any* two functions f_1 and f_2 such that $\widehat{f_1} = -f_1$ and $\widehat{f_2} = f_2$, then one can show, by computing f_- and f_+, that $\lambda f_1 = f_-$ and $\mu b = f_+$. This allows us to break down the problem of computing the magic function into two smaller problems, namely, computing its constituent Fourier ± 1-eigenfunctions.

Before discussing the other properties we seek in our magic function—or its constituent Fourier eigenfunctions—we develop some very useful machinery that allows us to show that radial functions are Schwartz. We usually treat radial functions as $\mathbb{R} \to \mathbb{C}$ functions, because all information about the input that is necessary to compute the corresponding output is captured by a (non-negative) real number: its norm. The decaying property (3.1.1) of Schwartz functions is something that, at first glance, makes it a bit tricky to ignore the dimension of the domain when dealing with radial Schwartz functions. Ideally, we would want to only show that the decaying property holds for the corresponding $\mathbb{R} \to \mathbb{C}$ function. We can show this for a class of radial Schwartz functions that only depend on $\|x\|^2$.

Proposition 3.1.6. *Let $f : \mathbb{R} \to \mathbb{C}$ be a Schwartz function. For all $d \in \mathbb{N}$, the function*

$$f_d : \mathbb{R}^d \to \mathbb{C} : x \mapsto f\left(\|x\|^2\right)$$

is Schwartz.

We do not prove this result here, but we do mention that there is a `sorry`-free Lean proof of it in our repository.

cross reference

The point of Proposition 3.1.6 is that it gives us a criterion to show that radial functions in higher dimensions that are functions not of the norm but of the norm *squared* are Schwartz, purely by considering the corresponding function that takes in a one-dimensional input. This will be instrumental in our argument.

With this, we end our discussion of radial Schwartz functions. The key takeaway is that while Schwartzness is a necessary condition for our magic function to satisfy, we can also impose the condition of radiality to simplify our construction. We will now take a closer look at Cohn and Elkies's groundbreaking result (Theorem 2.2.5) to determine further properties for the magic function to satisfy.

3.2 A Closer Examination of the Cohn-Elkies Linear Programming Bound

So far, we have examined the statement of Theorem 2.2.5 in detail: it immediately tells us that we want the magic function to be Schwartz and satisfy the conditions (CE1)-(CE3), and upon

noticing that these conditions only depend on the norm and that radial functions are very well-behaved, we have narrowed our search to radial Schwartz function obeying (CE1)-(CE3). It turns out that we can learn even more about the magic function when we examine the *proof* of Theorem 2.2.5 when we specialise to the case where the function f is optimal. Our examination of the proof of Theorem 2.2.5 is based on an insightful discussion in [1, p. 8].

Specifically, let f be a (radial) Schwartz function satisfying (CE1), (CE2) and (CE3). What it means for f to be optimal is that there exists a sphere packing $\mathcal{P}(X)$ in $\mathbb{R}^d$ such that the Cohn-Elkies bound indexed by f is precisely the density of this sphere packing. This would make $\mathcal{P}(X)$ an optimal sphere packing in $\mathbb{R}^d$ and f an optimal function.

Since it is enough to prove the upper-bound property for periodic sphere packings, we can simplify our search for the right f by assuming the Cohn-Elkies bound corresponding to f is the density of a *periodic* packing. In other words, we can assume there exists some lattice $\Lambda \subset \mathbb{R}^d$ such that the set of centres X is periodic with respect to Λ. This turns out to be helpful because we can then use the exact forms of the inequalities in the proof to deduce properties that f must have if it is optimal, corresponding to some optimal periodic packing.

In our argument, we fix an arbitrary Λ-periodic sphere packing $\mathcal{P}(X)$ of separation 1 and show the inequality (2.2.1). In the case where f is optimal, in the sense that the upper-bound is achieved, we must have that (2.2.1) is, in fact, an **equality**. The same must be true of the equivalent inequality, (2.2.2). This tells us that the intermediate inequalities (2.2.3) and (2.2.4) must *also* be equalities, because the chain of inequalities begins and ends at the same quantity. In particular, we can take a closer look at (2.2.3): the way we prove it is by writing

$$\left|{}^{X}\!/_{\Lambda}\right| \cdot f(0) = \sum_{x \in {}^{X}\!/_{\Lambda}} f(x-x) = \sum_{x \in X} \sum_{\substack{y \in {}^{X}\!/_{\Lambda} \\ y=x}} \geq \sum_{x \in X} \sum_{y \in {}^{X}\!/_{\Lambda}} f(x-y)$$

The terms we discard to prove the inequality are non-positive, as they are of the form $f(x-y)$ for $y \neq x$ (meaning $\|y-x\| \geq 1$, allowing us to apply (CE2)). If this inequality is an equality, then all the terms we discard must not merely be non-positive: they must, in fact, be zero. That is, we need

$$f(x-y) = 0 \text{ for all } \textbf{distinct } x \in X \text{ and } y \in {}^{X}\!/_{\Lambda} \tag{3.2.1}$$

By definition of ${}^{X}\!/_{\Lambda}$, *every* element of Λ is expressible as $x-y$ for some $x \in X$ and $y \in {}^{X}\!/_{\Lambda}$, because X consists of *all* Λ-translates of y. So, all non-zero lattice points are expressible as $x-y$ for x and y as in (3.2.1). We can therefore conclude that **an optimal function f with Cohn-Elkies bound equal to the density of a periodic sphere packing must vanish at all non-zero lattice points**.

It turns out that examining (CE2) gives us an *even stronger* condition on f. First, note that we must have $0 \leq f(0)$: the bound

$$\frac{f(0)}{\widehat{f}(0)} \cdot \mathrm{Vol}\left(B_d\left(0, \frac{1}{2}\right)\right)$$

is greater than or equal to a non-negative constant, and both $\mathrm{Vol}\left(B_d\left(0, \frac{1}{2}\right)\right)$ (as a volume) and $\widehat{f}(0)$ (by (CE3)) are non-negative, meaning $f(0)$ cannot possibly be negative. Indeed, this is true regardless of whether f is optimal. Since (CE2) tells us that f is non-positive at points with norm at least 1, we can conclude that f not only has zeroes but **double zeroes** at all lattice points with norm at least 1: the behaviour of f, viewed as an $\mathbb{R} \to \mathbb{R}$ function of the norm r of a point on $\mathbb{R}^d$, is such that sign-changes, if any, from non-negative to non-positive cannot occur at zeroes ≥ 1, and thereafter, there are no more sign changes.

Putting these conclusions about single and double zeroes together with our observation about splitting radial Schwartz functions into their constituent ± 1-Fourier eigenfunctions, we can conclude that we need to find **Fourier eigenfunctions with double zeroes at lattice points**. It is no accident that this is precisely the title of [12, Section 4].

3.3 The Properties Desired of Viazovska's Fourier Eigenfunctions

We begin by summarising the properties we would like the magic function to have. We then examine which of these properties come from the eigenfunctions. Finally, we will mention tools that are used to show that both its ± 1-Fourier eigenfunctions satisfy the conditions we list below.

For the remainder of this thesis, we will fix the following notation.

Notation. Going forward, the magic function for 8-dimensional sphere packing shall be denoted g, its $+1$-eigenfunction shall be denoted a, and its -1-eigenfunction shall be denoted b.

We now list the properties we would like g to have.

1. g needs to be a Schwartz function.
2. It suffices for g to be radial.
3. g needs to satisfy the Cohn-Elkies conditions (CE1), (CE2) and (CE3).
4. g needs to have single zeroes at all non-zero points in Λ_8.
5. g needs to have double zeroes at all but finitely many points in Λ_8.
6. The Cohn-Elkies Linear Programming Bound indexed by g must be equal to the density of the E_8 sphere packing. That is, we need

$$\frac{g(0)}{\widehat{g}(0)} \cdot \mathrm{Vol}\left(B_8\left(0, \frac{1}{2}\right)\right) = \frac{\pi^4}{384}.$$

Of these properties, the following would be inherited from a and b:

1. Schwartzness
2. Radiality
3. Having single zeroes at all non-zero points in Λ_8
4. Having double zeroes at all but finitely many points in Λ_8

That is, if we can construct a and b such that they satisfy the above properties, then g will satisfy them as well. The remaining properties will have to do with the coefficients of the linear combination of a and b that makes up g.

Our use of the theory of modular forms will primarily be to define the eigenfunctions. We do not go into much detail about the theory of modular forms in this exposition: the formalisation

of these details is being handled by other collaborators, and the informal and formal aspects of it are beyond the scope of this M4R. We will also avoid discussing motivational elements that come from the theory of modular forms, and instead refer the reader to a beautiful reverse-engineered exposition of Viazovska's construction by Henry Cohn published in the proceedings of the 2022 ICM [1].

We structure the remainder of this thesis as follows. We will begin by informally discussing the construction of Viazovska's magic function, first discussing a general bounding technique that is necessary to prove that both eigenfunctions are Schwartz, then proving Schwartzness and double zeroes, before discussing the construction of the magic function as a linear combination of the two eigenfunctions. We will primarily be following the contents of the blueprint [13], though the structure of our exposition will differ slightly: instead of discussing both eigenfunctions separately, we will prove their properties in parallel, to highlight the similarities in the two constructions that allow us to reuse formalised code.

Chapter 4

Viazovska's Magic Function, Informally

In this chapter, we will construct the $+1$-eigenfunction a, the -1-eigenfunction b, and the magic function g. The theory developed in Chapter 3 tells us what properties we would like all three functions to satisfy, and in Section 3.3, we summarised those properties concisely. Over the course of this chapter, we prove that they do, indeed, satisfy them. The content is based heavily on Viazovska's original paper [12] and a more detailed version of her proof that she wrote for the project blueprint [13]. Note that the current version of the blueprint may contain modifications that have been included to make it correspond more closely with the current state of the formalisation effort.

We begin by defining the functions in question. In each subsequent section of this chapter, we will prove a certain property for each of the eigenfunctions. Finally, in Section 4.5, we will prove that g does, indeed, satisfy the properties outlined in Chapter 3.

4.1 Defining Viazovska's Fourier Eigenfunctions

The $\pm$-eigenfunctions of g—and, by extension, g itself—are defined in terms of modular and quasimodular forms (recall the definitions of these terms from Section 2.3). Specifically, the $+1$-eigenfunction a is defined in terms of the so-called ϕ- and ψ-functions, which are in turn defined in terms of the Eisenstein series (cf. `sorry`) and the discriminant form (cf. `sorry`), while the -1-eigenfunction b is defined in terms of the Jacobi Theta functions (cf. `sorry`).

Define in Chapter 2 and cross-ref here

We begin by defining the $+1$-eigenfunction a.

4.1.1 The $+1$-Eigenfunction

We begin by defining the ϕ-functions.

Definition 4.1.1 (The ϕ-Functions). Define the functions $\phi_0, \phi_{-2}, \phi_{-4} : \mathbb{H} \to \mathbb{C}$ by

$$\phi_{-4} := \frac{E_4^2}{\Delta} \tag{4.1.1}$$

$$\phi_{-2} := \frac{E_4 \left(E_2 E_4 - E_6\right)}{\Delta} \tag{4.1.2}$$

$$\phi_0 := \frac{\left(E_2 E_4 - E_6\right)^2}{\Delta} \tag{4.1.3}$$

These functions admit important transformation properties that are necessary to prove that the $+1$-eigenfunction is made up of integrals of holomorphic functions. This fact will in turn allow us to apply the Cauchy-Goursat Theorem (and variants thereof) that will allow us to shift contours of integration.

Lemma 4.1.2. *For all $z \in \mathbb{H}$,*

$$\phi_0(z+1) = \phi_0(z) \tag{4.1.4}$$

$$\phi_0\left(\frac{-1}{z}\right) = \phi_0(z) - \frac{12i}{\pi} \cdot \frac{1}{z} \cdot \phi_{-2}(z) - \frac{36}{\pi^2} \cdot \frac{1}{z^2} \cdot \phi_{-4}(z) \tag{4.1.5}$$

We do not prove these here, but mention that they both follow from the weight k slash action formulae on E_k for $k \in \{2, 4, 6\}$. When $k = 4$ and $k = 6$, we have weight k invariance, because E_4 and E_6 are modular forms, but when $k = 2$, we need to use Lemma 2.3.11. A detailed proof of these transformations can be found in [13].

We now define the $+1$-eigenfunction a.

Definition 4.1.3 (Viazovska's $+1$-Fourier Eigenfunction). Define $a_{\text{rad}} : \mathbb{R} \to \mathbb{C}$ by

$$a_{\text{rad}}(r) := I_1(r) + I_2(r) + I_3(r) + I_4(r) + I_5(r) + I_6(r) \tag{4.1.6}$$

where, for all $r \in \mathbb{R}$,

$$I_1(r) := \int_{-1}^{-1+i} \phi_0\left(\frac{-1}{z+1}\right) (z+1)^2 \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.7}$$

$$I_2(r) := \int_{-1+i}^{i} \phi_0\left(\frac{-1}{z+1}\right) (z+1)^2 \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.8}$$

$$I_3(r) := \int_{1}^{1+i} \phi_0\left(\frac{-1}{z-1}\right) (z-1)^2 \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.9}$$

$$I_4(r) := \int_{1+i}^{i} \phi_0\left(\frac{-1}{z-1}\right) (z-1)^2 \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.10}$$

$$I_5(r) := -2 \int_{0}^{i} \phi_0\left(\frac{-1}{z}\right) z^2 \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.11}$$

$$I_6(r) := 2 \int_{i}^{i\infty} \phi_0(z) \, e^{\pi i r z} \, \mathrm{d}z \tag{4.1.12}$$

Define the $+1$-Fourier eigenfunction $a : \mathbb{R}^8 \to \mathbb{C}$ by

$$a(x) := a_{\text{rad}}\left(\|x\|^2\right) \tag{4.1.13}$$

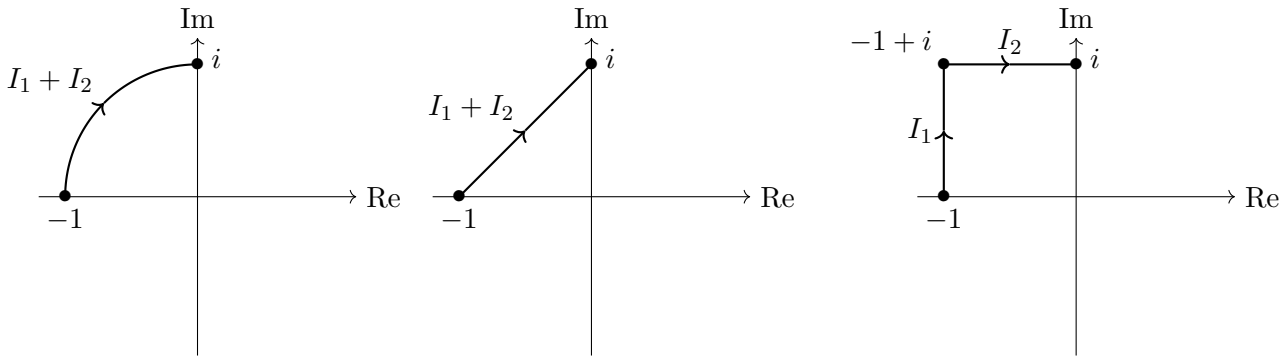
It is immediate from (4.1.13) that a is radial. All of its properties are determined by its radial part a_{rad}. There are similar definitions in Lean.

There is an important remark that must be made about the definitions in (4.1.7)-(4.1.12): in the original paper [12], the integrals I_1 and I_2 are combined, as are I_3 and I_4, and expressed in

the following manner:

$$I_1(r) + I_2(r) = \int_{-1}^{i} \phi_0\left(\frac{-1}{z+1}\right) (z+1)^2 \, e^{\pi i r z} \, \mathrm{d}z$$
$$I_3(r) + I_4(r) = \int_{1}^{i} \phi_0\left(\frac{-1}{z-1}\right) (z-1)^2 \, e^{\pi i r z} \, \mathrm{d}z$$

with the contours not specified. The most 'classical' choice would be quarter-circular contours, though the same results can be achieved working with straight and rectangular contours.



(a) Quarter-Circular Contour **(b)** Straight Line Contour **(c)** Rectangular Contour

Figure 4.1: Different contours along which we can integrate the integrand of I_1 and I_2 to get an integral equal to $I_1 + I_2$

The reason the choice of contours does not matter is that in the integrands of $I_1, \ldots, I_5$, we multiply terms of the form $\phi_0\left(\frac{-1}{z}\right)$ by z^2. If we apply (4.1.5) and multiply through, it is clear that we are removing any singularities introduced by $\frac{1}{z^2}$ and $\frac{1}{z}$ factors. We can then use the fact that $\Delta(z) \neq 0$ for all $z \in \mathbb{H}$ to conclude that the integrands are holomorphic up to these removable singularities.

The choice of rectangular contours (as in Section 4.1.1) as opposed to quarter-circles or straight lines for $I_1 + I_2$ and $I_3 + I_4$ is motivated by the versions of the Cauchy-Goursat Theorem that have been formalised in Lean. See Section 5.3 for more.

We are now ready to define the -1-eigenfunction b.

4.1.2 The -1-Eigenfunction

Recall the H-functions defined as the fourth powers of the Theta functions in Definition 2.3.15. We begin by defining the h function, in terms of which we define the ψ-functions.

Definition 4.1.4 (The h-Function). Define the function $h : \mathbb{H} \to \mathbb{C}$ by

$$h(z) := 128 \frac{H_3(z) + H_4(z)}{H_2(z)^2} \tag{4.1.14}$$

where H_2, H_3 and H_4 are as defined in Definition 2.3.15.

In [12], the ψ-functions are defined in terms of the h-function via slash actions.

Definition 4.1.5 (The ψ-Functions). Define the functions $\psi_I, \psi_S, \psi_T : \mathbb{H} \to \mathbb{C}$ by where h is as defined in Definition 4.1.4.

By merely unfolding definitions, it is possible to show that the ψ-functions can be expressed directly in terms of the H-functions.

Lemma 4.1.6.

It will be useful, particularly to prove Schwartzness, to express the ψ-functions in an alternate form, as fractions with the discriminant in the denominator.

Proposition 4.1.7 (The ψ-Fucntions). *We can express ψ_I, ψ_S, ψ_T in the following manner:*

$$\psi_I = \frac{H_4^3\left(2H_4^2 + 5H_4H_2 + 5H_2^2\right)}{2\Delta} \tag{4.1.15}$$

$$\psi_S = \frac{-H_2^3\left(2H_2^3 + 5H_2H_4 + 5H_4^2\right)}{2\Delta} \tag{4.1.16}$$

$$\psi_T = \psi_I - \psi_S \tag{4.1.17}$$

where Δ is the discriminant form.

This proposition can be proved by clearing denominators in

The ψ-functions satisfy the following relations.

Lemma 4.1.8. *For all $z \in \mathbb{H}$,*

$$\psi_T(z) = (z+1)^2\,\psi_S\left(\frac{-1}{z+1}\right)$$

$$\psi_I(z) = z^2\psi_S\left(\frac{-1}{z}\right)$$

These relations can be deduced directly from Definition 4.1.5.

We are now ready to define the -1-eigenfunction, denoted b.

Definition 4.1.9 (Viazovska's -1-Fourier Eigenfunction). Define $b_{\text{rad}} : \mathbb{R} \to \mathbb{C}$ by

$$b_{\text{rad}}(r) := J_1(r) + J_2(r) + J_3(r) + J_4(r) + J_5(r) + J_6(r) \tag{4.1.18}$$

where, for all $r \in \mathbb{R}$,

$$J_1(r) := \int_{-1}^{-1+i} \psi_T(z)\, e^{\pi i r z}\, \mathrm{d}z \tag{4.1.19}$$

$$J_2(r) := \int_{-1+i}^{i} \psi_T(z)\, e^{\pi i r z}\, \mathrm{d}z \tag{4.1.20}$$

$$J_3(r) := \int_{1}^{1+i} \psi_T(z)\, e^{\pi i r z}\, \mathrm{d}z \tag{4.1.21}$$

$$J_4(r) := \int_{1+i}^{i} \psi_T(z)\ e^{\pi i r z}\ \mathrm{d}z \tag{4.1.22}$$

$$J_5(r) := -2\int_{0}^{i} \psi_I(z)\ e^{\pi i r z}\ \mathrm{d}z \tag{4.1.23}$$

$$J_6(r) := 2\int_{i}^{i\infty} \psi_S(z)\ e^{\pi i r z}\ \mathrm{d}z \tag{4.1.24}$$

Define the -1-Fourier eigenfunction $a : \mathbb{R}^8 \to \mathbb{C}$ by

$$b(x) := b_{\mathrm{rad}}\left(\|x\|^2\right) \tag{4.1.25}$$

Note that by applying the relations in Lemma 4.1.8, we can express $J_1, \ldots, J_5$ in the following manner:

$$\begin{aligned}
J_1(r) &= \int_{-1}^{-1+i} \psi_S\left(\frac{-1}{z+1}\right)(z+1)^2\ e^{\pi i n z}\ \mathrm{d}z \\
J_2(r) &= \int_{-1+i}^{i} \psi_S\left(\frac{-1}{z+1}\right)(z+1)^2\ e^{\pi i n z}\ \mathrm{d}z \\
J_3(r) &= \int_{1}^{1+i} \psi_S\left(\frac{-1}{z+1}\right)(z+1)^2\ e^{\pi i n z}\ \mathrm{d}z \\
J_4(r) &= \int_{1+i}^{i} \psi_S\left(\frac{-1}{z+1}\right)(z+1)^2\ e^{\pi i n z}\ \mathrm{d}z \\
J_5(r) &= -2\int_{0}^{i} \psi_I(z)\ z^2\, e^{\pi i n z}\ \mathrm{d}z
\end{aligned}$$

With $J_1, \ldots, J_5$ expressed in this manner, and J_6 expressed as in (4.1.24), there is a marked visual similarity between the J_j and the I_j. A consequence of this is that similar strategies can be used to prove properties about both a and b.

4.2 Establishing the Schwartzness Property

The magic function is a linear combination of a and b, which are each defined as compositions of a_{rad} and b_{rad} with the norm-squared function. From Proposition 3.1.6, we know that it is enough to establish that a_{rad} and b_{rad} are Schwartz to establish that a and b are Schwartz. In particular, this means the smoothness and decaying conditions need to be satisfied with respect to $\mathbb{R}$ inputs instead of $\mathbb{R}^8$ inputs, a substantial simplification. We can further simplify the problem by taking advantage of linearity.

We know that the Schwartz space is a $\mathbb{C}$-vector space, making it closed under addition. To show that a_{rad} and b_{rad} are Schwartz functions, we show that their constituent integrals $I_1, \ldots, I_6$ and $J_1, \ldots, J_6$ are Schwartz. We need to show both smoothness and rapid decay. Smoothness is fairly straightforward. Rapid decay, on the other hand, requires an additional ingredient.

It turns out that we can establish a an upper-bound for all functions of the form $\frac{f}{\Delta}$, where Δ is the discriminant form and f admits a Fourier expansion whose coefficients grow polynomially.

Theorem 4.2.1 ([13, Lemma 7.4]). *Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic. Denote by $c_f(n)$ its*

nth Fourier coefficient, with $c_f(n_0) \neq 0$, so that

$$f(z) = \sum_{n=n_0}^{\infty} c_f(n)\, e^{i\pi n z}$$

If there exists $k \in \mathbb{N}$ such that $c_f(n) = \mathrm{O}\big(n^k\big)$ as $n \to \infty$, then there exists a constant $C_f > 0$ such that for all $z \in \mathbb{H}$ with $\mathrm{Im}(z) > 1/2$,

$$\left|\frac{f(z)}{\Delta(z)}\right| \leq C_f\, e^{-\pi(n_0-2)\,\mathrm{Im}(z)}$$

Proof. Fix $z \in \mathbb{H}$ and assume $\mathrm{Im}(z) > 1/2$. Recall from Theorem 2.3.13 that Δ can be expressed as a (convergent) infinite product. We can hence write

$$\left|\frac{f(z)}{\Delta(z)}\right| = \left|\frac{\sum_{n=n_0}^{\infty} c_f(n)\, e^{\pi i n z}}{e^{2\pi i z}\prod_{n=1}^{\infty}(1-e^{2\pi i n z})^{24}}\right| = \left|e^{\pi i(n_0-2)z}\right| \cdot \frac{\left|\sum_{n=n_0}^{\infty} c_f(n)\, e^{\pi i(n-n_0)z}\right|}{\prod_{n=1}^{\infty}|1-e^{2\pi i n z}|^{24}}$$

Noting that $\left|e^{iz}\right| = e^{-\mathrm{Im}(z)}$ and $\mathrm{Im}(z) > \frac{1}{2}$, we can see that

$$\left|e^{\pi i(n_0-2)z}\right| \cdot \frac{\left|\sum_{n=n_0}^{\infty} c_f(n)\, e^{\pi i(n-n_0)z}\right|}{\left|\prod_{n=1}^{\infty} 1-e^{2\pi i n z^{24}}\right|} \leq e^{-\pi(n-n_0)\,\mathrm{Im}(z)} \cdot \frac{\sum_{n=0}^{\infty}|c_f(n)|\; e^{-\pi(n-n_0)/2}}{\left|\prod_{n=1}^{\infty} 1-e^{2\pi i n z^{24}}\right|}$$

It has been verified formally that the absolute value of a convergent infinite product is the product of the absolute values, and moreover, that the product of the absolute values is convergent. It has also been verified formally that the infinite product is monotonic on convergent infinite products whose terms are nonnegative. Hence,

$$\left|\prod_{n=1}^{\infty}\left(1-e^{2\pi i n z}\right)^{24}\right| = \prod_{n=1}^{\infty}|1-e^{2\pi i n z}|^{24} \geq \prod_{n=1}^{\infty}\left(1-e^{-2\pi n\,\mathrm{Im}(z)}\right)^{24} \geq \prod_{n=1}^{\infty}\left(1-e^{-\pi n}\right)^{24}$$

We note that the third and fourth products are convergent because they are expressible, via the product formula, as $e^{2\pi\,\mathrm{Im}(z)}\Delta(i\cdot\mathrm{Im}(z))$ and $e^{\pi}\Delta(i/2)$ respectively. Hence, defining

$$C_f := \frac{\sum_{n=0}^{\infty}|c_f(n)|\; e^{-\pi(n-n_0)/2}}{\prod_{n=1}^{\infty}(1-e^{-\pi n})^{24}}$$

we can see that $\left|\frac{f(z)}{\Delta(z)}\right| \leq C_f\, e^{-\pi(n_0-2)\,\mathrm{Im}(z)}$, as desired. □

The purpose of the above is to bound the ϕ- and ψ-functions using Theorem 4.2.1. Since these functions are defined as sums and products of the Eisenstein series and the H-functions, whose Fourier series have the properties that

1. the coefficients grow polynomially
2. there is an index n_0 below which all Fourier coefficients are zero

it is enough to show that sums and products of functions exhibiting this property inherit it. Note that this also explicitly establishes a fact we mentioned in passing in Section 2.3.3, where we argued that the Fourier coefficients of H_3 and H_4 grow polynomially because those of Θ_3 and Θ_4 do.

Proposition 4.2.2. *Let $f_1, f_2 : \mathbb{H} \to \mathbb{C}$ be functions with (absolutely convergent) Fourier expansions*

$$f_1(z) = \sum_{n=n_1}^{\infty} c_1(n)\, e^{\pi i n z}$$
$$f_2(z) = \sum_{n=n_2}^{\infty} c_2(n)\, e^{\pi i n z}$$

such that for $i \in \{1,2\}$, $c_i(n_i) \neq 0$ and $\exists k_i \in \mathbb{N}$ such that $c_i(n) = \mathrm{O}\left(n^{k_i}\right)$ as $n \to \infty$. Then, their product $f_1 f_2$ is expressible as an absolutely convergent Fourier series

$$f_1(z) f_2(z) = \sum_{n=n_1+n_2}^{\infty} c(n)\, e^{\pi i n z}$$

such that $c(n_1+n_2) \neq 0$ and $\exists k \in \mathbb{N}$ such that $c(n) = \mathrm{O}\left(n^k\right)$ as $n \to \infty$.

Proof. Fix $z \in \mathbb{H}$. Then, due to absolute convergence, we can write

$$\begin{aligned} f_1(z) f_2(z) &= \left(\sum_{n=n_1}^{\infty} c_1(n)\, e^{\pi i n z}\right)\left(\sum_{m=n_2}^{\infty} c_2(m)\, e^{\pi i m z}\right) \\ &= \sum_{n=n_1}^{\infty}\sum_{m=n_2}^{\infty} c_1(n) c_2(m)\, e^{\pi i (n+m) z} \end{aligned}$$

The smallest value of $m+n$ is clearly n_1+n_2. Denoting this by n_0, we can write

$$f_1(z) f_2(z) = \sum_{\ell=n_0}^{\infty} c(n)\, e^{\pi i \ell z}$$

where for each $\ell \geq n_0$, $c(\ell)$ is a sum of finitely many terms of the form $c_1(n)c_2(m)$, with $n+m=\ell$. Now, we know there exist positive numbers $C, D \in \mathbb{R}$ and $N, M \in \mathbb{N}$ such that for all $n \geq N$ and $m \geq M$, $|c_1(n)| \leq C\left|n^{k_1}\right| = Cn^{k_1}$ and $|c_2(m)| \leq D\left|m^{k_2}\right| = Dm^{k_2}$. Defining

$$C' := \sum_{n=n_1}^{N} |c_1(n)| \qquad D' := \sum_{m=n_2}^{M} |c_2(m)|$$

we can bound $|c(\ell)|$ above by a sum of finitely many terms of the form $C'\,|c_2(m)|$, $D'\,|c_1(n)|$, and $|c_1(n)| \cdot |c_2(m)|$, with $n \geq N$ and $m \geq M$. Take $k = k_1 + k_2$. Then, $|c_1(n)| \leq Cn^{k_1} \leq C\,(m+n)^{k_1}$ and $|c_2(n)| \leq Dm^{k_1} \leq D\,(m+n)^{k_2}$, so $c(\ell) = \mathrm{O}\left((m+n)^{k_1+k_2}\right) = \mathrm{O}\left(\ell^k\right)$. □

The analogous result for sums is clear, with $n_0 \geq \min(n_1, n_2)$ and $k = \max(k_1, k_2)$. Note that for sums, n_0 may not be exactly $\min(n_1, n_2)$ because the Fourier coefficients of smallest index may cancel each other out. For the remainder of this thesis, we use the following notation.

Notation. For a function f with a Fourier expansion, denote by
- $n_0(f)$ the smallest index n such that $c_f(n) \neq 0$ (if it exists)
- $c_f(n)$ the nth Fourier coefficient of f

We will not use this notation for functions for which $n_0(f)$ does not exist.

In the following subsections, we apply the above results and show that a_{rad} and b_{rad} are Schwartz functions. In each case, since the bound in Theorem 4.2.1 is given in terms of n_0, we compute the values of n_0 explicitly.

4.2.1 The $+1$-Eigenfunction

We begin by proving that $I_1, \dots, I_6$ are smooth. The key to proving this is the Leibniz Integral Rule[1], which states that under mild conditions, the derivative with respect to one variable of the integral with respect to the other variable of a function of two variables is given by the integral of the corresponding partial derivative, which implies the analogous differentiability criterion.

Lemma 4.2.3. *For all $1 \leq j \leq 6$ and $k \in \mathbb{N}$, $I_j(r)$ is k times differentiable.*

Proof. Fix $1 \leq j \leq 6$. We know, from Definition 4.1.3, that

$$I_j(r) = \int_{X_j} g_j(z)\ e^{\pi i r z}\ \mathrm{d}z$$

for intervals X_j and holomorphic functions $g_j : \mathbb{H} \to \mathbb{C}$. The Leibniz Integral Rule then tells us that for all $k \in \mathbb{N}$, the kth derivative of I_j at some $r \in \mathbb{R}$ is given by

$$\int_{X_j} g_j(z)\ (\pi i z)^k\ e^{\pi i r z}\ \mathrm{d}z \tag{4.2.1}$$

In particular, I_j is smooth (in r) for all j. □

We are now ready to prove that $I_1, \dots, I_6$ and their derivatives satisfy the decaying property. As a first step, we show that we can apply Theorem 4.2.1.

Lemma 4.2.4. *There exist real numbers $C_0, C_{-2}, C_{-4} > 0$ such that*

$$|\phi_0(z)| \leq C_0 e^{-2\pi \operatorname{Im}(z)} \tag{4.2.2}$$
$$|\phi_{-2}(z)| \leq C_{-2} \tag{4.2.3}$$
$$|\phi_{-4}(z)| \leq C_{-4} e^{2\pi \operatorname{Im}(z)} \tag{4.2.4}$$

for all $z \in \mathbb{H}$ with $\operatorname{Im}(z) > \frac{1}{2}$.

Proof. Fix $z \in \mathbb{H}$ and assume that $\operatorname{Im}(z) > 1/2$. Since the Fourier coefficients of E_2, E_4 and E_6 grow polynomially (see Theorems 2.3.8 and 2.3.10), by Proposition 4.2.2, the Fourier coefficients of the numerators of ϕ_0, ϕ_{-2} and ϕ_{-4} grow polynomially as well. All that remains is to compute n_0 for the numerators of ϕ_0, ϕ_{-2} and ϕ_{-4}. Denote these N_0, N_{-2} and N_{-4} respectively. Note that $n_0(E_2) = n_0(E_4) = n_0(E_6) = 0$, with $c_{E_2}(0) = c_{E_4}(0) = C_{E_6}(0) = 1$.

- $N_0 = 4$. Recall that the numerator of ϕ_0 is $(E_2E_4 - E_6)^2$. Proposition 4.2.2 then tells us that $n_0(E_2E_4) = 0$. So, $n_0(E_2E_4 - E_6) \geq 0$. In fact, the 0th coefficients of both E_2E_4 and E_6 are 1, so they cancel. Hence, $n_0(E_2E_4 - E_6) = 2$. Hence, by $n_0\Big((E_2E_4 - E_6)^2\Big) = 4$.

[1] Also known as Leibniz's technique for 'differentiating under the integral sign'

- $N_{-2} = 2$. Recall that the numerator of ϕ_{-2} is $E_4\,(E_2E_4 - E_6)$. $n_0(E_2E_4 - E_6) = 2$ as shown above. Hence, $n_0(E_4\,(E_2E_4 - E_6)) = 2$.
- $N_{-4} = 0$. Recall that the numerator of ϕ_0 is E_4^2. Hence, $n_0\big(E_4^2\big) = 0$.

Substituting these values into Theorem 4.2.1 then gives us the desired bounds. □

We can now bound I_1, I_3 and I_5.

Lemma 4.2.5. *There exists a positive real number C_0 such that for all $r \in \mathbb{R}$,*

$$|I_1(r)|\,, |I_3(r)|\,, |I_5(r)| \leq \int_1^\infty C_0\, e^{-2\pi s}\, e^{-\pi r/s}\, \mathrm{d}s$$

Proof. For conciseness, we only bound $|I_1|$ explicitly. Parametrise $z = -1 + it$ in (4.1.7). Then, for all $r \in \mathbb{R}$, we can write

$$I_1(r) = -i \int_0^1 \phi_0\Big(\frac{-1}{it}\Big)\, t^2\, e^{-\pi i r}\, e^{\pi r t}\, \mathrm{d}t$$

Writing $s = \frac{1}{t}$ and simplifying, we get that

$$I_1(r) = -i \int_1^\infty \phi_0(is)\; s^{-4}\, e^{-\pi i r}\, e^{-\pi r/s}\, \mathrm{d}t$$

Applying the triangle inequality, multiplicativity and monotonicity, we get

$$|I_1(r)| \leq \int_1^\infty \Big|\phi_0(is)\; s^{-4}\, e^{-\pi i r}\, e^{-\pi r/s}\Big|\, \mathrm{d}t \leq \int_1^\infty |\phi_0(is)|\; e^{-\pi i r/s}$$

Since $s > \frac{1}{2}$ inside the integral, we know from Lemma 4.2.4 that $\exists C_0 > 0$ such that

$$|I_1(r)| \leq \int_1^\infty C_0\, e^{-2\pi s}\, e^{-\pi r/s}\, \mathrm{d}s$$

as required. The bounds on $|I_3|$ and $|I_5|$ are computed similarly. □

Now that we have bounds on the integrals with bounded vertical contours, we compute bounds on the integrals with bounded horizontal contours.

Lemma 4.2.6. *There exists a positive real number C_1 such that for all $r \in \mathbb{R}$,*

$$|I_2(r)|\,, |I_4(r)| \leq C_1\, e^{-\pi r}$$

Proof. For conciseness, we only bound $|I_2|$ explicitly. Parametrise $z = -1 + t + i$ in (4.1.8). Then, for all $r \in \mathbb{R}$, we can write

$$I_2(r) = \int_0^1 \phi_0\Big(\frac{-1}{t+i}\Big)\; (t+i)^2\; e^{-\pi i r}\, e^{\pi i r t}\, e^{-\pi r}\; \mathrm{d}t$$

Applying the triangle inequality, multiplicativity and monotonicity, we get

$$|I_2(r)| \leq \int_0^1 \Big|\phi_0\Big(\frac{-1}{t+i}\Big)\Big| \cdot 2e^{-\pi r}\, dt$$

It is therefore enough to show that Theorem 4.2.1 applies. To that end, we manipulate the expression inside ϕ_0:

$$\frac{-1}{t+i} = \frac{t}{t^2+1} + \frac{1}{t^2+1}i$$

From this, we can deduce that its imaginary part is greater than $\frac{1}{2}$ for all $t \in (0,1)$. Then, Theorem 4.2.1 tells us that $\exists C_0 > 0$ such that

$$\left|\phi_0\left(\frac{-1}{t+i}\right)\right| \leq C_0\, e^{-2\pi\cdot\frac{1}{t^2+1}} \leq C_0\, e^{-2\pi\cdot\frac{1}{2}} = C_0\, e^{-\pi}$$

Then, taking $C_1 := 2\, C_0\, e^{-\pi}$ and applying monotonicity,

$$|I_2(r)| \leq \int_0^1 C_1\, e^{-\pi r}\ \mathrm{d}t = C_1\, e^{-\pi r}$$

as required. The bound on $|I_4|$ is computed similarly. □

Finally, we bound the integral with the unbounded vertical contour.

Lemma 4.2.7. *There exists a positive real number C_2 such that for all $r \in \mathbb{R}$,*

$$|I_6(r)| \leq C_2\, \frac{e^{-\pi(r+2)}}{r+2}$$

Proof. Parametrise $z = it$ in (4.1.12). Then, for all $r \in \mathbb{R}$, we can write

$$I_6(r) = -2i \int_1^\infty \phi_0(it)\ e^{-\pi r t}\ \mathrm{d}t$$

Applying the triangle inequality, multiplicativity and monotonicity, we get

$$|I_6(r)| \leq 2 \int_1^\infty |\phi_0(it)|\ e^{-\pi r t}\ \mathrm{d}t$$

Let C_0 be as in Theorem 4.2.1 and define $C_1 := 2C_0$. Then,

$$|I_6(r)| \leq 2 \int_1^\infty C_0\, e^{-2\pi t}\, e^{-\pi r t}\ \mathrm{d}t = C_1 \int_1^\infty e^{-(2\pi+\pi r)t}\ \mathrm{d}t = C_1 \frac{e^{-\pi(r+2)}}{\pi\,(r+2)}$$

Defining $C_2 := \frac{C_1}{\pi}$ then yields the desired result. □

The formal proofs of the bounds in Lemmas 4.2.5 to 4.2.7 are complete up to proofs that they are bounded by integrable functions, which is necessary to apply monotonicity of the integral due to the definition of the integral in `mathlib`. The evaluation of the final integral in Lemma 4.2.7 is also currently a `sorry`.

Update

We now demonstrate how bounds on higher derivatives of I_1 can be reduced to the case we saw in Lemma 4.2.5.

Lemma 4.2.8. *For all $k \in \mathbb{N}$, there exists a positive real number $C_0^{(k)}$ such that for all $r \in \mathbb{R}$,*

$$\left|I_1^{(k)}(r)\right| \leq \int_1^\infty C_0^{(k)}\, e^{-2\pi s}\, e^{-\pi r/s}\, \mathrm{d}s$$

Proof. Fix $k \in \mathbb{N}$ and $r \in \mathbb{R}$. As we showed in (4.2.1), the kth derivative of I_1 at r can be expressed as

$$I_1^{(k)}(r) = \int_{-1}^{-1+i} \phi_0\left(\frac{-1}{z+1}\right)\,(z+1)^2\,(\pi i z)^k\, e^{\pi i r z}\, \mathrm{d}z$$

Parametrise $z = -1 + it$. Then,

$$I_1^{(k)}(r) = -i\int_0^1 \phi_0\left(\frac{-1}{it}\right)\, t^2\,(-\pi i - \pi t)^k\, e^{-\pi i r}\, e^{\pi r t}\, \mathrm{d}z$$

Writing $s = \frac{1}{t}$ and simplifying, we get that

$$I_1^{(k)}(r) = -i\int_1^\infty \phi_0(is)\; s^{-4}\left(-\pi i - \frac{\pi}{s}\right)^k\, e^{-\pi i r}\, e^{-\pi r/s}\, \mathrm{d}t$$

Observe that for all $s \in [1, \infty)$,

$$\left|-\pi i - \frac{\pi}{s}\right| = \pi\left|i + \frac{1}{s}\right| \leq \pi\sqrt{2}$$

Combining this with the bound computed in the proof of Lemma 4.2.5, we know $\exists C_0 > 0$ such that

$$\left|I_1^{(k)}(r)\right| \leq \int_1^\infty C_0\left(\pi\sqrt{2}\right)^k\, e^{-2\pi s}\, e^{-\pi r/s}\, \mathrm{d}s$$

Defining $C_0^{(k)} := C_0\left(\pi\sqrt{2}\right)^k$ then gives the desired result. □

Analogous results can be proved for the other I_j, with any arising constants subsumed into the constants defined in Lemmas 4.2.5 to 4.2.7. Therefore, if we can show that the functions (without the constants) on the right-hand sides of Lemmas 4.2.5 to 4.2.7 decay faster than any inverse power of r, we will have the result.

This fact is obvious for $|I_2|$, $|I_4|$ and $|I_6|$, but less so for $|I_1|$, $|I_3|$ and $|I_5|$. In these cases, the result is actually a consequence of a deeper result involving modified Bessel functions of the second kind.

An important comment we make here is that the fact that the functions that (up to constants) bound the I_j and their derivatives decay rapidly also establishes that the integrals I_j converge absolutely. In particular, their integrands are integrable, which will be a stepping stone to applying results like Fubini's Theorem later on.

We now show that similar results hold for the -1-eigenfunction.

4.2.2 The -1-Eigenfunction

A theme we will see repeatedly is that there are analogues for the -1-eigenfunction of many results on the $+1$-eigenfunction. The proof of the Schwartzness property is no exception to this rule.

Given that the contours used are similar, the main step in which the proof that b_{rad} is Schwartz differs from the proof that a_{rad} is Schwartz is the step in which we show that Theorem 4.2.1 is applicable.

4.3 Establishing the Eigenfunction Property

In this section, we show that the $\pm$-Eigenfunctions are, indeed, $\pm$-Eigenfunctions of the Fourier transform.

In the previous section, we did not work with a and b directly as it was sufficient to work with a_{rad} and b_{rad} instead. In this section, however, we will need to use the following formula for the n-dimensional Fourier transform of the n-dimensional Gaussian.

Theorem 4.3.1 (Fourier Transform of a Gaussian). *Fix $n \in \mathbb{N}$ and $b \in \mathbb{C}$, with $\operatorname{Re}(b) > 0$. If $F : \mathbb{R}^n \to \mathbb{C}$ is given by*

$$F(x) = e^{-b\|x\|^2}$$

then the Fourier transform of F is given by

$$\widehat{F}(\omega) = \left(\frac{\pi}{b}\right)^{n/2} e^{-\pi^2\|\omega\|^2/b}$$

A formal proof exists in `mathlib`.

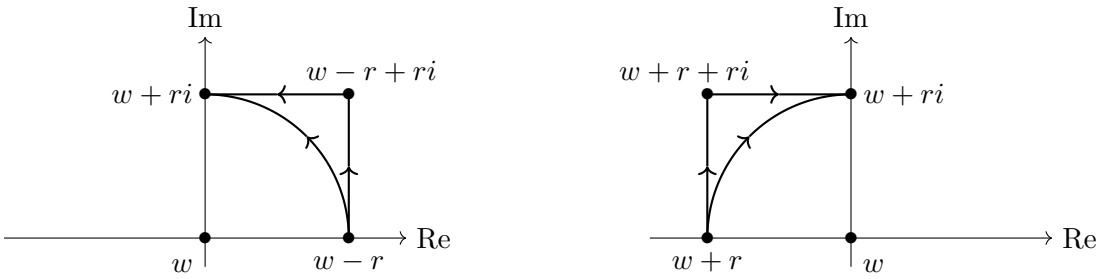
We will also need the following version of the Cauchy-Goursat Theorem, which allows us to deform contours of integration in the complex plane.

Theorem 4.3.2 (Cauchy-Goursat: Squares and Circles). *Fix $w \in \mathbb{C}$ and $r > 0$. Let γ be the quarter-circle parametrised by $\gamma(t) = w + r\cos(t) + |r|\, i\sin(t)$ for $0 \le t \le \pi/2$. For any $f : \mathbb{C} \to \mathbb{C}$ that is holomorphic in the region enclosed by γ and the line segments from $w + r$ to $w + r + ir$ and $w + r + ir$ to $w + ir$, we have*

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{w+r}^{w+r+ir} f(z)\,\mathrm{d}z + \int_{w+r+ir}^{w+ir} f(z)\,\mathrm{d}z$$

While this is an immediate consequence of the more general (and well-known) Cauchy-Goursat Theorem from complex analysis, there are numerous challenges involved in formalising this and other versions of the theorem. We will discuss it in Section 5.3. We also note that the above result implies an analogous result for

We now prove that a is indeed a $+1$-eigenfunction of the Fourier transform.



(a) Contours for which Theorem 4.3.2 holds. **(b)** Contours for which an analogous result holds.

Figure 4.2: The contour deformations permitted by Theorem 4.3.2.

4.3.1 The $+1$-Eigenfunction

The Fourier transform acts very interestingly on a. Recall from Theorem 3.1.3 that the Fourier transform is a linear isomorphism of Schwartz spaces. Since $I_1, \dots, I_6$ are Schwartz, so are their compositions with the norm-squared function. Hence, for all $x \in \mathbb{R}^8$,

$$\mathcal{F}(a(x)) = \mathcal{F}\left(\sum_{j=1}^{6} I_j\left(\|x\|^2\right)\right) = \sum_{j=1}^{6} \mathcal{F}\left(I_j\left(\|x\|^2\right)\right)$$

The strategy to show that $\mathcal{F}(a) = a$ will be to show that $\mathcal{F}$ acts on the $I_j\left(\|x\|^2\right)$ in the following manner:[2]

$$\mathcal{F}\left(I_1\left(\|x\|^2\right) + I_2\left(\|x\|^2\right)\right) = I_3\left(\|x\|^2\right) + I_4\left(\|x\|^2\right) \tag{4.3.1}$$
$$\mathcal{F}\left(I_3\left(\|x\|^2\right) + I_4\left(\|x\|^2\right)\right) = I_1\left(\|x\|^2\right) + I_2\left(\|x\|^2\right) \tag{4.3.2}$$
$$\mathcal{F}\left(I_5\left(\|x\|^2\right)\right) = I_6\left(\|x\|^2\right) \tag{4.3.3}$$
$$\mathcal{F}\left(I_6\left(\|x\|^2\right)\right) = I_5\left(\|x\|^2\right) \tag{4.3.4}$$

Since, in addition to being Schwartz, all the $I_j\left(\|x\|^2\right)$ (and their sums) are radial, Proposition 3.1.5 tells us that (4.3.2) and (4.3.4) follow from (4.3.1) and (4.3.3) respectively. We now prove (4.3.1) and (4.3.3).

As a preliminary step, though, we need to show integrability.

Proposition 4.3.3. *Fix $1 \le j \le 6$, and write $I_j(r) = \int_{X_j} f(r, z)\, \mathrm{d}z$. Then, the Fourier Integral*

$$\mathcal{F}\left(I_j\left(\|x\|^2\right)\right)(\xi) = \int_{\mathbb{R}^8} \int_{X_j} \left| f\left(\|x\|^2, z\right) e^{-2\pi i \langle x, \xi \rangle} \right| \mathrm{d}z\, \mathrm{d}x$$

converges for all $\xi \in \mathbb{R}^8$.

Proof. Fix $\xi \in \mathbb{R}^8$. Note that we can disregard the $e^{-2\pi i \langle x, \xi \rangle}$ factor since it has absolute value 1: the only reason we mention it is to emphasise that we are proving that the Fourier integral

[2] Note that we are abusing notation by denoting the function $x \mapsto I_j\left(\|x\|^2\right) \in \mathcal{S}(\mathbb{R}^8, \mathbb{C})$ by $I_j\left(\|x\|^2\right)$.

converges absolutely. Effectively, we need to show that the function

$$(x, z) \mapsto f\left(\|x\|^2, z\right)$$

admits an absolutely convergent integral over $\mathbb{R}^8 \times X_j$ with respect to the product measure.

It has been formally verified that proving this is equivalent to proving the following two facts.

1. **The integral over X_j of the function $z \mapsto f\left(\|x\|^2, z\right)$ is absolutely convergent for almost every $x \in \mathbb{R}^8$.**

   This is actually true for all x, and follows from the arguments in Section 4.2.1.

2. **The integral over $\mathbb{R}^8$ of the function $x \mapsto \int_{X_j} \left|f\left(\left\|x^2\right\|, z\right)\right| \mathrm{d}z$ is absolutely convergent.**

   By inspection, the arguments in Section 4.2.1 bound the function $r \mapsto \int_{X_j} |f(r, z)| \mathrm{d}z$ by a Schwartz function. Composing with the norm squared preserves Schwartzness by Proposition 3.1.6, so we can bound the function $x \mapsto \int_{X_j} \left|f\left(\|x\|^2, z\right)\right| \mathrm{d}z$ by a Schwartz function. It has been formally verified that Schwartz functions are integrable, so the integral over $\mathbb{R}^8$ of the function $x \mapsto \int_{X_j} \left|f\left(\|x\|^2, z\right)\right| \mathrm{d}z$ must be absolutely convergent.

We can therefore conclude that the Fourier integral converges absolutely. □

Note that the above proposition is stronger than showing merely that the $I_j\left(\|x\|^2\right)$ are integrable, as that does not imply that the integrands of the $I_j\left(\|x\|^2\right)$ are integrable with respect to the product measure on $\mathbb{R}^8 \times X_j$. Proposition 4.3.3, however, does imply this, and we will need this to swap integrals using Fubini's theorem to prove the eigenfunction property.

Lemma 4.3.4. *The Fourier transform maps $I_1\left(\|x\|^2\right) + I_2\left(\|x\|^2\right)$ to $I_3\left(\|x\|^2\right) + I_4\left(\|x\|^2\right)$.*

Proof. Since $\mathcal{F}$ is linear, we treat I_1 and I_2 separately. Observe that □

4.3.2 The -1-Eigenfunction

4.4 Establishing the Double Zeroes Property

The way we prove that a and b have double zeroes at E_8 lattice points with norm $> \sqrt{2}$ (or at normalised E_8 lattice points with norm > 1) is by showing that a_{rad} and b_{rad} agree, for $r > 2$, with functions that have double zeroes at *all* even integers. It will then follow that a and b have double zeroes at all points on the E_8 lattice with norm $> \sqrt{2}$, since all elements of Λ_8 have norm of the form $\sqrt{2n}$ for some $n \in \mathbb{N}$ (cf. `sorry`).

ADD RESULT TO E8 SECTION

The strategy to prove these two equalities will be to perform a change of contours using a version of the Cauchy-Goursat Theorem and use the relations and transformation rules between the ϕ- and ψ-functions to combine integrals so that the result is exactly a_{rad} or b_{rad}.

The version of the Cauchy-Goursat Theorem we use is the following.

Theorem 4.4.1 (Cauchy-Goursat for Unbounded Contours). *Suppose $f : \mathbb{C} \to \mathbb{C}$ is a function such that $f(z) \to 0$ as $\mathrm{Im}(z) \to \infty$. Then, for all $x_1, y_1, x_2, y_2 \in \mathbb{R}$, if f is holomorphic at z for all $z \in \mathbb{C}$ with $x_1 < \mathrm{Re}(z) < x_2$ and $y_2 < \mathrm{Im}(z)$, then*

$$\int_{x_1+iy_1}^{x_1+i\infty} f(z)\,\mathrm{d}z = \int_{x_1+iy_1}^{x_1+iy_2} f(z)\,\mathrm{d}z + \int_{x_1+iy_2}^{x_2+iy_2} f(z)\,\mathrm{d}z + \int_{x_2+iy_2}^{x_2+i\infty} f(z)\,\mathrm{d}z$$

provided that f is integrable on the unbounded vertical contours.

We discuss the informal and formal proofs of this theorem in Section 5.3.

For both a_{rad} and b_{rad}, we begin by their alternate expressions for them. We then make estimates to prove that the integrals in these expressions converge. We finally manipulate the expressions and apply Theorem 4.4.1 to show that they do, indeed, agree with a_{rad} and b_{rad} on inputs > 2.

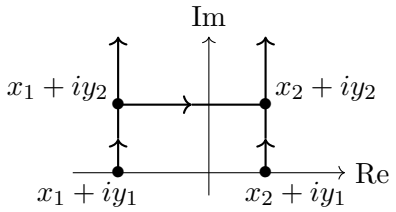


Figure 4.3: Visualising the contours in Theorem 4.4.1.

4.4.1 The $+1$-Eigenfunction

We begin by defining the integral by which we represent a_{rad}.

Definition 4.4.2 (Alternate Representation of a). Define $d : (2, \infty) \to \mathbb{C}$ by

$$d(r) = -4\sin^2\left(\frac{\pi r}{2}\right) \int_0^{i\infty} \phi_0\left(\frac{-1}{z}\right) z^2\, e^{\pi i r z}\,\mathrm{d}z$$

for all $r \in (2, \infty)$.

It is clear that we can parametrise the integral in d by $z = it$ for $t \in (0, \infty)$, and write

$$d(r) = 4i\sin^2\left(\frac{\pi r}{2}\right) \int_0^{\infty} \phi_0\left(\frac{i}{t}\right) t^2\, e^{-\pi r t}\,\mathrm{d}t \tag{4.4.1}$$

We begin by showing that this integral converges for $r > 2$. We do this by estimating the integrand.

Lemma 4.4.3. *$\exists C_0 > 0$ such that for $t \in (0, 2)$, $\left|\phi_0\left(\frac{i}{t}\right)\right| \le C_0\, e^{-2\pi/t}$.*

Proof. The result then follows immediately from (4.2.2), with $z = i/t$. □

We can hence conclude that for $t \in (0, 2)$, the integrand in (4.4.1) is bounded:

$$\left|\phi_0\left(\frac{i}{t}\right) t^2\, e^{-\pi r t}\right| \le 4C_0\, e^{-2\pi/t}\, e^{-\pi r t} \le 4C_0$$

We can also estimate the integrand for $t \ge 2$.

Lemma 4.4.4. *$\exists C > 0$ such that for $t \geq 2$, $\left|\phi_0\left(\frac{i}{t}\right)\right| \leq Ct^{-2}e^{2\pi t}$.*

Proof. From (4.1.5), we know that for all $t \geq 2$,

$$\left|\phi_0\left(\frac{i}{t}\right)\right| = \left|\phi_0\left(\frac{-1}{it}\right)\right| \leq \left|\phi_{0(it)}\right| + \frac{12}{\pi t}\left|\phi_{-2}(it)\right| + \frac{36}{\pi^2 t^2}\left|\phi_{-4}(it)\right|$$

Estimating each of these terms using Lemma 4.2.4, we know $\exists C_0, C_{-2}, C_{-4} > 0$ such that

$$|\phi_0(it)| + \frac{12}{\pi t}|\phi_{-2}(it)| + \frac{36}{\pi^2 t^2}|\phi_{-4}(it)| \leq C_0 e^{-2\pi t} + \frac{12}{\pi t}C_{-2} + \frac{36}{\pi^2 t^2}C_{-4}e^{2\pi t}$$

For $t \geq 2$, $C_0 e^{-2\pi t}$ and $\frac{12}{\pi t}C_{-2}$ are clearly bounded by constants, and the growth of the above expression is dominated by $t^{-2}e^{2\pi t}$. Hence, we can conclude that $\exists C > 0$ such that for $t \geq 2$, $\left|\phi_0\left(\frac{i}{t}\right)\right| \leq Ct^{-2}e^{2\pi t}$, as required. □

We can hence conclude that for $t \geq 2$, the integrand in (4.4.1) is bounded by an integrable function:

$$\left|\phi_0\left(\frac{i}{t}\right) t^2 e^{-\pi r t}\right| \leq C\left(t^{-2}e^{2\pi t}\right)\left(t^2 e^{-\pi r t}\right) = Ce^{\pi t(2-r)}$$

Here, we require $r > 2$ so that the exponent is negative. Since d was defined precisely on such r, we can conclude that the integral in the definition of d converges absolutely.

Arguing as above yields another important result.

Lemma 4.4.5. *For all $r > 2$ and $z \in \mathbb{H}$, as $\mathrm{Im}(z) \to \infty$,*

$$\phi_0\left(\frac{-1}{z}\right) z^2 e^{\pi i r z} \to 0$$

Proof sketch. We do not offer more than a sketch here because the proof is almost identical to that of Lemma 4.4.4. The idea is to apply (4.1.5), multiply through, apply Lemma 4.2.4 to bound the expression in absolute value, and use the fact that $r > 2$ to conclude that the bound decays exponentially as $\mathrm{Im}(z) \to \infty$. □

The function $z \mapsto \phi_0\left(\frac{-1}{z}\right) z^2 e^{\pi i r z}$ is also holomorphic on $\mathbb{H}$, a fact that is again seen by applying (4.1.5) and the fact that the numerators of the ϕ-function are holomorphic and the denominators are non-vanishing on $\mathbb{H}$.

We are now ready to show the following.

Proposition 4.4.6. *For all $r > 2$, $d(r) = a_{\mathrm{rad}}(r)$.*

Proof. Fix $r > 2$. We begin by performing the following manipulation:

$$-4\sin^2\left(\frac{\pi r}{2}\right) = -4\left(\frac{e^{i\pi r/2} - e^{-i\pi r/2}}{2i}\right)^2 = \left(e^{i\pi r/2} - e^{-i\pi r/2}\right)^2 = e^{i\pi r} - 2 + e^{-i\pi r}$$

Therefore, for all $r > 2$,

$$d(r) = -4\sin^2\Big(\frac{\pi r}{2}\Big)\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z = \big(e^{i\pi r} - 2 + e^{-i\pi r}\big)\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z$$
$$= \int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi ir(z+1)}\,\mathrm{d}z - 2\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z + \int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi ir(z-1)}\,\mathrm{d}z$$

By a simple change of variables, we can rewrite the first and third integrals as follows. We can also split the second integral into two parts as shown below.

$$\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi ir(z-1)}\,\mathrm{d}z = \int_{-1}^{-1+i\infty}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi ir(z+1)}\,\mathrm{d}z = \int_1^{1+i\infty}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$-2\int_0^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z = -2\int_0^{i}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z - 2\int_i^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z$$
$$= I_5(r) - 2\int_i^{i\infty}\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\,\mathrm{d}z$$

We can now apply Theorem 4.4.1 to the first and third integrals, noting that the required integrability conditions do hold because the integrals making up d and a_{rad} converge absolutely.

$$\int_{-1}^{-1+i\infty}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z = \int_{-1}^{-1+i}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$+\int_{-1+i}^{i}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$+\int_{i}^{i\infty}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$= I_1(r) + I_2(r) + \int_i^{i\infty}\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$\int_{1}^{1+i\infty}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z = \int_{1}^{1+i}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$+\int_{1+i}^{i}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$+\int_{i}^{i\infty}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z$$
$$= I_3(r) + I_4(r) + \int_i^{i\infty}\phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz}\,\mathrm{d}z$$

Hence, we can express $d(r)$ as the sum of the following six integrals:

$$d(r) = I_1(r) + I_2(r) + I_3(r) + I_4(r) + I_5(r)$$
$$+\int_i^{i\infty}\Big(\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2\,e^{\pi irz} + \phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2\,e^{\pi irz} - 2\phi_0\Big(\frac{-1}{z}\Big)\,z^2\,e^{\pi irz}\Big)\,\mathrm{d}z$$

One can show, by applying (4.1.5) and simplifying, that

$$\phi_0\Big(\frac{-1}{z+1}\Big)\,(z+1)^2 + \phi_0\Big(\frac{-1}{z-1}\Big)\,(z-1)^2 + \phi_0\Big(\frac{-1}{z}\Big)\,z^2 = 2\phi_0$$

Hence, the sixth integral above is precisely I_6, proving that $d(r) = a_{\mathrm{rad}}(r)$ for all $r > 2$. □

The factor of $-4\sin^2(\pi r/2)$ in the definition of d then ensures that a_{rad} has double zeroes at all even integers except possibly 0 and ± 2, which allows us to conclude that a has double zeroes at all points in $\mathbb{R}^8$—particularly those lying on Λ_8—with norm of the form $\sqrt{2n}$ for $n \in \mathbb{N} \setminus \{0, 1\}$.

4.4.2 The -1-Eigenfunction

4.5 The Magic of g

Chapter 5

Viazovska's Magic Function, Formally

5.1 Project Design

5.2 A Metaprogramming Approach

In this section, we discuss an unexpected biproduct of this project: the development of a normalisation-simplification automation for performing computations in the complex numbers.

Lean, like other interactive theorem provers, primarily interacts with its users through **tactics**. Fundamentally, the proof of a theorem in Lean is given by a **proof term**, which can be thought of as a concise expression that captures the information of how the hypotheses or inputs of the theorem are transformed into its conclusion by giving exactly the conclusion into which these inputs are transformed. A tactic is a command that, when invoked by a Lean user, performs a step in the construction of the proof term for a theorem.

The most basic tactics can be thought of as being 'syntax sugar' rather than invocations of computation or reasoning algorithms. Consider the following code.

```
example (P Q R : Prop) : P ∧ (Q ∧ R) ↔ (P ∧ Q) ∧ R := by
  constructor
  · intro h
    constructor
    · constructor
      · exact h.1
      · exact h.2.1
    · exact h.2.2
  · intro h
    constructor
    · exact h.1.1
    · constructor
      · exact h.1.2
      · exact h.2
```

Listing 5.1: A tactic-mode proof of the associativity of ∧

This proof demonstrates how the `constructor`, `intro` and `exact` tactics work. These tactics give the Lean compiler the following instructions:

- `constructor`: "Prove the goal by proving the two statements it consists of." It works on conjunctions and biconditionals, that is, if the goal is of the form `A ∧ B`, then `constructor`

replaces it with two goals, namely, `A` and `B`, and if the goal is of the form `A ↔ B`, then `constructor` replaces it with two goals, namely, `A → B` and `B → A`.

- `intro`: "Prove the goal by introducing the assumption term and proving the conclusion term." It works on implications and universal quantifications, that is, if the goal is of the form `A → B`, then `intro h` introduces an assumption `h` of `A` and replaces the goal with `B`, and if the goal is of the form `∀ (x : A), B`, then `intro a` introduces term `a` of type `A` and replaces the goal with `B`.

- `exact`: "Prove the goal with the following." It works on any goal where the proof of that goal is already known, that is, if the goal is `B` and some proof `h` of `B` is already known, then `exact h` proves the goal with `h`.

In addition, the terms `h.1`, `h.2.1`, etc are shorthand for "the first constituent of `h`", "the first constituent of (the second constituent of `h`)", etc, where by "first constituent" and "second constituent", we mean the terms respectively to the left and right of the ∧ symbol.

What the tactics used in Listing 5.1 are actually doing is constructing the following **term-mode proof** of the same result.

```
example (P Q R : Prop) : P ∧ (Q ∧ R) ↔ (P ∧ Q) ∧ R :=
  ⟨fun h ↦ ⟨⟨h.1, h.2.1⟩, h.2.2⟩,
   fun h ↦ ⟨h.1.1, ⟨h.1.2, h.2⟩⟩⟩
```

Listing 5.2: A term-mode proof of the associativity of ∧

This proof is significantly shorter than the tactic-mode proof see in Listing 5.1. While the code is arguably less readable than the tactic-mode proof, it is not too difficult to dissect:

- The `constructor` occurrences in Listing 5.1 correspond to the *anonymous constructors* ⟨,⟩ in Listing 5.2.

- The `intro h` occurrences in Listing 5.1 correspond to the function definitions `fun h ↦` in Listing 5.2.

- The `exact` occurrences in Listing 5.1 correspond to the terms inside the anonymous constructors in Listing 5.2.

In particular, the proof in Listing 5.2 consists solely of functions and constructors. No tactics occur anywhere in the argument (note the absence of the `by` keyword, which marks the beginning of a tactic-mode proof).

It turns out that Listing 5.2 still contains some syntax sugar. It is possible to use helper lemmas like `Iff.intro` and `And.intro` to avoid using the anonymous constructors, but a proof term completely devoid of the constructor syntax would look like the following.

```
example (P Q R : Prop) : P ∧ (Q ∧ R) ↔ (P ∧ Q) ∧ R := {
  mp := fun a ↦ {
    left := {
      left := And.casesOn a fun left right ↦ And.casesOn right fun _ _ ↦ left
      right := And.casesOn a fun _ right ↦ And.casesOn right fun left _ ↦ left
    }
    right := And.casesOn a fun _ right ↦ And.casesOn right fun _ right ↦ right
  }
  mpr := fun a ↦ {
    left := And.casesOn a fun left _right ↦ And.casesOn left fun left _ ↦ left
```

```
    right := {
      left := And.casesOn a fun left _ ↦ And.casesOn left fun _ right ↦ right
      right := And.casesOn a fun left right ↦ And.casesOn left fun _ _ ↦ right
    }
  }
}
```

Listing 5.3: A proof term for the associativity of ∧

Proof terms, as we can see from Listing 5.3, are often long and do not clearly communicate the mathematical ideas they represent. Tactics overcome this by constructing proof terms without revealing them to the user. Indeed, there are tactics that serve as more than just syntax sugar: for example, results in intuitionistic propositional logic (such as the associativity of ∧) can be proved by the tactic `itauto`. That is, the following code compiles.

```
example (P Q R : Prop) : P ∧ (Q ∧ R) ↔ (P ∧ Q) ∧ R := by itauto
```

Listing 5.4: A one-line tactic proof for the associativity of ∧

Other tactics like `tauto` and `simp` also work. The proof term generated by such a tactic can be viewed by typing `show_term`.[1] For a more detailed explanation of how proof terms and tactics work, see [20, particularly Chapters 3 and 5].

Metaprogramming is the science of writing tactics in Lean. While syntax-sugar tactics are incredibly useful (compare the readability of Listings 5.1 to 5.3), automation tactics often go an even longer way in keeping the focus of nontrivial mathematical proofs on precisely their nontrivial aspects. Given how computationally involved the construction of Viazovska's Magic Function is (as seen in Chapter 4), the author, after a discussions with Macbeth, realised that the most efficient approach to formalising some of the computational aspects of Viazovska's argument was to write a tactic. The first version of this tactic, developed as a collaboration between Macbeth, Xie and the author, with inputs from Mehta, was called `norm_numI`.

In the forthcoming subsections, we explore the motivation and technique used to develop `norm_numI`, and briefly discuss how the tactic maybe further developed and the scope of its applicability expanded.

5.2.1 Complex Computations are Complex

Computations in general are quite challenging to perform in interactive theorem provers. This is because such languages are designed for *proof* rather than *computation*. Indeed, tactics that simplify goals do not do so merely by simplifying expressions: they construct proofs that the simplified expression is, indeed, equal to the original expression. Existing tactics like `norm_num`, `simp` and `field_simp` do not always do this successfully when the expressions in question are in $\mathbb{C}$. `simp` and `field_simp` are very general tactics that work in a wide variety of settings. They both work by constructing a special set of equality or biconditional lemmas by sifting through the library and performing repeated rewrite operations to transform the goal into a simpler form. `field_simp` is specifically designed to simplify expressions in fields, and can handle operations like clearing denominators. However, it does not have access to the particularities of the field in question (such as the fact that $i^2 = -1$ in $\mathbb{C}$). The tactic that we will be most interested in, specifically because it is designed to handle simplifications of *numerals* in *specific* settings, is `norm_num`.

[1] For the curious reader, Listing 5.3 was generated by repeatedly typing `show_term by tauto` inside each field.

`norm_num` is a tactic that handles expressions involving numerals. It works best in $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. For example, it handles the following.

```
example : (1 : ℕ) + 2 + 3 + 4 = 10 := by norm_num
example : (-2 : ℚ) * (3 + 8/9) = -70/9 := by norm_num
example : (-9 : ℤ) + 5 * (6 - 20) = -79 := by norm_num
```

Listing 5.5: `norm_num` simplifying expressions in $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$

It is worth mentioning, however, that `norm_num` often has difficulties in $\mathbb{R}$. This is due to the immense technical detail baked into the very definition of $\mathbb{R}$ in Lean, which allows for the existence of transcendental numbers. In the following example, none of `norm_num`, `field_simp`, `ring` and `simp` can prove the result in one line, because they are unable to treat π as more than a symbol. Indeed, the entire proof rests on a `mathlib` result, `Real.pi_gt_three`.

```
example : (π - 1) / (π - 1) = 1 := by
  have hₐ : (1 : ℝ) < 3 := by norm_num
  have h₂ : 1 ≠ π := ne_of_lt <| hₐ.trans pi_gt_three
  have h₃ : π - 1 ≠ 0 := sub_ne_zero_of_ne h₂.symm
  field_simp [h₃]
```

Listing 5.6: An expression in $\mathbb{R}$ not handled immediately by simplification tactics

Observe, however, that `norm_num` *is* able to prove the inequality $1 < 3$ despite it being an expression in $\mathbb{R}$. The reason is that `norm_num` can navigate the canonical inclusions from $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ into $\mathbb{R}$, meaning that it can simplify expressions in $\mathbb{R}$ that come from expressions it can simplify in $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{Q}$.[2] It cannot, however, show that $1 < \pi$, because it does not treat π *as a numeral*. In $\mathbb{C}$, `norm_num` faces this challenge not only with real transcendental numbers like π but also with the imaginary constant i. Consider the following example.[3]

```
example : (1 + I) * (1 + I * I * I) = 2 := by
  simp only [I_mul_I, neg_mul, one_mul, mul_add, mul_one, mul_neg, add_mul,
    neg_add_rev, neg_neg]
  ring
```

Listing 5.7: A nontrivial computation in $\mathbb{C}$, done formally

Again, `simp`, `field_simp`, `ring` and `norm_num` all fail, because i, like π in Listing 5.6, is not handled as a numeral. Observe, however, that $(1+i)(1+i\cdot i\cdot i)$ lies in $\mathbb{Z}[i]$. This means that if it is expressed as $a+bi$, with a and b both being (not necessarily simplified) real expressions, then in fact, a and b are both images of expressions in $\mathbb{Z}$. This means that `norm_num` would be able to individually handle both a and b, resulting in a simplified expression of the form $a'+b'i$, with a' and b' being simplified. This suggests that the key to writing a tactic that can simplify expressions like those in Listing 5.7 is to find a way to separate them into their real and imaginary parts, which in turn involves navigating the fact that $i^2 = -1$.

5.2.2 Parsing and Normalisation

The core idea behind `norm_num` is that it simplifies expressions by computing normal forms. In its most basic form, `norm_num` attempts to prove equalities of by putting the left and right hand sides in unique normal forms that can simply be inspected to check if the two sides are equal. For example, in the natural numbers, simple arithmetic facts are true by *reflexivity*, that is, they are

[2] We say `norm_num` can handle coercions.

[3] Note that in Lean, the imaginary constant is denoted by an uppercase `I` instead of a lowercase i. We will adhere to standard mathematical conventions and use a lowercase i when referring to the imaginary constant in informal contexts.

proved by the tactic `rfl`, which proves definitional equality. Hence, the right normal form for numerical expressions in $\mathbb{N}$ is to just compute them and express them as a single natural number (ie, as the right number of applications of the successor function to 0). Then, by inspection, two expressions are equal if and only if their normal forms—that is, their simplifications into single natural numbers—are equal. Inequalities work similarly.

While the working of `norm_num` may appear trivial in $\mathbb{N}$, its versatility becomes clearer in *semirings*. Recall that a semiring is an algebraic structure that is a commutative, additive monoid and a multiplicative monoid, with the quintessential example being $\mathbb{N}$. If A is any semiring, there is a natural semiring homomorphism $\uparrow\colon \mathbb{N} \to A$ given by

$$\begin{aligned} \uparrow 0 &:= 0 \\ \uparrow 1 &:= 1 \\ \uparrow 2 &:= 1 + 1 \\ &\vdots \end{aligned}$$

A **numeral** in A is then any element of the image of $\uparrow$, and `norm_num` puts expressions involving numerals in normal forms by recognising them as numerals, computing the normal form of their pre-images in $\uparrow$, and pushing the image back through $\uparrow$. The nontriviality of `norm_num` for numerical expressions in A is then not how it computes normal forms in $\mathbb{N}$ but how it navigates $\uparrow$. For more on how `norm_num` works in semirings, follow the tutorial in [21, `Metaprogramming/NormNum`]. We will not discuss it in any more detail here, but will instead discuss the working of `norm_numI`—specifically, what the desired normal form is and how it is computed.

Since $\mathbb{R}$ is a subfield of $\mathbb{C}$, the constraints that the standard `norm_num` faces in $\mathbb{R}$ are also constraints it can reasonably be expected to face in $\mathbb{C}$. The goal of `norm_numI` is not to overcome *these* constraints, but to overcome the *additional* constraints that come from not treating i as a numeral. The target normal form for an expression in $\mathbb{C}$ is therefore given by separating it into its real and imaginary parts, both of which are real expressions, and normalising them as much as possible.

The key to `norm_numI` is the `parse` function. It separates an expression $z \in \mathbb{C}$ into its real and imaginary parts by performing a recursive pattern-match. For example, if the outermost operation is addition—ie, if $z = z_1 + z_2$—then it calls itself on both z_j, obtaining real and imaginary parts $a_j, b_j \in \mathbb{R}$ and proofs that $z_j = a_j + b_j i$, and returns the expression $(a_1 + a_2) + (b_1 + b_2)\, i$ as well as a proof that $z = (a_1 + a_2) + (b_1 + b_2)\, i$, which it obtains via a helper lemma `split_add`. It performs similar recursive actions if z is of the form $z_1 \cdot z_2$, z_1^{-1}, z_1/z_2, $-z_1$, $z_1 - z_2$, $\overline{z_1}$, z_1^n for some $n \in \mathbb{N}$, i, or a decimal/natural number. The recursion is guaranteed to terminate, because an expression that is fed into the function cannot contain infinitely many characters.

Example 5.2.1. The expression $z = (1 + i)\,(1 + i \cdot i \cdot i)$ (cf. Listing 5.7) would be parsed in the following manner.

1. To parse z, write $z = z_1 \cdot z_2$, where $z_1 = 1 + i$ and $z_2 = 1 + i \cdot i \cdot i$.
2. To parse z_1, write $z_1 = z_{11} + z_{12}$ where $z_{11} = 1$ and $z_{12} = i$.
3. z_{11} is parsed as $1 + 0i$.
4. z_{12} is parsed as $0 + 1i$.
5. By `split_add`, $z_1 = z_{11} + z_{12}$ is parsed as
$$(1 + 0) + (0 + 1)\, i$$
6. To parse z_2, write $z_2 = z_{21} + z_{22}$, where $z_{21} = 1$ and $z_{22} = i \cdot i \cdot i$.

7. z_{21} is parsed as $1 + 0i$.
8. To parse z_{22}, write $z_{22} = z_{221} \cdot z_{222}$, where $z_{221} = i \cdot i$ and $z_{222} = i$.
9. To parse z_{221}, write $z_{221} = z_{2211} \cdot z_{2212}$, where $z_{2211} = i$ and $z_{2212} = i$.
10. z_{2211} is parsed as $0 + 1i$.
11. z_{2212} is parsed as $0 + 1i$.
12. By `split_mul`, $z_{221} = z_{2211} \cdot z_{2212}$ is parsed as

$$(0 \cdot 0 - 1 \cdot 1) + (0 \cdot 1 + 0 \cdot 1)\, i$$

13. z_{222} is parsed as $0 + 1i$.
14. By `split_mul`, $z_{22} = z_{221} \cdot z_{222}$ is parsed as

$$((0 \cdot 0 - 1 \cdot 1) \cdot 0 - (0 \cdot 1 + 1 \cdot 0) \cdot 1) + ((0 \cdot 0 - 1 \cdot 1) \cdot 1 + 0 \cdot (0 \cdot 1 + 1 \cdot 0))\, i$$

15. By `split_add`, $z_2 = z_{21} + z_{22}$ is parsed as

$$\begin{aligned}&(1 + ((0 \cdot 0 - 1 \cdot 1) \cdot 0 - (0 \cdot 1 + 1 \cdot 0) \cdot 1)) \\ &+ (0 + ((0 \cdot 0 - 1 \cdot 1) \cdot 1 + 0 \cdot (0 \cdot 1 + 1 \cdot 0)))\, i\end{aligned}$$

16. By `split_mul`, $z = z_1 + z_2$ is parsed as

$$\begin{aligned}&\Big((1 + 0) \cdot (1 + ((0 \cdot 0 - 1 \cdot 1) \cdot 0 - (0 \cdot 1 + 1 \cdot 0) \cdot 1)) \\ &\quad - (0 + 1) \cdot (0 + ((0 \cdot 0 - 1 \cdot 1) \cdot 1 + 0 \cdot (0 \cdot 1 + 1 \cdot 0))) \Big) \\ &+ \Big((1 + 0) \cdot (0 + ((0 \cdot 0 - 1 \cdot 1) \cdot 1 + 0 \cdot (0 \cdot 1 + 1 \cdot 0))) \\ &\quad + (1 + ((0 \cdot 0 - 1 \cdot 1) \cdot 0 - (0 \cdot 1 + 1 \cdot 0) \cdot 1)) \cdot (0 + 1) \Big) i\end{aligned}$$

Note that in Lean, the result of parsing does not appear as an expression of the form $a + bi$, but rather, as a `structure` with fields `re` and `im` for the real and imaginary parts. However, the underlying idea is no different from what we have described.

```
{
  re :=
    (1 + 0) * (1 + ((0 * 0 - 1 * 1) * 0 - (0 * 1 + 1 *
0) * 1)) -
      (0 + 1) * (0 + ((0 * 0 - 1 * 1) * 1 + (0 * 1 + 1
* 0) * 0)),
  im :=
    (1 + 0) * (0 + ((0 * 0 - 1 * 1) * 1 + (0 * 1 + 1 *
0) * 0)) +
      (0 + 1) * (1 + ((0 * 0 - 1 * 1) * 0 - (0 * 1 + 1
* 0) * 1)) }
```

Figure 5.1: The Lean output of the steps shown in Example 5.2.1.

Clearly, despite being mathematically valid, the result of parsing can be long and uninformative, making it an unsuitable choice of normal form for our purposes. However, by separating complex expressions into their real and imaginary parts, `parse` perfectly sets up a very simple normalisation procedure that will put our expression in an appropriate normal form. Since we know of such a procedure for real expressions—namely, `norm_num`—and since `parse` expresses any complex expression as a combination of two real expressions, the normalisation procedure simply makes calls to `norm_num` to express each of them in a *real* normal form. The result is a complex number in a normal form $a + bi$ (or, in Lean notation, `{re := a, im := b}`) with a and b both simplified to the greatest extent possible (as expressions in $\mathbb{R}$) by applying `norm_num`.

Note that `norm_numI` is currently implemented as a `conv` tactic rather than a full tactic, meaning that it is only capable of modifying expressions (and providing a proof that the modification is valid). It is not currently capable of proving goals, which are necessarily logical statements, such as equalities. This means that it needs to be used as follows.

```
example : (1 + I) * (1 + I * I * I) = 2 := by
  conv_lhs => norm_numI
  conv_rhs => norm_numI
```

Listing 5.8: Using `norm_numI` as a `conv` tactic

Unpacking the code,

- `conv_lhs => norm_numI` applies the parsing-normalisation procedure outlined above to convert the expression `(1 + I) * (1 + I * I * I)` on the left-hand side to `{re := 2, im := 0}`.
- `conv_rhs => norm_numI` applies the parsing-normalisation procedure outlined above to convert the expression `2` on the right-hand side to `{re := 2, im := 0}`.

Since the two sides are then exactly the same, the goal is proved. Note that both lines are necessary: 2 on its own has not been separated into real and imaginary parts. The imaginary part needs to be explicitly shown to be 0, which the `parse` function does.

After Macbeth, Xie and the author's initial success with this `conv` tactic, Macbeth proceeded to create an *extension* of the existing `norm_num` tactic that uses the parsing technique outlined above to handle complex expressions. This tactic is still being developed, but is being tested on active code from the project with immensely promising results.

5.2.3 Scope for Further Development

The benefits of having such a tactic cannot be overstated. There are numerous instances across the project where the collaborators have had to repeatedly prove computational facts in $\mathbb{C}$, such as $1+i \neq 0$, or clear complex denominators to separate expressions into their real and imaginary parts. It is precisely this that motivated the development of `norm_numI`, and indeed, `norm_numI` is proving to be a viable solution.

The applicability of `norm_numI`—or, at the very least, of the underlying idea, that the right normal form for expressions in $\mathbb{C}$ is to separate them into real and imaginary parts and express those in a normal form—extends well beyond expressions in $\mathbb{C}$. A key motivation for Xie, one of the co-creators of `norm_numI` and an ardent algebraist, was to create a similar tactic that would normalise and simplify expressions in quaternion algebras. A discussion at a London Learning Lean event hosted at the London Institute of Mathematical Sciences in March 2025 sparked speculation among well-regarded members of the Lean Community that similar ideas might be applicable in other field extensions (with $\mathbb{C}$ regarded as $\mathbb{R}[X]\big/(X^2+1)$ and i as the image of X via the quotient map). The idea would be that the separation into components corresponds to taking advantage of some linear independence criterion, with algebraic dependences not handled by the simplification step being captured by the helper lemmas in the normalisation step.

There are numerous technical difficulties with implementing such a tactic that can work in other contexts, chief amongst them the fact that there would need to be a single modification to the tactic capable of handling very differently behaved algebraic structures (for instance, it would need to work the same way in quadratic, cubic, quartic and higher degree field extensions). It would hence need to have some awareness of the behaviours of each field in which it is applied, which is technically challenging. Nevertheless, the development of `norm_numI` and the subsequent `norm_num` modification marks a significant, and long overdue, step in the right direction, so that fewer nontrivial proofs are needed to prove trivial facts.

5.3 The Cauchy-Goursat Theorem

There are some areas of mathematics that are notoriously difficult to formalise. Algebra, for example, tends to be easier to formalise than analysis. Within analysis, it tends to be particularly difficult to formalise geometric ideas. The Jordan Curve Theorem, for example, tends to be a particularly difficult theorem to formalise. It was not until 2007 that this theorem, proposed in the late 19th Century, was formalised by Tom Hales [22] in HOL Light, and to this day, no formalisation exists in Lean. The author had the privilege of meeting Hales in Pittsburgh, USA, in March 2025 to discuss the formalisation of 8-dimensional sphere packing in Lean, and the very first question Hales asked was what the strategy was to overcome the challenges of not having a Lean formalisation of the Jordan Curve Theorem. It turns out that there is a workaround, which we explore in this section.

Before we discuss the workaround, we briefly discuss the statement of the Jordan Curve Theorem and its relevance to this project. The Jordan Curve Theorem essentially states that a simple closed curve $C \subset \mathbb{R}^2$, given as the image of a continuous injection from $\mathbb{S}^1$, divides $\mathbb{R}^2 \setminus C$ into a bounded region, known as the *interior* of C, and an unbounded region, known as the *exterior* of C. The relevance of the Jordan Curve Theorem is that in Viazovska's construction—specifically, in the proof that a and b satisfy the double zero property—it becomes necessary to deform contours. While the contours in question are not closed, the proof that the deformation is possible (under a vanishing condition) follows from limiting applications of the Cauchy-Goursat Theorem, which states that integral of $\mathbb{C} \to \mathbb{C}$ function around a closed contour is zero if the function is holomorphic in the interior. The point is that the Jordan Curve Theorem is necessary to define the interior, and without some version of the Jordan Curve Theorem, it becomes challenging to state the all-important holomorphicity condition of the Cauchy-Goursat Theorem.

While it may not appear, at first, that there *is* a workaround, it turns out that weaker versions of the Cauchy-Goursat Theorem, where the interior is more easily defined, are sufficient for our purposes. We outline the bounded and unbounded versions of these below.

5.3.1 The Cauchy-Goursat Theorem for Bounded Rectangular Contours

5.3.2 The Cauchy-Goursat Theorem for Unbounded Rectangular Contours

5.3.3 Scope for Further Development

Bibliography

[1] H. Cohn. The work of Maryna Viazovska. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 82–105. EMS Press, 2023. Presented at ICM, July 6–14, 2022.

[2] T. C. Hales. A Proof of the Kepler Conjecture. *Annals of Mathematics*, 162(3):1065–1185, 2005.

[3] E. Klarreich. Sphere Packing Solved in Higher Dimensions. *Quanta Magazine*, March 2016.

[4] H. Cohn. A Conceptual Breakthrough in Sphere Packing. *Notices of the American Mathematical Society*, 64(02):102–115, Feb. 2017.

[5] A. Thue. Om nogle geometrisk-taltheoretiske Theoremer. *Forhandlingerne ved de Skandinaviske Naturforskeres*, 14, 1892. Zbl 24.0259.01.

[6] T. C. Hales. Cannonballs and Honeycombs. *Notices of the American Mathematical Society*, 47(4):440–449, Apr. 2000.

[7] J. Kepler. *Strena seu de nive sexangula*. Francofurti ad Moenum : apud Godefridum Tampach, 1611. ETH-Bibliothek Zürich, Rar 4342: 2, https://doi.org/10.3931/e-rara-478.

[8] T. C. Hales et al. A Formal Proof of the Kepler Conjecture. *Forum of Mathematics, Pi*, 5: e2, 2017.

[9] H. Cohn and N. Elkies. New Upper Bounds on Sphere Packings I. *Annals of Mathematics*, 157(2):689–714, 2003.

[10] T. C. Hales. Introduction to the Flyspeck Project. In T. Coquand, H. Lombardi, and M.-F. Roy, editors, *Mathematics, Algorithms, Proofs*, volume 5021 of *Dagstuhl Seminar Proceedings (DagSemProc)*, pages 1–11, Dagstuhl, Germany, 2006. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[11] W. Gowers, B. Green, F. Manners, and T. Tao. On a conjecture of Marton. *Annals of Mathematics*, 201(2):515 – 549, 2025.

[12] M. S. Viazovska. The sphere packing problem in dimension 8. *Annals of Mathematics*, 185 (3):991–1015, 2017.

[13] C. Birkbeck, S. Hariharan, S. Lee, G. Ma, B. Mehta, and M. Viazovska. Sphere Packing in Lean - Project Blueprint, 2025. https://thefundamentaltheor3m.github.io/Sphere-Packing-Lean/blueprint/index.html.

[14] C. Birkbeck, S. Hariharan, S. Lee, G. Ma, B. Mehta, and M. Viazovska. Sphere Packing in Lean - Documentation, 2025. https://thefundamentaltheor3m.github.io/Sphere-Packing-Lean/docs/.

[15] T. F. Görbe. Exceptionally Beautiful Symmetries. https://tamasgorbe.com/symmetry.

[16] J.-P. Serre. *A Course in Arithmetic*. Springer-Verlag, New York, 1973.

[17] L. Schwartz. *Théorie des distributions*. Hermann, Paris, 1978. Nouvelle édition.

[18] F. Diamond and J. Shurman. *A First Course in Modular Farms*. Number 228 in Graduate Texts in Mathematics. Springer New York, NY, New York, first edition.

[19] K. Buzzard. Families of modular forms. *Journal de Théorie des Nombres de Bordeaux*, 13 (1):43–52, 2001.

[20] J. Avigad, L. de Moura, S. Kong, and S. Ullrich. *Theorem Proving in Lean 4*. https://leanprover.github.io/theorem_proving_in_lean4/title_page.html.

[21] H. Macbeth. metaprogramming - tutorials for metaprogramming in lean. https://github.com/hrmacbeth/metaprogramming.

[22] T. C. H. and. The jordan curve theorem, formally and informally. *The American Mathematical Monthly*, 114(10):882–894, 2007.