Geometry of numbers, class group statistics and free path lengths

SAMUEL HOLMIN

Doctoral Thesis
Stockholm, Sweden 2015

KTH
TRITA-MAT-A 2015:15
Institutionen för Matematik
ISRN KTH/MAT/A–15/15-SE
100 44 Stockholm
ISBN 978-91-7595-797-5
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i matematik fredagen den 15 januari 2016 kl 13.00 i sal F3, Kungl Tekniska högskolan, Lindstedtsvägen 26, Stockholm.

Tryck: Universitetsservice US AB

Abstract

This thesis contains four papers, where the first two are in the area of geometry of numbers, the third is about class group statistics and the fourth is about free path lengths. A general theme throughout the thesis is lattice points and convex bodies.

In Paper A we give an asymptotic expression for the number of integer matrices with primitive row vectors and a given nonzero determinant, such that the Euclidean matrix norm is less than a given large number. We also investigate the density of matrices with primitive rows in the space of matrices with a given determinant, and determine its asymptotics for large determinants.

In Paper B we prove a sharp bound for the remainder term of the number of lattice points inside a ball, when averaging over a compact set of (not necessarily unimodular) lattices, in dimensions two and three. We also prove that such a bound cannot hold if one averages over the space of all lattices.

In Paper C, we give a conjectural asymptotic formula for the number of imaginary quadratic fields with class number h, for any odd h, and a conjectural asymptotic formula for the number of imaginary quadratic fields with class group isomorphic to G, for any finite abelian p-group G where p is an odd prime. In support of our conjectures we have computed these quantities, assuming the generalized Riemann hypothesis and with the aid of a supercomputer, for all odd h up to a million and all abelian p-groups of order up to a million, thus producing a large list of "missing class groups." The numerical evidence matches quite well with our conjectures.

In Paper D, we consider the distribution of free path lengths, or the distance between consecutive bounces of random particles in a rectangular box. If each particle travels a distance R, then, as $R \to \infty$ the free path lengths coincides with the distribution of the length of the intersection of a random line with the box (for a natural ensemble of random lines) and we determine the mean value of the path lengths. Moreover, we give an explicit formula for the probability density function in dimension two and three. In dimension two we also consider a closely related model where each particle is allowed to bounce N times, as $N \to \infty$, and give an explicit formula for its probability density function.

Sammanfattning

Denna avhandling innehåller fyra artiklar, varav de första två är i ämnet geometrisk talteori, den tredje handlar om klassgruppstatistik, och den fjärde handlar om fria väglängder. Ett generellt tema genom avhandlingen är gitterpunkter och konvexa kroppar.

I Artikel A ger vi ett asymptotiskt uttryck för antalet heltalsmatriser med primitiva radvektorer och en given determinant, sådana att den euklidiska matrisnormen är mindre än ett givet stort tal. Vi undersöker också tätheten av matriser med primitiva radvektorer i rummet av matriser med en given determinant, och avgör dess asymptotiska beteende för stora determinanter.

I Artikel B bevisar vi en skarp övre gräns på feltermen för antalet gitterpunkter inuti en boll då vi tar medelvärdet över en kompakt mängd av (inte nödvändigtvis unimodulära) gitter, i dimension två och tre. Vi bevisar även att en sådan övre gräns inte kan hålla om vi tar medelvärdet över rummet av alla gitter.

I Artikel C ger vi en förmodad asymptotisk formel för antalet imaginära kvadratiska kroppar med klasstal h, för udda h, och en förmodad asymptotisk formel för antalet imaginära kvadratiska kroppar med klassgrupp isomorf med G, för ändliga abelska p-grupper G där p är ett udda primtal. För att stödja vår förmodan så har vi beräknat dessa kvantiteter, under antagandet av den generaliserade Riemannhypotesen och med hjälp av en superdator, för all udda h upp till en miljon G och alla abelska p-grupper av ordning upp till en miljon, och vi har därmed producerat en stor lista på "saknade klassgrupper". De numeriska resultaten matchar våra förmodanden väl.

I Artikel D betraktar vi fördelningen av fria väglängder, dvs sträckan mellan på varandra följande studsar av slumpmässiga partiklar i en rektangulär låda. Om varje partikel färdas en sträcka R, så överensstämmer fördelningen av fria väglängder då $R \to \infty$ med fördelningen av längden av snittet av en slumpmässig linje med lådan (för en naturlig ensemble av slumpmässiga linjer), och vi beräknar medelvärdet av fria väglängderna. Vi ger ett explicit uttryck för täthetsfunktionen i dimension två och tre. I dimension två betraktar vi även en relaterad modell där varje partikel tillåts studsa N gånger, och vi ger ett explicit uttryck för täthetsfunktionen då $N \to \infty$.

Contents

Acknowledgements

First and foremost I would like to express my gratitude to my advisor Pär Kurlberg. You have always been generous with your time and your constant enthusiasm and encouragement made this thesis possible. I feel truly fortunate to have had you as my advisor.

Secondly I would like to thank my friends and coworkers at KTH for making it fun to go to work. In particular I want to thank my officemates Andreas Minne, Erik Aas, Sebastian Öberg and Erik Duse for your helpful comments and all our fun discussions.

Finally I want to thank my family for being there.

1 Introduction

This thesis consists of this introduction and four papers. As indicated by the title of the thesis, the first two papers are in the area of geometry of numbers, the third paper is about class group statistics, and the fourth paper is about free path lengths. Lattices and convex bodies are a general theme throughout the thesis.

In this introduction we give an informal overview of the results obtained in the papers contained in this thesis, intended to be accessible to a general audience. For the sake of exposition, we will deviate from the formulations used in the papers, and instead present the results with a more geometrical flavor.

1.1 Overview of Paper A

Consider a parallelogram with integer coordinates which cannot be decomposed into smaller parallelograms with integer coordinates. We will call such an object a **primitive parallelogram**; see Figure 1.1 for an illustration. How many primitive parallelograms are there with an area of 10?

There are infinitely many such primitive parallelograms: in fact, starting with a single primitive parallelogram, we can produce another one with the same area by for example shifting it an integer distance up or to the right, or by shearing it (see Figure 1.2), and by repeating either of these operations we can produce arbitrarily many different parallelograms, all of which are primitive and have the same area.

Thus, in order to make the counting problem interesting we need to impose some restriction not only on the location of the parallelograms but also on their *size*. A natural restriction is to consider all primitive parallelograms

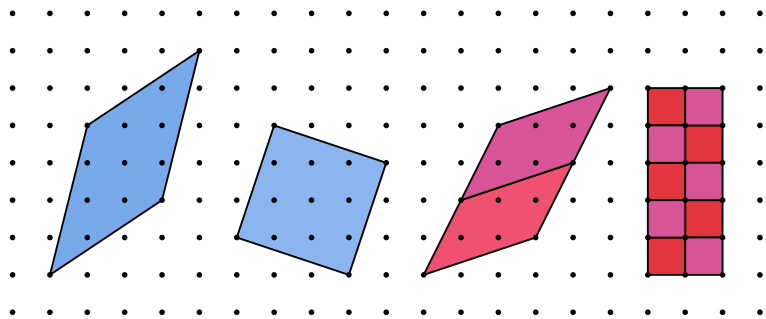


Figure 1.1 – Four parallelograms of area 10. The two blue parallelograms are primitive, but the two red parallelograms are not, since they can be partitioned into several smaller parallelograms.

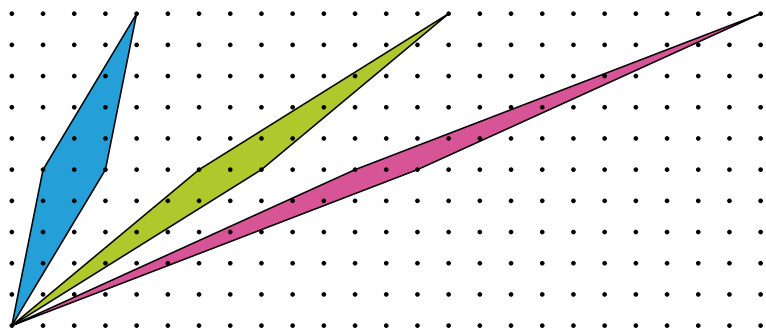


Figure 1.2 – Three primitive parallelograms of area 10, where the green parallelogram is obtained by applying the shear transformation $(x, y) \mapsto (x + y, y)$ to the vertices of the blue parallelogram, and the pink parallelogram is obtained by applying the same shearing transformation to the green parallelogram.

with the origin of the plane as a vertex, and such that

$$\sqrt{a^2 + b^2} \tag{1.1.1}$$

is bounded by some large number T, say 100, where a and b are the side-lengths of the parallelogram.

The first result of Paper A implies that the number of such primitive parallelograms is approximately

$$2.4 \cdot T^2$$

for large values of T. Indeed, by a brute force computer calculation one can find that the correct number for $T = 100$ is 24000, which coincides ex-

actly with our approximation (typically however, the approximation will be slightly off). In fact, in Paper A we prove an analogous result for the generalization of the above problem to n dimensions, where we replace primitive parallelograms of area 10 with n-dimensional primitive *parallelepipeds* with a given positive volume k, such that the origin is a vertex of the parallelepiped and such that $\sqrt{a_1^2 + \cdots + a_n^2} \leq T$, where $a_1, a_2, \ldots, a_n$ are the side-lengths of the parallelepiped (the n-dimensional generalization of a parallelogram).[1]

Next, suppose we choose a parallelogram with integer coordinates and area 10 at random.[2] What is the probability that the parallelogram we choose is primitive? Equivalently, what is the proportion of primitive parallelograms out of the set of all integer parallelograms with area 10? It follows from the results of Paper A that the probability in this case is

$$22.222\ldots\%.$$

A brute force computer search reveals that the proportion for $T \leq 100$ is $24000/107816 = 0.22260\ldots$, which indeed is close to the value above. In Paper A we determine the probability in the generalized case in n dimensions and a given positive volume k. Let us denote this probability by $D_n(k)$.

We may ask which values can occur for the probability $D_n(k)$. We prove in Paper A that in two dimensions, probabilities arbitrarily close to any given probability between 0% and 100% occur. Let us now focus on $n \geq 3$ dimensions. The probability $D_n(k)$ is maximized and equal to 1 for $k = 1$ only, and we prove that $D_n(k)$ is close to 1 precisely if k has no small divisors (for example if k is a large prime, say 31337). Although we have until now assumed that the volume k is positive, one can make sense of the value $D_n(0)$, and we prove that the probability $D_n(k)$ is minimized and equal to

$$\frac{1}{\zeta(n-1)^n} \tag{1.1.2}$$

for $k = 0$ only, where ζ is the Riemann zeta function, and we prove that $D_n(k)$ is close to this minimum value for precisely those values of k which are divisible by all small numbers (for example factorials, say $k = 7! = 5040$).

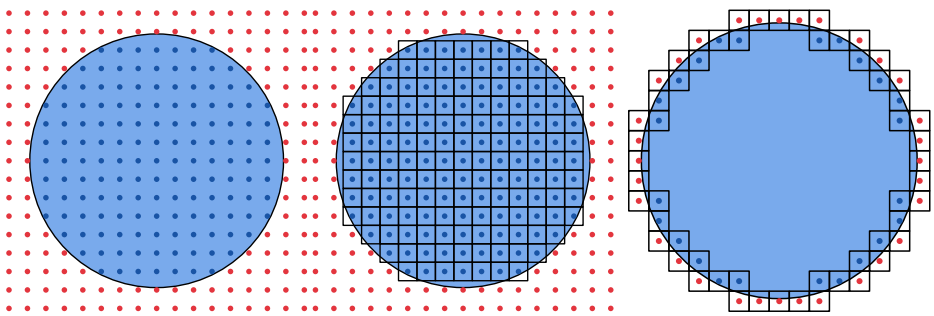
[1] In Paper A we actually compute a different number $N'_{n,k}(T)$, which is precisely a factor $n!/2$ larger than the number we are talking about in this section.

[2] To make this rigorous, we may think of the randomization process as selecting uniformly at random one parallelogram out of all parallelograms of area 10 with the origin as a vertex which satisfy the condition (1.1.1) for some fixed large T.

Finally, we prove that for $n \geq 4$ there are "gaps" for the probability $D_n(k)$ in the sense that not all probabilities between the minimum (1.1.2) and the maximum 1 can be attained; for example, in dimension four the probability values are never betwen 74% and 81%, even though the minimum value is about 48%. No statement was made[3] in Paper A about whether there are any gaps in dimension $n = 3$.

1.2 Overview of Paper B

Consider the lattice of all points with integer coordinates in the plane and draw a large circle centered at the origin. How many lattice points are there inside the circle?



(a) (b) (c)

Figure 1.3 – Approximating the number of lattice points inside a circle by the area of the circle.

It is easy to see that the area of the circle is an approximation of the number of lattice points inside the circle: namely, if we draw a square of area 1 around each lattice point inside the circle (see Figure 1.3b for an illustration), then the number of lattice points inside the circle is equal to the area of the union of all these squares, and since this jagged shape approximates the circle, the conclusion follows. Thus we are lead to the natural follow-up question of how good this approximation is. Let t be the radius of the circle and let us write $N(t)$ for the number of lattice points inside the circle. Then, as we have shown, $N(t)$ is approximately πt^2. Define

[3] At the time of this writing, I have proved that there are no gaps in dimension $n = 3$.

the **discrepancy**

$$E(t) := \left| N(t) - \pi t^2 \right|$$

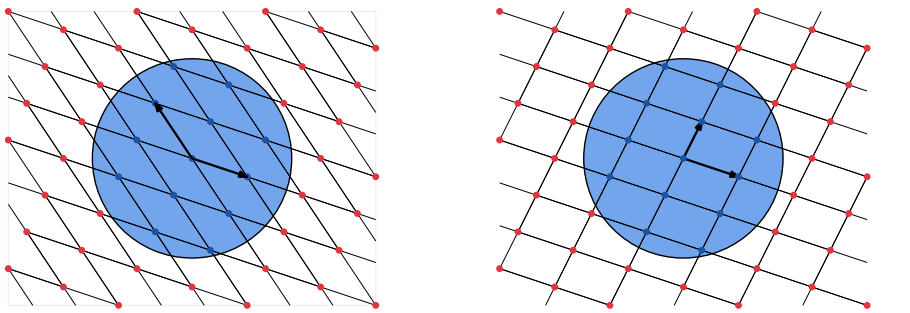
to be the difference between the actual number of lattice points inside the circle and our approximation. How large is $E(t)$ for large radii t? This is known as the **Gauss circle problem**. We can argue as follows to get an upper bound on the discrepancy $E(t)$.

Consider the set of squares of area 1 centered at a lattice point such that the square touches the boundary of the circle, as in Figure 1.3c. If we want to adjust our approximation πt^2 to become the correct value $N(t)$, then we need to add, for every square containing a blue lattice point in Figure 1.3c the missing area in that square (the white part), and we need to subtract, for every square containing a red lattice point in Figure 1.3c the area of the blue part in that square. Thus, in the worst case, we would have to add (or subtract) no more than the area of all the squares in Figure 1.3c. The number of squares in Figure 1.3c is, up to a constant, approximately equal to the circumference $2\pi t$ of the circle, and therefore $E(t)$ should be of the order t for large t; since this is much smaller than πt^2 for large t, this justifies our calling the latter an approximation for $N(t)$.

However, the actual size of the discrepancy $E(t)$ should be much smaller, by the following heuristic. When we add together all the small "adjustments" described in the previous paragraph, we should intuitively expect that many of the positive adjustments will be cancelled out by negative adjustments. If we cheat and pretend that the adjustments are "random" and independent of each other, then the situation becomes similar to a **one-dimensional random walk**: take a large number of steps of length 1, and at each step either move forwards or backwards at random. How far from our starting point should we expect to end up at the end of the random walk? The answer turns out to be roughly the square root of the number of steps, and we should therefore expect the discrepancy $E(t)$ to be of the order $t^{1/2}$ for large t. How small can we make the exponent in the discrepancy? Landau has proven that the exponent cannot be $1/2$ or smaller, but Hardy has conjectured that it can be made arbitrarily close to $1/2$. The best result to date is the exponent $131/208 \approx 0.6298\ldots$, due to Huxley.

In Paper B, we consider the generalization of the Gauss circle problem where we replace the lattice of integer points with a random lattice; see Figure 1.4a. A **lattice** in the plane is the set of points that can be reached

from the origin by taking steps of $v, -v, w$, and $-w$ where v and w are two non-parallel vectors; the vectors v and w are called **basis vectors** for the lattice. For an illustration, see Figure 1.4a, where the two black arrows are the basis vectors and the lattice is the set of blue and red dots. (Note that different basis vectors can yield the same lattice; see Figure 1.4b.) For a general lattice, the number of lattice points inside a circle is approximately equal to the area of the circle, divided by the area of the parallelogram spanned by the two basis vectors. We may again ask what the discrepancy is between the number of points inside the circle and this approximation.



(a) Lattice points inside a circle (b) Same lattice, different basis vectors

Figure 1.4

A random lattice may be generated by choosing uniformly at random one basis vector each from two bounded regions in the plane, such that no choice can yield two vectors that are parallel or arbitrarily close to parallel, and such that arbitrarily short basis vectors cannot be chosen; see Figure 1.5 for an example. In Paper B, we prove that the *expected* discrepancy for a random lattice is of the order $t^{1/2}$ for large t.[4]

We also consider the analogous problem in three dimensions. Similar to the arguments in two dimensions, we can show that the number of points from a given lattice inside a sphere of large radius t is approximately proportional to the volume $\frac{4}{3}\pi t^3$ of the sphere. A naïve argument shows, as before, that the discrepancy can be bounded, up to a constant, by the surface area of the sphere, which is of the order t^2, but a heuristic argument suggests that the actual size of the discrepancy should be roughly the square root

[4]In Paper B, we actually use a more natural but less intuitive method of generating random lattices, but the proof also works for the simpler model that we describe here.

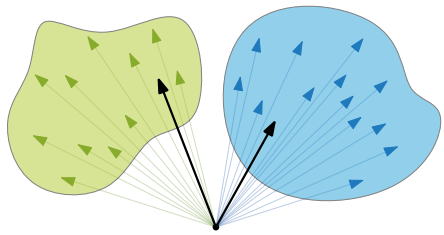


Figure 1.5 – Generating a random lattice by choosing one blue basis vector and one green basis vector.

of this. We prove in Paper B that the *expected* discrepancy for a random lattice is t, up to a logarithmic factor.

One may ask if the conditions we placed on the two regions in Figure 1.5 are necessary. In Paper B, we prove that if we relax these conditions, then one can find pairs of regions such that the expected discrepancy must be strictly larger (in fact, of order $t^{1.5}$) in the three-dimensional case.

1.3 Overview of Paper C

A different way of counting the lattice points inside a circle is to add together, for each smaller circle centered at the origin, the number of points which lie exactly on that circle; see Figure 1.6.

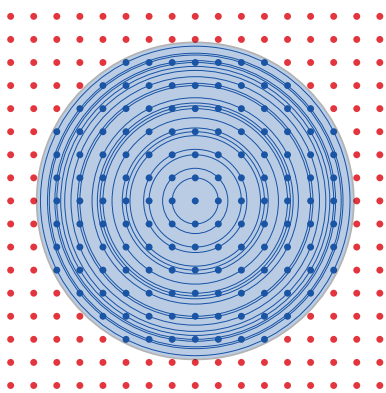


Figure 1.6 – Each lattice point inside the large blue circle is on the boundary of some concentric blue circle.

To be explicit, each lattice point (x, y) inside a circle of radius t centered at the origin satisfies the equation

$$x^2 + y^2 = m \tag{1.3.1}$$

for some integer $m \leq \sqrt{t}$. The number of integer points (x, y) which satisfy (1.3.1) is commonly denoted $r_2(m)$, and thus the number of lattice points inside the circle of radius t centered at the origin can also be given as

$$\sum_{m \leq \sqrt{t}} r_2(m),$$

which corresponds to summing the number of points on each blue circle in Figure 1.6.

More generally one could count lattice points inside an ellipse in an analogous fashion. If a, b, c are integers and $m \geq 0$, then the equation

$$ax^2 + bxy + cy^2 = m \tag{1.3.2}$$

describes an ellipse in the plane centered at the origin if and only if

$$b^2 - 4ac < 0.$$

We call $D := b^2 - 4ac$ the **discriminant** of the **quadratic form** $ax^2 + bxy + cy^2$ (or just **form** for short). If there exist integers x, y such that the equation (1.3.2) is satisfied, then the quadratic form is said to **represent** m.

This motivates (among many other reasons) the study of quadratic forms and the set of integers which a quadratic form represents. How many quadratic forms are there with a given negative discriminant D? When counting quadratic forms, we will dismiss those quadratic forms which are essentially just another quadratic form in disguise. For example, we will dismiss the quadratic form

$$10x^2 + 5xy + 15y^2, \tag{1.3.3}$$

as it can be be obtained from the quadratic form

$$f(x, y) := 2x^2 + xy + 3y^2 \tag{1.3.4}$$

by multiplying the latter form by 5. We will also dismiss the quadratic form

$$g(x,y) := 50x^2 - 5xy + 8y^2, \tag{1.3.5}$$

as it can be obtained from f by writing $g(x,y) = f(5x-y, 3y)$. The quadratic form

$$h(x,y) := 2x^2 + 5xy + 6y^2$$

can be obtained from f by writing $h(x,y) = f(x+y, y)$, but on the other hand, we can also obtain f from h by writing $f(x,y) = h(x-y, y)$, so we will dismiss neither, but instead consider $f(x,y)$ and $h(x,y)$ to be **equivalent**, and regard both of them to be the same form. The **opposite form** of f (note that we only flip the sign of the middle term),

$$F(x,y) := 2x^2 - xy + 3y^2, \tag{1.3.6}$$

can be obtained from f by writing $F(x,y) = f(x,-y)$, and conversely we can obtain f from F by writing $f(x,y) = F(x,-y)$, but the variable substitution $(x,y) \mapsto (x,-y)$ flips the orientation of the plane and as such we do not necessarily regard F as equivalent to f (unless a variable substitution between them which keeps the orientation of the plane also exists). The set of possible discriminants we are left with after dismissing all quadratic forms such as (1.3.3) and (1.3.5) are precisely[5] the (negative) **fundamental discriminants**. Two quadratic forms with the same negative fundamental discriminant represent exactly the same set of integers if and only if the forms are equivalent or opposite.

The **class number** of a negative fundamental discriminant D is the number of quadratic forms with discriminant D, where we count equivalent quadratic forms as the same quadratic form. The **Gauss class number problem** is to find all negative fundamental discriminants D with class number 1; this is a highly nontrivial problem which was not solved until the 1950s or 1960s.[6]

In Paper C we give, for any odd number h, a formula which we conjecture approximates the number of negative fundamental discriminants D such that the class number of D is h. (Our restriction to odd values simplifies

[5] See Proposition 7.1a in [Bue89].

[6] A proof was given by Heegner in 1952 which was not initially accepted, and it was later proved independently by Baker in 1966 and by Stark in 1967.

certain aspects of the technical arguments in the paper.) In support of our conjecture, we have computed the correct value (under the assumption of the generalized Riemann hypothesis) for all odd h up to a million with the aid of a supercomputer, and we find that our approximation is typically within 1% of the correct value within this range.

An ancient identity known as Brahmagupta's identity states that

$$(x^2 + ky^2) \cdot (z^2 + kw^2) = (xz + kyw)^2 + k(xw - yz)^2$$

for any value of k, or in other words, if we multiply the two quadratic forms $x^2 + ky^2$ and $z^2 + kw^2$ then we can write the result as the quadratic form $X^2 + kY^2$ where $X = xz + kyw$ and $Y = xw - yz$. Gauss managed to generalize this and prove that one can, for any quadratic forms f and g of the same negative fundamental discriminant, write

$$f(x, y) \cdot g(z, w) = Q(X, Y)$$

for some quadratic form $Q(X, Y)$, where X and Y are integer linear combinations of xz, xw, yz, yw. The set of quadratic forms with a given negative fundamental discriminant D together with this multiplication operation, and where we regard equivalent forms as the same form, is called the **class group** of D. For example, the discriminant of the quadratic form (1.3.4) is -23, and it can be shown that the class group of -23 has exactly three elements (and thus the class number of -23 is 3): the form f which we defined in (1.3.4), the form F which we defined in (1.3.6) and the form

$$i(x, y) := x^2 + xy + 6y^2.$$

The multiplication table of this group is given in Figure 1.7a. The multiplication table gives complete information about the structure of the class group. We note that the multiplication table in Figure 1.7a is identical (after a simple name-change) to the addition table of the group of integers $\{0, 1, 2\}$ modulo 3, and thus the two groups have the same structure.

It is natural to ask what the multiplication tables of the class groups look like. It is for example well-known that they are always symmetric about the main diagonal. In Paper C, we make a conjecture about which multiplication tables can be attained from the class groups of odd negative fundamental discriminants;[7] for technical reasons we restrict ourselves to

[7] In the paper, this conjecture is of course stated in terms of finite abelian groups rather than multiplication tables.

$\cdot$	i	f	F
i	i	f	F
f	f	F	i
F	F	i	f

(a) Multiplication table for the quadratic forms of discriminant -23.

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

(b) Addition table for the integers $0, 1, 2$ modulo 3.

Figure 1.7 – Two groups with the same structure.

the case where the size of the class group has only one prime divisor p, for odd primes p. We conjecture that a multiplication table is more likely to be attained (for some odd negative fundamental discriminant) if it has a high degree of "cyclicity," and therefore that most tables will not be attained. In support of our conjecture, we have calculated (under the assumption of the generalized Riemann hypothesis) with the aid of a supercomputer the structures of all class groups with an odd class number h up to a million. The data seems to support our conjecture, and in particular we have found (again, under the assumption of the generalized Riemann hypothesis) a large list of explicit examples of "missing class groups."

Remark 1.3.7. Distinguishing between equivalent and opposite forms makes it so that the inverse element of a quadratic form in the class group of a discriminant d becomes the opposite of the quadratic form, but the distinction is perhaps better motivated by the fact that this definition makes the class group of a negative *fundamental* discriminant isomorphic to the *ideal class group* of the quadratic field $\mathbb{Q}(\sqrt{d})$; see for instance Chapter 5 in [Cox89] for an exposition on ideal class groups.

1.4 Overview of Paper D

Consider a rectangular room containing only a point-shaped light bulb. Turn the light on, so that it sends out light rays in all directions. After a few minutes, the rays will have travelled a large distance (the same distance for each ray) and bounced a large number of times against the walls of the room; see Figure 1.8 for an illustration.

Between each pair of consecutive bounces of a given ray, the ray will have travelled a certain distance; we will refer to such a distance as a

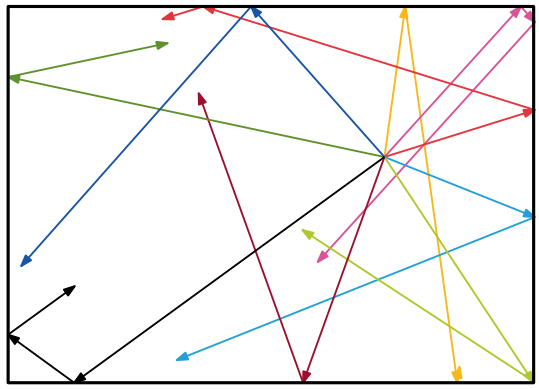


Figure 1.8 – Sending out a large number of rays for an equal distance each.

bounce length. In Paper D, we calculate the distribution of bounce lengths for the system of all rays; see Figure 1.9 for an example. We also calculate the distribution for the analogous problem in three dimensions. Moreover, we determine the expected value of the bounce lengths for the analogous problem in n dimensions for any $n \geq 2$, and find that it has the rather simple geometrical interpretation as the quotient

$$\frac{\text{volume of the room}}{\text{surface area of the room}}$$

multiplied by the constant $2\pi S_{n-1}/S_n$ where S_{n-1} is the surface area of the $(n-1)$-dimensional sphere in n-dimensional space.

Note that in the problem above, each ray contributes a different number of bounce lengths (for example, although they have travelled the same distance, the yellow ray in Figure 1.8 has bounced twice while the green ray has only bounced once). This means that the distribution of bounce lengths favors some directions more than others. It is thus interesting to ask what the distribution of bounce lengths of a *single ray* would be, if its starting direction is chosen uniformly at random (or equivalently, we could use many rays, but instead impose that they all travel for the same number of *bounces* rather than the same *distance*), so that all directions are treated equally. This second problem is less mathematically elegant than the first and we only have a result in two dimensions. We give explicitly the distribution for this problem in two dimensions in Paper D and find indeed that it differs from the distribution above. See Figure 1.10 for an example.

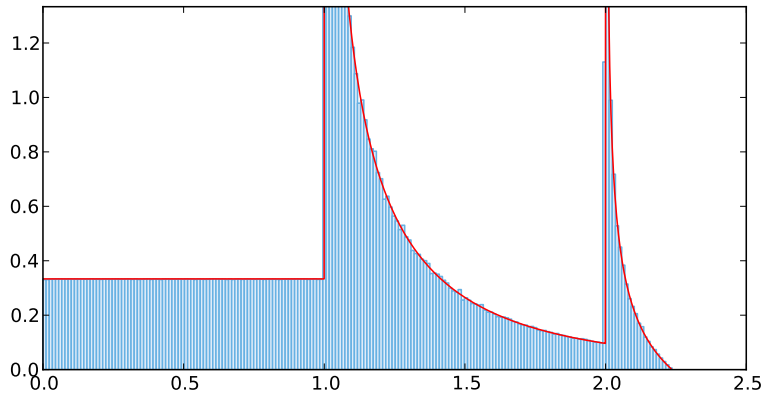


Figure 1.9 – Distribution of bounce lengths for a rectangle with side-lengths 1 and 2 (normalized to have area 1). The red curve is given by an explicit formula for the probability density function, and the blue histogram was obtained experimentally with a computer simulation by sending 100000 rays from the origin in uniformly random directions for a distance 1000 each. Note that the side-lengths of the rectangle can be recovered from the locations of the singularities of the curve.

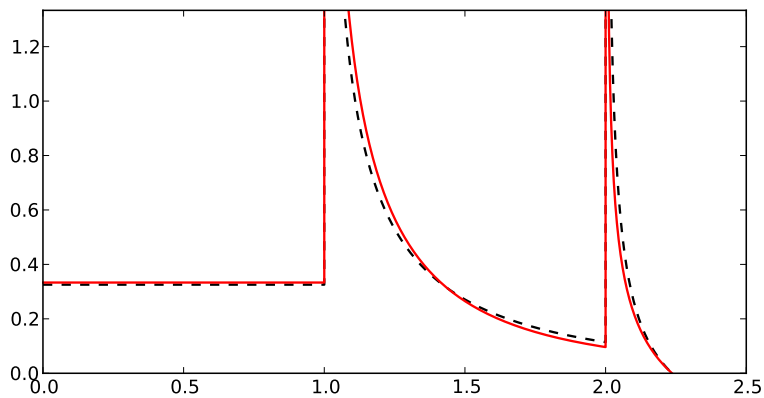


Figure 1.10 – Red curve: distribution of bounce lengths for a system of many particles in a rectangle with side-lengths 1 and 2. Black dashed curve: Distribution of bounce lengths for a system of a single random particle in the same rectangle.

References

[Bue89] Duncan A. Buell, *Binary quadratic forms*, Springer-Verlag, New York, 1989, Classical theory and modern computations. MR 1012948 (92b:11021)

[Cox89] David A. Cox, *Primes of the form* $x^2 + ny^2$, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1989, Fermat, class field theory and complex multiplication. MR 1028322 (90m:11016)

Paper A

Monatsh Math (2014) 173:209–230
DOI 10.1007/s00605-013-0518-x

Counting nonsingular matrices with primitive row vectors

Samuel Holmin

Received: 30 November 2012 / Accepted: 18 May 2013 / Published online: 4 June 2013

Abstract We give an asymptotic expression for the number of nonsingular integer $n \times n$-matrices with primitive row vectors, determinant k, and Euclidean matrix norm less than T, as $T \to \infty$. We also investigate the density of matrices with primitive rows in the space of matrices with determinant k, and determine its asymptotics for large k.

Keywords Matrices · Lattices · Primitive vectors · Asymptotics

Mathematics Subject Classification (2000) 11H06 Lattices and convex bodies

1 Introduction

An integer vector $v \in \mathbb{Z}^n$ is **primitive** if it cannot be written as an integer multiple $m \neq 1$ of some other integer vector $w \in \mathbb{Z}^n$. Let A be an integer $n \times n$-matrix with nonzero determinant k and primitive row vectors. We ask how many such matrices A there are of Euclidean norm at most T, that is, $\|A\| \leq T$, where $\|A\| := \sqrt{\sum a_{ij}^2} = \sqrt{\mathrm{tr}(A^t A)}$. Let $N'_{n,k}(T)$ be this number (the prime in the notation denotes the primitivity of the rows), and let $N_{n,k}(T)$ be the corresponding counting function for matrices with not necessarily primitive row vectors. We will determine the asymptotic behavior of $N'_{n,k}(T)$ for large T, and investigate the density $D_n(k) := \lim_{T\to\infty} N'_{n,k}(T)/N_{n,k}(T)$ of matrices with primitive vectors in the space

Communicated by A. Constantin.

S. Holmin (✉)
Department of mathematics, KTH, 100 44 Stockholm, Sweden
e-mail: holmin@kth.se

of matrices with nonzero determinant k. Since $N'_{n,k}$ and $N_{n,k}$ do not depend on the sign of k, we will without loss of generality assume that $k > 0$.

Let $M_{n,k}$ be the set of integer $n \times n$-matrices with determinant k. Then $N_{n,k}(T) = |B_T \cap M_{n,k}|$, where B_T is the (closed) ball of radius T centered at the origin in the space $M_n(\mathbb{R})$ of real $n \times n$-matrices equipped with the Euclidean norm. Throughout, we will assume that $n \geq 2$ and $k > 0$ unless stated otherwise.

Duke et al. [2] found that the asymptotic behavior of $N_{n,k}$ is given by

$$N_{n,k}(T) = c_{n,k} T^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(n+1)+\varepsilon}),$$

as $T \to \infty$, for a certain constant $c_{n,k}$ and all $\varepsilon > 0$, where the error term can be improved to $O(T^{4/3})$ for $n = 2$. The corresponding case for singular matrices was later investigated by Katznelson, who proved in [4] that

$$N_{n,0}(T) = c_{n,0} T^{n(n-1)} \log T + O(T^{n(n-1)}).$$

See the next page for the constants $c_{n,k}$ and $c_{n,0}$.

Let $M'_{n,k}$ be the set of matrices in $M_{n,k}$ with primitive row vectors. Then $N'_{n,k}(T) = |B_T \cap M'_{n,k}|$. Wigman [8] determined the asymptotic behavior of the counting function $|G_T \cap M'_{n,0}|$, where G_T is a ball of radius T in $M_n(\mathbb{R})$, under a slightly different norm than ours. The results can be transferred to our setting, whereby we have

$$\begin{aligned}
N'_{n,0}(T) &= c'_{n,0} T^{n(n-1)} \log T + O(T^{n(n-1)}), \quad n \geq 4, \\
N'_{3,0}(T) &= c'_{3,0} T^{3(3-1)} \log T + O(T^{3(3-1)} \log\log T), \\
N'_{2,0}(T) &= c'_{2,0} T^{2(2-1)} + O(T).
\end{aligned}$$

The case $n = 2$ above is equivalent to the **primitive circle problem**, which asks how many primitive vectors there are of length at most T in $\mathbb{Z}^2$ given any (large) T.

The main result in our paper is the following asymptotic expression for the number of nonsingular matrices with primitive row vectors and fixed determinant.

Theorem 1 *Let $k \neq 0$. Then*

$$N'_{n,k}(T) = c'_{n,k} T^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon}),$$

as $T \to \infty$ for a certain constant $c'_{n,k}$ and all $\varepsilon > 0$.

Section 3 is dedicated to the proof of this theorem.

The constant in Theorem 1 can be written as

$$c'_{n,k} = \frac{C_1}{|k|^{n-1}} \sum_{d_1 \cdots d_n = |k|} \prod_{i=1}^{n} \sum_{g | d_i} \mu(g) \left(\frac{d_i}{g}\right)^{i-1},$$

for $k \neq 0$, which may be compared to the constants obtained from [2], [4] and [8], namely

$$c_{n,k} = \frac{C_1}{|k|^{n-1}} \sum_{d_1 \cdots d_n = |k|} \prod_{i=1}^{n} d_i^{i-1}$$

$$c_{n,0} = C_0 \frac{n-1}{\zeta(n)}$$

$$c'_{n,0} = \begin{cases} C_0 \dfrac{n-1}{\zeta(n-1)^n \zeta(n)} & (n \geq 3) \\ \dfrac{\pi T^2}{\zeta(2)} & (n = 2) \end{cases}$$

where ζ is the Riemann zeta function, μ is the Möbius function, and C_0 and C_1 are constants defined as follows (these depend on n, but we will always regard n as fixed). Let ν be the normalized Haar measure on $\mathrm{SL}_n(\mathbb{R})$. The measure w below is obtained by averaging the $n(n-1)$-dimensional volume of $E \cap A_u$ over all classes $A_u := \{A \in M_n(\mathbb{R}) : Au = 0\}$ for nonzero $u \in \mathbb{R}^n$. In Appendix C we give a precise definition of w and calculate $w(B_1)$.

Write V_n for the volume of the unit ball in $\mathbb{R}^n$ and S_{n-1} for the surface area of the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$. Then

$$C_0 := w(B_1) = \frac{V_{n(n-1)} S_{n-1}}{2} = \frac{\pi^{n^2/2}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n(n-1)}{2} + 1\right)},$$

$$C_1 := \lim_{T \to \infty} \frac{\nu(B_T \cap \mathrm{SL}_n(\mathbb{R}))}{T^{n(n-1)}} = \frac{V_{n(n-1)} S_{n-1}}{2\zeta(2) \cdots \zeta(n)} = \frac{C_0}{\zeta(2) \cdots \zeta(n)}.$$

1.1 Density

It will be interesting to compare the growth of $N'_{n,k}$ to that of $N_{n,k}$. We define the **density** of matrices with primitive rows in the space $M_{n,k}$ to be

$$D_n(k) := \lim_{T \to \infty} \frac{N'_{n,k}(T)}{N_{n,k}(T)} = \frac{c'_{n,k}}{c_{n,k}}.$$

The asymptotics of $N_{n,0}$ and $N'_{n,0}$ are known from [4] and [8], and taking their ratio, we see that

$$D_n(0) = \frac{1}{\zeta(n-1)^n}$$

for $n \geq 3$. We will be interested in the value of $D_n(k)$ for large n and large k. The limit of $D_n(k)$ as $k \to \infty$ does not exist, but it does exist for particular sequences of k.

We say that a sequence of integers is **totally divisible** if its terms are eventually divisible by all positive integers smaller than m, for any m. We say that a sequence of integers is **rough** if its terms eventually have no divisors smaller than m (except

for 1), for any m. An equivalent formulation is that a sequence $(k_1, k_2, \ldots)$ is totally divisible if and only if $|k_i|_p \to 0$ as $i \to \infty$ for all primes p, and $(k_1, k_2, \ldots)$ is rough if and only if $|k_i|_p \to 1$ as $i \to \infty$ for all primes p, where $|m|_p$ denotes the p-adic norm of m.

We state our main results about the density D_n. We prove these in sect. 4.

Theorem 2 *Let $n \geq 3$ be fixed. Then D_n is a multiplicative function, and $D_n(p^m)$ is strictly decreasing as a function of m for any prime p. We have*

$$\frac{1}{\zeta(n-1)^n} = D_n(0) < D_n(k) < D_n(1) = 1$$

for all $k \neq 0, 1$. Now let $k_1, k_2, \ldots$ be a sequence of integers. Then

$$D_n(k_i) \to 1$$

if and only if $(k_1, k_2, \ldots)$ is a rough sequence, and

$$D_n(k_i) \to \frac{1}{\zeta(n-1)^n}$$

if and only if $(k_1, k_2, \ldots)$ is a totally divisible sequence. Moreover, $D_n(k) \to 1$ uniformly as $n \to \infty$.

Remark 3 Given an integer sequence $k_1, k_2, \ldots$, write $k_i = \pm \prod_p p^{m_p(i)}$ for the prime decomposition of k_i for nonzero k_i, and otherwise formally define $m_p(i) = \infty$ for all p if k_i is zero. For $n \geq 3$, it follows from Theorem 2 that the limit $\lim_{i\to\infty} D_n(k_i)$ exists and is equal to $\prod_p \lim_{i\to\infty} D_n(p^{m_p(i)})$ where the product extends over all primes p, whenever every sequence of prime exponents $(m_p(1), m_p(2), \ldots)$ is either eventually constant or tends to ∞.

We prove Theorem 2 for nonzero k_i, but it is interesting that this formulation holds for $k = 0$ also. The case of $k = 0$ was proved by Wigman [8], where he found that $D_n(0)$ equals $1/\zeta(n-1)^n$. We remark that Theorem 2 implies that

$$D_n(k_i) \to D_n(0)$$

if and only if $(k_1, k_2, \ldots)$ is totally divisible, for any fixed $n \geq 3$.

For completeness, let us state what happens in the rather different case $n = 2$.

Proposition 4 *Let $n = 2$. Then D_n is a multiplicative function, and $D_n(p^m)$ is strictly decreasing as a function of m for any prime p. We have*

$$D_2(k_i) \to 0$$

if and only if $\lim_{i\to\infty} \sum_{p|k_i} 1/p \to \infty$. Moreover,

$$D_2(k_i) \to 1$$

if and only if $\lim_{i\to\infty} \sum_{p|k_i} 1/p \to 0$. *The sums are taken over all primes* p *which divide* k_i.

In light of Remark 3, one may ask which values in the interval $[D_n(0), 1]$ can be obtained as partial limits of the function D_n. In this direction, we have the following result.

Proposition 5 *For* $n \geq 4$, *the set of values of* $D_n(k)$ *as* k *ranges over* $\mathbb{Z}$ *is not dense in in the interval* $[D_n(0), 1]$. *For* $n = 2$, *the set of values of* $D_2(k)$ *as* k *ranges over* $\mathbb{Z}$ *is dense in the interval* $[0, 1]$.

Section 4 is dedicated to the proofs of Theorem 2, Propositions 4 and 5.

1.2 Proof outline of Theorem 1

Our proof of Theorem 1 uses essentially the same approach as [2]. The set $M'_{n,k}$ is partitioned into a finite number of orbits $A\,\mathrm{SL}_n(\mathbb{Z})$, where $A \in M_{n,k}$ are matrices in Hermite normal form with primitive row vectors. We count the matrices in each orbit separately. The number of matrices in each orbit scales as a fraction $1/k^{n-1}$ of the number of matrices in $\mathrm{SL}_n(\mathbb{Z})$. We can view $\mathrm{SL}_n(\mathbb{Z})$ as a lattice in the space $\mathrm{SL}_n(\mathbb{R})$, and the problem is reduced to a lattice point counting problem. The lattice points inside the ball B_T are counted by evaluating the normalized Haar measure of $B_T \cap \mathrm{SL}_n(\mathbb{R})$.

2 Preliminaries

The **Riemann zeta function** ζ is given by

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - 1/p^s}$$

for $\mathrm{Re}\ s > 1$, where we use the convention that when an index p is used in a sum or product, it ranges over the set of primes.

The **Möbius function** μ is defined by $\mu(k) := (-1)^m$ if k is a product of m distinct prime factors (that is, k is **square-free**), and $\mu(k) := 0$ otherwise. We note that μ is a **multiplicative function**, that is, a function $f : \mathbb{N}^* \to \mathbb{C}$ defined on the positive integers such that $f(ab) = f(a)f(b)$ for all coprime a, b.

We will use the fact that $\mathrm{SL}_n(\mathbb{R}) = M_{n,1}$ has a normalized Haar measure ν which is bi-invariant (see [6]).

2.1 Lattice point counting

Let G be a topological group with a normalized Haar measure ν_G and a lattice $\Gamma \subseteq G$, and let G_T be an increasing family of bounded subsets of G for all $T \geq 1$. Under certain conditions (see for instance [3]), we have

$$|G_T \cap \Gamma| \sim \nu_G(G_T \cap G),$$

where we by $f(T) \sim g(T)$ mean that $f(T)/g(T) \to 1$ as $T \to \infty$. In this paper, we are interested in the lattice $\mathrm{SL}_n(\mathbb{Z})$ inside $\mathrm{SL}_n(\mathbb{R})$, and the following result will be crucial.

Theorem 6 ([2, Theorem 1.10]). *Let B_T be the ball of radius T in the space $M_n(\mathbb{R})$ of real $n \times n$-matrices under the Euclidean norm $\|A\| = \sqrt{\mathrm{tr}(A^t A)}$. Let ν be the normalized Haar measure of $\mathrm{SL}_n(\mathbb{R})$. Then*

$$|B_T \cap \mathrm{SL}_n(\mathbb{Z})| = \nu(B_T \cap \mathrm{SL}_n(\mathbb{R})) + O_\varepsilon(T^{n(n-1)-1/(n+1)+\varepsilon})$$

for all $\varepsilon > 0$, and the main term is given by

$$|B_T \cap \mathrm{SL}_n(\mathbb{Z})| \sim C_1 T^{n(n-1)}, \quad C_1 = \frac{1}{\zeta(2)\cdots\zeta(n)} \frac{\pi^{n^2/2}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n(n-1)}{2}+1\right)}.$$

In fact, a slightly more general statement is true. We can replace the balls B_T in Theorem 6 with balls under any norm on $M_n(\mathbb{R})$, and the asymptotics will still hold, save for a slighty worse exponent in the error term.

Theorem 7 ([3, Corollary 2.3]). *Let $\|\cdot\|'$ be any norm on the vector space $M_n(\mathbb{R})$, and let G_T be the ball of radius T in $M_n(\mathbb{R})$ under this norm. Let ν be the normalized Haar measure of $\mathrm{SL}_n(\mathbb{R})$. Then*

$$|G_T \cap \mathrm{SL}_n(\mathbb{Z})| = \nu(G_T \cap \mathrm{SL}_n(\mathbb{R})) + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon})$$

for all $\varepsilon > 0$.

We will be interested in the following particular case of Theorem 7. Let $A \in M_{n,k}$. Then $\|X\|' := \|A^{-1}X\|$ defines a norm on $M_n(\mathbb{R})$, and the ball of radius T in $M_n(\mathbb{R})$ under the norm $\|\cdot\|'$ is $A \cdot B_T$.

Corollary 8 *Let $A \in M_{n,k}$. Then*

$$|AB_T \cap \mathrm{SL}_n(\mathbb{Z})| = \nu(AB_T \cap \mathrm{SL}_n(\mathbb{R})) + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon})$$

for all $\varepsilon > 0$, using the notation from Theorem 6.

3 The number of matrices with primitive rows

In the present section, we will prove Theorem 1. We begin by noting that the common divisors of the entries of each row in an integer $n \times n$-matrix A are preserved under multiplication on the right by any matrix $X \in \mathrm{SL}_n(\mathbb{Z})$. In particular, if each row of A is primitive, then each row of AX is primitive, for any $X \in \mathrm{SL}_n(\mathbb{Z})$. So we get:

Lemma 9 *If $A \in M'_{n,k}$ then $AX \in M'_{n,k}$ for all $X \in \mathrm{SL}_n(\mathbb{Z})$. Thus $A \cdot \mathrm{SL}_n(\mathbb{Z}) \subseteq M'_{n,k}$.*

Consequently $M'_{n,k}$ may be written as a disjoint union of orbits of $\mathrm{SL}_n(\mathbb{Z})$:

$$M'_{n,k} = \bigcup_{A \in \mathcal{A}} A \,\mathrm{SL}_n(\mathbb{Z}),$$

for properly chosen subsets $\mathcal{A}$ of $M'_{n,k}$. In fact, as we will show in the following, the number of orbits is finite, and so we may take $\mathcal{A}$ to be finite.

A lower triangular integer matrix

$$C := \begin{pmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \ddots & 0 \\ \vdots & & \ddots & 0 \\ c_{n1} & \cdots & c_{n(n-1)} & c_{nn} \end{pmatrix}$$

is said to be in (lower) **Hermite normal form** if $0 < c_{11}$ and $0 \le c_{ij} < c_{ii}$ for all $j < i$. The following result is well-known.

Lemma 10 ([1, Theorem 2.4.3]). *Assume $k > 0$. Given an arbitrary matrix $A \in M_{n,k}$, the orbit $A\,\mathrm{SL}_n(\mathbb{Z})$ contains a unique matrix C in Hermite normal form.*

We may thus write

$$M'_{n,k} = \bigcup_{i=1}^{m} A_i \,\mathrm{SL}_n(\mathbb{Z}),$$

where $A_1, \ldots, A_m$ are the unique matrices in Hermite normal form with primitive row vectors and determinant k, and $m := |M'_{n,k}/\mathrm{SL}_n(\mathbb{Z})|$. By counting the number of matrices in Hermite normal form with determinant $k > 0$, we get

$$|M_{n,k}/\mathrm{SL}_n(\mathbb{Z})| = \sum_{d_1 \cdots d_n = k} d_1^0 d_2^1 \cdots d_n^{n-1},$$

where the sum ranges over all positive integer tuples $(d_1, \ldots, d_n)$ such that $d_1 \cdots d_n = k$.

Proposition 11 *Let $k > 0$. Then*

$$|M'_{n,k}/\mathrm{SL}_n(\mathbb{Z})| = \sum_{d_1 \cdots d_n = k} \prod_{i=1}^{n} \sum_{g \mid d_i} \mu(g) \left(\frac{d_i}{g}\right)^{i-1}$$

where the first sum ranges over all positive integer tuples $(d_1, \ldots, d_n)$ such that $d_1 \cdots d_n = k$.

Proof We want to count those matrices in Hermite normal form which are in $M'_{n,k}$, that is, $n \times n$-matrices in Hermite normal form with determinant k and all rows primitive. The number of such matrices is

$$\left|M'_{n,k}/\operatorname{SL}_n(\mathbb{Z})\right| = \sum_{d_1 \cdots d_n = k} \prod_{i=1}^{n} v_i(d_i),$$

where $v_i(d)$ is the number of primitive vectors $(x_1, \ldots, x_{i-1}, d)$ such that $0 \le x_1, \ldots, x_{i-1} < d$. There is a bijective correspondence between the primitive vectors $(x_1, \ldots, x_{i-1}, d)$ and the vectors $y = (y_1, \ldots, y_{i-1})$ such that $1 \le y_1, \ldots, y_{i-1} \le d$ and $\gcd(y)$ is coprime to d. Let $d = p_1^{a_1} \cdots p_j^{a_j}$ be the prime factorization of d. The number of vectors y which are divisible by some set of primes $P \subseteq \{p_1, \ldots, p_j\}$ is

$$\left(\frac{d}{\prod_{p \in P} p}\right)^{i-1},$$

so by the principle of inclusion/exclusion (see [7]), we have

$$\begin{aligned} v_i(d) &= \sum_{P \subseteq \{p_1, \ldots, p_j\}} (-1)^{|P|} \left(\frac{d}{\prod_{p \in P} p}\right)^{i-1} \\ &= \sum_{g \mid p_1 \cdots p_j} \mu(g) \left(\frac{d}{g}\right)^{i-1} = \sum_{g \mid d} \mu(g) \left(\frac{d}{g}\right)^{i-1}. \end{aligned}$$

□

We are now ready to derive the asymptotics of $N'_{n,k}(T)$.

Proof of Theorem 1 Let us write $A_1, \ldots, A_m$ for all the $n \times n$-matrices in Hermite normal form with determinant k, where $m := |M'_{n,k}/\operatorname{SL}_n(\mathbb{Z})|$, and let $1 \le i \le m$. Then

$$|B_T \cap A_i \operatorname{SL}_n(\mathbb{Z})| = \left|A_i (A_i^{-1} B_T \cap \operatorname{SL}_n(\mathbb{Z}))\right| = \left|A_i^{-1} B_T \cap \operatorname{SL}_n(\mathbb{Z})\right|,$$

which by Corollary 8 is equal to

$$\nu(A_i^{-1} B_T \cap \operatorname{SL}_n(\mathbb{R})) + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon})$$

for any $\varepsilon > 0$. Since $A_i/k^{1/n} \in \operatorname{SL}_n(\mathbb{R})$, we get by the invariance of the measure ν that

$$\begin{aligned} \nu(A_i^{-1} B_T \cap \operatorname{SL}_n(\mathbb{R})) &= \nu\left(\frac{A_i}{k^{1/n}} \left(A_i^{-1} B_T \cap \operatorname{SL}_n(\mathbb{R})\right)\right) \\ &= \nu\left(k^{-1/n} B_T \cap \frac{A_i}{k^{1/n}} \operatorname{SL}_n(\mathbb{R})\right) = \nu\left(B_{T/k^{1/n}} \cap \operatorname{SL}_n(\mathbb{R})\right). \end{aligned}$$

By Theorem 6, the last expression is equal to

$$C_1(T/k^{1/n})^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon}),$$

and thus

$$|B_T \cap A_i \operatorname{SL}_n(\mathbb{Z})| = \frac{C_1}{k^{n-1}} T^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon}). \tag{1}$$

Now,

$$N'_{n,k}(T) = \left|B_T \cap M'_{n,k}\right| = \left|B_T \cap \bigcup_{i=1}^{m} A_i \operatorname{SL}_n(\mathbb{Z})\right| = \sum_{i=1}^{m} |B_T \cap A_i \operatorname{SL}_n(\mathbb{Z})|,$$

so applying (1) we get

$$\begin{aligned} N'_{n,k}(T) &= \sum_{i=1}^{m} \frac{C_1}{k^{n-1}} T^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon}) \\ &= \left|M'_{n,k}/\operatorname{SL}_n(\mathbb{Z})\right| \frac{C_1}{k^{n-1}} T^{n(n-1)} + O_\varepsilon(T^{n(n-1)-1/(2n)+\varepsilon}), \end{aligned}$$

and we need only apply Proposition 11 to get an explicit constant for the main term. This concludes the proof. □

4 Density of matrices with primitive rows

Set

$$a_n(k) := |M_{n,k}/\operatorname{SL}_n(\mathbb{Z})| = \sum_{d_1 \cdots d_n = k} d_1^0 \cdots d_n^{n-1}, \tag{2}$$

$$a'_n(k) := |M'_{n,k}/\operatorname{SL}_n(\mathbb{Z})| = \sum_{d_1 \cdots d_n = k} \prod_{i=1}^{n} \sum_{g \mid d_i} \mu(g) \left(\frac{d_i}{g}\right)^{i-1}. \tag{3}$$

We would like to calculate the density of matrices with primitive rows in $M_{n,k}$ for $k \neq 0$, that is, the quantity

$$D_n(k) = \lim_{T\to\infty} \frac{N'_{n,k}(T)}{N_{n,k}(T)} = \frac{c'_{n,k}}{c_{n,k}} = \frac{|M'_{n,k}/\operatorname{SL}_n(\mathbb{Z})|}{|M_{n,k}/\operatorname{SL}_n(\mathbb{Z})|} = \frac{a'_n(k)}{a_n(k)}.$$

We will prove in Sect. 4.1 that a_n, a'_n and D_n are multiplicative functions, and therefore we need only understand their behavior for prime powers $k = p^m$. We will now prove a sequence of lemmas which we will finally use in Sect. 4.2 to prove Theorem 2.

Lemma 12 *The functions a'_n and a_n are connected via the identity*

$$a'_n(p^m) = \sum_{i=0}^{m} (-1)^i \binom{n}{i} a_n(p^{m-i})$$

for primes p and $m \geq 0$.

Proof $a_n(p^m)$ counts the number of $n \times n$-matrices in Hermite normal form with determinant p^m, whereas $a'_n(p^m)$ counts the number of such with primitive rows. If A is a matrix in $M_{n,k} \setminus M'_{n,k}$, then some set of rows, indexed by $S \subseteq [n] := \{1, \ldots, n\}$ (where $|S| \leq m$), are divisible by p. The number of such matrices is $a_n(p^{m-|S|})$, and thus by the inclusion/exclusion principle,

$$a'_n(p^m) = \sum_{\substack{S \subseteq [n] \\ |S| \leq m}} (-1)^{|S|} a_n(p^{m-|S|}) = \sum_{i=0}^{m} (-1)^i \binom{n}{i} a_n(p^{m-i}). \qquad \square$$

Lemma 13 *For any prime p and $m \geq 1$, the following recursion holds:*

$$a_n(p^m) = p^{n-1} a_n(p^{m-1}) + a_{n-1}(p^m),$$

or equivalently,

$$a_n(p^{m-1}) = \frac{a_n(p^m) - a_{n-1}(p^m)}{p^{n-1}}.$$

Proof We split the sum

$$a_n(p^m) = \sum_{d_1 \cdots d_n = p^m} d_1^0 \cdots d_n^{n-1}$$

into two parts, one part where d_n is divisible by p, and another part where it is not (so that $d_n = 1$). The terms corresponding to $d_n = 1$ sum to $a_{n-1}(p^m)$. Where d_n is divisible by p, we can write $d_n =: pe_n$ for some e_n. Let $e_i := d_i$ for all $i < n$. Thus,

$$\sum_{\substack{d_1 \cdots d_n = p^m \\ p | d_n}} d_1^0 \cdots d_n^{n-1} = \sum_{e_1 \cdots e_n = p^{m-1}} e_1^0 \cdots (pe_n)^{n-1} = p^{n-1} a_n(p^{m-1}).$$

Adding the two parts gives us $a_n(p^m) = p^{n-1} a_n(p^{m-1}) + a_{n-1}(p^m)$, from which the second claim in the lemma follows by rearrangement. $\square$

Lemma 14 *Let n and p be fixed, where $n \geq 3$ and p is a prime. Then*

$$D_n(p^m) \to \left(1 - \frac{1}{p^{n-1}}\right)^n$$

as $m \to \infty$.

Proof We apply the simple upper bound

$$a_{n-1}(p^m) = \sum_{d_1 \cdots d_{n-1} = p^m} d_1^0 \cdots d_n^{n-2} \le \sum_{d_1 \cdots d_{n-1} = p^m} (p^m)^{n-2} = (m+1)^{n-1}(p^m)^{n-2}$$

to the expression for $a_n(p^{m-1})$ in Lemma 13:

$$\begin{aligned} a_n(p^{m-1}) &= \frac{1}{p^{n-1}}(a_n(p^m) - a_{n-1}(p^m)) \\ &= \frac{1}{p^{n-1}} a_n(p^m) + O((p^m)^{n-2}(m+1)^{n-1}). \end{aligned}$$

Repeated application (at most n times) of this formula yields the asymptotics

$$a_n(p^{m-i}) = \frac{1}{(p^{n-1})^i} a_n(p^m) + O((p^m)^{n-2}(m+1)^{n-1})$$

for $1 \le i \le n$.

Now let $m \to \infty$, so that we may assume m to be larger than n. The sum in Lemma 12 then extends up to $i = n$ (because the factors $\binom{n}{i}$ vanish for larger i), so

$$\begin{aligned} a'_n(p^m) &= \sum_{i=0}^{n} (-1)^i \binom{n}{i} a_n(p^{m-i}) \\ &= \sum_{i=0}^{n} (-1)^i \binom{n}{i} \frac{1}{(p^{n-1})^i} a_n(p^m) + O((p^m)^{n-2}(m+1)^{n-1}). \end{aligned}$$

We divide by $a_n(p^m)$ on both sides and use the fact that $a_n(p^m) \ge (p^m)^{n-1}$, so that

$$\begin{aligned} D_n(p^m) &= \sum_{i=0}^{n} (-1)^i \binom{n}{i} \frac{1}{(p^{n-1})^i} + O\left(\frac{(p^m)^{n-2}(m+1)^{n-1}}{(p^m)^{n-1}}\right) \\ &= \sum_{i=0}^{n} \binom{n}{i} \left(\frac{-1}{p^{n-1}}\right)^i + O\left(\frac{(m+1)^{n-1}}{p^m}\right) \\ &= \left(1 - \frac{1}{p^{n-1}}\right)^n + O\left(\frac{(m+1)^{n-1}}{p^m}\right). \end{aligned}$$

As $m \to \infty$, the second term on the right vanishes. □

4.1 Multiplicativity and monotonicity of the density function

In this section we will prove the following proposition.

Proposition 15 *The function D_n is multiplicative, and $D_n(p^m)$ is strictly decreasing as a function of m for any fixed prime p and dimension $n \geq 2$.*

We may rewrite (2) as

$$a_n = (\cdot)^{n-1} * \cdots * (\cdot)^0$$

where $(\cdot)^i$ is the function $x \mapsto x^i$ and $*$ denotes the Dirichlet convolution. Similarly, we may rewrite (3) as

$$a_n' = (\mu * (\cdot)^{n-1}) * \cdots * (\mu * (\cdot)^0), \tag{4}$$

so by the commutativity and associativity of the Dirichlet convolution we have

$$a_n' = \mu^{*n} * a_n,$$

where μ^{*n} denotes the convolution of μ with itself n times (so that $\mu^{*1} = \mu$). Since the Dirichlet inverse of μ is the constant function 1, we have also the relation

$$a_n = 1^{*n} * a_n'.$$

As μ and $(\cdot)^i$ are multiplicative functions, it follows that a_n, a_n' and D_n are multiplicative as well.

Now, we want to show that $D_n(p^m) = a_n'(p^m)/a_n(p^m)$ is strictly decreasing as a function of m, for fixed $n \geq 2$ and primes p, or equivalently that

$$\frac{a_n'(p^m)}{a_n(p^m)} > \frac{a_n'(p^{m+1})}{a_n(p^{m+1})} \tag{5}$$

for all $m \geq 0$. The inequality (5) is equivalent to

$$\frac{a_n'(p^m)}{(1^{*n} * a_n')(p^m)} > \frac{a_n'(p^{m+1})}{(1^{*n} * a_n')(p^{m+1})}$$

for all $m \geq 0$, which is equivalent to

$$\frac{a_n'(p^m)}{\sum_{i=0}^{m} 1^{*n}(p^i) a_n'(p^{m-i})} > \frac{a_n'(p^{m+1})}{\sum_{i=0}^{m+1} 1^{*n}(p^i) a_n'(p^{m+1-i})},$$

or, after taking the reciprocal of both sides,

$$\sum_{i=0}^{m} 1^{*n}(p^i) \frac{a_n'(p^{m-i})}{a_n'(p^m)} < \sum_{i=0}^{m+1} 1^{*n}(p^i) \frac{a_n'(p^{m+1-i})}{a_n'(p^{m+1})}.$$

Since the last term ($i = m + 1$) on the right hand side is positive, this inequality holds if

$$\frac{a'_n(p^{m-i})}{a'_n(p^m)} \leq \frac{a'_n(p^{m+1-i})}{a'_n(p^{m+1})}$$

for all $i \leq m$. We can rearrange this inequality as

$$\frac{a'_n(p^{m+1})}{a'_n(p^m)} \leq \frac{a'_n(p^{m+1-i})}{a'_n(p^{m-i})},$$

which states that $a'_n(p^{m+1})/a'_n(p^m)$ is a non-increasing function of m, for fixed $n \geq 2$ and p prime. We will therefore be done if we can prove that

$$a'_n(p^{m+1})a'_n(p^{m+1}) \geq a'_n(p^m)a'_n(p^{m+2}) \tag{6}$$

for all $m \geq 0$, or equivalently, that the function $m \mapsto a'_n(p^m)$ is logarithmically concave:

We say that a sequence $u : \mathbb{N}_0 \to \mathbb{R}$ is **logarithmically concave** if

$$u_r^2 - u_{r-1}u_{r+1} \geq 0$$

for all $r \geq 1$. We note that a sequence u of positive real numbers is logarithmically concave if and only if $u_1/u_0 \geq u_2/u_1 \geq u_3/u_2 \geq \cdots$, that is, if and only if $(u_1/u_0, u_2/u_1, u_3/u_2, \ldots)$ is a non-increasing sequence. Also note that if u is positive and logarithmically concave, then the inequality $u_{i+1}/u_i \geq u_{j+1}/u_j$ implies the inequality $u_{i+1}u_j - u_{j+1}u_i \geq 0$ for all indices $i < j$.

Let $\star$ denote the discrete convolution, so that $(u \star v)_r = \sum_{j=0}^{r} u_{r-j}v_j$ for all $r \geq 0$ given any sequences $u, v : \mathbb{N}_0 \to \mathbb{R}$. We will need the following fact, which follows from the proof of Theorem 1 in [5].

Theorem 16 ([5, Theorem 1]). *Let $u, v : \mathbb{N}_0 \to \mathbb{R}$ be sequences such that $u_0 = v_0 = 1$, and let $w = u \star v$. Then we may write $w_r^2 - w_{r-1}w_{r+1} = \mathrm{I} + \mathrm{II} + \mathrm{III}$, where*

$$\mathrm{I} = \sum_{0 \leq i < j \leq r-1} (v_j v_{i+1} - v_{j+1}v_i)(u_{r-j}u_{r-i-1} - u_{r-1-j}u_{r-i}),$$

$$\mathrm{II} = \sum_{j=0}^{r-1} v_j(u_{r-j}u_r - u_{r-1-j}u_{r+1}),$$

$$\mathrm{III} = v_r u_r + \sum_{j=0}^{r-1} u_j(v_r v_{r-j} - v_{r+1}v_{r-1-j}),$$

for all $r \geq 1$. In particular, if u, v are positive and logarithmically concave sequences, then so is w, since all factors in the sums in I, II, III *are non-negative for such u, v.*

Fix n and p. Then since $(\mu * (\cdot)^i)(p^m) = \sum_{r=0}^{m} \mu(p^{m-r}) p^{ri} = (M \star P_i)(m)$ where M is the sequence $(1, -1, 0, 0, 0, \ldots)$ and where P_i is the sequence $(1, p^i, p^{2i}, p^{3i}, \ldots)$, Eq. (4) implies that the function $m \mapsto a_n'(p^m)$ can be written as

$$(M \star P_{n-1}) \star \cdots \star (M \star P_0).$$

Lemma 17 *Let $0 \le i < j$. Then $(M \star P_i) \star (M \star P_j)$ is positive and logarithmically concave if and only if $i > 0$.*

Proof Write $u := M \star P_i$ and $v := M \star P_j$ where $i < j$. We have $u_0 = 1$ and $u_r = p^{ir} - p^{i(r-1)}$ for all $r \ge 1$. Thus $u_1 u_r - u_0 u_{r+1} = (p^i - 1)(p^{ir} - p^{i(r-1)}) - (p^{i(r+1)} - p^{ir}) = p^{i(r-1)} - p^{ir} = -u_r$ for all $r \ge 1$, and $u_s u_r - u_{s-1} u_{r+1} = 0$ when $s \ge 2, r \ge 1$ or $s = 1, r = 0$. Likewise $v_s v_r - v_{s-1} v_{r+1}$ is $-v_r$ if $s = 1, r \ge 1$, and 0 otherwise.

Let $w := u \star v = (M \star P_i) \star (M \star P_j)$. By Theorem 16 we can write $w_r^2 - w_{r-1} w_{r+1} = \mathrm{I} + \mathrm{II} + \mathrm{III}$, where

$$\begin{aligned} \mathrm{I} &= (-v_{r-1})(-u_{r-1}), \\ \mathrm{II} &= v_{r-1}(-u_r), \\ \mathrm{III} &= v_r u_r + u_{r-1}(-v_r), \end{aligned}$$

for all $r \ge 1$, and therefore

$$\begin{aligned} w_r^2 - w_{r-1} w_{r+1} &= u_{r-1} v_{r-1} + u_r v_r - u_r v_{r-1} - u_{r-1} v_r \\ &= (u_r - u_{r-1})(v_r - v_{r-1}). \end{aligned}$$

Thus, since $(u_0, u_1, \ldots)$ is a non-decreasing sequence for $i > 0$, and likewise $(v_0, v_1, \ldots)$ is a non-decreasing sequence for $j > 0$, we get $w_r^2 - w_{r-1} w_{r+1} \ge 0$ for all $r \ge 1$ for $i > 0$. Also, the sequence w is positive for $i, j > 0$ since it is then the convolution of two positive sequences. If $i = 0$, then the inequality $w_r^2 - w_{r-1} w_{r+1} \ge 0$ fails for $r = 1$ since then $u_1 - u_0 = (p^0 - 1) - 1 < 0$ and $v_1 - v_0 = (p^j - 1) - 1 > 0$. □

We will prove Proposition 15 by induction on n. The base case is the following proposition, which we will prove in Appendix B.

Proposition 18 *For $n = 4, 5$ and any fixed prime p, the function $m \mapsto a_n'(p^m)$ is logarithmically concave.*

It happens that $a_n'(p^m)$, as a function of m, is not logarithmically concave for $n = 2$ or $n = 3$ for all p [it fails the inequality (6) for $r = 1$ when $p = 2$], so we will also need the following proposition, which we prove in Appendix A.

Proposition 19 *For $n = 2, 3$ and any fixed prime p, the function $m \mapsto D_n(p^m)$ is strictly decreasing.*

The proofs of Propositions 19 and 18 consist of explicitly evaluating $a_n(p^m)$ and $a_n'(p^m)$ for the values of n in question, both of which are polynomials in p with exponents in m, and verifying Eqs. (5) and (6), respectively.

Proof of Proposition 15 By Propositions 18 and 19, it suffices to consider $n > 5$. By Proposition 18 and Theorem 16, it follows that $A'_n(m) := a'_n(p^m)$ is logarithmically concave for all $n > 5$ and any p, since for any even $n > 5$, we can write

$$A'_n = A'_4 \star [(M \star P_4) \star (M \star P_5)] \star \cdots \star [(M \star P_{n-2}) \star (M \star P_{n-1})],$$

and for any odd $n > 5$, we can write

$$A'_n = A'_5 \star [(M \star P_5) \star (M \star P_6)] \star \cdots \star [(M \star P_{n-2}) \star (M \star P_{n-1})],$$

and in both cases we have written A'_n as the convolution of positive and logarithmically concave sequences, by Lemma 17. We have thus proven the inequality (6), and this concludes the proof of Proposition 15. □

4.2 Asymptotics of the density function

In this section we prove Theorem 2 and thus derive the asymptotics of $D_n(k)$. Fix $n \geq 3$. For any nonzero integer k_i, write $k_i = \prod_p p^{m_p(i)}$ as a product of prime powers, where all but finitely many of the exponents $m_p(i)$ are zero. Then since D_n is multiplicative, we have

$$D_n(k_i) = \prod_p D_n(p^{m_p(i)}).$$

Now, by Lemma 14 and Proposition 15, we get

$$1 \geq \prod_p D_n(p^{m_p(i)}) > \prod_p \left(1 - \frac{1}{p^{n-1}}\right)^n = \frac{1}{\zeta(n-1)^n} > 0,$$

so it follows by dominated convergence that

$$\lim_{i\to\infty} \prod_p D_n(p^{m_p(i)}) = \prod_p \lim_{i\to\infty} D_n(p^{m_p(i)}), \tag{7}$$

whenever $(k_1, k_2, \ldots)$ is a sequence of nonzero integers such that the limit $\lim_{i\to\infty} D_n(p^{m_p(i)})$ exists for each prime p.

Let $(k_1, k_2, \ldots)$ be a sequence of nonzero integers. It now follows from (7), Proposition 15 and the fact that $D_n(1) = 1$, that

$$D_n(k_i) \to 1$$

if and only if $m_p(i) \to 0$ as $i \to \infty$ for all p, that is, if and only if $(k_1, k_2, \ldots)$ is a rough sequence. Likewise it follows, using Lemma 14, that

$$D_n(k_i) \to \frac{1}{\zeta(n-1)^n}$$

if and only if $m_p(i) \to \infty$ for all p, that is, if and only if $(k_1, k_2, \ldots)$ is a totally divisible sequence. Since $D_n(0) = 1/\zeta(n-1)^n$, we may allow the elements of the sequence $(k_1, k_2, \ldots)$ to also assume the value 0.

Finally, it follows that $D_n(k) \to 1$ as $n \to \infty$ uniformly with respect to k since

$$D_n(k) \geq \frac{1}{\zeta(n-1)^n} \to 1$$

as $n \to \infty$ because $\zeta(n-1) = 1 + O(2^{-n})$ for $n \geq 3$. We have thus proved all parts of Theorem 2. □

We conclude this section by proving Proposition 4, which tells us the asymptotics of $D_2(k)$ for $n = 2$.

Proof of Proposition 4 If $m = 0$, we have $D_2(p^m) = 1$. Assume $m > 0$. The 2×2-matrices in Hermite normal form with determinant p^m and primitive rows are of the form $\begin{pmatrix} 1 & 0 \\ x & p^m \end{pmatrix}$ where $0 \leq x < p^m$, $p \nmid x$. Thus $a_2'(p^m) = p^m(1 - 1/p)$. Moreover,

$$a_2(p^m) = \sum_{d_1 d_2 = p^m} d_2 = \sum_{i+j=m} p^i = \sum_{i=0}^{m} p^i = \frac{p^{m+1} - 1}{p - 1} = p^m \frac{1 - 1/p^{m+1}}{1 - 1/p},$$

so

$$D_2(p^m) = \frac{(1 - 1/p)^2}{1 - 1/p^{m+1}} \tag{8}$$

for all $m \geq 1$. We see immediately that $D_2(p^m)$ is strictly decreasing as a function of m, for any fixed p. Therefore

$$\left(1 - \frac{1}{p}\right)^2 \leq D_2(p^m) \leq 1 - \frac{1}{p}.$$

Since D_2 is multiplicative, we get

$$\left[\prod_{p|k} \left(1 - \frac{1}{p}\right)\right]^2 \leq D_2(k) \leq \prod_{p|k} \left(1 - \frac{1}{p}\right).$$

The left and right sides both tend to 0 if and only if $\lim_{i\to\infty} \sum_{p|k_i} 1/p \to \infty$, and they both converge to 1 if and only if $\lim_{i\to\infty} \sum_{p|k_i} 1/p \to 0$. □

4.3 The image of the density function

Proof of Proposition 5 for $n \geq 4$. By Proposition 15, the function D_n is multiplicative, and $D_n(p^m)$ is strictly decreasing as a function of m for any fixed p, n. Thus we get $D_n(k) \leq D_n(2)$ whenever k is divisible by 2. When k is not divisible by 2, we get

$$D_n(k) \geq \prod_{p\geq 3} \lim_{m\to\infty} D_n(p^m) = \prod_{p\geq 3} \left(1 - 1/p^{n-1}\right)^n$$
$$= \frac{1}{\left(1-1/2^{n-1}\right)^n} \prod_p \left(1-1/p^{n-1}\right)^n = \frac{1}{\left(1-1/2^{n-1}\right)^n} \frac{1}{\zeta(n-1)^n}.$$

by Lemma 14. We will show that this value is larger than $D_n(2)$, which will prove that the image of $D_n : \mathbb{Z} \to \mathbb{R}$ is not dense in $[D_n(0), 1]$. By Eq. (2) we have $a_n(2) = \sum_{i=1}^n 2^{i-1} = 2^n - 1$ and by Lemma 12 we have $a_n'(2) = a_n(2) - n a_n(1) = (2^n - 1) - n$, so $D_n(2) = 1 - n/(2^n - 1)$.

Thus it suffices to prove

$$\frac{1}{\left(1-1/2^{n-1}\right)^n} \frac{1}{\zeta(n-1)^n} > 1 - \frac{n}{2^n - 1}. \tag{9}$$

This inequality can be verified numerically for $n = 4, 5$. Let us now assume $n \geq 6$. The inequality (9) is is equivalent to

$$-\log(1 - n/(2^n - 1)) - n \log\left(1 - 1/2^{n-1}\right) > n \log \zeta(n-1).$$

By Taylor expansion, the first term on the left hand side is $> n/(2^n - 1) > n/2^n$, and the second term on the left hand side is $> n/2^{n-1}$. Thus the inequality above follows from $1/2^n + 1/2^{n-1} \geq \log \zeta(n-1)$, or equivalently $e^{3/2^n} \geq \zeta(n-1)$. We bound the left hand side from below by $1 + 3/2^n$, and we bound the right hand side from above by $1 + 1/2^{n-1} + \int_2^\infty \frac{dx}{x^{n-1}}$. Thus the inequality follows from $1 + 3/2^n \geq 1 + 1/2^{n-1} + 1/((n-2)2^{n-2})$ or equivalently $3 \geq 2 + 4/(n-2)$, which is true for all $n \geq 6$. □

Proof of Proposition 5 for $n = 2$. It suffices to show that the set of values of $-\log(D_2(k))$ as k ranges over positive square-free integers is dense in $[0, \infty)$. By the identity (8) we have

$$D_2(p) = \frac{(1-1/p)^2}{1-1/p^2} = \frac{1-1/p}{1+1/p} = \frac{p-1}{p+1} = 1 - \frac{2}{p+1}.$$

Let $k > 0$ be squarefree, and let P_0 be the set of primes dividing k. Then

$$-\log D_2(k) = -\log \prod_{p\in P_0} D_2(p) = \sum_{p\in P_0} \left(-\log\left(1 - \frac{2}{p+1}\right)\right).$$

The terms $d_p := -\log(1 - \frac{2}{p+1})$ are positive, decreasing, and tend to zero as $p \to \infty$. By Taylor expansion, the sum $\sum_p d_p$ over all primes is larger than $\sum_p \frac{2}{p+1}$, which diverges since $\sum_p 1/p$ diverges.

Now, given any $x \in [0, \infty)$ and any $\varepsilon > 0$, we can find a k such that $-\log D_2(k)$ is within a distance ε from x as follows. Let p_0 be the smallest prime such that

$d_{p_0} < \varepsilon$, and let P_0 be the smallest set of consecutive primes, starting with p_0, such that $\sum_{p\in P_0} d_p \geq x$. Then the sum $\sum_{p\in P_0} d_p = -\log D_2(k)$ is at a distance at most $d_{p_0} < \varepsilon$ from x since d_p is decreasing, where $k = \prod_{p\in P_0} p$, and we are done. □

Acknowledgments The author was partially supported by the Swedish Research Council. I would like to thank my advisor Pär Kurlberg for suggesting this problem to me and for all his help and encouragement.

Appendix A: Proof of Proposition 19

We prove Proposition 19. Recall from Eq. (8) that $D_2(p^m) = (1-1/p)^2/(1-1/p^{m+1})$ if $m > 0$, and otherwise $D_2(p^0) = 1$. Thus we see immediately that $D_2(p^m)$ is strictly decreasing as a function of m.

The case $n = 3$ remains. By Eq. (2) we get

$$\begin{aligned} a_3(p^m) &= \sum_{j_1+j_2+j_3=m} p^{j_2+2j_3} = \sum_{j_3=0}^{m} p^{2j_3} \sum_{j_2=0}^{m-j_3} p^{j_3} = \sum_{j_3=0}^{m} p^{2j_3} \frac{1-p^{m-j_3+1}}{1-p} \\ &= \frac{1}{1-p} \sum_{j_3=0}^{m} (p^{2j_3} - p^{m+j_3+1}) = \frac{1}{1-p} \left(\frac{1-p^{2(m+1)}}{1-p^2} - p^{m+1} \frac{1-p^{m+1}}{1-p} \right) \\ &= \frac{1-p^{2(m+1)} - (1+p)p^{m+1} + (1+p)p^{2(m+1)}}{(1-p)(1-p^2)} \\ &= \frac{1-p^{m+1}-p^{m+2}+p^{(m+1)+(m+2)}}{(1-p)(1-p^2)} = \frac{(p^{m+1}-1)(p^{m+2}-1)}{(p-1)(p^2-1)}. \end{aligned}$$

for all $m \geq 1$.

Let us write $I(P) := 1$ if the condition P is true, and $I(P) := 0$ if the condition P is false. By Eq. (3) we get for all $m \geq 1$ that

$$\begin{aligned} a_3'(p^m) &= \sum_{\substack{j_1,j_2,j_3\geq 0: \\ p^{j_1}p^{j_2}p^{j_3}=p^m}} \prod_{i=1}^{3} \sum_{\substack{r\geq 0: \\ p^r | p^{j_i}}} \mu(p^r)(p^{j_i-r})^{i-1} \\ &= \sum_{\substack{j_2,j_3\geq 0: \\ j_2+j_3=m}} \left(p^{j_2} - p^{j_2-1} I(j_2>0) \right) \left(p^{2j_3} - p^{2(j_3-1)} I(j_3>0) \right). \end{aligned} \tag{10}$$

We expand the product in the summand and split the sum into several geometric series which we sum individually. We get

$$\begin{aligned} &\sum_{j_2=0}^{m} \left(p^{2m-j_2} - p^{2m-j_2-1} I(j_2>0) - p^{2m-j_2-2} I(m>j_2) + p^{2m-j_2-3} I(0<j_2<m) \right) \\ &= p^{2m} \left(\frac{1-p^{-(m+1)}}{1-p^{-1}} - p^{-1} \left(\frac{1-p^{-(m+1)}}{1-p^{-1}} - 1 \right) - p^{-2} \frac{1-p^{-m}}{1-p^{-1}} + p^{-3} \left(\frac{1-p^{-m}}{1-p^{-1}} - 1 \right) \right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{p^{2m}}{1-p^{-1}} \left(1-p^{-(m+1)} - p^{-1}\left(1-p^{-(m+1)} - (1-p^{-1})\right)\right.\\
&\quad \left. -p^{-2}(1-p^{-m}) + p^{-3}\left(1-p^{-m}-(1-p^{-1})\right)\right)\\
&= \frac{p^{2m}}{1-p^{-1}} \left((1-p^{-2})^2 - p^{-m-1}(1-p^{-1})^2\right) = \frac{p^{2m}(1-p^{-2})^2 - p^{m-1}(1-p^{-1})^2}{1-p^{-1}}\\
&= (p^{2m}(p+1)^2 - p^{m+1})\frac{p-1}{p^3}.
\end{aligned}
$$

Since $D_3(1) = 1$ and $D_3(p^m) < 1$ for all $m > 0$ (the diagonal matrix with diagonal entries $1, 1, p^m$ is in Hermite normal form, but its last row is not primitive), it suffices to show that $D_3(p^m) > D_3(p^{m+1})$ for all $m \geq 1$. To see this, we note that

$$
\begin{aligned}
&\frac{p^{2m}(p+1)^2 - p^{m+1}}{(p^{m+1}-1)(p^{m+2}-1)} > \frac{p^{2m+2}(p+1)^2 - p^{m+2}}{(p^{m+2}-1)(p^{m+3}-1)}\\
&\quad \Longleftrightarrow\ p^{2m}(p+1)^2[(p^{m+3}-1) - p^2(p^{m+1}-1)] - p^{m+1}[(p^{m+3}-1)\\
&\qquad -p(p^{m+1}-1)] > 0\\
&\quad \Longleftrightarrow\ p^{2m}(p+1)^2(p^2-1) - p^{m+1}(p^{m+2}+1)(p-1) > 0\\
&\quad \Longleftrightarrow\ p^{2m}(p+1)^3 - p^{2m}(p^3 + p^{1-m}) > 0,
\end{aligned}
$$

where the last inequality is true since $(p+1)^3 > p^3 + 1 \geq p^3 + p^{1-m}$ for all $m \geq 1$ and all $p \geq 2$. This concludes the proof of Proposition 19. □

Appendix B: Proof of Proposition 18

B1: The case $n = 4$

We prove Proposition 18 for $n = 4$. By Eq. (3), we can write

$$
\begin{aligned}
a_4'(p^m) &= \sum_{\substack{j_1,j_2,j_3,j_4 \geq 0:\\ p^{j_1}p^{j_2}p^{j_3}p^{j_4} = p^m}} \prod_{i=1}^{4} \sum_{\substack{r \geq 0:\\ p^r \mid p^{j_i}}} \mu(p^r)(p^{j_i - r})^{i-1}\\
&= \sum_{j_2+j_3+j_4=m} \left(p^{j_2} - p^{j_2-1} I(j_2 > 0)\right)\left(p^{2j_3} - p^{2(j_3-1)} I(j_3 > 0)\right)\\
&\qquad \left(p^{3j_4} - p^{3(j_4-1)} I(j_4 > 0)\right),
\end{aligned}
$$

where $I(P)$ is defined as in (10). We evaluate this sum in the same way that we evaluated $a_3'(p^m)$ in Appendix A: We expand the product in the summand and eliminate the symbols $I(P)$ by splitting the sum into several geometric series over different ranges, corresponding to the conditions $j_2 > 0$, and so on, and compute each geometric series individually. We assume $m \geq 1$ to guarantee that $\sum_{j_2=1}^{m}$, for instance, is never an empty sum. Thus, by a tedious but straightforward calculation, we get

$$a_4'(p^m)=\frac{(p-1)p^{m-6}}{p+1}\Big(p^{2m}-p^{m+1}-4p^{m+2}-6p^{m+3}-4p^{m+4}-p^{m+5}+3p^{2m+1}$$
$$+6p^{2m+2}+7p^{2m+3}+6p^{2m+4}+3p^{2m+5}+p^{2m+6}+p^3\Big). \tag{11}$$

Using this, one may show that

$$a_4'(p^{m+1})^2-a_4'(p^m)a_4'(p^{m+2})$$
$$=(p-1)^4p^{3m-7}\left((p+1)^2\left(p^2+p+1\right)^3p^{2m}-\left(p^2+p+1\right)^3p^m+(p+1)^2p\right)$$
$$=(p-1)^4p^{3m-7}\left(((p+1)^2p^m-1)\left(p^2+p+1\right)^3p^m+(p+1)^2p\right), \tag{12}$$

which we see is positive for all $p\geq 2$ and all $m\geq 1$. Moreover, using $a_4'(p^0)=1$, we get

$$a_4'(p^1)^2-a_4'(p^0)a_4'(p^2)=(p-1)(p+2)\left(p^3-3\right),$$

which is positive for all $p\geq 2$. Thus we have proved the inequality (6) for all $m\geq 0$, which completes the proof of Proposition 18 for the case $n=4$. □

Equations (11) and (12) may be verified with a computer algebra system, for instance with the Mathematica code provided at http://www.math.kth.se/~holmin/files/x/a4prime_is_logconcave.

B2: The case $n=5$

We prove Proposition 18 for $n=5$. We repeat the procedure above. We evaluate

$$a_5'(p^m)=\sum_{j_2+j_3+j_4+j_5=m}\left(p^{j_2}-p^{j_2-1}I(j_2>0)\right)\left(p^{2j_3}-p^{2(j_3-1)}I(j_3>0)\right)$$
$$\times\left(p^{3j_4}-p^{3(j_4-1)}I(j_4>0)\right)\left(p^{4j_5}-p^{4(j_5-1)}I(j_4>0)\right).$$

As before, we expand the product in the summand, and split the sum into several geometric series. This yields $a_5'(p^m)=\frac{p-1}{p^{10}(p+1)(p^2+p+1)}\Big(p^{4m}-p^{m+6}+p^{2m+3}+5p^{2m+4}+11p^{2m+5}+14p^{2m+6}+11p^{2m+7}+5p^{2m+8}+p^{2m+9}-p^{3m+1}-5p^{3m+2}-15p^{3m+3}-30p^{3m+4}-45p^{3m+5}-51p^{3m+6}-45p^{3m+7}-30p^{3m+8}-15p^{3m+9}-5p^{3m+10}-p^{3m+11}+4p^{4m+1}+10p^{4m+2}+20p^{4m+3}+31p^{4m+4}+40p^{4m+5}+44p^{4m+6}+40p^{4m+7}+31p^{4m+8}+20p^{4m+9}+10p^{4m+10}+4p^{4m+11}+p^{4m+12}\Big)$, valid for $m\geq 1$.

We get $a_5'(p)-a_5'(1)a_5'(p^2)=(p-1)\left((p-1)p\left(p^2+p+3\right)(p(p+2)+2)-10\right)$, which we see is positive, and thus we have proved the inequality (6) for $m=0$.

For $m \geq 1$, we get $a_5'(p^{m+1})^2 - a_5'(p^m)a_5(p^{m+2}) = \frac{(p-1)^4 p^{3m-13}}{p^2+p+1}\Big(-p^{3m} + p^{4m} - p^{m+2} - 4p^{m+3} - 10p^{m+4} - 16p^{m+5} - 19p^{m+6} - 16p^{m+7} - 10p^{m+8} - 4p^{m+9} - p^{m+10} + 2p^{2m+1} + 10p^{2m+2} + 34p^{2m+3} + 80p^{2m+4} + 143p^{2m+5} + 201p^{2m+6} + 224p^{2m+7} + 201p^{2m+8} + 143p^{2m+9} + 80p^{2m+10} + 34p^{2m+11} + 10p^{2m+12} + 2p^{2m+13} - 8p^{3m+1} - 32p^{3m+2} - 88p^{3m+3} - 188p^{3m+4} - 328p^{3m+5} - 480p^{3m+6} - 600p^{3m+7} - 646p^{3m+8} - 600p^{3m+9} - 480p^{3m+10} - 328p^{3m+11} - 188p^{3m+12} - 88p^{3m+13} - 32p^{3m+14} - 8p^{3m+15} - p^{3m+16} + 6p^{4m+1} + 23p^{4m+2} + 64p^{4m+3} + 143p^{4m+4} + 266p^{4m+5} + 423p^{4m+6} + 584p^{4m+7} + 706p^{4m+8} + 752p^{4m+9} + 706p^{4m+10} + 584p^{4m+11} + 423p^{4m+12} + 266p^{4m+13} + 143p^{4m+14} + 64p^{4m+15} + 23p^{4m+16} + 6p^{4m+17} + p^{4m+18} + p^6 + 2p^5 + p^4\Big)$. The first factor is obviously positive for $p \geq 2$, and the second factor may be rearranged as $(752p^{4m+9} - 646p^{3m+8}) + (706p^{4m+10} - 600p^{3m+9}) + (706p^{4m+8} - 600p^{3m+7}) + (584p^{4m+11} - 480p^{3m+10}) + (584p^{4m+7} - 480p^{3m+6}) + (423p^{4m+12} - 328p^{3m+11}) + (423p^{4m+6} - 328p^{3m+5}) + (266p^{4m+13} - 188p^{3m+12}) + (266p^{4m+5} - 188p^{3m+4}) + (143p^{4m+14} - 88p^{3m+13}) + (143p^{4m+4} - 88p^{3m+3}) + (64p^{4m+15} - 32p^{3m+14}) + (64p^{4m+3} - 32p^{3m+2}) + (23p^{4m+16} - 8p^{3m+15}) + (23p^{4m+2} - 8p^{3m+1}) + (6p^{4m+17} - p^{3m+16}) + (6p^{4m+1} - p^{3m}) + (224p^{2m+7} - 19p^{m+6}) + (201p^{2m+8} - 16p^{m+7}) + (201p^{2m+6} - 16p^{m+5}) + (143p^{2m+9} - 10p^{m+8}) + (143p^{2m+5} - 10p^{m+4}) + (80p^{2m+10} - 4p^{m+9}) + (80p^{2m+4} - 4p^{m+3}) + (34p^{2m+11} - p^{m+10}) + (34p^{2m+3} - p^{m+2}) + 10p^{2m+12} + 10p^{2m+2} + 2p^{2m+13} + 2p^{2m+1} + 2p^5 + p^{4m+18} + p^{4m} + p^6 + p^4$, where every term is positive for all $p \geq 2$, and we have thus proved the inequality (6) for $m \geq 1$. This concludes the proof of Proposition 18 for $n = 5$. □

The computations of $a_5'(p^m)$ and $a_5'(p^{m+1})^2 - a_5'(p^m)a_5'(p^{m+2})$ may be verified with the Mathematica code provided at http://www.math.kth.se/~holmin/files/x/a5prime_is_logconcave.

Appendix C: Calculation of a measure

In [4] the asymptotics

$$N_{n,0}(T) = \frac{n-1}{\zeta(n)} w(B) T^{n(n-1)} \log T + O(T^{n(n-1)})$$

are given, where B is the unit ball in $M_n(\mathbb{R})$. The measure w on $M_n(\mathbb{R})$ is defined in [4] as follows. Let $A_u := \{A \in M_n(\mathbb{R}) : Au = 0\}$ be the space of matrices annihilating the nonzero vector $u \in \mathbb{R}^n \setminus \{0\}$. We define for (Lebesgue measurable) subsets $E \subseteq M_n(\mathbb{R})$ the measure $w_u(E) := \mathrm{vol}(E \cap A_u)$ where vol is the standard $n(n-1)$-dimensional volume on A_u, and define the measure $w(E) := (1/2) \int_{S^{n-1}} w_u(E)\, d\nu(u)$, where ν is the standard Euclidean surface measure on the $(n-1)$-dimensional sphere S^{n-1}.

We shall now calculate $w(B)$. The set $B \cap A_u$ is the unit ball in the $n(n-1)$-dimensional vector space A_u. Its volume does not depend on $u \neq 0$, and if $u = (0, \ldots, 0, 1)$, then $B \cap A_u$ is the unit ball in $\mathbb{R}^{n(n-1)}$, when identifying $M_n(\mathbb{R})$ with $\mathbb{R}^{n^2}$. Denote by $V_{n(n-1)}$ the volume of the unit ball in $\mathbb{R}^{n(n-1)}$. Thus $w_u(B) = V_{n(n-1)}$, independently of $u \neq 0$, and

$$w(B) = V_{n(n-1)} \frac{1}{2} \int_{\mathrm{S}^{n-1}} d\nu(u) = \frac{V_{n(n-1)} S_{n-1}}{2},$$

where S_{n-1} is the surface area of the sphere S^{n-1}. The volume and surface area of the unit ball is well known, and we may explicitly calculate

$$C_0 := w(B) = \frac{\pi^{n^2/2}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n(n-1)}{2} + 1\right)}.$$

Recalling from Theorem 6 the expression for C_1, we get the following relation.

$$C_1 = \frac{1}{\zeta(2) \cdots \zeta(n)} \frac{\pi^{n^2/2}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n(n-1)}{2} + 1\right)} = \frac{1}{\zeta(2) \cdots \zeta(n)} C_0.$$

References

1. Cohen, H.: A Course in Computational Algebraic Number Theory, Graduate Texts in Mathematics, vol. 138. Springer-Verlag, Berlin (1993)
2. Duke, W., Rudnick, Z., Sarnak, P.: Density of integer points on affine homogeneous varieties. Duke Math. J. **71**(1), 143–179 (1993). doi:10.1215/S0012-7094-93-07107-4
3. Gorodnik, A., Nevo, A.: The Ergodic Theory of Lattice Subgroups, Annals of Mathematics Studies, vol. 172. Princeton University Press, Princeton (2010)
4. Katznelson, Y.R.: Singular matrices and a uniform bound for congruence groups of $\mathrm{SL}_n(\mathbb{Z})$. Duke Math. J. **69**(1), 121–136 (1993). doi:10.1215/S0012-7094-93-06906-2
5. Menon, K.V.: On the convolution of logarithmically concave sequences. Proc. Am. Math. Soc. **23**, 439–441 (1969)
6. Siegel, C.L.: A mean value theorem in geometry of numbers. Ann. Math. **46**(2), 340–347 (1945)
7. Stanley, R.P.: Enumerative Combinatorics: Volume 1 (Cambridge Studies in Advanced Mathematics), vol. 49. Cambridge University Press, Cambridge (1997)
8. Wigman, I.: Counting singular matrices with primitive row vectors. Monatsh. Math. **144**(1), 71–84 (2005). doi:10.1007/s00605-004-0250-7

Paper B

The number of points from a random lattice that lie inside a ball

Samuel Holmin *

January 18, 2015

We prove a sharp bound for the remainder term of the number of lattice points inside a ball, when averaging over a compact set of (not necessarily unimodular) lattices, in dimensions two and three. We also prove that such a bound cannot hold if one averages over the space of all lattices.

1 Introduction

Let Ω be the (closed) standard unit ball in $\mathbb{R}^n$. A **lattice** in $\mathbb{R}^n$ is a set of the form $X \cdot \mathbb{Z}^n \subseteq \mathbb{R}^n$ for some $X \in \mathrm{GL}_n(\mathbb{R})$. The set of all lattices may be identified with the space $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$, and we equip it with a measure μ induced by the Haar measure on $\mathrm{GL}_n(\mathbb{R})$. Let $N_X(t)$ be the number of points from the lattice $X\mathbb{Z}^n$ inside the ball $t\Omega$ of radius t. We have $N_X(t) = \#(X\mathbb{Z}^n \cap t\Omega) = \#(\mathbb{Z}^n \cap t\Omega_X)$, where $\Omega_X := X^{-1}\Omega$. Let $E_X(t) := N_X(t) - \mathrm{vol}(t\Omega_X)$. Consider the set of unit cubes centered at the set of integer points $u \in \mathbb{Z}^n$. Since $N_X(t)$ equals the number of cubes whose center is inside $t\Omega_X$, which coincides with the volume of the union of these cubes, we can write

$$N_X(t) = \mathrm{vol}(t\Omega_X) + \sum_{\text{cubes } T \text{ intersecting } \partial(t\Omega)} Y_T,$$

where Y_T equals $\mathrm{vol}(T \setminus t\Omega_X)$ if the center of T is inside $t\Omega$, and Y_T equals $-\mathrm{vol}(T \cap t\Omega_X)$ otherwise. There are approximately $\mathrm{vol}(\partial(t\Omega_X)) = t^{n-1}\,\mathrm{vol}(\partial(\Omega_X))$ correction terms Y_T, each bounded, so it follows that $N_X(t)$ is asymptotic to $t^n\,\mathrm{vol}(\Omega_X)$. Heuristically, if the correction terms Y_T were i.i.d. random variables, the central limit theorem would imply that the standard deviation of the remainder term $E_X(t) = \sum_T Y_T$ is approximately proportional to $\sqrt{\mathrm{vol}(\partial(t\Omega_X))}$ for large t. This suggests that $|E_X(t)|$ should be of the order $t^{(n-1)/2}$ for fixed X.

Let $\delta > 0$ be a small arbitrary constant. For the integer lattice $\mathbb{Z}^2$, Hardy conjectured that $|E_{\mathbb{Z}^2}(t)| = O(\sqrt{\mathrm{vol}(\partial(t\Omega))} \cdot t^{\delta}) = O(t^{1/2+\delta})$ as $t \to \infty$ [Har17]. It is known that

*The author was partially supported by the Swedish Research Council.

$|E_X(t)| \neq O(t^{1/2})$ for every lattice in $\mathbb{R}^2$, due to Nowak [Now85a], and the best known upper bound is $|E_X(t)| = O(t^{131/208+\delta})$, where $131/208 \approx 0.62981$, due to Huxley [Hux03]. Hardy's conjecture holds on average in the sense that $\sqrt{\frac{1}{t}\int_0^t |E_X(\tau)|^2\,d\tau} = \Theta(t^{1/2})$, due to Bleher [Ble92].

In three dimensions, it is known that $|E_X(t)| \neq O(t)$, due to Nowak [Now85b], and the best known upper bound for arbitrary lattices in $\mathbb{R}^3$ is $|E_X(t)| = O(t^{63/43+\delta})$, where $63/43 \approx 1.465$, due to Müller [Mül99], with the improvement $|E_{\mathbb{Z}^3}(t)| = O(t^{21/16+\delta})$ for the integer lattice $\mathbb{Z}^3$, where $21/16 = 1.3125$, due to Heath-Brown [HB99]. On average, we have $\sqrt{\frac{1}{t}\int_t^{2t} |E_X(\tau)|^2\,d\tau} = O(t^{1+\delta})$, see [ISS02]. (Note: Several of the cited results above were given in a more precise form; for instance, the latter bound was given as $O(t \log t)$.)

In higher dimensions, the following is known. For every lattice X in $n \geq 3$ dimensions, we have $|E_X(t)| \neq o(t^{(n-1)/2})$ (this result is due to Landau [Lan24]). It is not known for any $n \geq 2$ if there exists for each $\delta > 0$ some X such that $|E_X(t)| = O(t^{(n-1)/2+\delta})$, but Schmidt proved in [Sch60] that $|E_X(t)| = O(t^{n/2+\delta})$ for almost every lattice, when $n \geq 2$. The best general bound for $n \geq 5$ is $|E_X(t)| = O(t^{n-2})$, due to Götze [Göt04], and this bound is attained by the integer lattices (to be specific, $|E_{\mathbb{Z}^n}(t)| \neq o(t^{n-2})$ for every $n \geq 4$, see Krätzel [Krä00]). See [IKKN06] for an excellent survey on results about lattice points in convex domains.

The main result of this paper is that the bound $O(t^{(n-1)/2+\delta})$ holds on average in dimensions two and three, when averaging over any compact set of lattices:

Theorem 1. *Fix a compact subset L_0 of $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ and denote by $\mathbb{E}_0[f(X)] := \int_{L_0} f(X)\,d\mu(X)$ the mean of a function f over the set L_0. Then there exists some $\alpha > 0$ such that*

$$\sqrt{\mathbb{E}_0\Big[|E_X(t)|^2\Big]} = O(t(\log t)^\alpha)$$

as $t \to \infty$ for dimension $n = 3$, and

$$\sqrt{\mathbb{E}_0\Big[|E_X(t)|^2\Big]} = O(t^{1/2}) \tag{2}$$

as $t \to \infty$ for dimension $n = 2$.

The majority of this paper will focus on the three-dimensional case, as it is the more difficult case. Our bound (2) in two dimensions is an improvement of Theorem 1.1(ii) in [PT02], which had an additional factor t^δ. The corresponding statement of Theorem 1 for orthogonal lattices (that is, lattices $X\mathbb{Z}^n$ where X is a diagonal matrix), but with an additional factor t^δ, was proved by Hofmann, Iosevich, Weidinger in [HIW04], and our proof of Theorem 1 is inspired by theirs.

The assumption in Theorem 1 that L_0 is compact cannot be removed when $n = 3$: as Corollary 4 below shows, if we average over the set $L_{a,b} = \{X \in \mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z}) : 0 < a \leq |\det X| \leq b < \infty\}$, which is not compact, then we get both a lower and an upper bound with an exponent strictly larger than what Theorem 1 guarantees. The failure of

the heuristic in this case may be explained by the fact that $L_{a,b}$ contains lattices with arbitrarily short lattice vectors.

Proposition 3. *For any fixed $n \geq 3$, we have*

$$\sqrt{\mathbb{E}_1\left[|E_X(t)|^2\right]} = \Theta(\sqrt{\mathrm{vol}(t\Omega)}) = \Theta(t^{n/2})$$

as $t \to \infty$, where $\mathbb{E}_1[f(X)] := \int_{\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})} f(X)\, d\mu_1(X)$ is the mean of f over the set of all lattices in $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$, and where μ_1 is the normalized Haar measure on $\mathrm{SL}_n(\mathbb{R})$.

Corollary 4. *Fix $0 < a < b$. For any fixed $n \geq 3$, we have*

$$\sqrt{\mathbb{E}_{a,b}\left[|E_X(t)|^2\right]} = \Theta(t^{n/2})$$

as $t \to \infty$, where $\mathbb{E}_{a,b}[f(X)] := \int_{L_{a,b}} f(X)\, d\mu(X)$ is the mean of f over $L_{a,b} = \{X \in \mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z}) : a \leq |\det X| \leq b\}$.[*]

This paper is organized as follows. Sections 3 through section 6 are dedicated to the proof of Theorem 1 for $n = 3$. We sketch in section 7 how the given proof may be modified for the slightly easier case $n = 2$. Proposition 3 is an easy consequence of the mean value formulas of Siegel and Rogers; we prove Proposition 3 and Corollary 4 in section 8.

Remark 5. The actual measure used in Theorem 1 is not important; the proof holds for any measure of the form $f(X)\, dX$ and any compact set L_0 of $\mathrm{GL}_n(\mathbb{R})$, where dX is the Euclidean measure on the entries of the matrix X and $f : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^+$ is a function which is bounded above and below in $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ throughout L_0.

For instance, one may use the following natural measure for generating random lattices close to a given lattice. Fix a matrix $X_0 \in \mathrm{GL}_n(\mathbb{R})$. We generate random vectors $x_1, \ldots, x_n$, where each vector x_i is generated by a uniform probability measure on vectors sufficiently close to the ith column of X_0, and then we let $x_1, \ldots, x_n$ be the basis vectors of our random lattice. This corresponds to taking $f(X) = 1$ for all X and taking $L_0 := \{X_0 + tE : |t| \leq \varepsilon\}$, where E is the $n \times n$-matrix of all ones, and $\varepsilon > 0$ is sufficiently small such that L_0 does not contain any singular matrices.

2 Notation

Throughout this paper, we will assume that the parameter $t > 1$ is large. We will write $f(t) \lessapprox g(t)$ if there exists a constant $c > 0$ and an integer $m \geq 0$ such that $|f(t)| \leq |cg(t)(\log t)^m|$ for all sufficiently large t. We see that $\lessapprox$ is a transitive relation. As customary, we will write $f(t) \ll g(t)$ if there exists a constant c such that $|f(t)| \leq |cg(t)|$ for all sufficiently large t.

[*]Note that averaging over the whole set $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ does not make sense, since $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ has infinite covolume and consequently the expected value of any constant would be infinite.

Given a function $f : \mathbb{R}^k \to \mathbb{R}$ for some k, we write $\widehat{f}(\xi) = \int_{\mathbb{R}^k} f(x) e^{-2\pi i x \cdot \xi}\, dx$ for its Fourier transform.

We will write $\mathbb{Z}^n(a)$ for the set of all nonzero integer vectors $k = (k_1, \ldots, k_n)$ such that $|k_i| \le a$ for each $1 \le i \le n$. For a vector k and a matrix X, we will write $\|k\|_X := \left\|(X^{-1})^{\mathsf{T}} k\right\|$. Finally, we will frequently use the notation $\widetilde{k} := (N^{-1})^{\mathsf{T}} k$ where N is a given upper triangular matrix which will be clear from context.

3 Decomposition of the Haar measure

Let μ be the Haar measure on $\mathrm{GL}_3(\mathbb{R})$. The measure μ induces a measure on the quotient space $\mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z})$, and we will abuse notation by denoting both of these measures by the symbol μ. Let $\mathcal{F} \subseteq \mathrm{GL}_3(\mathbb{R})$ be a fundamental domain relative to $\mathrm{GL}_3(\mathbb{Z})$. If $f : \mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z}) \to \mathbb{R}$ is an integrable function, we shall write $f(X) := f(X \cdot \mathrm{GL}_3(\mathbb{Z}))$ for $X \in \mathrm{GL}_3(\mathbb{R})$, and then

$$\int_{\mathrm{GL}_3(\mathbb{R})/\,\mathrm{GL}_3(\mathbb{Z})} f(X)\, d\mu(X) = \int_{\mathcal{F} \subseteq \mathrm{GL}_3(\mathbb{R})} f(X)\, d\mu(X),$$

where in the right-hand side we are integrating with respect to the measure on $\mathrm{GL}_3(\mathbb{R})$.

We will use the Iwasawa decomposition $\mathrm{GL}_3(\mathbb{R}) = \mathcal{K} \cdot \mathcal{A} \cdot \mathcal{N}$ where $\mathcal{K} = \mathrm{O}_3(\mathbb{R})$ is the group of orthogonal matrices, $\mathcal{A}$ is the group of diagonal matrices with positive diagonal entries, and $\mathcal{N}$ is the group of upper triangular matrices with ones on the diagonal. If $X \in \mathrm{GL}_3(\mathbb{R})$, then there is a unique $(K, A, N) \in \mathcal{K} \times \mathcal{A} \times \mathcal{N}$ such that $X = KAN$. Let $\mathcal{N}^+$ be the set of all matrices $N \in \mathcal{N}$ such that all entries of N above the diagonal belong to the interval $[1, 2)$. (We will later use the fact that the entries of $N \in \mathcal{N}^+$ are not close to zero.) By performing Euclid's algorithm on the columns of N using elementary column operations, one can show that there exists for any $X = KAN$ some matrix $U \in \mathrm{GL}_3(\mathbb{Z})$ such that $XU \in \mathcal{K} \cdot \mathcal{A} \cdot \mathcal{N}^+$, which shows that the set $\mathcal{K} \cdot \mathcal{A} \cdot \mathcal{N}^+ \subseteq \mathrm{GL}_3(\mathbb{R})$ contains a fundamental domain $\mathcal{F}^+$ relative to $\mathrm{GL}_3(\mathbb{Z})$.

The Haar measure μ on $\mathrm{GL}_3(\mathbb{R})$ can be expressed in terms of the left-invariant Haar measures on $\mathcal{K}, \mathcal{A}$ and $\mathcal{N}$ as follows. Let $\mathcal{R} := \mathcal{A} \cdot \mathcal{N}$ be the group of upper triangular matrices with positive diagonal elements. The Haar measure on $\mathcal{A}$ is $dA = db_1\, db_2\, db_3\, /(b_1 b_2 b_3)$ where b_1, b_2, b_3 are the diagonal elements of $A \in \mathcal{A}$, and the Haar measure on $\mathcal{N}$ is $dN = d\eta_1\, d\eta_2\, d\eta_3$ where η_1, η_2, η_3 are the entries of $N \in \mathcal{N}$ above the diagonal. Write $\mu_{\mathcal{K}}$ for the (appropriately normalized) Haar measure on $\mathcal{K}$. Theorem 8.32 from [Kna02] implies that for any integrable function f, we have

$$\int_{\mathrm{GL}_3(\mathbb{R})} f(X)\, d\mu(X) = \int_{\mathcal{N}} \int_{\mathcal{A}} \int_{\mathcal{K}} f(KAN) \frac{\Delta_{\mathcal{R}}(AN)}{\Delta_{\mathrm{GL}_3(\mathbb{R})}(AN)} \frac{\Delta_{\mathcal{N}}(N)}{\Delta_{\mathcal{R}}(N)}\, d\mu_{\mathcal{K}}(K)\, dA\, dN$$

where $X = KR = KAN$, and $\Delta_G : G \to \mathbb{R}^+$ is the modular function associated with a topological group G. Let us write $\Delta(A, N) := \frac{\Delta_{\mathcal{R}}(AN)}{\Delta_{\mathrm{GL}_3(\mathbb{R})}(AN)} \frac{\Delta_{\mathcal{N}}(N)}{\Delta_{\mathcal{R}}(N)}$. The modular functions can be computed (in fact, one may show that $\Delta_{\mathrm{GL}_3(\mathbb{R})} = \Delta_{\mathcal{N}} = 1$, and $\Delta_{\mathcal{R}}(R) = b_1^2 b_3^{-2}$ where b_1, b_2, b_3 are the diagonal elements of R), but all we will need is that Δ is bounded

when restricted to a compact set, which follows from the fact that the modular functions are continuous and positive (see [Kna02]).

For our purposes, the parametrization

$$N = \begin{pmatrix} 1 & \eta_1 & \eta_2 \\ 0 & 1 & \eta_3 \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{N}^+, \qquad \eta_i \in [1,2), \tag{6}$$

$$A = \begin{pmatrix} 1/\sqrt{a_1} & 0 & 0 \\ 0 & 1/\sqrt{a_2} & 0 \\ 0 & 0 & 1/\sqrt{a_3} \end{pmatrix} \in \mathcal{A}, \qquad a_i \in (0,\infty),$$

will be useful. (The forthcoming expression (16) will take on a simpler form.) We get the Jacobian $\left|\frac{\partial(b_1,b_2,b_3)}{\partial(a_1,a_2,a_3)}\right| = 2^{-3}(a_1a_2a_3)^{-2}$. Writing $\Delta(a,\eta) := \Delta(A,N)$, and letting f be a non-negative integrable function on $\mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z})$, we obtain

$$\int\limits_{\mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z})} f(X)\,d\mu(X) = \int_{\mathcal{F}^+} f(X)\,d\mu(X) \leq$$

$$\int_{\mathcal{K}\cdot\mathcal{A}\cdot\mathcal{N}^+} f(X)\,d\mu(X) = \iiint\limits_{\substack{K\in\mathcal{K} \\ a\in(0,\infty)^3 \\ \eta\in[1,2)^3}} f(KAN)\frac{\Delta(a,\eta)}{2^3(a_1a_2a_3)^2}\,da\,d\eta\,d\mu_{\mathcal{K}}(K),$$

where $da = da_1\,da_2\,da_3$ and $d\eta = d\eta_1\,d\eta_2\,d\eta_3$ are the standard Lebesgue measures.

Integrating over the compact set $L_0 \subseteq \mathrm{GL}_3(\mathbb{R})/\mathrm{GL}_3(\mathbb{Z})$ with respect to the measure μ corresponds to integrating over the compact set

$$L_0' := L_0 \cdot \mathrm{GL}_3(\mathbb{Z}) \cap \mathcal{F}^+ \subseteq \mathrm{GL}_3(\mathbb{R}) \tag{7}$$

with respect to the measure $da\,d\eta\,d\mu_{\mathcal{K}}(K)$. For each $i = 1,2,3$, let ψ_i be the characteristic function of the smallest closed interval contained in $(0,\infty)$ which contains all values that a_i assumes when $X = KAN$ ranges over the compact set L_0'. Since $g(X) := |E_X(t)|^2$ is rotation invariant (that is, $g(KX) = g(X)$ for all $K \in \mathcal{K}, X \in \mathrm{GL}_3(\mathbb{R})$) and non-negative, we have

$$\int_{L_0} |E_X(t)|^2\,d\mu(X) \leq \int_{[1,2)^3}\int_{(0,\infty)^3} |E_{AN}(t)|^2 \frac{\Delta(a,\eta)}{2^3(a_1a_2a_3)^2}\psi_1(a_1)\psi_2(a_2)\psi_3(a_3)\,da\,d\eta\,.$$

The support of $\psi_1\psi_2\psi_3$ is contained in $(0,\infty)^3$, so for simplicity of notation, we will allow the inner integral to range over all of $\mathbb{R}^3$. Since $\Delta(a,\eta)/(2^3(a_1a_2a_3)^2)$ and $4\pi|\det A|^2$ are bounded above and below throughout the support of $\psi_1\psi_2\psi_3$, a bound of the right-hand side above will be equivalent, up to constants, to a bound of

$$\int_{[1,2)^3}\int_{\mathbb{R}^3} |E_{AN}(t)|^2 \frac{\Delta(a,\eta)}{2^3(a_1a_2a_3)^2}\frac{2^3(a_1a_2a_3)^2}{\Delta(a,\eta)}4\pi|\det A|^2\psi_1(a_1)\psi_2(a_2)\psi_3(a_3)\,da\,d\eta$$
$$= \int_{[1,2)^3}\int_{\mathbb{R}^3} |E_{AN}(t)|^2\psi(a)\,da\,d\eta, \tag{8}$$

where we have defined

$$\psi(a) := 4\pi|\det A|^2\psi_1(a_1)\psi_2(a_2)\psi_3(a_3).$$

(It is convenient to introduce the factor $4\pi|\det A|^2$ as it will later be cancelled by a factor appearing from $|E_{AN}(t)|^2$.) Thus, in order to bound $\int_{L_0}|E_X(t)|^2\,d\mu(X)$, it suffices to bound (8).

4 Setup

We define a smoothed version of

$$N_X(t) = \sum_{k\in\mathbb{Z}^3}\chi_{t\Omega_X}(k)$$

by

$$N_X^\varepsilon(t) := \sum_{k\in\mathbb{Z}^3}\chi_{t\Omega_X} * \rho_\varepsilon(k) \tag{9}$$

where $\rho : \mathbb{R}^3 \to \mathbb{R}$ is a mollifier and $\rho_\varepsilon(x) := \varepsilon^{-3}\rho(x/\varepsilon)$ for a parameter $\varepsilon = \varepsilon(t) > 0$. (Recall that a mollifier is a smooth, non-negative function with compact support and unit mass.) We define $\rho(x) := \rho_0(x_1)\rho_0(x_2)\rho_0(x_3)$ where $\rho_0 : \mathbb{R} \to \mathbb{R}$ is an even mollifier such that $|\widehat{\rho_0}(y)| \ll e^{-\sqrt{y}}$ for large y; see [Ing33] for the construction of such a function ρ_0. We obtain the asymptotics

$$|\widehat{\rho}(x)| \ll e^{-\sqrt{|x_1|}-\sqrt{|x_2|}-\sqrt{|x_3|}} \ll e^{-\sqrt{\|x\|}} \tag{10}$$

as $\|x\| \to \infty$, by the inequality $(\sqrt{|x_1|}+\sqrt{|x_2|}+\sqrt{|x_3|})^4 \geq x_1^2+x_2^2+x_3^2$. Note that the Fourier transform $\widehat{\rho}$ is real-valued since ρ is an even function.

Since the convolution $\chi_{t\Omega_X} * \rho_\varepsilon$ is smooth, we may apply the Poisson summation formula to the sum (9), and since both of the functions $\chi_{t\Omega_X}$ and ρ_ε have compact support, the convolution theorem $\widehat{\chi_{t\Omega_X} * \rho_\varepsilon} = \widehat{\chi_{t\Omega_X}} \cdot \widehat{\rho_\varepsilon}$ holds. Moreover, $\widehat{\chi_{t\Omega_X}}(0,0,0) = \int_{t\Omega_X} 1 = t^3 \operatorname{vol}\Omega_X$ and $\widehat{\rho_\varepsilon}(0,0,0) = \int \rho_\varepsilon = 1$, so we get

$$N_X^\varepsilon(t) = t^3\operatorname{vol}\Omega_X + \sum_{k\neq(0,0,0)}\widehat{\chi_{t\Omega_X}}(k)\widehat{\rho_\varepsilon}(k) =: t^3\operatorname{vol}\Omega_X + E_X^\varepsilon(t).$$

We first show that the function N_X^ε approximates N_X well:

Lemma 11. *There exists a constant $R > 0$ such that*

$$N_X^\varepsilon(t - R\varepsilon) \leq N_X(t) \leq N_X^\varepsilon(t + R\varepsilon),$$

where R only depends on the mollifier ρ.

Proof. Let R be the radius of a ball centered at the origin which contains the support of ρ, so that the support of ρ_ε is contained in a ball of radius εR. Consider

$$\chi_{t\Omega_X} * \rho_\varepsilon(k) = \int \rho_\varepsilon(x)\chi_{t\Omega_X}(k-x)\,dx\,.$$

The integral ranges over all $x \in \operatorname{supp}\rho_\varepsilon$, so we may assume that $\|x\| \le \varepsilon R$ inside the integral. If k is inside $t\Omega_X$ and at a distance at least εR from the boundary $\partial(t\Omega_X)$, then $\chi_{t\Omega_X}(k-x) = 1$, so the integral becomes $\int \rho_\varepsilon(x)\,dx = 1$, which agrees with $\chi_{t\Omega_X}(k) = 1$. If on the other hand k is outside $t\Omega_X$ and at a distance at least εR from the boundary $\partial(t\Omega_X)$, then $\chi_{t\Omega_X}(k-x) = 0$, so the integral vanishes and again agrees with $\chi_{t\Omega_X}(k) = 0$. Finally, if k is at a distance at most εR from the boundary $\partial(t\Omega_X)$, then since $0 \le \chi_{t\Omega_X} \le 1$ and ρ_ε is nonnegative, the integral is bounded below by 0 and above by $\int \rho_\varepsilon = 1$. We have thus proved that $\chi_{t\Omega_X} * \rho_\varepsilon$ equals $\chi_{t\Omega_X}$ at all points at a distance at least εR from the boundary of $t\Omega_X$, and at all other points it assumes a value in $[0,1]$. This proves the lemma, since $N_X(t)$ counts the number of lattice points inside $t\Omega_X$, while $N_X^\varepsilon(t - R\varepsilon)$ counts each of these with a weight at most 1, and $N_X^\varepsilon(t + R\varepsilon)$ counts all the same lattice points, plus a few more with various weights in $[0,1]$. □

Using the lemma, we arrive at:

Claim 12. To prove Theorem 1 for $n = 3$ it suffices to prove that

$$\int_{[1,2)^3}\int_{\mathbb{R}^3} |E_{AN}^\varepsilon(t)|^2 \psi(a)\,da\,d\eta \lessapprox t^2 \tag{13}$$

for all $\varepsilon = \varepsilon(t)$ such that $\varepsilon \ge 1/t$ for all sufficiently large t.

Proof. Lemma 11 implies that

$$\begin{aligned} E_X(t) &\le E_X^{\varepsilon_0}(t + R\varepsilon_0) + \operatorname{vol}(\Omega_X)((t+R\varepsilon_0)^3 - t^3),\\ -E_X(t) &\le -E_X^{\varepsilon_0}(t - R\varepsilon_0) + \operatorname{vol}(\Omega_X)(t^3 - (t-R\varepsilon_0)^3),\end{aligned}$$

for any $\varepsilon_0 > 0$. Choosing $\varepsilon_0 := 2/t$ we get

$$\begin{aligned} |E_X(t)| &\le \max(|E_X^{\varepsilon_0}(t+R\varepsilon_0) + O(t)|, |E_X^{\varepsilon_0}(t - R\varepsilon_0) + O(t)|)\\ &\ll |E_X^{\varepsilon_0}(t+R\varepsilon_0)| + |E_X^{\varepsilon_0}(t-R\varepsilon_0)| + t.\end{aligned}$$

The asymptotic constant depends on the determinant of X, but if we restrict X to the compact set L_0' (see (7)), then the determinant of X is bounded by a constant which only depends on the fixed set L_0. By (8) we have

$$\begin{aligned}\int_{L_0} |E_X(t)|^2\,d\mu(X) &\ll \int_{[1,2)^3}\int_{\mathbb{R}^3} |E_{AN}(t)|^2\psi(a)\,da\,d\eta\\ \ll \int_{[1,2)^3}\int_{\mathbb{R}^3} |E_X^{\varepsilon_0}(t+R\varepsilon_0)|^2\psi(a)\,da\,d\eta &+ \int_{[1,2)^3}\int_{\mathbb{R}^3} |E_X^{\varepsilon_0}(t-R\varepsilon_0)|^2\psi(a)\,da\,d\eta + t^2,\end{aligned}$$

and noting that $\varepsilon_0 \geq 1/(t + R\varepsilon_0)$ and $\varepsilon_0 \geq 1/(t - R\varepsilon_0)$ for all sufficiently large $t \pm R\varepsilon_0$, the hypothesis (13) implies that the right-hand side above is

$$\lessapprox (t + R\varepsilon_0)^2 + (t - R\varepsilon_0)^2 + t^2 \ll t^2,$$

and thus $\sqrt{\int_{L_0} |E_X(t)|^2 \, d\mu(X)} \lessapprox t$ follows. □

For the remainder of the section we will assume that $\varepsilon \geq 1/t$ for all sufficiently large t. We will now estimate the behavior of E_X^ε. Consider the Fourier transform of the characteristic function χ_Ω of the standard unit ball Ω in $\mathbb{R}^3$. Taking advantage of the fact that χ_Ω is a radial function and hence that its Fourier transform is radial as well, an easy calculation shows that (see equation 10 in chapter 6.4 in [SS03])

$$\widehat{\chi_\Omega}(k) = \frac{2}{\|k\|} \int_0^1 \sin(2\pi\|k\|r) r \, dr,$$

which can be integrated by parts to get

$$\widehat{\chi_\Omega}(k) = -\frac{\cos(2\pi\|k\|)}{\pi\|k\|^2} + \frac{\sin(2\pi\|k\|)}{2\pi^2\|k\|^3}.$$

Since $\Omega_X = X^{-1} \cdot \Omega$ we get

$$\widehat{\chi_{\Omega_X}}(k) = \int_{X^{-1}\cdot\Omega} e^{2\pi i x\cdot k} \, dx = \int_\Omega e^{2\pi i X^{-1} y \cdot k} \left|\det X^{-1}\right| dy$$
$$= \left|\det X^{-1}\right| \widehat{\chi_\Omega}((X^{-1})^\mathsf{T} k) = |\det X|^{-1} \left(-\frac{\cos(2\pi\|k\|_X)}{\pi\|k\|_X^2} + \frac{\sin(2\pi\|k\|_X)}{2\pi^2\|k\|_X^3} \right),$$

recalling the definition

$$\|k\|_X = \|(X^{-1})^\mathsf{T} k\|.$$

Recall that $E_X^\varepsilon(t) = \sum_{k\neq(0,0,0)} \widehat{\chi_{t\Omega_X}}(k)\widehat{\rho_\varepsilon}(k)$. It is straightforward to show that $\widehat{\chi_{t\Omega_X}}(k) = t^3\widehat{\chi_{\Omega_X}}(tk)$ and $\widehat{\rho_\varepsilon}(k) = \widehat{\rho}(\varepsilon k)$. Hence we can write

$$E_X^\varepsilon(t) = S_1 + S_2 :=$$
$$-|\det X|^{-1} t \sum_{k\neq(0,0,0)} \frac{\cos(2\pi\|tk\|_X)}{\pi\|k\|_X^2} \widehat{\rho}(\varepsilon k) + |\det X|^{-1} \sum_{k\neq(0,0,0)} \frac{\sin(2\pi\|tk\|_X)}{2\pi^2\|k\|_X^3} \widehat{\rho}(\varepsilon k),$$

where both sums S_1, S_2 are real since $\widehat{\rho}$ is real-valued. For $X = AN, A \in \mathcal{A}, N \in \mathcal{N}^+$, we have $|\det X|^{-1} \ll 1$, so for such X we get

$$|S_2| \ll \sum_{k\neq(0,0,0)} \frac{|\widehat{\rho}(\varepsilon k)|}{\|k\|^3}.$$

We use the fact that $|\widehat{\rho}(\varepsilon k)|$ decreases as $1/\|\varepsilon k\|^N \leq t^N/\|k\|^N$ for any $N > 0$, provided that $\varepsilon \geq 1/t$. Then we get $|S_2| \ll \sum_{k\neq 0} t^N/\|k\|^{3+N} = t^N \sum_{k\neq 0} 1/\|k\|^{3+N} \ll t^N$, where

the final sum converges to a constant by integral comparison for any $N > 0$. Choosing $N = 1/2$ gives us $|S_2| \ll t^{1/2}$.

Consequently we have

$$|E_X^\varepsilon(t)|^2 = (S_1 + S_2)^2 \ll S_1^2 + S_2^2 \ll S_1^2 + t,$$

and thus, to prove Theorem 1 for $n = 3$, by Claim 12 it will suffice to prove that $\int_{[1,2)^3} \int_{\mathbb{R}^3} S_1^2 \psi(a)\, da\, d\eta \lessapprox t^2$, where

$$S_1^2 = |\det X|^{-2} t^2 \sum_{k,l \neq (0,0,0)} \frac{\cos(2\pi\|tk\|_X)\cos(2\pi\|tl\|_X)}{\pi^2 \|k\|_X^2 \|l\|_X^2} \widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)$$

and $X = AN$, using the parametrization (6). Write the product $\cos(2\pi\|tk\|_X)\cos(2\pi\|tl\|_X)$ as $(e^\alpha + e^{-\alpha})(e^\beta + e^{-\beta})/4 = \frac{1}{4}(e^{\alpha+\beta} + e^{\alpha-\beta} + e^{-\alpha+\beta} + e^{-\alpha-\beta})$ where $\alpha := 2\pi i t\|k\|_X$ and $\beta := 2\pi i t\|l\|_X$. We split the integral into a sum of four integrals and treat each case separately, that is, we will prove

$$t^2 \int_{[1,2)^3} \int_{\mathbb{R}^3} \sum_{k,l \neq (0,0,0)} |\det A|^{-2} \frac{e^{2\pi i t \Phi_{k,l}(AN)}}{4\pi^2 \|k\|_{AN}^2 \|l\|_{AN}^2} \widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)\psi(a)\, da\, d\eta \lessapprox t^2$$

where $\Phi_{k,l}(X) = \pm\|k\|_X \pm \|l\|_X$, for all four different combinations of sign choices.

We cancel the factor t^2 on both sides and exchange the order of integration and summation (noting that the sum is uniformly convergent by the rapid decay of $\widehat{\rho}$). Thus, recalling that $\psi(a) = 4\pi|\det A|^2 \psi_1(a_1)\psi_2(a_2)\psi_3(a_3)$, we arrive at:

Claim 14. To prove Theorem 1 for $n = 3$ it suffices to prove that

$$\sum_{k,l \neq (0,0,0)} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^2 \|l\|^2} |I_{k,l}(t)| \lessapprox 1, \tag{15}$$

for all $\varepsilon = \varepsilon(t)$ such that $\varepsilon \geq 1/t$ for all sufficiently large t, where

$$\begin{aligned} I_{k,l}(t) &:= \int_{[1,2)^3} \int_{\mathbb{R}^3} e^{2\pi i t \Phi_{k,l}(AN)} \psi_{k,l}(AN)\, da\, d\eta, \\ \Phi_{k,l}(AN) &:= \pm\|k\|_{AN} \pm \|l\|_{AN}, \\ \psi_{k,l}(AN) &:= \left(\frac{\|k\|}{\|k\|_{AN}}\right)^2 \left(\frac{\|l\|}{\|l\|_{AN}}\right)^2 \psi_1(a_1)\psi_2(a_2)\psi_3(a_3), \end{aligned}$$

for all four choices of signs in the definition of $\Phi_{k,l}$.

Consider $\Phi_{k,l}(AN)$ for $A \in \mathcal{A}, N \in \mathcal{N}^+$. Write $\widetilde{k} := (N^{-1})^{\mathsf{T}} k$ and $\widetilde{l} := (N^{-1})^{\mathsf{T}} l$. Then $\|k\|_{AN} = \|(A^{-1})^{\mathsf{T}}(N^{-1})^{\mathsf{T}} k\| = \|A^{-1}\widetilde{k}\|$. Similarly $\|l\|_A = \|A^{-1}\widetilde{l}\|$. Using the parametrization (6), we get

$$A^{-1} = \begin{pmatrix} \sqrt{a_1} & 0 & 0 \\ 0 & \sqrt{a_2} & 0 \\ 0 & 0 & \sqrt{a_3} \end{pmatrix}, \qquad (N^{-1})^{\mathsf{T}} = \begin{pmatrix} 1 & 0 & 0 \\ -\eta_1 & 1 & 0 \\ \eta_1\eta_3 - \eta_2 & -\eta_3 & 1 \end{pmatrix}$$

and therefore

$$\Phi_{k,l}(AN) = \pm\sqrt{a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2} \pm \sqrt{a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2}. \tag{16}$$

where $\widetilde{k}_i^2$ denotes the square of the ith component of the vector $\widetilde{k} = (N^{-1})^{\mathsf{T}}k$, and similarly for $\widetilde{l}_i^2$. Note that our choice of parametrization (6) of the entries of A turned the expressions inside the square roots in the exponent $\Phi_{k,l}(AN)$ into linear forms of a_1, a_2, a_3.

Since $\|X^{-1}\|_{\mathrm{op}}\|k\| \le \|Xk\| \le \|X\|_{\mathrm{op}}\|k\|$ where $\|X\|_{\mathrm{op}}$ is the operator norm of the matrix X for any X, it follows that $\|k\|_{AN} \ll \|k\| \ll \|k\|_{AN}$ and likewise for l, when $AN \in \mathcal{A}\cdot\mathcal{N}^+$. Hence $\psi_{k,l}(AN)$ can be bounded above and below by constants uniform in k and l (but depending on L_0), and thus $|I_{k,l}(t)| \ll \int|\psi_{k,l}| \ll 1$.

We now show that we may neglect the terms in the sum (15) for which either $\|k\|$ or $\|l\|$ is large, where the notion of "large" is given by the following definition.

Definition 17. We set $\mathcal{U}(t) := 32t(\log t)^2$ for all $t > 1$. Note that $\mathcal{U}(t) \lessapprox t$ and $\log(\mathcal{U}(t)) \lessapprox 1$.

Lemma 18. *Assuming that $\varepsilon \ge 1/t$ for all sufficiently large t, we have*

$$\sum_{\substack{k,l\neq(0,0,0)\\ \|k\|\ge\mathcal{U}(t) \text{ or } \|l\|\ge\mathcal{U}(t)}} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \ll 1$$

where the analogous bound holds if we interchange k and l.

Proof. It suffices to bound the sum

$$\sum_{\substack{k,l\neq(0,0,0)\\ \|k\|\ge\mathcal{U}(t)}} = \sum_{\substack{k,l\neq(0,0,0)\\ \|k\|,\|l\|\ge\mathcal{U}(t)}} + \sum_{\substack{k,l\neq(0,0,0)\\ \|k\|\ge\mathcal{U}(t)>\|l\|}} . \tag{19}$$

Using the bounds $|I_{k,l}(t)| \ll 1$, $|\widehat{\rho}(\varepsilon l)| \ll 1$, and finally $|\widehat{\rho}(\varepsilon k)| \ll e^{-\sqrt{\|\varepsilon k\|}}$ from (10), and assuming that $\varepsilon \ge 1/t$, the second sum on the right above can be written as

$$\sum_{\substack{k,l\neq(0,0,0)\\ \|k\|\ge\mathcal{U}(t)>\|l\|}} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \ll \sum_{\substack{l\neq(0,0,0)\\ \|l\|\le\mathcal{U}(t)}} \frac{1}{\|l\|^2} \sum_{\substack{k\neq(0,0,0)\\ \|k\|\ge\mathcal{U}(t)}} \frac{e^{-\sqrt{\|k/t\|}}}{\|k\|^2}. \tag{20}$$

The first sum on the right-hand side of (20) is

$$\ll \int_1^{\mathcal{U}(t)} \frac{1}{r^2} r^2\, dr \ll \mathcal{U}(t) \ll t(\log t)^2.$$

The second sum on the right-hand side of (20) is

$$\ll \int_{\mathcal{U}(t)/2}^{\infty} e^{-\sqrt{r/t}}\, dr \ll \left(-2te^{-\sqrt{r/t}}\left(\sqrt{r/t}+1\right)\right)\Big|_{r=\mathcal{U}(t)/2} \ll$$
$$te^{-\sqrt{16(\log t)^2}}\sqrt{16(\log t)^2} \ll te^{-4\log t}(\log t)^2 = t^{-3}(\log t)^2. \tag{21}$$

Thus the right-hand side of (20) is

$$\ll t^{-2}(\log t)^4 \ll 1.$$

The first sum on the right-hand side of (19) can be written as

$$\sum_{\substack{k,l\neq(0,0,0)\\ \|k\|,\|l\|\geq\mathcal{U}(t)}} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \ll \sum_{\substack{k\neq(0,0,0)\\ \|k\|\geq\mathcal{U}(t)}} \frac{e^{-\sqrt{\|k/t\|}}}{\|k\|^2} \sum_{\substack{l\neq(0,0,0)\\ \|l\|\geq\mathcal{U}(t)}} \frac{e^{-\sqrt{\|l/t\|}}}{\|l\|^2},$$

which by our previous calculation is $\ll (t^{-3}(\log t)^2)^2 \ll 1$. □

Remark 22. If one only wants to prove a weaker version of Theorem 1 with a bound of the form $O(t^{(n-1)/2+\delta})$ for some $\delta > 0$, with no log factors, it suffices to take $\mathcal{U}(t) = t^{1+\delta'}$ for some sufficiently small $\delta' > 0$, and to use the elementary estimate $\widehat{\rho}(x) \ll 1/\|x\|^2$ for the Fourier transform of ρ in the proof of Lemma 18.

The lemma above shows that we may restrict ourselves to summing only over the integer vectors $k, l \neq (0,0,0)$ bounded in norm by $\mathcal{U}(t)$, and thus it is enough to sum over $k, l \neq (0,0,0)$ such that $|k_i|, |l_j| \leq \mathcal{U}(t)$ for all $i, j \in \{1,2,3\}$. Thus we have:

Claim 23. To prove Theorem 1 for $n = 3$ it suffices to prove that

$$\sum_{k,l\in\mathbb{Z}^3(\mathcal{U}(t))} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \lessapprox 1 \tag{24}$$

where the sum extends over all nonzero integer vectors $k, l \in \mathbb{Z}^3$ with entries bounded by $\mathcal{U}(t)$.

5 Neglecting integer vectors with vanishing coordinates

In order to bound the sum on the left-hand side of (24), we will need to take advantage of nontrivial bounds of the oscillating integral $I_{k,l}(t)$. We will derive such a bound in Section 6, but for technical reasons, in order to use that bound, we need the first two coordinates of k and l to be nonzero. In the present section, we will prove that we can neglect the part of the sum where some of k_1, k_2, l_1, l_2 are zero.

We begin by showing that the terms for which both some coordinate of k and some coordinate of l is zero can be neglected:

Lemma 25. *We have*

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1=l_1=0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \lessapprox 1.$$

The same bound holds if we exchange k_1 for any other component of k, and l_1 for any other component of l.

Proof. We use the trivial bound $|I_{k,l}(t)| \ll 1$ and split the sum into one over k and one over l. The sum over k satisfies

$$\sum_{\substack{k\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1=0}} \frac{1}{\|k\|^2} = \sum_{\substack{|k_2|,|k_3|\le\mathcal{U}(t)\\ (k_2,k_3)\neq(0,0)}} \frac{1}{\|(k_2,k_3)\|^2} \ll \int_1^{\mathcal{U}(t)} \frac{1}{r^2} r\,dr \ll \log(\mathcal{U}(t)) \lessapprox 1$$

where in the second sum we are only summing over integer vectors in $\mathbb{Z}^2$. The same bound holds for the sum over l, so the statement of the lemma follows. □

We now need a lemma on oscillating integrals; see the corollary of Proposition 2 in chapter VIII in [Ste93].

Lemma 26 (van der Corput lemma)**.** *Let $\phi, \psi_0 : [a,b] \to \mathbb{R}$ be smooth functions defined on some interval $[a,b]$, and suppose that ϕ' is monotonic and that there exists a constant $c_0 > 0$ such that $\phi'(x) \ge c_0$ for all x. Then*

$$\left|\int_a^b e^{it\phi(x)}\psi_0(x)\,dx\right| \le \frac{C}{c_0 t}\left(|\psi_0(b)| + \int_a^b |\psi_0'(x)|\,dx\right)$$

for all $t > 0$, where C is an absolute constant.

We prove in the following two lemmas that we can also neglect the terms for which precisely one of k and l has a zero in the first two coordinates.†

Lemma 27. *We have*

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1=0\\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \lessapprox 1.$$

The same bound holds if we exchange the roles of k and l.

Proof. Assume that $k_1 = 0, k \neq (0,0,0)$ and $l_1, l_2, l_3 \neq 0$. Consider $\Phi_{k,l}(AN)$, given by equation (16). The partial derivative with respect to a_1 is

$$\frac{\partial}{\partial a_1}\Phi_{k,l}(AN) = \pm\frac{\widetilde{k}_1^2}{2\|k\|_{AN}} \pm \frac{\widetilde{l}_1^2}{2\|l\|_{AN}}.$$

Now, since $\widetilde{k}_1 = k_1 = 0$ and $\widetilde{l}_1 = l_1 \neq 0$, we get

$$\frac{\partial}{\partial a_1}\Phi_{k,l}(AN) = \pm\frac{\widetilde{l}_1^2}{2\|l\|_{AN}} \gg \frac{l_1^2}{\|l\|} \gg \frac{|l_1|}{\|l\|}.$$

†This does not imply an analogous statement for the third coordinate because the proof depends on a bound of the integral $I_{k,l}(t)$, and our choice of decomposition $\mathcal{KAN}^+$ of our integration domain is not symmetric in the coordinates.

Moreover, the second derivative with respect to a_1 is

$$\left(\frac{\partial}{\partial a_1}\right)^2 \Phi_{k,l}(AN) = \mp \frac{l_1^4}{4\|l\|_{AN}^3},$$

which is either always positive or always negative, depending on the sign $\pm$ in the definition of $\Phi_{k,l}$. Thus the map $\phi(a_1) := \Phi_{k,l}(AN)$ for fixed a_2, a_3 is such that $|\phi'(a_1)| \gg |l_1|/\|l\|$ and ϕ' is monotonic. Writing $[b_1, b_2]$ for the support of the characteristic function ψ_1, we can apply the van der Corput Lemma 26 to the integral

$$\int_{b_1}^{b_2} e^{2\pi i t \Phi_{k,l}(AN)} \psi_0(a_1)\, da_1$$

where we have defined $\psi_0(a_1) := \frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2}$. The function ψ_0 is bounded since $\|k\|_{AN} \gg \|k\|$ and $\|l\|_{AN} \gg \|l\|$. Its derivative, by the assumption that $\widetilde{k}_1 = k_1 = 0, \widetilde{l}_1 = l_1 \neq 0$, is

$$\begin{aligned}\psi_0'(a_1) &= \frac{d}{da_1} \frac{\|k\|^2\|l\|^2}{(a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2)(a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2)} \\ &= -\frac{\|k\|^2\|l\|^2 l_1^2}{(a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2)(a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2)^2} = -\frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2} \frac{l_1^2}{\|l\|_{AN}^2},\end{aligned}$$

which is also bounded. Thus the van der Corput Lemma gives us the bound

$$\left| \int_{\mathbb{R}} e^{2\pi i t \Phi_{k,l}(AN)} \psi_{k,l}(AN)\, da_1 \right| \ll \frac{1}{t} \frac{\|l\|}{|l_1|},$$

where the asymptotic constant is independent of k, l. Integrating in the rest of the variables yields by compactness

$$|I_{k,l}(t)| \ll \int_{[1,2)^3} \int_{\mathbb{R}^2} \frac{1}{t} \frac{\|l\|}{|l_1|} \psi_2(a_2)\psi_3(a_3)\, da_2\, da_2\, d\eta \ll \frac{1}{t}\frac{\|l\|}{|l_1|}.$$

Using this bound, it now follows that

$$\sum_{\substack{k,l \in \mathbb{Z}^3(\mathcal{U}(t)) \\ k_1=0 \\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2} |I_{k,l}(t)| \ll \frac{1}{t} \sum_{\substack{k \in \mathbb{Z}^3(\mathcal{U}(t)) \\ k_1=0}} \frac{1}{\|k\|^2} \sum_{\substack{l \in \mathbb{Z}^3(\mathcal{U}(t)) \\ l_1,l_2,l_3\neq 0}} \frac{1}{\|l\||l_1|}.$$

The sum over k has logarithmic behavior in $\mathcal{U}(t)$ since we are summing over a two-dimensional space. We will split the sum over l into one over l_1, and one over (l_2, l_3). We have $\|l\| \geq \|(0, l_2, l_3)\| \geq \|(l_2, l_3)\|$, so

$$\begin{aligned}\sum_{\substack{k,l \in \mathbb{Z}^3(\mathcal{U}(t)) \\ k_1=0 \\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2} |I_{k,l}(t)| &\lessapprox \frac{1}{t} \sum_{1\leq|l_1|\leq\mathcal{U}(t)} |l_1|^{-1} \sum_{1\leq|l_2|,|l_3|\leq\mathcal{U}(t)} \|(l_2,l_3)\|^{-1} \\ &\ll \frac{1}{t}\int_1^{\mathcal{U}(t)} \frac{1}{x}\, dx \int_1^{\mathcal{U}(t)} \frac{1}{r} r\, dr \ll \frac{1}{t} \cdot \log(\mathcal{U}(t)) \cdot \mathcal{U}(t) \lessapprox 1. \end{aligned} \tag{28}$$

This completes the proof that the sum over $k_1 = 0$ can be neglected. □

Lemma 29. *We have*

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \lessapprox 1.$$

The same bound holds if we exchange the roles of k and l.

Proof. Assume that $k_2 = 0, k \neq (0,0,0)$ and $l_1, l_2, l_3 \neq 0$. We write

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| =$$

$$\int_{[1,2)^3} \sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0}} \frac{1}{\|k\|^2\|l\|^2}\left|\int_{\mathbb{R}^3} e^{2\pi i\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta\,. \tag{30}$$

We will split the latter sum into two parts: one in which $|l_2 - 2\eta_1 l_1| \geq 1$, and one in which $|l_2 - 2\eta_1 l_1| < 1$. We will bound the sum over $|l_2 - 2\eta_1 l_1| \geq 1$ by mimicking the proof of Lemma 27, with the difference that we consider instead the directional derivative of $\Phi_{k,l}(AN)$ with respect to the direction $(-\eta_1^2, 1, 0)$.

We deal first with the part of the sum (30) with $|l_2 - 2\eta_1 l_1| \geq 1$. We change the order of integration inside the integral $I_{k,l}(t)$ such that the innermost integral is taken with respect to a_2, and perform a one-variable substitution from a_2 to $u := -\eta_1^2 a_1 + a_2$ inside this integral. Recalling the expression (16), it now follows, since $k_2 = 0$, that

$$\begin{aligned}\frac{\partial}{\partial u}\Phi_{k,l}(AN) &= -\eta_1^2\frac{\partial}{\partial a_1}\Phi_{k,l}(AN) + \frac{\partial}{\partial a_2}\Phi_{k,l}(AN) = \pm\frac{-\eta_1^2\widetilde{k}_1^2 + \widetilde{k}_2^2}{2\|k\|_{AN}} \pm \frac{-\eta_1^2\widetilde{l}_1^2 + \widetilde{l}_2^2}{2\|l\|_{AN}}\\ &= \pm\frac{-\eta_1^2 k_1^2 + (-\eta_1 k_1 + k_2)^2}{2\|k\|_{AN}} \pm \frac{-\eta_1^2 l_1^2 + (-\eta_1 l_1 + l_2)^2}{2\|l\|_{AN}}\\ &= \pm\frac{-\eta_1^2 l_1^2 + (-\eta_1 l_1 + l_2)^2}{2\|l\|_{AN}} = \pm\frac{-2\eta_1 l_1 l_2 + l_2^2}{2\|l\|_{AN}} = \pm\frac{l_2(l_2 - 2\eta_1 l_1)}{2\|l\|_{AN}}\end{aligned}$$

and

$$\left(\frac{\partial}{\partial u}\right)^2\Phi_{k,l}(AN) = \mp\frac{(l_2(l_2 - 2\eta_1 l_1))^2}{4\|l\|_{AN}^3}.$$

Whenever $|l_2 - 2\eta_1 l_1| \geq 1$ holds, we get a bound of the form $\left|\frac{\partial}{\partial u}\Phi_{k,l}\right| \gg |l_2|/\|l\|$ with $u \mapsto \frac{\partial}{\partial u}\Phi_{k,l}$ monotonic. Since $\psi_1\psi_2$ is the characteristic function of a rectangle, it follows that the support of $u \mapsto \psi_{k,l}(AN)$ is some interval $[b_1, b_2]$, which is bounded in length (independent of k and l). The function $u \mapsto \psi_{k,l}(AN)$ restricted to the interval $[b_1, b_2]$ coincides with the function $u \mapsto \frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2}$ because $\psi_1\psi_2\psi_3$ is a characteristic function.

The function $u \mapsto \psi_{k,l}(AN)$ is bounded, and so is

$$\frac{\partial}{\partial u}\psi_{k,l}(AN) = \frac{\partial}{\partial u}\frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2} =$$
$$-\frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2}\cdot\frac{(-\eta_1^2\widetilde{k}_1^2+\widetilde{k}_2^2)}{\|k\|_{AN}^2} - \frac{\|k\|^2\|l\|^2}{\|k\|_{AN}^2\|l\|_{AN}^2}\cdot\frac{(-\eta_1^2\widetilde{l}_1^2+\widetilde{l}_2^2)}{\|l\|_{AN}^2}$$

on the interval $[b_1, b_2]$ since $\left|-\eta_1^2\widetilde{l}_1^2+\widetilde{l}_2^2\right| \ll \|\widetilde{l}\|^2 \ll \|l\|_{AN}^2$ and $-\eta_1^2\widetilde{k}_1^2+\widetilde{k}_2^2 = 0$. Thus, whenever $|l_2 - 2\eta_1 l_1| \geq 1$ holds, the van der Corput Lemma 26 gives us the bound

$$\left|\int_{\mathbb{R}} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da_2\right| = \left|\int_{\mathbb{R}} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,du\right| \ll \frac{1}{t}\frac{\|l\|}{|l_2|},$$

and estimating trivially in the remaining variables a_1, a_3 yields

$$\left|\int_{\mathbb{R}^3} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| \ll \frac{1}{t}\frac{\|l\|}{|l_2|}.$$

This bound yields

$$\int_{[1,2)^3}\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0\\ |l_2-2\eta_1 l_1|\geq 1}}\frac{1}{\|k\|^2\|l\|^2}\left|\int_{\mathbb{R}^3} e^{2\pi i\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right|d\eta \ll$$
$$\frac{1}{t}\int_{[1,2)^3}\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0\\ |l_2-2\eta_1 l_1|\geq 1}}\frac{1}{\|k\|^2\|l\|^1|l_2|}\,d\eta \leq \frac{1}{t}\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0}}\frac{1}{\|k\|^2\|l\||l_2|} \lessapprox 1,$$

where the last bound is completely analogous to the bound (28).

It remains to bound the part of the sum (30) with $|l_2 - 2\eta_1 l_1| < 1$. When $|l_2 - 2\eta_1 l_1| < 1$, there are at most two values that l_2 may assume when η_1, l_1 are held fixed, and using

$\|(l_1,l_2,l_3)\| \geq \|(l_1,0,l_3)\| = \|(l_1,l_3)\|$, we get

$$\int_{[1,2)^3} \sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0\\ |l_2-2\eta_1 l_1|<1}} \frac{1}{\|k\|^2\|l\|^2}\left|\int_{\mathbb{R}^3} e^{2\pi i\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta \ll$$

$$\int_{[1,2)^3} \sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0\\ |l_2-2\eta_1 l_1|<1}} \frac{1}{\|k\|^2\|l\|^2}\, d\eta \ll$$

$$\int_{[1,2)^3} \sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2,l_3\neq 0\\ |l_2-2\eta_1 l_1|<1}} \frac{1}{\|(k_1,k_3)\|^2\|(l_1,l_3)\|^2}\, d\eta \ll$$

$$\sum_{1\leq|k_1|,|k_3|\leq\mathcal{U}(t)}\ \sum_{1\leq|l_1|,|l_3|\leq\mathcal{U}(t)} \frac{1}{\|(k_1,k_3)\|^2\|(l_1,l_3)\|^2} \lessapprox 1,$$

and we are done. □

Putting the lemmas together, we have thus demonstrated:

Claim 31. To prove Theorem 1 for $n = 3$ it suffices to prove that

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \lessapprox 1,$$

where k_3, l_3 may assume both zero and nonzero values.

Proving the inequality in Claim 31 is the heart of the proof of Theorem 1; we will do this in the next section.

6 Concluding the proof of Theorem 1

Recall that $\widetilde{k} = (N^{-1})^{\mathsf{T}}k, \widetilde{l} = (N^{-1})^{\mathsf{T}}l$. We now define $\gamma := -\eta_1$. Then we have $\widetilde{k}_1 = k_1, \widetilde{k}_2 = \gamma k_1 + k_2$ and $\widetilde{l}_1 = l_1, \widetilde{l}_2 = \gamma l_1 + l_2$, and thus

$$\begin{aligned} \widetilde{k}_1\widetilde{l}_2 - \widetilde{k}_2\widetilde{l}_1 &= k_1l_2 - k_2l_1,\\ \widetilde{k}_1\widetilde{l}_2 + \widetilde{k}_2\widetilde{l}_1 &= k_1l_2 + k_2l_1 + 2\gamma k_1l_1. \end{aligned} \tag{32}$$

The crucial ingredient in the proof of the inequality in Claim 31 is the following inequality, and the uniformity of the bound is essential, as we will apply it to all terms of an infinite sum.

Lemma 33. *Assume that* $\left|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2\right| \neq 0$. *Then*

$$\left|\int_{\mathbb{R}^3} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| \leq \frac{C}{t}\frac{\|k\|^{3/2}\|l\|^{3/2}}{\left|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2\right|}$$

for all $t > 0$, *where* C *is a constant which does not depend on* k, l, N *(but which does depend on the already fixed cutoff function* ψ*).*

We will postpone the proof of Lemma 33 until we need to use it; Lemma 33 compels us to split the sum in Claim 31 into parts as follows. We write

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \leq$$
$$\int_{[1,2)^3}\Big(\sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3\Big)\frac{1}{\|k\|^2\|l\|^2}\left|\int_{\mathbb{R}^3} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta, \tag{34}$$

where $\sum_1$ is the sum over $|k_1l_2 + k_2l_1 + 2\gamma k_1l_1| < 1/2$; $\sum_2$ is the sum over $k_1l_2 - k_2l_1 = 0$; $\sum_3$ is the sum over $|k_1l_2 + k_2l_1 + 2\gamma k_1l_1| \geq 1/2$ and $k_1l_2 - k_2l_1 \neq 0$, and where all sums range over $k, l \in \mathbb{Z}^3(\mathcal{U}(t))$ such that $k_1, k_2, l_1, l_2 \neq 0$.

The following lemma shows that we may neglect the sums $\sum_1$ and $\sum_2$:

Lemma 35. *For any* $|\gamma| \geq 1$, *we have*

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \frac{1}{\|k\|^2\|l\|^2} \lessapprox 1,$$

where the asymptotic constant is independent of γ, *and*

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0\\ |k_1l_2-k_2l_1|=0}} \frac{1}{\|k\|^2\|l\|^2} \lessapprox 1.$$

Proof. We obtain the second sum by substituting $k_2 \mapsto -k_2$ and $\gamma = 0$ in the first sum. Thus it suffices to bound the first sum in the cases $|\gamma| \geq 1$ and $\gamma = 0$. We will treat both

cases simultaneously. We have

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \frac{1}{\|k\|^2\|l\|^2} \ll$$

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \frac{1}{(\|(k_1,k_2)\|+|k_3|)^2(\|(l_1,l_2)\|+|l_3|)^2} \ll$$

$$\sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \int_0^{\mathcal{U}(t)}\int_0^{\mathcal{U}(t)} \frac{dk_3\,dl_3}{(\|(k_1,k_2)\|+|k_3|)^2(\|(l_1,l_2)\|+|l_3|)^2} \ll$$

$$\sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \frac{1}{\|(k_1,k_2)\|\|(l_1,l_2)\|} \le \sum_{\substack{1\le|a|,|b|,|x|,|y|\le\mathcal{U}(t)\\ bx-ay=[2\gamma ab]}} \frac{1}{\|(a,b)\|\|(x,y)\|} \le$$

$$\sum_{r=1}^{\mathcal{U}(t)} \sum_{\substack{1\le|a|,|b|\le\mathcal{U}(t)\\ \gcd(a,b)=1}} \sum_{\substack{1\le|x|,|y|\le\mathcal{U}(t)\\ bx-ay=[2\gamma r^2ab]/r}} \frac{1}{r\|(a,b)\|\|(x,y)\|},$$

where we have used the notation $[x]$ for the integer nearest to $x\in\mathbb{R}$, where we round away from zero if there is an ambiguity.

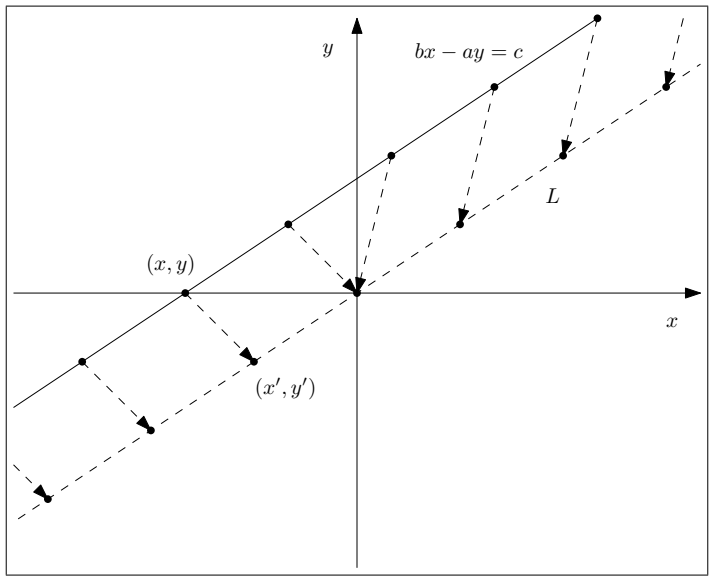
Consider the innermost sum, in which a,b,r are fixed, and let $c:=[2\gamma r^2ab]/r$. Now, since $\gcd(a,b)=1$, the equation $bx-ay=c$ has the set of solutions $(x,y)=(x_0,y_0)+m(a,b), m\in\mathbb{Z}$, granted there exists some solution $(x_0,y_0)\in\mathbb{Z}^2$. For each solution (x,y) we will define (x',y') to be the integer vector on the line L spanned by (a,b) which is closest to (x,y) among all vectors (x',y') with $\|(x',y')\|\le\|(x,y)\|$; if there is an ambiguity, choose the shorter vector (x',y'). See Figure 1. We see that the set of solutions $(x,y)\in\mathbb{Z}^2$ maps to the set of vectors $(x',y')=m(a,b), m\in\mathbb{Z}$, with at most two vectors (x,y) mapping to any given (x',y'). Now we will bound $1/\|(x,y)\|$ by $1/\|(x',y')\|=1/(m\|(a,b)\|)$ if $m\neq 0$, and otherwise we will use the bound $1/\|(x,y)\|\le 1/D$, where D is the distance between the line $bx-ay=c$ and the origin in $\mathbb{R}^2$. Note that the case $m=0$ cannot occur if $\gamma=0$ since we are summing over nonzero vectors only; but if $|\gamma|\ge 1$, we get $D=|c|/\|(a,b)\|\ge|2rab|/\|(a,b)\|$. We also have $|m|\le\sqrt{2}\cdot\mathcal{U}(t)$. Thus the last sum above can be bounded by

$$\sum_{r=1}^{\mathcal{U}(t)} \sum_{\substack{1\le|a|,|b|\le\mathcal{U}(t)\\ \gcd(a,b)=1}} \left(2\frac{1}{r\|(a,b)\|}\frac{\|(a,b)\|}{2rab}+4\sum_{m=1}^{\sqrt{2}\mathcal{U}(t)}\frac{1}{r\|(a,b)\|\|m(a,b)\|}\right) \ll$$

$$\sum_{r=1}^{\mathcal{U}(t)}\sum_{a=1}^{\mathcal{U}(t)}\sum_{b=1}^{\mathcal{U}(t)}\frac{1}{r^2}\frac{1}{a}\frac{1}{b}+\sum_{r=1}^{\mathcal{U}(t)}\sum_{1\le|a|,|b|\le\mathcal{U}(t)}\sum_{m=1}^{\sqrt{2}\mathcal{U}(t)}\frac{1}{r}\frac{1}{m}\frac{1}{\|(a,b)\|^2}\lessapprox 1,$$

where all the individual sums in the last expression have at worst logarithmic behavior in $\mathcal{U}(t)$, so we are done. □

Figure 1: In the proof of Lemma 35, each integer point (x, y) on the line $bx - ay = c$ is mapped to the closest integer point (x', y') on the parallel line L with shorter or equal length.



It remains to deal with the third part of (34), and for this we will need to use the integral bound from Lemma 33. First let us prove Lemma 33.

Proof of Lemma 33. We will prove the bound for the inner integral with respect to a_1 and a_2. Then the result follows by the compactness of the integration domain. Recalling (16), the integral we need to bound is

$$\int_{\mathbb{R}^2} \exp\Big(2\pi i t\Big(\pm\sqrt{a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2} \pm \sqrt{a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2}\Big)\Big) \times$$
$$\times \left(\frac{\|k\|}{\sqrt{a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2}}\right)^2 \left(\frac{\|l\|}{\sqrt{a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2}}\right)^2 \psi_1(a_1)\psi_2(a_2)\, da_1\, da_2\,.$$

We perform a variable substitution from (a_1, a_2) to (x, y) where $x := a_1\widetilde{k}_1^2 + a_2\widetilde{k}_2^2 + a_3\widetilde{k}_3^2$, $y := a_1\widetilde{l}_1^2 + a_2\widetilde{l}_2^2 + a_3\widetilde{l}_3^2$, which yields the Jacobian $1/|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|$. The integral above becomes

$$\frac{1}{|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|} \int_{\mathbb{R}^2} e^{2\pi i t(\pm\sqrt{x}\pm\sqrt{y})} \frac{\|k\|^2}{x} \frac{\|l\|^2}{y} \Psi_{k,l,N}(x, y)\, dx\, dy\,. \tag{36}$$

where we define $\Psi_{k,l,N}(x,y) := \psi_1(a_1)\psi_2(a_2)$ (noting that a_1, a_2 may be expressed in terms of x, y when a_3, η, k, l are held fixed). Since a_1, a_2, a_3 are bounded above and below throughout the support of $\psi_1\psi_2\psi_3$, it follows that $|x| \ll \|\widetilde{k}\|^2 \ll \|k\|^2$, and similarly $|x| \gg \|\widetilde{k}\|^2 \gg \|k\|^2$, throughout the support of $\Psi_{k,l,N}$. Likewise $|y| \ll \|l\|^2$ and $|y| \gg \|l\|^2$ throughout the support of $\Psi_{k,l,N}$.

We will assume without loss of generality that $\|k\| \geq \|l\|$, and use integration by parts on the inner integral of (36) with respect to x; if instead $\|k\| \leq \|l\|$ were the case, we repeat the following argument but integrate by parts instead with respect to y. An antiderivative of $e^{2\pi it\sqrt{x}}$ with respect to x is $\dfrac{e^{2\pi it\sqrt{x}}}{\pi it}\left(\sqrt{x} - \dfrac{1}{2\pi it}\right)$. Since $\psi_1\psi_2$ is the characteristic function of a rectangle, it follows that $x \mapsto \Psi_{k,l,N}(x,y)$ is the characteristic function of some interval $[b_1(y), b_2(y)]$, where the length of the interval is $\ll \|k\|^2$. Thus

$$\int_{\mathbb{R}} e^{2\pi it(\pm\sqrt{x}\pm\sqrt{y})} \frac{\|k\|^2}{x}\frac{\|l\|^2}{y}\Psi_{k,l,N}(x,y)\,dx =$$
$$\left[\frac{e^{2\pi it(\pm\sqrt{x}\pm\sqrt{y})}}{\pm\pi it}\left(\sqrt{x} - \frac{1}{2\pi it}\right)\frac{\|k\|^2}{x}\frac{\|l\|^2}{y}\right]_{x=b_1(y)}^{b_2(y)} -$$
$$\int_{b_1(y)}^{b_2(y)} \frac{e^{2\pi it(\pm\sqrt{x}\pm\sqrt{y})}}{\pm\pi it}\left(\sqrt{x} - \frac{1}{2\pi it}\right)\left(-\frac{\|k\|^2}{x^2}\right)\frac{\|l\|^2}{y}\,dx\,.$$

Using the bounds $\|k\|^2 \ll |x| \ll \|k\|^2$, we can bound the above expression by

$$2\sup_{x\in[b_1(y),b_2(y)]}\left(\frac{e^{2\pi it(\pm\sqrt{x}\pm\sqrt{y})}}{\pm\pi it}\left(\sqrt{x} - \frac{1}{2\pi it}\right)\frac{\|k\|^2}{x}\frac{\|l\|^2}{y}\right) +$$
$$|b_2(y) - b_1(y)| \times \sup_{x\in[b_1(y),b_2(y)]}\left(\frac{e^{2\pi it(\pm\sqrt{x}\pm\sqrt{y})}}{\pm\pi it}\left(\sqrt{x} - \frac{1}{2\pi it}\right)\left(-\frac{\|k\|^2}{x^2}\right)\frac{\|l\|^2}{y}\right) \ll$$
$$\sup_{x\in[b_1(y),b_2(y)]}\left(\frac{1}{t}\sqrt{x}\frac{\|k\|^2}{x}\frac{\|l\|^2}{y}\right) + |b_2(y) - b_1(y)| \times \sup_{x\in[b_1(y),b_2(y)]}\left(\frac{1}{t}\frac{\sqrt{x}}{x}\frac{\|k\|^2}{x}\frac{\|l\|^2}{y}\right) \ll$$
$$\frac{1}{t}\sqrt{\|k\|^2}\frac{\|k\|^2}{\|k\|^2}\frac{\|l\|^2}{y} + \|k\|^2\frac{1}{t}\frac{1}{\sqrt{\|k\|^2}}\frac{\|k\|^2}{\|k\|^2}\frac{\|l\|^2}{y} = 2\frac{1}{t}\|k\|\frac{\|l\|^2}{y}.$$

We finally integrate with respect to y, and use the bounds $\|l\|^2 \ll |y| \ll \|l\|^2$. Write $D := \{y \in \mathbb{R} : \Psi_{k,l,N}(x,y) = 1 \text{ for some } x \in \mathbb{R}\}$ for the domain of integration. Thus (36) is bounded by

$$\frac{1}{|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|}\|l\|^2\sup_{y\in D}\left(\frac{1}{t}\|k\|\frac{\|l\|^2}{y}\right) \ll \frac{1}{|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|}\frac{\|k\|\|l\|^2}{t} =$$
$$\frac{1}{|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|}\frac{\|k\|^{3/2}\|l\|^{3/2}}{t}\frac{\|l\|^{1/2}}{\|k\|^{1/2}} \leq \frac{1}{|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2|}\frac{\|k\|^{3/2}\|l\|^{3/2}}{t},$$

where the last inequality follows from our assumption $\|k\| \geq \|l\|$. □

Applying Lemma 33, and recalling (32), it now only remains to bound

$$\int_{[1,2)^3} \sum\nolimits_3 \frac{1}{\|k\|^2\|l\|^2} \left| \int_{\mathbb{R}^3} e^{2\pi i t \Phi_{k,l}(AN)} \psi_{k,l}(AN)\, da \right| d\eta \ll$$
$$\int_{[1,2)^3} \sum\nolimits_3 \frac{1}{\|k\|^2\|l\|^2} \frac{1}{t} \frac{\|k\|^{3/2}\|l\|^{3/2}}{\left|\widetilde{k}_1^2\widetilde{l}_2^2 - \widetilde{k}_2^2\widetilde{l}_1^2\right|}\, d\eta =$$
$$\int_{[1,2)^3} \sum\nolimits_3 \frac{1}{\|k\|^2\|l\|^2} \frac{1}{t} \frac{\|k\|^{3/2}\|l\|^{3/2}}{|k_1l_2 - k_2l_1||k_1l_2 + k_2l_1 + 2\gamma k_1l_1|}\, d\eta\,.$$

The integrand only depends on $\eta_1 = -\gamma$. Integrating with respect to η_2 and η_3, the expression above becomes

$$\int_{(-2,-1]} \sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t)) \\ k_1,k_2,l_1,l_2\neq 0 \\ |k_1l_2+k_2l_1+2\gamma k_1l_1|\geq 1/2 \\ k_1l_2-k_2l_1\neq 0}} \frac{1}{\|k\|^2\|l\|^2} \frac{1}{t} \frac{\|k\|^{3/2}\|l\|^{3/2}}{|k_1l_2 - k_2l_1||k_1l_2 + k_2l_1 + 2\gamma k_1l_1|}\, d\gamma\,.$$

We split the sum into one over k_3, l_3 and one over the other coordinates. We use the fact that $\|k\| \geq |k_3|$ if $k_3 \neq 0$, and otherwise $\|k\| \geq 1$, and likewise for l. Thus the above

expression is bounded by

$$\frac{1}{t}\int_{-2}^{-1} \sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|\ge 1/2\\ k_1l_2-k_2l_1\neq 0}} \frac{1}{|k_1l_2-k_2l_1||k_1l_2+k_2l_1+2\gamma k_1l_1|}\,d\gamma \times$$

$$\times\left(1+\sum_{1\le|k_3|\le\mathcal{U}(t)}\frac{1}{|k_3|^{1/2}}\right)\left(1+\sum_{1\le|l_3|\le\mathcal{U}(t)}\frac{1}{|l_3|^{1/2}}\right)\ll$$

$$\frac{\left((\mathcal{U}(t))^{1/2}\right)^2}{t}\int_{-2}^{-1} \sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|\ge 1/2\\ k_1l_2-k_2l_1\neq 0}} \frac{1}{|k_1l_2-k_2l_1||k_1l_2+k_2l_1+2\gamma k_1l_1|}\,d\gamma \lessapprox$$

$$\int_{-2}^{-1} \sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|\ge 1/2\\ k_1l_2-k_2l_1\neq 0}} \frac{1}{|k_1l_2-k_2l_1||k_1l_2+k_2l_1+2\gamma k_1l_1|}\,d\gamma \le \tag{37}$$

$$\int_{-2}^{-1}\sum_{r=1}^{\mathcal{U}(t)} \sum_{\substack{1\le|k_1|,|k_2|,|l_1|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1r|\ge 1/(2r)\\ k_1l_2-k_2l_1\neq 0\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|k_1l_2-k_2l_1||k_1l_2+k_2l_1+2\gamma k_1l_1r|}\,d\gamma \le$$

$$\int_{-2}^{-1}\sum_{r=1}^{\mathcal{U}(t)}\sum_{1\le|w|\le 2\mathcal{U}(t)^2}\sum_{\substack{1\le|k_1|,|l_1|\le\mathcal{U}(t)\\ \gcd(k_1,l_1)=1}}\sum_{\substack{1\le|k_2|,|l_2|\le\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1r|\ge 1/(2r)\\ k_1l_2-k_2l_1=w}} \frac{1}{r^2|w||k_1l_2+k_2l_1+2\gamma k_1l_1r|}\,d\gamma\,. \tag{38}$$

Consider the innermost sum, where k_1, l_1, γ, w, r are fixed. Since $\gcd(k_1,l_1)=1$ inside the sum, it follows that the equation $k_1l_2-k_2l_1=w$ has the set of solutions $(k_2,l_2)=(x_0,y_0)+m(k_1,l_1), m\in\mathbb{Z}$, granted there exists some solution $(x_0,y_0)\in\mathbb{Z}^2$. Therefore $k_1l_2+k_2l_1+2\gamma k_1l_1r$ assumes the values $c_0+2k_1l_1m$ for $m\in\mathbb{Z}$ as (k_2,l_2) varies, where $c_0:=k_1y_0+l_1x_0+2\gamma k_1l_1r$ is constant. In particular, $k_1l_2+k_2l_1+2\gamma k_1l_1r$ assumes consecutive values spaced a distance $2|k_1l_1|$ apart, with at most two values smaller than $2|k_1l_1|$ in absolute value, and the number of values it assumes is $\le 2\mathcal{U}(t)$. It follows that

the expression (38) above is

$$\ll \int_{-2}^{-1} \sum_{r=1}^{\mathcal{U}(t)} \sum_{1\le |w|\le 2\mathcal{U}(t)^2} \sum_{\substack{1\le |k_1|,|l_1|\le \mathcal{U}(t)\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|} \times$$

$$\times \left(\sum_{1\le |m|\le \mathcal{U}(t)} \frac{1}{2|mk_1l_1|} + \sum_{\substack{1\le |k_2|,|l_2|\le \mathcal{U}(t)\\ \frac{1}{2r}\le |k_1l_2+k_2l_1+2\gamma k_1l_1r|<2|k_1l_1|\\ k_1l_2-k_2l_1=w}} \frac{1}{|k_1l_2+k_2l_1+2\gamma k_1l_1r|} \right) d\gamma\,.$$

We expand this into a sum of two terms. We have

$$\int_{-2}^{-1} \sum_{r=1}^{\mathcal{U}(t)} \sum_{1\le |w|\le 2\mathcal{U}(t)^2} \sum_{\substack{1\le |k_1|,|l_1|\le \mathcal{U}(t)\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|} \sum_{1\le |m|\le \mathcal{U}(t)} \frac{1}{2|mk_1l_1|}\, d\gamma \lessapprox 1,$$

which takes care of the first term. It remains to bound

$$\int_{-2}^{-1} \sum_{\substack{1\le r,|k_1|,|l_1|\le \mathcal{U}(t)\\ 1\le |w|\le 2\mathcal{U}(t)^2\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|} \sum_{\substack{1\le |k_2|,|l_2|\le \mathcal{U}(t)\\ \frac{1}{2r}\le |k_1l_2+k_2l_1+2\gamma k_1l_1r|<2|k_1l_1|\\ k_1l_2-k_2l_1=w}} \frac{1}{|k_1l_2+k_2l_1+2\gamma k_1l_1r|}\, d\gamma\,.$$

We may without loss of generality assume that $k_1l_2+k_2l_1+2\gamma k_1l_1r$ is positive in the innermost sum, since we obtain the opposite case by switching the signs of k_1, k_2, w. Moreover, we may extend the sum to range over all $(k_2,l_2)\in\mathbb{Z}^2$. It thus suffices to bound

$$\sum_{\substack{1\le r,|k_1|,|l_1|\le \mathcal{U}(t)\\ 1\le |w|\le 2\mathcal{U}(t)^2\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|} \int_{-2}^{-1} \sum_{\substack{(k_2,l_2)\in\mathbb{Z}^2\\ \frac{1}{2r}\le (k_1l_2+k_2l_1+2\gamma k_1l_1r)<2|k_1l_1|\\ k_1l_2-k_2l_1=w}} \frac{1}{(k_1l_2+k_2l_1+2\gamma k_1l_1r)}\, d\gamma\,.$$

In the innermost sum, which is a sum over precisely one pair (k_2,l_2), and where k_1,l_1,γ,w,r are fixed, denote by $f(\gamma)$ the unique positive value in $[1/(2r),|2k_1l_1|)$ which $k_1l_2+k_2l_1+2\gamma k_1l_1r$ assumes as (k_2,l_2) varies, if it exists, or let $f(\gamma)$ be undefined otherwise. Then $f(\gamma)=c+2\gamma k_1l_1r \pmod{2|k_1l_1|}$ on its domain of definition, where $c=k_1y_0+l_1x_0$ is a constant, so $f(\gamma)$ coincides with a sawtooth wave with slope $2k_1l_1r$ and period $1/r$, except that it is undefined where the sawtooth wave has a value in $[0,1/(2r))$. Now we can partition $(-2,1]\cap \mathrm{dom}(f)$ into at most $r+1$ subintervals I_m such that f is linear on each. The integral of $1/f(\gamma)$ with respect to γ on any such subinterval I_m is

$$\int_{I_m} \frac{d\gamma}{f(\gamma)} = \left[\frac{\log|k_1l_2+k_2l_1+2\gamma k_1l_1r|}{2k_1l_1r}\right]_{\gamma=\inf I_m}^{\sup I_m} \ll$$
$$\frac{\log\left|(2+4r)\mathcal{U}(t)^2\right| + \left|\log\frac{1}{2r}\right|}{|2k_1l_1r|} \lessapprox \frac{\log r}{|k_1l_1r|},$$

where the asymptotic constants are independent of m. We now get

$$\sum_{\substack{1\le r,|k_1|,|l_1|\le \mathcal{U}(t)\\ 1\le|w|\le 2\mathcal{U}(t)^2\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|}\sum_{m=1}^{r+1}\int_{I_m}\frac{d\gamma}{f(\gamma)} \lessapprox$$

$$\sum_{\substack{1\le r,|k_1|,|l_1|\le \mathcal{U}(t)\\ 1\le|w|\le 2\mathcal{U}(t)^2\\ \gcd(k_1,l_1)=1}} \frac{1}{r^2|w|}\frac{(r+1)\log r}{|k_1 l_1 r|} \lessapprox 1,$$

and this completes the proof of Theorem 1 for $n = 3$. □

7 Proof of Theorem 1 for $n = 2$

We will sketch how the proof of Theorem 1 for the case $n = 3$ may be modified for the case $n = 2$.

By a decomposition of the measure on $\mathrm{GL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z})$ analogous to equation (8), it suffices to prove that

$$\sqrt{\int_{[1,2)^2}\int_{\mathbb{R}^2}|E_{AN}(t)|^2\psi(a)\,da\,d\eta} \ll t^{1/2},$$

where $\psi(a) := 4\pi|\det A|^2\psi_1(a_1)\psi_2(a_2)$ for the characteristic functions ψ_1, ψ_2 of two closed intervals contained in $(0,\infty)$, and where we use the parametrization $N = \begin{pmatrix}1 & \eta_1\\ 0 & 1\end{pmatrix}, \eta_1 \in [1,2)$, $A = \begin{pmatrix}1/\sqrt{a_1} & 0\\ 0 & 1/\sqrt{a_2}\end{pmatrix}, a_i \in (0,\infty)$.

The analog of Claim 12 in two dimensions is that it suffices to prove

$$\int_{[1,2)^2}\int_{\mathbb{R}^2}|E^\varepsilon_{AN}(t)|^2\psi(a)\,da\,d\eta \ll t,$$

for all $\varepsilon = \varepsilon(t)$ such that $\varepsilon \ge 1/t^{1/2}$, where $E^\varepsilon_X(t) := \sum_{k\neq(0,0)}\widehat{\chi_{t\Omega_X}}(k)\widehat{\rho_\varepsilon}(k)$ and $\rho_\varepsilon(x) = \varepsilon^{-2}\rho_0(x_1/\varepsilon)\rho_0(x_2/\varepsilon)$, and where as before $\rho_0 : \mathbb{R}\to\mathbb{R}$ is an even mollifier such that $|\widehat{\rho_0}(y)| \ll e^{-\sqrt{y}}$ for large y.

Next, to estimate the behavior of E^ε_X, we begin by considering the Fourier transform of the characteristic function χ_Ω of the standard unit ball in $\mathbb{R}^2$. It equals (see equation 11 in chapter 6.4 of [SS03])

$$\widehat{\chi_\Omega}(k) = 2\pi\int_0^1 J_0(2\pi\|k\|r)r\,dr,$$

where we have written J_α for the Bessel function of the first kind of order α. Integrating the Taylor series of J_0 (see equation 9.1.10 of [AS64]) term by term, we obtain

$$\widehat{\chi_\Omega}(k) = \frac{J_1(2\pi\|k\|)}{\|k\|}.$$

Using the asymptotics $J_1(x) = \sqrt{\frac{2}{\pi x}}\cos(x - 3\pi/4) + O(x^{-3/2})$ for large x (see equation 9.2.1 of [AS64]), we obtain

$$\widehat{\chi_\Omega}(k) = \frac{\cos(2\pi\|k\| - \frac{3\pi}{4})}{\pi\|k\|^{3/2}} + O(\|k\|^{-5/2}),$$

so it follows, as before, that

$$\widehat{\chi_{\Omega_X}}(k) = |\det X|^{-1}\frac{\cos(2\pi\|k\|_X - \frac{3\pi}{4})}{\pi\|k\|_X^{3/2}} + O(\|k\|^{-5/2})$$

where we have defined $\|k\|_X := \|(X^{-1})^{\mathsf{T}}k\|$.

Since $E_X^\varepsilon(t) = \sum_{k\neq(0,0)}\widehat{\chi_{t\Omega_X}}(k)\widehat{\rho_\varepsilon}(k) = \sum_{k\neq(0,0)} t^2\widehat{\chi_{\Omega_X}}(tk)\widehat{\rho}(\varepsilon k)$, we obtain, as before,

$$\begin{aligned} E_X^\varepsilon(t) &= |\det X|^{-1}\sum_{k\neq(0,0)}\left(\frac{t^2}{t^{3/2}}\frac{\cos(2\pi\|tk\|_X - \frac{3\pi}{4})}{\pi\|k\|_X^{3/2}} + \frac{t^2}{t^{5/2}}O(\|k\|^{-5/2})\right)\widehat{\rho}(\varepsilon k) \\ &= |\det X|^{-1}t^{1/2}\left(\sum_{k\neq(0,0)}\frac{\cos(2\pi t\|k\|_X - \frac{3\pi}{4})}{\pi\|k\|_X^{3/2}}\widehat{\rho}(\varepsilon k)\right) + O(1). \end{aligned}$$

Writing $\cos(x) = (e^{ix} + e^{-ix})/2$ and squaring E_X^ε, it follows, analogous to Claim 14, since $\widehat{\rho}$ is real-valued, that it suffices to show that

$$\sum_{k,l\neq(0,0)}\frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^{3/2}\|l\|^{3/2}}|I_{k,l}(t)| \ll 1, \tag{39}$$

for all $\varepsilon = \varepsilon(t)$ such that $\varepsilon \geq 1/t^{1/2}$, where

$$\begin{aligned} I_{k,l}(t) &:= \int_{[1,2)^2}\int_{\mathbb{R}^2} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\,d\eta, \\ \Phi_{k,l}(AN) &:= \pm\|k\|_{AN} \pm \|l\|_{AN}, \\ \psi_{k,l}(AN) &:= \left(\frac{\|k\|}{\|k\|_{AN}}\right)^{3/2}\left(\frac{\|l\|}{\|l\|_{AN}}\right)^{3/2}\psi_1(a_1)\psi_2(a_2), \end{aligned}$$

for all four choices of signs in the definition of $\Phi_{k,l}$.

The rest of the proof consists of bounding different parts of the sum (39). Doing this for $n = 2$ amounts to repeating the arguments for $n = 3$ with the difference that now k, l instead range over $\mathbb{Z}^2$ and that the exponents of $\|k\|$ and $\|l\|$ in (39) are $3/2$ instead of 2. Many of the bounds are improved in the case $n = 2$, the majority of them becoming $o(1)$, but we remark that we cannot do better than $\Theta(1)$, since with $k = l = (0, 1)$ and choosing $\Phi_{k,l}(AN) = \|k\|_{AN} - \|l\|_{AN}$ we get $I_{k,l}(t) = \Theta(1)$. In contrast, most of the bounds fail for $n \geq 4$ using the exact method above; the technical reason being that the exponents of $\|k\|, \|l\|$ for $k, l \in \mathbb{Z}^n$ in the analog of (39) become $(n + 1)/2$, whereas we would need the exponents to be roughly of the order n to get our desired bounds. We will now proceed to bound the sum (39).

7.1 Neglecting integer vectors with large coordinates

We imitate the proof of Lemma 18. Let as before $\mathcal{U}(t) = 32t\log^2 t$. Since $\varepsilon = \varepsilon(t) \geq 1/t^{1/2} \geq 1/t$ we have

$$\sum_{\substack{k\neq(0,0)\\ \|k\|\geq\mathcal{U}(t)}} \frac{e^{-\sqrt{\|\varepsilon k\|}}}{\|k\|^{3/2}} \ll \int_{\mathcal{U}(t)/2}^{\infty} \frac{e^{-\sqrt{r/t}}}{r^{3/2}} r\,dr \leq \int_{\mathcal{U}(t)/2}^{\infty} e^{-\sqrt{r/t}}\,dr \ll t^{-3}\log^2 t$$

where the last inequality is the same as inequality (21). We also have

$$\sum_{l\neq(0,0)} \frac{1}{\|l\|^{3/2}} \ll \int_1^{\mathcal{U}(t)} \frac{1}{r^{3/2}} r\,dr \leq \int_1^{\mathcal{U}(t)} dr \ll \mathcal{U}(t) \ll t\log^2 t.$$

As in the proof of Lemma 18, it follows that

$$\sum_{\substack{k,l\neq(0,0)\\ \|k\|\geq\mathcal{U}(t)}} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^{3/2}\|l\|^{3/2}} |I_{k,l}(t)| \ll \sum_{l\neq(0,0)} \frac{1}{\|l\|^{3/2}} \sum_{\substack{k\neq(0,0)\\ \|k\|\geq\mathcal{U}(t)}} \frac{e^{-\sqrt{\|\varepsilon k\|}}}{\|k\|^{3/2}} \ll 1$$

due to the rapid decay of ρ. Consequently it suffices to restrict the sum (39) to the terms for which $|k_j|, |l_j| \leq \mathcal{U}(t)$ for $j = 1, 2$.

7.2 Neglecting integer vectors with vanishing coordinates

We have

$$\sum_{\substack{k\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_1=0}} \frac{1}{\|k\|^{3/2}} \ll \int_1^{\mathcal{U}(t)} \frac{1}{r^{3/2}}\,dr \ll 1,$$

which, as in Lemma 25, implies that we may neglect the terms of the sum (39) where both some coordinate of k and some coordinate of l is zero.

Next, to show that we may neglect the terms for which precisely one of k_1 or l_1 is zero, we need to modify the proof of Lemma 27. The proof may be repeated verbatim up until ψ_0 is defined, which should be changed to $\psi_0(a_1) := \frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|_{AN}^{3/2}\|l\|_{AN}^{3/2}}$, the derivative of which is bounded since it is

$$\psi_0'(a_1) = -\frac{3}{4}\frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|_{AN}^{3/2}\|l\|_{AN}^{3/2}} \frac{l_1^2}{\|l\|_{AN}^2} \ll \frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|^{3/2}\|l\|^{3/2}} \frac{\|l\|^2}{\|l\|^2} = 1.$$

This yields, as in the proof of Lemma 27, that $|I_{k,l}(t)| \ll \frac{1}{t}\frac{\|l\|}{|l_1|}$ where the asymptotic constant is independent of k, l, and thus

$$\sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_1=0, k_2, l_1, l_2\neq 0}} \frac{|\widehat{\rho}(\varepsilon k)\widehat{\rho}(\varepsilon l)|}{\|k\|^{3/2}\|l\|^{3/2}} |I_{k,l}(t)| \ll \frac{1}{t} \sum_{\substack{k\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_1=0}} \frac{1}{\|k\|^{3/2}} \sum_{\substack{l\in\mathbb{Z}^2(\mathcal{U}(t))\\ l_1, l_2\neq 0}} \frac{1}{|l_1|\|l\|^{1/2}} \ll$$

$$\frac{1}{t}\cdot 1 \cdot \log(\mathcal{U}(t)) \cdot \sqrt{\mathcal{U}(t)} \ll 1.$$

Next, to show that we may neglect the terms for which precisely one of k_2 or l_2 is zero, we need to modify the proof of Lemma 29. Again the proof can be repeated verbatim up until the point where we need to show that the partial derivative $\frac{\partial}{\partial u}\frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|_{AN}^{3/2}\|l\|_{AN}^{3/2}}$ is bounded, where $u = -\eta_1^2 a_1 + a_2$. Indeed it is, for

$$\frac{\partial}{\partial u}\frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|_{AN}^{3/2}\|l\|_{AN}^{3/2}} = -\frac{3}{4}\frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|_{AN}^{3/2}\|l\|_{AN}^{3/2}}\frac{(-\eta_1^2 l_1^2 + \widetilde{l}_2^2)}{\|l\|_{AN}^2} \ll \frac{\|k\|^{3/2}\|l\|^{3/2}}{\|k\|^{3/2}\|l\|^{3/2}}\frac{\|l\|^2}{\|l\|^2} = 1.$$

This yields, as in Lemma 29, that

$$\int_{[1,2)^2} \sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0\\ |l_2-2\eta_1 l_1|\geq 1}} \frac{1}{\|k\|^{3/2}\|l\|^{3/2}}\left|\int_{\mathbb{R}^2} e^{2\pi i\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta \ll$$

$$\frac{1}{t}\int_{[1,2)^2} \sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0\\ |l_2-2\eta_1 l_1|\geq 1}} \frac{1}{\|k\|^{3/2}\|l\|^{1/2}|l_2|}\,d\eta \leq \frac{1}{t}\sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0}} \frac{1}{\|k\|^{3/2}\|l\|^{1/2}|l_2|} \ll 1.$$

Moreover, since there is at most one value that l_2 may assume in the region $|l_2 - 2\eta_1 l_1| < 1$, we also have

$$\int_{[1,2)^2} \sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0\\ |l_2-2\eta_1 l_1|< 1}} \frac{1}{\|k\|^{3/2}\|l\|^{3/2}}\left|\int_{\mathbb{R}^2} e^{2\pi i\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta \ll$$

$$\int_{[1,2)^2} \sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0\\ |l_2-2\eta_1 l_1|< 1}} \frac{1}{\|k\|^{3/2}\|l\|^{3/2}}\,d\eta \ll \int_{[1,2)^2} \sum_{\substack{k,l\in\mathbb{Z}^2(\mathcal{U}(t))\\ k_2=0\\ l_1,l_2\neq 0\\ |l_2-2\eta_1 l_1|< 1}} \frac{1}{k_1^{3/2}l_1^{3/2}}\,d\eta \ll$$

$$\sum_{k_1=1}^{\infty}\frac{1}{k_1^{3/2}}\sum_{l_1=1}^{\infty}\frac{1}{l_1^{3/2}} \ll 1.$$

Thus it remains only to prove that

$$\sum_{\substack{k,l\in\mathbb{Z}^3(\mathcal{U}(t))\\ k_1,k_2,l_1,l_2\neq 0}} \frac{1}{\|k\|^2\|l\|^2}|I_{k,l}(t)| \ll 1. \tag{40}$$

7.3 Concluding the proof of Theorem 1 for $n = 2$

We can bound (40) by

$$\int_{[1,2)^2}\Big(\sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3\Big)\frac{1}{\|k\|^{3/2}\|l\|^{3/2}}\left|\int_{\mathbb{R}^2} e^{2\pi i t\Phi_{k,l}(AN)}\psi_{k,l}(AN)\,da\right| d\eta,$$

where $\sum_1$ is the sum over $|k_1l_2 + k_2l_1 + 2\gamma k_1l_1| < 1/2$; $\sum_2$ is the sum over $k_1l_2 - k_2l_1 = 0$; $\sum_3$ is the sum over $|k_1l_2 + k_2l_1 + 2\gamma k_1l_1| \geq 1/2$ and $k_1l_2 - k_2l_1 \neq 0$, and where all sums range over all $k, l \in \mathbb{Z}^2(\mathcal{U}(t))$ such that $k_1, k_2, l_1, l_2 \neq 0$. Recall that $\gamma = -\eta_1$.

In order to show that we may neglect the sums $\sum_1$ and $\sum_2$, we need to modify the proof of Lemma 35. The sum $\sum_1$ is

$$\ll \sum_{\substack{1\leq|k_1|,|k_2|,|l_1|,|l_2|\leq\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|<1/2}} \frac{1}{\|(k_1,k_2)\|^{3/2}\|(l_1,l_2)\|^{3/2}} \leq$$

$$\sum_{r=1}^{\mathcal{U}(t)} \sum_{\substack{1\leq|a|,|b|\leq\mathcal{U}(t)\\ \gcd(a,b)=1}} \sum_{\substack{1\leq|x|,|y|\leq\mathcal{U}(t)\\ bx-ay=[2\gamma r^2ab]/r}} \frac{1}{r^{3/2}\|(a,b)\|^{3/2}\|(x,y)\|^{3/2}},$$

As in the proof of Lemma 35, the sum over (x, y) can be bounded, up to constants, by a sum where we replace each vector (x, y) by $m(a, b)$ for $|m| \leq \sqrt{2}\mathcal{U}(t)$, $m \neq 0$, with one additional term where we replace $1/\|(x, y)\|$ by $\|(a, b)\|/|2rab|$, and thus the expression above is

$$\ll \sum_{r=1}^{\mathcal{U}(t)} \sum_{\substack{1\leq|a|,|b|\leq\mathcal{U}(t)\\ \gcd(a,b)=1}} \left(\frac{\|(a,b)\|^{3/2}}{r^{3/2}\|(a,b)\|^{3/2}|2rab|^{3/2}} + \sum_{m=1}^{\sqrt{2}\mathcal{U}(t)} \frac{1}{r^{3/2}\|(a,b)\|^{3/2}\|m(a,b)\|^{3/2}} \right) \ll$$

$$\sum_{r=1}^{\mathcal{U}(t)} \sum_{a=1}^{\mathcal{U}(t)} \sum_{b=1}^{\mathcal{U}(t)} \frac{1}{r^3} \frac{1}{a^{3/2}} \frac{1}{b^{3/2}} + \sum_{r=1}^{\mathcal{U}(t)} \sum_{1\leq|a|,|b|\leq\mathcal{U}(t)} \sum_{m=1}^{\sqrt{2}\mathcal{U}(t)} \frac{1}{r^{3/2}m^{3/2}\|(a,b)\|^3} \ll 1,$$

and thus $\sum_1 \ll 1$. As in the proof of Lemma 35, the above argument can be repeated verbatim (with the substitutions $k_2 \mapsto -k_2$ and $\gamma \mapsto 0$) to prove that also $\sum_2 \ll 1$.

It thus remains only to deal with $\sum_3$. Lemma 33 still holds for $n = 2$ (when integrating instead over $\mathbb{R}^2$). Applying Lemma 33 to the sum $\sum_3$, we get

$$\int_{[1,2)} \sum\nolimits_3 \frac{1}{\|k\|^{3/2}\|l\|^{3/2}} \left| \int_{\mathbb{R}^2} e^{2\pi i t\Phi_{k,l}(AN)} \psi_{k,l}(AN)\, da \right| d\eta \leq$$

$$\frac{C}{t} \int_{-2}^{-1} \sum_{\substack{1\leq|k_1|,|k_2|,|l_1|,|l_2|\leq\mathcal{U}(t)\\ |k_1l_2+k_2l_1+2\gamma k_1l_1|\geq1\\ k_1l_2-k_2l_1\neq0}} \frac{1}{|k_1l_2 - k_2l_1||k_1l_2 + k_2l_1 + 2\gamma k_1l_1|}\, d\gamma,$$

where C is a constant which does not depend on any of k, l, N (but which does depend on ψ), but this is precisely $\frac{C}{t}$ multipled by the expression (37) on page 22, which we have already proved is $\lessapprox 1$ as part of the proof for $n = 3$. Thus the expression above is $\lessapprox \frac{1}{t}$, so it is $\ll 1$. This completes the proof of Theorem 1 for $n = 2$. □

8 Proof of Proposition 3 and Corollary 4

Denote by

$$\mathbb{E}_1[f(X)] := \int_{\mathrm{SL}_n(\mathbb{R})/\,\mathrm{SL}_n(\mathbb{Z})} f(X)\, d\mu_1(X)$$

the mean value of f over the set of all lattices with unit determinant, where μ_1 is the normalized Haar measure on $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$. We quote the mean value formulas of Siegel and Rogers (see [Sie45] and Theorem 4 in [Rog55]).

Theorem 41 (Siegel's mean value formula)**.** *Suppose that $n \geq 2$. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be an integrable function, and let $\Lambda := X\mathbb{Z}^n$ for $X \in \mathrm{SL}_n(\mathbb{R})$. Then*

$$\mathbb{E}_1\left[\sum_{u\in\Lambda} \rho(u)\right] = \int_{\mathbb{R}^n} \rho(x)\,dx + \rho(0).$$

Theorem 42 (Rogers's mean value formula)**.** *Suppose that $n \geq 3$. Let $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a non-negative Borel-measurable function, and let $\Lambda := X\mathbb{Z}^n$ for $X \in \mathrm{SL}_n(\mathbb{R})$. Then*

$$\mathbb{E}_1\left[\sum_{u,v\in\Lambda} \rho(u,v)\right] = \iint_{\mathbb{R}^n\times\mathbb{R}^n} \rho(x,y)\,dx\,dy + \rho(0,0) +$$
$$2\sum_{q=1}^{\infty} \sum_{\substack{r\geq 1\\ \gcd(q,r)=1}} \frac{1}{q^n} \int_{\mathbb{R}^n} \left(\rho\Big(x, \frac{q}{r}x\Big) + \rho\Big(\frac{q}{r}x, x\Big)\right) dx\,.$$

Proof of Proposition 3. Taking $\rho(u) := \chi_{t\Omega}(u)$ in Siegel's mean value formula, we obtain

$$\mathbb{E}_1[N_X(t)] = \mathrm{vol}(t\Omega) + 1,$$

and taking $\rho(u,v) := \chi_{t\Omega}(u)\chi_{t\Omega}(v)$ in Rogers's mean value formula, we obtain

$$\mathbb{E}_1\Big[N_X(t)^2\Big] = \mathrm{vol}(t\Omega)^2 + 1 + 4\sum_{q=1}^{\infty} \sum_{\substack{r\geq 1\\ \gcd(q,r)=1}} \frac{1}{q^n} \int_{\mathbb{R}^n} \chi_{t\Omega}(x)\chi_{t\Omega}\Big(\frac{q}{r}x\Big)\,dx,$$

so that

$$\mathbb{E}_1\Big[N_X(t)^2\Big] - (\mathrm{vol}(t\Omega)^2 + 1) = 4 \sum_{\substack{q,r\geq 1\\ \gcd(q,r)=1}} \frac{1}{(qr)^n} \int_{\mathbb{R}^n} \chi_{t\Omega}(qx)\chi_{t\Omega}(rx)\,dx =$$
$$4 \sum_{\substack{q,r\geq 1\\ \gcd(q,r)=1}} \frac{1}{(qr)^n}\,\mathrm{vol}\Big(\frac{t}{\max(q,r)}\Omega\Big) = \sum_{\substack{q,r\geq 1\\ \gcd(q,r)=1}} \frac{4\,\mathrm{vol}(t\Omega)}{(qr)^n \max(q,r)^n} =: c_n\,\mathrm{vol}(t\Omega),$$

where $c_n \geq 4$ is a constant (which is clearly convergent for $n \geq 2$). Thus we have

$$\mathbb{E}_1\Big[E_X(t)^2\Big] = \mathbb{E}_1\Big[(N_X(t) - \mathrm{vol}(t\Omega))^2\Big] =$$
$$\mathbb{E}_1\Big[N_X(t)^2\Big] - 2\,\mathrm{vol}(t\Omega)\mathbb{E}_1[N_X(t)] + \mathrm{vol}(t\Omega)^2 =$$
$$c_n\,\mathrm{vol}(t\Omega) + 1 - 2\,\mathrm{vol}(t\Omega) = 1 + (c_n - 2)\,\mathrm{vol}(\Omega)t^n = \Theta(t^n),$$

so $\sqrt{\mathbb{E}_1\Big[|E_X(t)|^2\Big]} = \Theta(t^{n/2})$. This completes the proof of Proposition 3. □

Proof of Corollary 4. We identify $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ with $\mathrm{GL}_n^+(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$, where $\mathrm{GL}_n^+(\mathbb{R})$ is the subset of $\mathrm{GL}_n(\mathbb{R})$ consisting of matrices with positive determinant, and use the decomposition $\mathrm{GL}_n^+(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z}) = (\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})) \cdot \mathcal{D}$, where $\mathcal{D} = \{rI : r > 0\}$ is the set of positive multiples of the identity matrix I. We identify the Haar measure on $\mathrm{GL}_n^+(\mathbb{R})/\mathrm{SL}_n(\mathbb{R})$ with the Haar measure μ on $\mathrm{GL}_n(\mathbb{R})$, which is well-known to be bi-invariant. The Haar measure dr/r on $\mathcal{D}$ is bi-invariant as well since $\mathcal{D}$ is commutative. Thus the modular functions on these topological groups are identically 1 (see [Kna02]). Consequently, Theorem 8.32 from [Kna02] implies that

$$\int_{a\le|\det X|\le b} |E_X(t)|^2 \, d\mu(X) = \int_{\substack{rI\in\mathcal{D}\\ a\le r^n\le b}} \int_{\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})} |E_{rX}(t)|^2 \, d\mu_1(X) \frac{dr}{r}.$$

We have $E_{rX}(t) = E_X(t/r)$ for any $r > 0$, so the inner integral can be written as $\mathbb{E}_1\left[|E_X(t/r)|^2\right]$. Using the bounds from Proposition 3 on the inner integral, and bounding the outer integral trivially, we get

$$\int_{L_{a,b}} |E_X(t)|^2 \, d\mu(X) = \Theta(t^n). \qquad \square$$

Acknowledgements

I would like to thank my advisor Pär Kurlberg for suggesting this problem to me and for all his help and encouragement.

References

[AS64] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[Ble92] Pavel Bleher. On the distribution of the number of lattice points inside a family of convex ovals. *Duke Math. J.*, 67(3):461–481, 1992.

[Göt04] Friedrich Götze. Lattice point problems and values of quadratic forms. *Invent. Math.*, 157(1):195–226, 2004.

[Har17] G. H. Hardy. The Average Order of the Arithmetical Functions P(x) and delta(x). *Proc. London Math. Soc.*, S2-15(1):192, 1917.

[HB99] D. R. Heath-Brown. Lattice points in the sphere. In *Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997)*, pages 883–892. de Gruyter, Berlin, 1999.

[HIW04] S. Hofmann, A. Iosevich, and D. Weidinger. Lattice points inside random ellipsoids. *Michigan Math. J.*, 52(1):13–21, 2004.

[Hux03] M. N. Huxley. Exponential sums and lattice points. III. *Proc. London Math. Soc. (3)*, 87(3):591–609, 2003.

[IKKN06] A. Ivić, E. Krätzel, M. Kühleitner, and W. G. Nowak. Lattice points in large regions and related arithmetic functions: recent developments in a very classic topic. In *Elementare und analytische Zahlentheorie*, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main, 20, pages 89–128. Franz Steiner Verlag Stuttgart, Stuttgart, 2006.

[Ing33] A. E. Ingham. A Note on Fourier Transforms. *J. London Math. Soc.*, S1-9(1):29, 1933.

[ISS02] Alexander Iosevich, Eric Sawyer, and Andreas Seeger. Mean square discrepancy bounds for the number of lattice points in large convex bodies. *J. Anal. Math.*, 87:209–230, 2002. Dedicated to the memory of Thomas H. Wolff.

[Kna02] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, second edition, 2002.

[Krä00] Ekkehard Krätzel. *Analytische Funktionen in der Zahlentheorie*, volume 139 of *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*. B. G. Teubner, Stuttgart, 2000.

[Lan24] E. Landau. Über die anzahl der gitterpunkte in gewissen bereichen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1924:137–150, 1924.

[Mül99] Wolfgang Müller. Lattice points in large convex bodies. *Monatsh. Math.*, 128(4):315–330, 1999.

[Now85a] Werner Georg Nowak. An Ω-estimate for the lattice rest of a convex planar domain. *Proc. Roy. Soc. Edinburgh Sect. A*, 100(3-4):295–299, 1985.

[Now85b] Werner Georg Nowak. On the lattice rest of a convex body in $\mathbf{R}^s$. *Arch. Math. (Basel)*, 45(3):284–288, 1985.

[PT02] Y. Petridis and J. A. Toth. The remainder in Weyl's law for random two-dimensional flat tori. *Geom. Funct. Anal.*, 12(4):756–775, 2002.

[Rog55] C. A. Rogers. Mean values over the space of lattices. *Acta Math.*, 94:249–287, 1955.

[Sch60] Wolfgang M. Schmidt. A metrical theorem in geometry of numbers. *Trans. Amer. Math. Soc.*, 95:516–529, 1960.

[Sie45] Carl Ludwig Siegel. A mean value theorem in geometry of numbers. *Ann. of Math. (2)*, 46:340–347, 1945.

[SS03] Elias M. Stein and Rami Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.

[Ste93] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

Paper C

MISSING CLASS GROUPS AND CLASS NUMBER STATISTICS FOR IMAGINARY QUADRATIC FIELDS

S. HOLMIN, N. JONES, P. KURLBERG, C. MCLEMAN AND K. PETERSEN

ABSTRACT. The number $\mathcal{F}(h)$ of imaginary quadratic fields with class number h is of classical interest: Gauss' class number problem asks for a determination of those fields counted by $\mathcal{F}(h)$. The unconditional computation of $\mathcal{F}(h)$ for $h \leq 100$ was completed by M. Watkins, using ideas of Goldfeld and Gross-Zagier; Soundararajan has more recently made conjectures about the order of magnitude of $\mathcal{F}(h)$ as $h \to \infty$ and determined its average order. In the present paper, we refine Soundararajan's conjecture to a conjectural asymptotic formula and also consider the subtler problem of determining the number $\mathcal{F}(G)$ of imaginary quadratic fields with class group isomorphic to a given finite abelian group G. Using Watkins' tables, one can show that some abelian groups do *not* occur as the class group of any imaginary quadratic field (for instance $(\mathbb{Z}/3\mathbb{Z})^3$ does not). This observation is explained in part by the Cohen-Lenstra heuristics, which have often been used to study the distribution of the *p-part* of an imaginary quadratic class group. We combine heuristics of Cohen-Lenstra together with our refinement of Soundararajan's conjecture to make precise predictions about the asymptotic nature of the *entire* imaginary quadratic class group, in particular addressing the above-mentioned phenomenon of "missing" class groups, for the case of p-groups as p tends to infinity. Furthermore, conditionally on the Generalized Riemann Hypothesis, we extend Watkins' data, tabulating $\mathcal{F}(h)$ for odd $h \leq 10^6$ and $\mathcal{F}(G)$ for G a p-group of odd order with $|G| \leq 10^6$. (In order to do this, we need to examine the class numbers of all negative prime fundamental discriminants $-q$, for $q \leq 1.1881 \cdot 10^{15}$.) The numerical evidence matches quite well with our conjectures.

1. INTRODUCTION

Given a fundamental discriminant $d < 0$, let $H(d)$ denote the ideal class group of the imaginary quadratic field $\mathbb{Q}(\sqrt{d})$, and let $h(d) := |H(d)|$ denote the class number. A basic question is:

Question 1.1. Which finite abelian groups G occur as $H(d)$ for some negative fundamental discriminant d?

Equivalently, which finite abelian groups G do *not* occur as $H(d)$? The case where $G \simeq (\mathbb{Z}/2\mathbb{Z})^r$ has classical connections via genus theory to Euler's "idoneal numbers," and it follows from work of Chowla [8] that for every $r \gg 1$, the group $(\mathbb{Z}/2\mathbb{Z})^r$ does not occur as the class group of any imaginary quadratic field. Later work of various authors ([6], [45], [17]) has shown that $(\mathbb{Z}/n\mathbb{Z})^r$ does not occur as an imaginary quadratic class group for $r \gg 1$ and $2 \leq n \leq 6$ (in fact, Heath-Brown showed that groups with exponent 2^a or $3 \cdot 2^a$ occur only finitely many times.) Moreover, $(\mathbb{Z}/n\mathbb{Z})^r$ does not occur for $n > 6$ and $r \gg_n 1$ *assuming* the Generalized Riemann Hypothesis (cf. [6, 45]); in fact they show that the exponent of $H(d)$ tends to infinity as $d \to -\infty$.

Due to the possible existence of Siegel zeroes, the unconditional results mentioned above are ineffective. To find explicit examples of missing class groups, one can undertake a brute-force search using tables of M. Watkins [44], who used the ideas of Goldfeld and Gross-Zagier to give an unconditional resolution of Gauss' class number problem for class numbers $h \leq 100$. Such a search reveals that none of the groups

$$\left(\frac{\mathbb{Z}}{3\mathbb{Z}}\right)^3, \quad \frac{\mathbb{Z}}{9\mathbb{Z}} \times \left(\frac{\mathbb{Z}}{3\mathbb{Z}}\right)^2, \quad \left(\frac{\mathbb{Z}}{3\mathbb{Z}}\right)^4$$

occur as the class group of an imaginary quadratic field.

It is also natural to ask how common the groups that do occur are:

Question 1.2. Given a finite abelian group G, for how many fundamental discriminants $d < 0$ is $H(d) \simeq G$?

2010 *Mathematics Subject Classification.* 11R29, 11Y40.

Key words and phrases. Class numbers, class groups, Cohen-Lenstra heuristics.

In order to address this question, we are led to investigate a closely related issue:

Question 1.3. Given an integer $h > 0$, for how many fundamental discriminants $d < 0$ is $|H(d)| = h$?

Questions 1.1, 1.2, and 1.3 appear to be beyond the realm of what one can provably answer in full with current technology. In this paper, we combine the heuristics of Cohen-Lenstra with results on the distribution of special values of Dirichlet L-functions to give a conjectural asymptotic answer to Question 1.3, for h odd. (For this we only use the Cohen-Lenstra heuristic to predict divisibility properties of class numbers.) Further, using this conjectured asymptotic answer, we use the Cohen-Lenstra heuristic to predict the p-group decomposition of $H(d)$ and obtain a conjectured asymptotic answer to Question 1.2 in the case where G is a p-group for an odd prime p. (We believe that similar results hold for composite class number, though here one must be careful in how limits are taken; for instance with some groups of order $p_1^{n_1}p_2^{n_2}$, p_1 fixed and p_2 tending to infinity is very different from p_1 and p_2 both tending to infinity.) In particular, regarding Question 1.1, we establish a precise condition on the *shape* of an abelian p-group which governs whether or not it should occur as an imaginary quadratic class group for infinitely many primes p. For instance, our conjecture predicts that the group

$$\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^3$$

should appear as a class group for only finitely many primes p (in fact, quite likely for no primes p at all; cf. Conjecture 1.10 in Section 1.3.1), whereas the two groups

$$\frac{\mathbb{Z}}{p^3\mathbb{Z}}, \quad \frac{\mathbb{Z}}{p^2\mathbb{Z}} \times \frac{\mathbb{Z}}{p\mathbb{Z}}$$

should occur as a class group for infinitely many primes p.

Given a positive integer h we set

$$\mathcal{F}(h) := |\{\text{fundamental discriminants } d < 0 \ : \ h(d) = h\}|. \tag{1.1}$$

Thus for instance $\mathcal{F}(1) = 9$, which is the statement of the Baker-Stark-Heegner theorem on Gauss' class number 1 problem for imaginary quadratic fields. Given a fixed finite abelian group G, we consider the refined counting function

$$\mathcal{F}(G) := |\{\text{fundamental discriminants } d < 0 \ : \ H(d) \simeq G\}|,$$

so that $\mathcal{F}(h) = \sum_{|G|=h} \mathcal{F}(G)$, where the sum runs over isomorphism classes of finite abelian groups of order h. The Cohen-Lenstra heuristics suggest that, for any finite abelian group G of odd order h, the expected number of imaginary quadratic fields with class group G is given by

$$\mathcal{F}(G) \approx P(G) \cdot \mathcal{F}(h), \tag{1.2}$$

where

$$P(G) := \left(\frac{1}{|\operatorname{Aut}(G)|}\right) \Big/ \left(\sum_{\substack{\text{abel. groups } G' \\ \text{s.t. } |G'|=|G|}} \frac{1}{|\operatorname{Aut}(G')|}\right). \tag{1.3}$$

The first factor $P(G)$ may be evaluated explicitly, whereas the second factor $\mathcal{F}(h)$ is more delicate. K. Soundararajan has conjectured (see [40, p. 2]) that

$$\mathcal{F}(h) \asymp \frac{h}{\log h} \qquad (h \text{ odd}). \tag{1.4}$$

We refine Soundararajan's heuristic, sharpening (1.4) to a conjectural asymptotic formula, which involves certain constants associated to a random Euler product. Let $\mathbb{Y} = \{\mathbb{Y}(p) : p \text{ prime}\}$ denote a collection of independent identically distributed random variables satisfying

$$\mathbb{Y}(p) := \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

and let

$$L(1,\mathbb{Y}) := \prod_{p} \left(1 - \frac{\mathbb{Y}(p)}{p}\right)^{-1}$$

denote the corresponding random Euler product, which converges with probability one. Define the constant

$$\mathfrak{C} := 15 \prod_{\substack{\ell=3 \\ \ell \text{ prime}}}^{\infty} \prod_{i=2}^{\infty} \left(1 - \frac{1}{\ell^i}\right) \approx 11.317, \tag{1.5}$$

as well as the factor (defined for odd h)

$$\mathfrak{c}(h) := \prod_{p^n \| h} \prod_{i=1}^{n} \left(1 - \frac{1}{p^i}\right)^{-1}.$$

Conjecture 1.4. *We have*

$$\mathcal{F}(h) \sim \frac{\mathfrak{C}}{15} \cdot \mathfrak{c}(h) \cdot h \cdot \mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2 \log(\pi h / L(1,\mathbb{Y}))}\right) \sim \mathfrak{C} \cdot \mathfrak{c}(h) \cdot \frac{h}{\log(\pi h)} \tag{1.6}$$

as $h \longrightarrow \infty$ through odd values. (Here $\mathbb{E}$ denotes expected value.)

Conjecture 1.4 is developed from the Cohen-Lenstra heuristics together with large-scale distributional considerations of the special value $L(1,\chi_d)$. The former can be viewed as a product over non-archimedean primes; the latter as an archimedean factor — in a sense our prediction is a "global" (or adelic) generalization of the Cohen-Lenstra heuristic, somewhat similar to the Siegel mass formula.

More precisely, motivated by the Cohen-Lenstra heuristic we introduce a correction factor that considers divisibility of h by a random odd positive integer (for instance a random class number is divisible by 3 with conjectural probability

$$1 - \prod_{i=1}^{\infty} \left(1 - \frac{1}{3^i}\right) \approx 43\%,$$

and this suggests a correction factor of $\left(1 - \prod_{i=1}^{\infty} \left(1 - \frac{1}{3^i}\right)\right) / (1/3)$ whenever 3 divides h). We remark that the Cohen-Lenstra heuristics have often been applied to give a probabilistic model governing the *p-part* of a class group, for a *fixed* prime p (see for instance [10, Section 9]). By contrast, the precise asymptotic predicted by Conjecture 1.4 involves applying these considerations for *all* primes p (including the archimedean prime).

The relevant information about the distribution of $L(1,\chi_d)$ is implicit in the following theorem, which gives the analogue of [40, Theorem 1] averaged over odd values of h.

Theorem 1.5. *Assume the Generalized Riemann Hypothesis. Then for any $\varepsilon > 0$, we have*

$$\sum_{\substack{h \le H \\ h \text{ odd}}} \mathcal{F}(h) = \frac{15}{4} \cdot \frac{H^2}{\log H} + O\left(H^2 (\log H)^{-3/2+\varepsilon}\right),$$

as $H \longrightarrow \infty$.

Remark 1.6. In fact, our analysis (cf. Section 4) yields the more accurate approximation

$$\begin{aligned} \mathcal{F}(h) &\sim \frac{\mathfrak{C}}{15} \cdot \mathfrak{c}(h) \cdot h \cdot \mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2 \log(\pi h / L(1,\mathbb{Y}))}\right) \\ &= \mathfrak{C} \cdot \mathfrak{c}(h) \cdot \frac{h}{\log(\pi h)} \cdot \left(1 + \frac{c_1}{\log(\pi h)} + \frac{c_2}{\log^2(\pi h)} + \frac{c_3}{\log^3(\pi h)} + o\left(\frac{1}{\log^3(\pi h)}\right)\right), \end{aligned} \tag{1.7}$$

where

$$\begin{aligned} c_1 &:= \frac{\pi^2}{15} \mathbb{E}\left(\frac{\log L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right) \approx -0.578, \\ c_2 &:= \frac{\pi^2}{15} \mathbb{E}\left(\frac{\log^2 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right) \approx 0.604, \\ c_3 &:= \frac{1}{c_0} \mathbb{E}\left(\frac{\log^3 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right) \approx -0.526. \end{aligned} \tag{1.8}$$

Without this higher order expansion we have a relative error of size $O(1/\log h)$; since we only have data for odd $h \leq 10^6$, the higher order expansion is essential to get a convincing fit to the observed data.

1.1. **Numerical evidence for Conjecture 1.4.** With the aid of a supercomputer and assuming GRH, we have computed $\mathcal{F}(h)$ and $\mathcal{F}(G)$ for all odd $h < 10^6$ and all p-groups G of odd size at most 10^6. For the correctness of the computation, and to obtain some important speedups, we use GRH in three ways. First, we use a recent result by Lamzouri, Li, and Soundararajan [25] in order to give an upper bound on the negative prime fundamental discrimants $d < 0$ for which $h(d) < 10^6$. Consequently it is enough to examine $h(-q)$ for all primes $q \equiv 3 \mod 4$, and $q \leq 1.1881 \cdot 10^{15}$. In particular, we must extend the class number computation [23], where Jacobson, Ramachandran, and Williams determine $h(-d)$ for $d < 10^{11}$, and $-d$ a fundamental discriminant. In order to avoid the costly full computation of the class number (especially for $-d > 10^{14}$), we use the Dirichlet class number formula $h(d) = L(1, \chi_d) \cdot \sqrt{|d|}/\pi$ in order to compute a lower bound on $h(d)$ by approximating $L(1, \chi_d)$. Assuming GRH, $L(1, \chi_d)$ is well approximated by a short truncated Euler product; to choose parameters we use some explicit GRH-conditional bounds due to Bach [1] together with a simple, but quite important, improvement (cf. Proposition 6.1.) Finally, for class groups that are far from cyclic (these are quite rare), we compute the full class group using PARI's `quadclassunit0`, an implementation of Buchmann-McCurley's sub-exponential, and GRH-conditional, algorithm. For more details regarding the computation, see Section 6.

The numerics give us quite convincing evidence in support of Conjecture 1.4. Below we give some samples[1] of computed values $\mathcal{F}(h)$ (conditional on the GRH) compared to the values predicted by Conjecture 1.4, rounded to the nearest integer. We also list the relative error $(\mathcal{F}(h) - \mathrm{pred}(h))/\mathrm{pred}(h)$ given as a percentage, where

$$\mathrm{pred}(h) := \mathfrak{C} \cdot \mathfrak{c}(h) \cdot \frac{h}{\log(\pi h)} \cdot \left(1 + \frac{c_1}{\log(\pi h)} + \frac{c_2}{\log^2(\pi h)} + \frac{c_3}{\log^3(\pi h)}\right). \tag{1.9}$$

h	10001	10003	10005	10007	10009	10011	10013	10015
$\mathcal{F}(h)$	10641	12154	20661	10536	10329	15966	12221	12975
pred(h)	10598	12116	21074	10383	10385	16144	12038	12993
Relative error	+0.41%	+0.31%	−1.96%	+1.48%	−0.54%	−1.10%	+1.52%	−0.14%
h	100001	100003	100005	100007	100009	100011	100013	100015
$\mathcal{F}(h)$	94623	85792	164289	86770	111948	142512	87138	108993
pred(h)	94213	85641	164806	86620	111210	142989	86577	108820
Relative error	+0.43%	+0.18%	−0.31%	+0.17%	+0.66%	−0.33%	+0.65%	+0.16%
h	999985	999987	999989	999991	999993	999995	999997	999999
$\mathcal{F}(h)$	1064529	1095135	771805	791007	1093645	914482	733397	1815672
pred(h)	1063376	1098842	769673	788871	1093732	911447	730673	1825811
Relative error	+0.11%	−0.34%	+0.28%	+0.27%	−0.01%	+0.33%	+0.37%	−0.56%

For large h the prediction seems fairly good as the relative error very often is smaller than 1%. To gain further insight, we study the fluctuations in the difference between the observed data and the predictions, normalized by dividing by the square root of the prediction (it is perhaps not a priori obvious, but with this normalization the resulting standard deviation is close to one in many circumstances). More precisely, we make a histogram of the values of

$$r(h) := \frac{\mathcal{F}(h) - \mathrm{pred}(h)}{\sqrt{\mathrm{pred}(h)}}$$

for various subsets of the (odd) integers. For notational convenience, we shall let μ and σ denote the mean and standard deviation, respectively, of the observed data in each plot.

[1]The complete list of computed values of $\mathcal{F}(h)$ is given in [20].

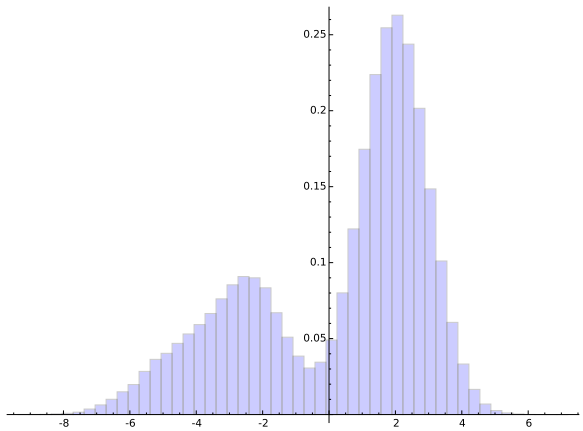


FIGURE 1. Histogram for $r(h)$, as h ranges over odd integers in $[500000, 1000000]$. $(\mu, \sigma) = (0.291561, 2.685280)$.

Interestingly, the probability distribution appears to be bimodal. A closer inspection of the table above indicates a small positive bias for h that are divisible by three. Separating out (odd) h according to divisibility by three, or not, results in the following two histograms:

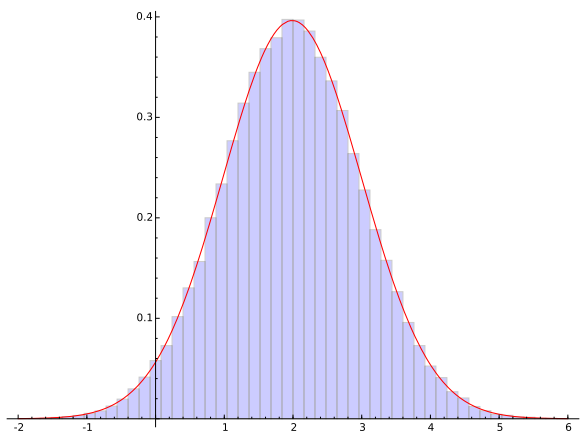


FIGURE 2. Histogram for $r(h)$, as $h \not\equiv 0 \bmod 3$ ranges over odd integers in $[500000, 1000000]$. $(\mu, \sigma) = (1.987995, 1.006428)$.

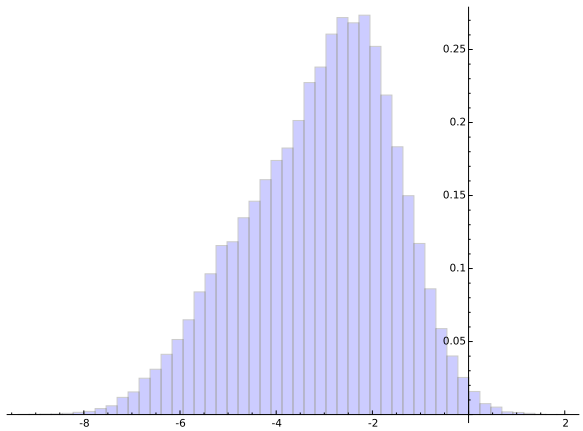


FIGURE 3. Histogram for $r(h)$, as $h \equiv 0 \mod 3$ ranges over odd integers in $[500000, 1000000]$. $(\mu, \sigma) = (-3.101265, 1.529449)$.

The curve (red in color printouts and online) in the first plot is a Gaussian probability density function with mean and standard deviation fitted to the data — the first plot appears to be Gaussian, whereas the second clearly is not.

Also note that (after our normalization), the effect of three divisibility is quite pronounced — the shift in the mean value is of order of magnitude a standard deviation.

By further separating $h \equiv 0 \mod 3$ into subsets according to the exact power of three that divides h, we obtain distributions that appear Gaussian; for comparison, we again plot a (red) curve giving the probability density function for a Gaussian random variable with the same mean and standard deviation as the observed data. (Note that there is a significant shift in the mean, whereas the standard deviation is close to one.)

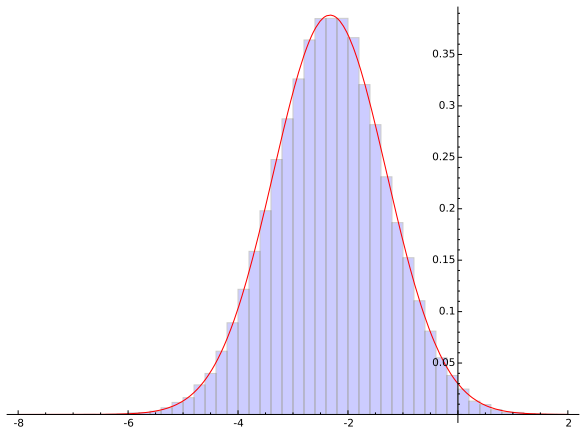


FIGURE 4. Histogram for $r(h)$, for odd h in (500000,1000000), $3||h$. $(\mu, \sigma) = (-2.326289, 1.027387)$.

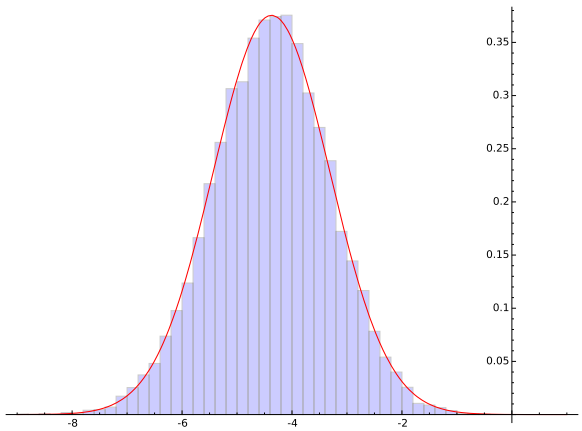


FIGURE 5. Histogram for $r(h)$, for odd h in (500000,1000000), $3^2||h$. $(\mu, \sigma) = (-4.372185, 1.062480)$.

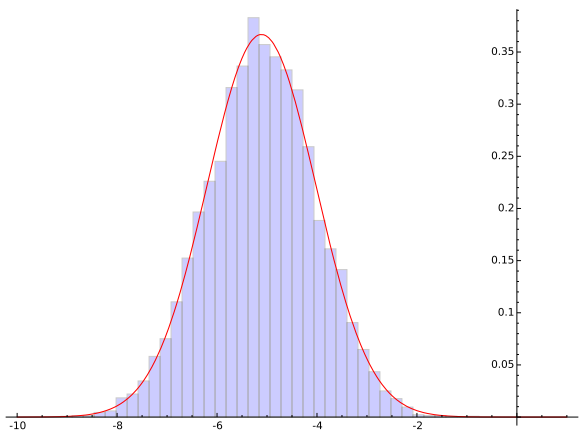


FIGURE 6. Histogram for $r(h)$, for odd h in (500000,1000000), $3^3||h$. $(\mu, \sigma) = (-5.110585, 1.087463)$.

The exact nature of this "three divisibility bias" is unclear, but inspired by the slow convergence in the Davenport-Heilbronn asymptotic[2]

$$\sum_{\substack{-X<d<0 \\ d \text{ fund. disc.}}} |H(d)[3]| \sim C \cdot X \tag{1.10}$$

(here $H(d)[3]$ denotes the 3-torsion subgroup of $H(d)$) we can slightly adjust $\mathfrak{c}(h)$ to remove most of this bias and obtain a more accurate prediction $\text{pred}'(h)$. (Essentially we examine the exact power of three divisibility

[2]In fact, a negative second order correction to (1.10) of size $X^{5/6}$ was recently obtained by T. Taniguchi and F. Thorne [42] and also independently by M. Bhargava, A. Shankar and J. Tsimmerman [4].

of h and adjust to the data, see Section 4.2 for more details.) With this adjustment, the fluctuations for

$$r'(h) := \frac{\mathcal{F}(h) - \mathrm{pred}'(h)}{\sqrt{\mathrm{pred}'(h)}}$$

(for the full set of odd h) is quite close to a Gaussian.

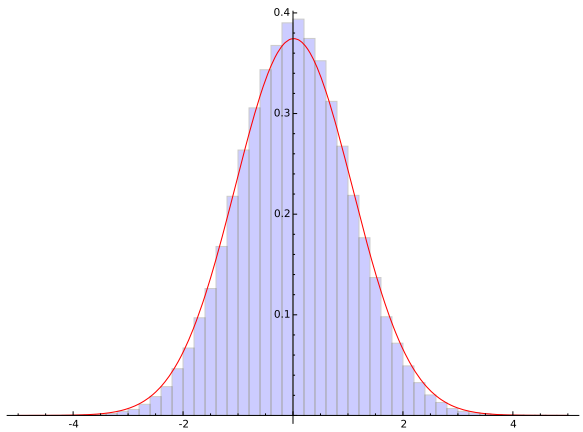


FIGURE 7. Histogram for $r'(h)$, for all odd h in (500000,1000000). $(\mu, \sigma) = (0.013214, 1.065277)$.

However, compared to the fitted Gaussian, the histogram is slightly more peaked, and has less mass in the tails. If we remove integers being divisible by 3^4 this effect is reduced and we get an improved fit to a Gaussian.

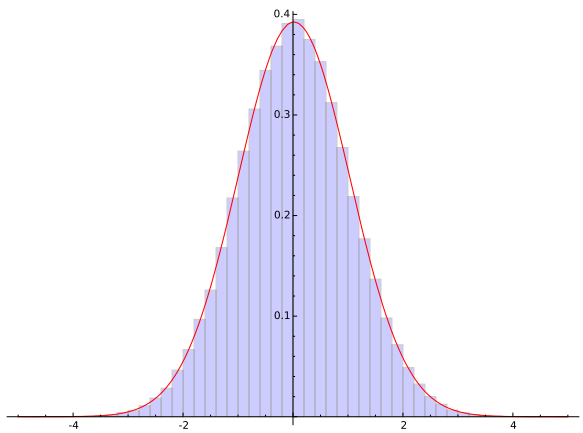


FIGURE 8. Histogram for $r'(h)$, for all odd h in (500000,1000000) and h not divisible by 81. $(\mu, \sigma) = (0.016292, 1.016726)$.

1.2. **Groups occuring as class groups.** We now return to our discussion of the quantity $\mathcal{F}(G)$. To make precise what we mean by the *shape* of an abelian p-group, recall the bijection

$$\{\text{partitions of } n\} \;\leftrightarrow\; \{\text{abelian groups of order } p^n\}$$

$$\lambda = (n_1, n_2, \dots, n_r) \mapsto G_\lambda(p) := \bigoplus_{i=1}^{r} \mathbb{Z}/p^{n_i}\mathbb{Z}.$$

Using (1.2) in conjunction with Conjecture 1.4, and evaluating each factor asymptotically, we are led to the following conjecture. Given a partition

$$\lambda = (n_1, n_2, \dots, n_r), \quad n_1 \geq n_2 \geq \cdots \geq n_r \geq 1, \quad n_1 + n_2 + \cdots + n_r = n$$

of n, define the *cyclicity index* of λ by

$$\mathrm{c}(\lambda) := \sum_{i=1}^{r} (3-2i)n_i = n_1 - \sum_{i=2}^{r} (2i-3)n_i. \tag{1.11}$$

Note that $\mathrm{c}(\lambda) \in [1-(n-1)^2, n]$ and $G_\lambda(p)$ is cyclic if and only if $\mathrm{c}(\lambda) = n$; thus $\mathrm{c}(\lambda)$ provides a measure of how much $G_\lambda(p)$ deviates from being cyclic.

Conjecture 1.7. *Fix $n \in \mathbb{N}$ and a partition λ of n. Then $\mathcal{F}(G_\lambda(p)) > 0$ for infinitely many primes p if and only if $\mathrm{c}(\lambda) \geq 0$. More precisely, if $\mathrm{c}(\lambda) > 0$ then as $p \to \infty$ we have*

$$\mathcal{F}\left(G_\lambda(p)\right) \sim \frac{\mathfrak{C}}{n} \cdot \frac{p^{\mathrm{c}(\lambda)}}{\log p},$$

where $\mathfrak{C}$ is as in (1.5)*. If $\mathrm{c}(\lambda) = 0$ then as $x \to \infty$ we have*

$$\sum_{\substack{p \leq x \\ p \text{ prime}}} \mathcal{F}\left(G_\lambda(p)\right) \sim \frac{\mathfrak{C}}{n} \cdot \frac{x}{(\log x)^2}.$$

If $\mathrm{c}(\lambda) < 0$ then

$$p \gg_\lambda 1 \implies \mathcal{F}(G_\lambda(p)) = 0.$$

Definition 1.8. We say that a partition λ of n is *attainable* if $\mathrm{c}(\lambda) \geq 0$.

Thus, Conjecture 1.7 implies that $G_\lambda(p)$ occurs as a class group for infinitely many primes p if and only if λ is attainable. What is the relative proportion of attainable partitions among all partitions? The following table suggests that the relative proportion decreases with n.

n	4	5	6	7	8	9	10	11	12	...	100
#{attainable partitions of n}	3	3	5	5	7	7	9	9	13	...	4742
#{partitions of n}	5	7	11	15	22	30	42	56	77	...	190 569 292
Ratio	0.6	0.43	0.45	0.33	0.32	0.23	0.21	0.16	0.17	...	0.000025

Our next theorem confirms this.

Theorem 1.9. *For a positive integer n, we have*

$$\frac{\#\{\textit{attainable partitions of } n\}}{\#\{\textit{partitions of } n\}} \;\ll\; n^{3/4} e^{(2-\sqrt{\frac{2}{3}}\pi)\sqrt{n}}.$$

In particular,

$$\lim_{n\to\infty} \frac{\#\{\textit{attainable partitions of } n\}}{\#\{\textit{partitions of } n\}} = 0.$$

1.3. **Numerical investigations of attainable groups.** For families of p-groups with $\mathrm{c}(\lambda) > 0$, we expect that many (if not all) groups should occur; in fact $\mathcal{F}(G_\lambda(p))$ should grow with p. On the other hand, there should be very few (if any at all) in case $\mathrm{c}(\lambda) < 0$ — we call these groups "sporadic".

In this section, we present numerical evidence supporting Conjecture 1.7 based on our numerical computation of $\mathcal{F}(G)$, conditional on GRH, for all p-groups G of odd size at most 10^6. (See Section 6 for details regarding the computation.)

1.3.1. *Numerics on $\mathcal{F}(G_\lambda(p))$.* We give in the tables below[3] the value of $\mathcal{F}(G_\lambda(p))$ (conditional on GRH) for each odd prime p and each partition λ of some $n \geq 3$, such that $|G_\lambda(p)| < 10^6$. To be precise: The second column in each table contains all partitions of n for some fixed n, ordered by decreasing cyclicity index $\mathrm{c}(\lambda)$, which itself is given in the leftmost column. The top row contains a list of all primes p such that $p^n < 10^6$, and under each p we list the values of $\mathcal{F}(G_\lambda(p))$ corresponding to the partition λ in the same row. Whenever a partition is omitted from a table, then it is implied that all omitted values of $\mathcal{F}(G_\lambda(p))$ are zero. Groups occuring in rows corresponding to negative cyclicity index ("sporadic groups") are star/bold-marked for emphasis (also see Section 1.3.2.)

c(λ)	λ	$p=3$	5	7	11	13	17	19	23	29	31	37	41
3	(3)	88	279	607	1856	2904	5797	7963	12958	24407	29201	46981	62327
1	(2, 1)	5	11	13	19	17	25	22	29	35	26	39	37
−3	(1, 1, 1)	0	0	0	0	0	0	0	0	0	0	0	0

c(λ)	λ	$p=43$	47	53	59	61	67	71	73	79	83	89	97
3	(3)	71617	91690	127190	170444	186988	242464	283998	306567	382770	438976	533751	678610
1	(2, 1)	39	29	46	48	57	55	60	66	51	73	66	69
−3	(1, 1, 1)	0	0	0	0	0	0	0	0	0	0	0	0

c(λ)	λ	$p=3$	5	7	11	13	17	19	23	29	31
4	(4)	206	1093	3404	16290	29496	77693	116710	233027	548392	701408
2	(3, 1)	19	47	71	146	197	244	343	480	644	779
0	(2, 2)	3	0	0	0	2	1	2	1	0	1
−2	(2, 1, 1)	0	0	0	0	0	0	0	0	0	0
−8	(1, 1, 1, 1)	0	0	0	0	0	0	0	0	0	0

c(λ)	λ	$p=3$	5	7	11	13
5	(5)	549	4610	19430	147009	314328
3	(4, 1)	56	218	444	1347	1894
1	(3, 2)	8	5	8	13	9
−1	(3, 1, 1)	0	**1**∗	0	0	0
−3	(2, 2, 1)	0	0	0	0	0
−7	(2, 1, 1, 1)	0	0	0	0	0
−15	(1, 1, 1, 1, 1)	0	0	0	0	0

c(λ)	λ	$p=3$	5	7
6	(6)	1512	19469	116278
4	(5, 1)	177	1024	2887
2	(4, 2)	18	37	58
0	(4, 1, 1)	0	3	0
0	(3, 3)	2	2	3
−2	(3, 2, 1)	0	0	0
−6	(3, 1, 1, 1)	0	0	0
−6	(2, 2, 2)	0	0	0
−8	(2, 2, 1, 1)	0	0	0
−14	(2, 1, 1, 1, 1)	0	0	0
−24	(1, 1, 1, 1, 1, 1)	0	0	0

c(λ)	λ	$p=3$	5	7
7	(7)	3881	86038	711865
5	(6, 1)	571	4259	17057
3	(5, 2)	58	177	372
1	(5, 1, 1)	7	7	6
1	(4, 3)	8	11	7
−1	(4, 2, 1)	**1**∗	0	0
−3	(3, 3, 1)	**1**∗	0	0
−5	(4, 1, 1, 1)	0	0	0
−5	(3, 2, 2)	0	0	0
−7	(3, 2, 1, 1)	0	0	0
−11	(2, 2, 2, 1)	0	0	0
−13	(3, 1, 1, 1, 1)	0	0	0
−15	(2, 2, 1, 1, 1)	0	0	0
−23	(2, 1, 1, 1, 1, 1)	0	0	0
−35	(1, 1, 1, 1, 1, 1, 1)	0	0	0

c(λ)	λ	$p=3$	5
8	(8)	10712	379751
6	(7, 1)	1585	18956
4	(6, 2)	180	719
2	(6, 1, 1)	18	30
2	(5, 3)	15	24
0	(5, 2, 1)	4	1
0	(4, 4)	2	0
−2	(4, 3, 1)	**1**∗	0
−4	(5, 1, 1, 1)	0	0
⋮	⋮	⋮	⋮
−48	(1, 1, 1, 1, 1, 1, 1, 1)	0	0

[3]The complete list of all $\mathcal{F}(G_\lambda(p))$ is given in [21], and a complete list of all corresponding discriminants d and groups $H(d)$ is given in [22].

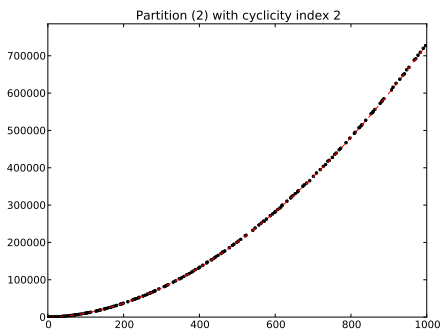
c(λ)	λ	$p=3$
9	(9)	28308
7	(8, 1)	4516
5	(7, 2)	454
3	(7, 1, 1)	42
3	(6, 3)	54
1	(6, 2, 1)	10
1	(5, 4)	4
−1	(5, 3, 1)	**1**∗
−3	(6, 1, 1, 1)	0
⋮	⋮	⋮
−63	(1, 1, 1, 1, 1, 1, 1, 1, 1)	0

c(λ)	λ	$p=3$
10	(10)	78657
8	(9, 1)	12433
6	(8, 2)	1446
4	(8, 1, 1)	160
4	(7, 3)	167
2	(7, 2, 1)	16
2	(6, 4)	14
0	(6, 3, 1)	1
0	(5, 5)	0
⋮	⋮	⋮
−80	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	0

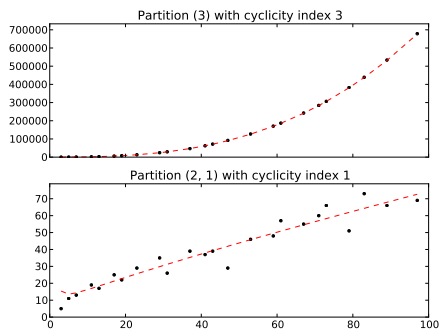
c(λ)	λ	$p=3$
11	(11)	216520
9	(10, 1)	35544
7	(9, 2)	3880
5	(9, 1, 1)	437
5	(8, 3)	460
3	(8, 2, 1)	58
3	(7, 4)	49
1	(7, 3, 1)	10
1	(6, 5)	9
−1	(8, 1, 1, 1)	0
−1	(7, 2, 2)	**1**∗
−1	(6, 4, 1)	**1**∗
−3	(7, 2, 1, 1)	0
⋮	⋮	⋮
−99	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	0

c(λ)	λ	$p=3$
12	(12)	603525
10	(11, 1)	98421
8	(10, 2)	10988
6	(10, 1, 1)	1291
6	(9, 3)	1265
4	(9, 2, 1)	220
4	(8, 4)	133
2	(8, 3, 1)	26
2	(7, 5)	17
0	(9, 1, 1, 1)	2
0	(8, 2, 2)	1
0	(7, 4, 1)	1
0	(6, 6)	2
−2	(8, 2, 1, 1)	**1**∗
−2	(7, 3, 2)	0
⋮	⋮	⋮
−120	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	0

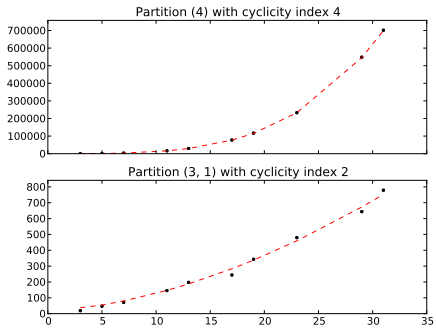
Below we plot, for p ranging over *odd primes*, observed values $\mathcal{F}(G_\lambda(p))$ (black dots) versus predicted values $P(G_\lambda(p)) \cdot \mathrm{pred}(|G_\lambda(p)|)$ (red dashed lines) for various partitions λ with positive cyclicity index $\mathrm{c}(\lambda) > 0$.



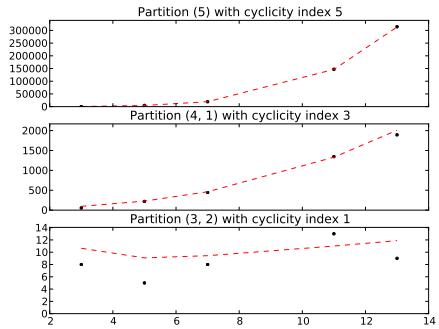
Partitions of 2



Partitions of 3



Partitions of 4



Partitions of 5

We remark that each vanishing entry in the tables above corresponds to a "missing" group. In particular we see that the group $(\mathbb{Z}/p\mathbb{Z})^3$ does not appear as the class group of a quadratic imaginary field for any prime $2 < p < 100$. Based on a combination of heuristics and numerics, it is reasonable to conjecture that $(\mathbb{Z}/p\mathbb{Z})^n$ does not occur for any odd prime p and any $n \geq 3$.

Conjecture 1.10. *For p odd, there are no elementary abelian p-groups of rank at least 3 which occur as the class group of an imaginary quadratic field.*

Indeed, by (1.2) and Conjecture 1.4, together with the observed (GRH-conditional) fact that no $(\mathbb{Z}/p\mathbb{Z})^n$ occurs as an imaginary quadratic class group for $p^n \leq 10^6$, we may bound the expected number of counterexamples by

$$\mathfrak{C} \sum_{\substack{p,n\geq 3 \\ p^n > 10^6}} \frac{\mathfrak{c}(p^n)}{np^{n^2-2n}\log p} \leq \mathfrak{C} \cdot \prod_{i=1}^{\infty}\left(1-\frac{1}{2^i}\right)^{-1} \sum_{\substack{p,n\geq 3 \\ p^n > 10^6}} \frac{1}{np^{n^2-2n}\log p}.$$

Since the right-hand sum can then be bounded by 10^{-4}, Conjecture 1.10 is heuristically justified.

Finally, we observe that none of the groups $G_\lambda(p)$ of odd size $< 10^6$ with $\mathrm{c}(\lambda) > 0$ are missing.

1.3.2. *Sporadic groups in negative cyclicity index case.* As just indicated with bold/star-marks in the tables, each of the groups

$$\frac{\mathbb{Z}}{5^3\mathbb{Z}} \times \left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)^2, \quad \frac{\mathbb{Z}}{3^4\mathbb{Z}} \times \frac{\mathbb{Z}}{3^2\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}}, \quad \left(\frac{\mathbb{Z}}{3^3\mathbb{Z}}\right)^2 \times \frac{\mathbb{Z}}{3\mathbb{Z}},$$
$$\frac{\mathbb{Z}}{3^4\mathbb{Z}} \times \frac{\mathbb{Z}}{3^3\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}}, \quad \frac{\mathbb{Z}}{3^5\mathbb{Z}} \times \frac{\mathbb{Z}}{3^3\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}}, \quad \frac{\mathbb{Z}}{3^7\mathbb{Z}} \times \left(\frac{\mathbb{Z}}{3^2\mathbb{Z}}\right)^2,$$
$$\frac{\mathbb{Z}}{3^6\mathbb{Z}} \times \frac{\mathbb{Z}}{3^4\mathbb{Z}} \times \frac{\mathbb{Z}}{3\mathbb{Z}}, \quad \frac{\mathbb{Z}}{3^8\mathbb{Z}} \times \frac{\mathbb{Z}}{3^2\mathbb{Z}} \times \left(\frac{\mathbb{Z}}{3\mathbb{Z}}\right)^2$$

occurs exactly once as an imaginary quadratic class group, even though $\mathrm{c}(\lambda) < 0$ for each corresponding partition λ. From the point of view of Conjecture 1.7, these examples may be regarded as "sporadic," since conjecturally they do not belong to an infinite family.

1.3.3. *Zero cyclicity index — the family $\mathcal{F}((\mathbb{Z}/p\mathbb{Z})^2)$.* The case of $\mathrm{c}(\lambda) = 0$ is intermediate in the sense that infinitely many groups in the family should occur, and infinitely many should not. Here the data is quite limited, and we restrict ourselves to the family $G = (\mathbb{Z}/p\mathbb{Z})^2$. The following table contains all odd primes p such that $p^2 < 10^6$, grouped according to the value of $\mathcal{F}((\mathbb{Z}/p\mathbb{Z})^2)$, assuming GRH.

n	All primes $p < 1000$ such that $\mathcal{F}((\mathbb{Z}/p\mathbb{Z})^2) = n$
0	11, 19, 37, 79, 89, 97, 103, 139, 151, 167, 181, 191, 193, 227, 229, 233, 241, 251, 271, 281, 283, 311, 313, 317, 349, 353, 359, 383, 401, 409, 433, 443, 463, 467, 479, 491, 499, 523, 563, 571, 587, 601, 619, 631, 643, 673, 701, 709, 733, 757, 769, 787, 809, 829, 877, 887, 907, 919, 929, 947, 953, 977, 983
1	3, 17, 23, 41, 43, 47, 61, 67, 73, 107, 109, 113, 127, 131, 137, 157, 163, 173, 179, 199, 239, 257, 263, 269, 277, 293, 307, 331, 337, 347, 367, 373, 379, 397, 419, 439, 457, 487, 503, 509, 521, 547, 557, 577, 599, 613, 617, 641, 653, 659, 677, 683, 691, 719, 727, 739, 743, 761, 797, 811, 821, 823, 839, 853, 857, 859, 863, 881, 937, 941, 971, 991, 997
2	5, 7, 29, 31, 53, 59, 71, 83, 101, 197, 211, 223, 389, 431, 449, 461, 569, 593, 607, 647, 661, 827, 883, 911
3	149, 421, 541, 751, 967
4	773
5	13

The limited data seems to support intermediate behaviour.

One may ask how well our prediction of $\mathcal{F}(G)$, using equation (1.2), holds up. The following graph compares the cumulants of the predictions with the observations.

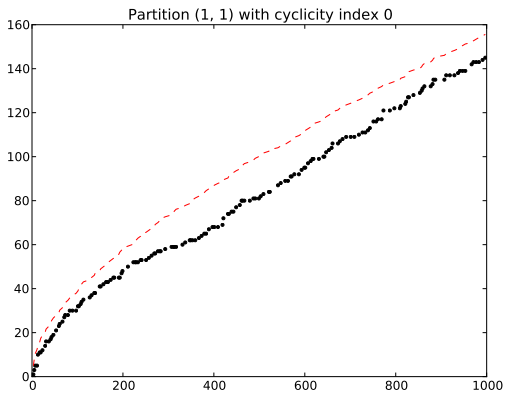


FIGURE 9. *Cumulative* observed values $\sum_{p<x} \mathcal{F}(G_{(1,1)}(p))$ (black dots) compared to *cumulative* predicted values $\sum_{p<x} P(G_{(1,1)}(p))\,\mathrm{pred}(p^2)$ (red dashed line), for each prime $x < 1000$.

1.4. **Related work.** Certain classes of finite abelian groups are already known *not* to occur as imaginary quadratic class groups. For instance, letting $H(d)[n]$ denote the n-torsion subgroup of $H(d)$, it is known that

$$|H(d)[2]| \ll |H(d)|^{o(1)}$$

(this is essentially genus theory together with Siegel's lower bound on the class number; if $H(d)$ has two rank r, then d has at least $r-1$ distinct prime factors). In particular, for any fixed $\epsilon > 0$ there are only finitely many imaginary quadratic class groups $H(d)$ satisfying $|H(d)[2]| \gg |H(d)|^{\epsilon}$. Weaker bounds are known for the size of the three torsion part; in [13] Venkatesh and Ellenberg (improving on Helfgott and Venkatesh [18] and Pierce [33]) show that

$$|H(d)[3]| \ll |d|^{1/3+\epsilon}.$$

From this and the (GRH-conditional) lower bound $d^{1/2} \ll |H(d)|$, one sees that, for any $\epsilon > 0$ there are only finitely many imaginary quadratic class groups $H(d)$ satisfying $|H(d)[3]| \gg |H(d)|^{2/3-\epsilon}$.

The problem of realizing a given abelian group as an imaginary quadratic class group may be viewed in the context of the following broader questions.

Question 1.11. Given a finite abelian group G, does there exist a number field K for which the ideal class group of K is isomorphic to G?

The answer to this problem is believed to be yes (one ought to be able to take K to be a real quadratic extension of $\mathbb{Q}$) but the problem is open in general, in spite of various partial results. G. Cornell [11] proved that every finite abelian group occurs as a subgroup of the ideal class group of some cyclotomic field, and Y. Yamamoto [47] proved that, for any $n \geq 1$, there are infinitely many imaginary quadratic fields whose class group contains $(\mathbb{Z}/n\mathbb{Z})^2$ as a subgroup. We note that Ozaki [30] has shown that any (possibly non-abelian) p-group occurs as the maximal unramified p-extension of some number field F.

Further broadening our perspective, we may also ask:

Question 1.12. Given an abelian group G, does there exist a Dedekind domain D for which the ideal class group of D is isomorphic to G?

In [9], Claborn answered this question in the affirmative; Leedham-Green subsequently showed that the Dedekind domain D can be taken to be a quadratic extension of a principal ideal ring.

Finally, we remark that the Cohen-Lenstra heuristics apply to a broader class of situations where finite abelian groups arise as co-kernels of random sub-lattices of $\mathbb{Z}^n$. For instance, [12] contains average results on the group of $\mathbb{Z}/p\mathbb{Z}$-rational points of an elliptic curve which are consistent with the Cohen-Lenstra heuristics (of course the rank can be at most two in this setting), and (in much the same spirit as our present consideration of missing class groups) [2] considers the question of which finite rank 2 abelian groups occur as the group of $\mathbb{Z}/p\mathbb{Z}$-rational points of some elliptic curve E over $\mathbb{Z}/p\mathbb{Z}$.

We conclude with some remarks regarding the numerical computations. Removing the assumption of GRH and making the computational results unconditional would be interesting, but probably very difficult since effective unconditional lower bounds on class numbers are quite weak (the best know bound, due to Oesterlé [29], is that $h(-q) \gg \log q$ for $-q$ a negative prime fundamental discriminant.) In particular, it would involve a major advance on Watkin's solution [44] to Gauss' class number problem for $h \leq 100$. (Of course, considering only odd h should be quite helpful.)

Determining $h(-d)$ for $d \in (0, D)$ and $-d$ ranging over fundamental discriminants is somewhat easier to do unconditionally, either by enumerating the primitive reduced quadratic forms in time $O(D^{3/2})$ (cf. [7]), or using GRH-conditional algorithms which, as suggested by A. Booker, can then be verified using the Eichler-Selberg trace formula. The latter algorithm, due to Jacobson, Ramachandran, and Williams [23], leads to a total running time of $O(D^{5/4})$, and allowed them to take $D = 10^{11}$. However, the verification step relies on knowing $h(-d)$ for all d in the relevant range, and seems difficult to adapt to a setting where only $h(-q)$ is known for $0 < q < D$ and $-q$ ranging over negative prime fundamental discriminants. On the other hand, Booker's algorithm [5] gives the correct value of $h(-d)$ in time $O(d^{1/4})$ if GRH is true (in time $O(d^{1/2})$ otherwise), and his algorithm can easily be restricted to prime discriminants. It would also be interesting to investigate the potential speedup from using Sutherland's primorial-steps algorithm (cf. [41, Ch. 4 and 11] — it exploits the smooth part of the class number, and results in better than $O(d^{1/7})$ median time to find $h(-d)$.

1.5. **Outline of the paper.** The organization is as follows: Section 2 covers the preliminary material on Cohen-Lenstra heuristics and the distribution of $L(1, \chi_d)$. In Section 3, we prove Theorem 1.5. In Section 4, we develop heuristics which lead to Conjectures 1.4 and 1.7. In Section 5, we discuss partition generating functions and give a proof of Theorem 1.9. In Section 6, we sketch the techniques used to obtain the numerical evidence.

1.6. **Acknowledgements.** P.K. was partially supported by grants from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the Swedish Research Council (621-2011-5498). S.H. was partially supported by a grant from the Swedish Research Council (621-2011-5498). K.P. was partially supported by Simons Foundation grant # 209266. The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at PDC Centre for High Performance Computing (PDC-HPC). We thank the PDC support (Magnus Helmersson, Jonathan Vincent, Radovan Bast,

Peter Gille, Jin Gong, Mattias Claesson) for their assistance concerning technical and implementational aspects in making the code run on the PDC-HPC resources.

We would like to thank Andrew Booker, Pete Clark, Henri Cohen, Noam Elkies, Farshid Hajir, Hendrik W. Lenstra, Steve Lester and Peter Sarnak for enlightening conversations on the topic.

2. Preliminaries

In this section, we briefly review relevant background material.

2.1. **Cohen-Lenstra heuristics.** When a finite abelian p-group G occurs in nature, its likelihood of occurrence is often found to be proportional to $1/|\operatorname{Aut}(G)|$. This suggests constructing a discrete probability measure μ on

$$\mathfrak{G}_p := \{\text{isomorphism classes of abelian } p\text{-groups}\}$$

by setting $\mu(\{G\}) := \dfrac{c}{|\operatorname{Aut}(G)|}$ for an appropriate positive constant c, if possible. The following lemma shows that this indeed the case, and is also useful for evaluating c.

Lemma 2.1. *We have that*

$$\sum_{\substack{G\in\mathfrak{G}_p\\ |G|=p^n}} \frac{1}{|\operatorname{Aut}(G)|} = \frac{1}{p^n}\prod_{i=1}^{n}\left(1-\frac{1}{p^i}\right)^{-1},$$

$$\sum_{\substack{G\in\mathfrak{G}_p\\ |G|\leq p^n}} \frac{1}{|\operatorname{Aut}(G)|} = \prod_{i=1}^{n}\left(1-\frac{1}{p^i}\right)^{-1}.$$

Proof. The first equation is [10, Cor 3.8, p. 40]; the second follows from the first by induction on n. □

Let us set

$$\eta_\infty(p) := \prod_{i=1}^{\infty}\left(1-\frac{1}{p^i}\right). \tag{2.1}$$

By taking $n \longrightarrow \infty$ in Lemma 2.1, we see that one must take $c = \eta_\infty(p)$ in order for $\mu(\mathfrak{G}_p) = 1$. In the Cohen-Lenstra model, the probability of G occurring as the p-part of a class group is thus given by

$$\mu(\{G\}) := \frac{\eta_\infty(p)}{|\operatorname{Aut}(G)|}. \tag{2.2}$$

Lemma 2.1 also has the following useful corollary. Here and later in the paper, we will also make use of the notation

$$\begin{aligned}
\mathfrak{D} &:= \{\text{negative fundamental discriminants}\},\\
\mathfrak{D}(x) &:= \{d\in\mathfrak{D}:\ -d\leq x\},\\
\mathfrak{D}' &:= \{q\in\mathfrak{D}:\ -q \text{ is prime}\},\\
\mathfrak{D}'(x) &:= \{q\in\mathfrak{D}':\ -q\leq x\}.
\end{aligned}$$

Recall that by genus theory, we have

$$h(d) \text{ is odd} \iff -d \text{ is prime}$$

for $d\in\mathfrak{D}$ with $d<-8$. This observation explains the following notation, wherein P denotes any property of positive odd integers.

$$\text{Prob}(h \text{ satisfies } P:\ h \text{ is an odd class number}) := \lim_{x\to\infty}\frac{\#\{q\in\mathfrak{D}'(x):\ h(q) \text{ satisfies } P\}}{\#\mathfrak{D}'(x)}. \tag{2.3}$$

Corollary 2.2. *Assuming the Cohen-Lenstra heuristics, for any $n \geq 0$ we have*

$$\mathrm{Prob}(p^n \nmid h : \ h \text{ is an odd class number}) = \prod_{i=n}^{\infty} \left(1 - \frac{1}{p^i}\right)$$

$$\mathrm{Prob}(p^n \parallel h : \ h \text{ is an odd class number}) = \frac{1}{p^n} \prod_{i=n+1}^{\infty} \left(1 - \frac{1}{p^i}\right).$$

Proof. The Cohen-Lenstra heuristics specify that

$$\mathrm{Prob}(p^n \nmid h : \ h \text{ is an odd class number}) = \mu(\{G \in \mathfrak{G}_p : |G| \leq p^{n-1}\}).$$

Together with Lemma 2.1, this gives the first equation, and the second equation follows from the first since $\mathrm{Prob}(p^n \parallel h : h \text{ is an odd class number})$ is equal to

$$\mathrm{Prob}(p^n \mid h : h \text{ is an odd class number}) - \mathrm{Prob}(p^{n+1} \mid h : h \text{ is an odd class number}). \qquad \square$$

2.2. **The class number formula and special values of L-functions.** Recall the class number formula, which in our context reads

$$L(1, \chi_d) = \frac{\pi h(d)}{\sqrt{|d|}} \qquad (d \in \mathfrak{D}, d < -8), \tag{2.4}$$

where $L(s, \chi_d) = \sum_{n=1}^{\infty} \chi_d(n) n^{-s}$ is the L-function attached to the Kronecker symbol $\chi_d := \left(\frac{d}{\cdot}\right)$. This formula connects the statistical study of class numbers to that of the special values $L(1, \chi_d)$. Building upon ideas that go back to P.D.T.A. Elliot, A. Granville and K. Soundararajan [15] proved that, on average over $d \in \mathfrak{D}$, $L(1, \chi_d)$ behaves like a random Euler product. More precisely, if $\mathbb{X}(p)$ denotes the random variable defined by

$$\mathbb{X}(p) := \begin{cases} 1 & \text{with probability } \frac{p}{2(p+1)} \\ 0 & \text{with probability } \frac{1}{p+1} \\ -1 & \text{with probability } \frac{p}{2(p+1)}, \end{cases}$$

and $L(1, \mathbb{X})$ denotes the random Euler product

$$L(1, \mathbb{X}) := \prod_{p} \left(1 - \frac{\mathbb{X}(p)}{p}\right)^{-1},$$

then [15, Theorem 2] (see also [40, p. 4]) implies that, for $|z| \leq \log x/(500(\log\log x)^2)$ and $\mathrm{Re}(z) > -1$, we have

$$\sum_{d \in \mathfrak{D}(x)} L(1, \chi_d)^z = |\mathfrak{D}(x)| \cdot \mathbb{E}(L(1, \mathbb{X})^z) + O\left(|\mathfrak{D}(x)| \exp\left(-\frac{\log x}{5 \log\log x}\right)\right), \tag{2.5}$$

where $\mathbb{E}$ denotes the expected value. This leads to the average result

$$\sum_{h \leq H} \mathcal{F}(h) = \frac{3\zeta(2)}{\zeta(3)} H^2 + O\left(H^2 (\log H)^{-1/2+\varepsilon}\right), \tag{2.6}$$

for any $\varepsilon > 0$ (see [40, Theorem 1]). In the interest of establishing the appropriate constant in Conjecture 1.4, we will next prove Theorem 1.5, which is an analogue of (2.6) averaged over *odd* values of h.

3. The average of $\mathcal{F}(h)$ over odd values of h

In this section we prove Theorem 1.5, that is we develop an asymptotic formula for $\sum_{\substack{h \leq H \\ h \text{ odd}}} \mathcal{F}(h)$. By genus theory, the restriction for $h \geq 3$ to be odd is equivalent to the condition that the associated discriminant d be *prime*. As an auxiliary result, we begin by proving the analogue of (2.5) over prime discriminants.

3.1. **The distribution of $L(1,\chi)$ over prime discriminants.** We now prove an asymptotic formula for the general moment of $L(1,\chi_q)$ averaged over $q \in \mathfrak{D}'(x)$. Our proof generally follows the methods used in [15, Theorem 2], but the restriction to prime discriminants demands that we use a different probabilistic model than the model $\mathbb{X}$ introduced earlier. Indeed, $\mathrm{Prob}(\mathbb{X}(p)=0) = 1/(p+1)$ corresponds to the probability that a random fundamental discriminant $d \in \mathfrak{D}$ is divisible by the prime p, and one computes

$$\mathrm{Prob}(p \mid d : d \in \mathfrak{D}) = \frac{|p\mathbb{Z}/p^2\mathbb{Z} - \{0\}|}{|\mathbb{Z}/p^2\mathbb{Z} - \{0\}|} = \frac{1}{p+1}.$$

On the other hand, the event $p \mid q$ can happen at most once for $q \in \mathfrak{D}'$, and so we replace $\mathbb{X}$ with $\mathbb{Y}$, where we recall that

$$\mathbb{Y}(p) := \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2. \end{cases} \tag{3.1}$$

The corresponding random Euler product is then

$$L(1,\mathbb{Y}) := \prod_p \left(1 - \frac{\mathbb{Y}(p)}{p}\right)^{-1}.$$

We will also make use of the following estimate for the remainder term in the Chebotarev density theorem for quadratic fields.

Proposition 3.1. *Assume the Generalized Riemann Hypothesis for Dedekind Zeta functions of quadratic number fields. Then for $d \in \mathbb{N}$ and any real non-principal Dirichlet character χ modulo d, we have*

$$\sum_{\substack{p \le x \\ \chi(p)=1}} 1 = \frac{1}{2}\,\mathrm{Li}(x) + O(x^{1/2}\log dx),$$

with an absolute implied constant.

Proof. This is a special case of a theorem of Lagarias-Odlyzko on the error term in the Chebotarev density theorem for general number fields; see [24, Theorem 1.3] and [39, Théorème 2]. □

As an immediate corollary, one deduces the following analogue of the Polya-Vinogradov Theorem, which gives square-root cancellation of characters sums over *prime* values.

Corollary 3.2. *Assume the Generalized Riemann Hypothesis for Dedekind Zeta functions of quadratic number fields. Then for $n \in \mathbb{N}$ which is not a square, we have*

$$\left|\sum_{q \in \mathfrak{D}'(x)} \chi_q(n)\right| \ll x^{1/2}\log(nx),$$

with an absolute implied constant.

The next theorem follows from Corollary 3.2, together with some technical lemmas from [15]. In particular, its proof will utilize several properties of the z-th divisor function $d_z(n)$ for $z \in \mathbb{C}$, which is characterized by the equation

$$\zeta(s)^z = \sum_{n=1}^{\infty} \frac{d_z(n)}{n^s} \qquad (\mathrm{Re}(s) > 1).$$

Further note that $d_z(n)$ is a multiplicative function, and for prime powers $n = p^a$ we have that

$$d_z(p^a) = \frac{\Gamma(z+a)}{a!\Gamma(z)} = \frac{z(z+1)(z+2)\dots(z+a-1)}{a!} \tag{3.2}$$

Theorem 3.3. *Assume the Generalized Riemann Hypothesis and let $\varepsilon > 0$. Then, uniformly for $|z| \le \log x/(500(\log\log x)^2)$, we have*

$$\sum_{q \in \mathfrak{D}'(x)} L(1,\chi_q)^z = |\mathfrak{D}'(x)| \cdot \mathbb{E}(L(1,\mathbb{Y})^z) + O_\varepsilon\left(x^{1/2+\varepsilon}\right).$$

Proof. By Lemma 2.3 of [15], for any $Z \in \mathbb{R}$ with $Z \geq \exp\left((\log q)^{10}\right)$ we have

$$L(1, \chi_q^z) = \sum_{n=1}^{\infty} \chi_q(n) \frac{d_z(n)}{n} e^{-n/Z} + O\left(\frac{1}{q}\right).$$

(Note that, since we are assuming GRH, we may ignore any possible exceptional discriminants.) Thus we have

$$\sum_{q \in \mathfrak{D}'(x)} L(1, \chi_q)^z = \sum_{n=1}^{\infty} \frac{d_z(n)}{n} e^{-n/Z} \sum_{q \in \mathfrak{D}'(x)} \chi_q(n) + O(\log\log x). \tag{3.3}$$

The main term in our asymptotic comes from the subsequence $n = m^2$; the other values of n contribute to the remainder term. Indeed, for $n = m^2$, we have

$$\sum_{q \in \mathfrak{D}'(x)} \chi_q(m^2) = |\mathfrak{D}'(x)| + O(\omega(m)),$$

and the contribution of these terms to (3.3) is thus

$$|\mathfrak{D}'(x)| \sum_{m=1}^{\infty} \frac{d_z(m^2)}{m^2} e^{-m^2/Z} + O\left(\log\log x + \sum_{m=1}^{\infty} \frac{|d_z(m^2)\omega(m)|}{m^2} e^{-m^2/Z}\right).$$

Using $\omega(m) \leq d(m)$ together with the bounds

$$\sum_{m=1}^{\infty} \frac{d_z(m^2) d(m)}{m^2} e^{-m^2/Z} \ll \log(|z|+2)^{4|z|+4} \ll_{\varepsilon} x^{\varepsilon}$$

and

$$\sum_{m=1}^{\infty} \frac{d_z(m^2)}{m^2} \left(1 - e^{-m^2/Z}\right) \leq \sum_{m=1}^{\infty} \frac{d_{(|z|+1)^2}(m)}{m^2} \left(\frac{m^2}{Z}\right)^{1/4} = \frac{\zeta(3/2)^{(1+|z|)^2}}{Z^{1/4}} \leq \frac{1}{x}$$

(see [15, p. 1014]), one finds that the contribution of the $n = m^2$ terms to (3.3) is thus

$$\begin{aligned}
|\mathfrak{D}'(x)| \sum_{m=1}^{\infty} \frac{d_z(m^2)}{m^2} + O_{\varepsilon}(x^{\varepsilon}) &= |\mathfrak{D}'(x)| \prod_p \left(\sum_{j=0}^{\infty} \frac{d_z(p^{2j})}{p^{2j}}\right) + O_{\varepsilon}(x^{\varepsilon}) \\
&= |\mathfrak{D}'(x)| \prod_p \left(\sum_{j=0}^{\infty} \binom{-z}{2j} \frac{1}{p^{2j}}\right) + O_{\varepsilon}(x^{\varepsilon}) \\
&= |\mathfrak{D}'(x)| \prod_p \frac{1}{2} \left(\left(1 + \frac{1}{p}\right)^{-z} + \left(1 - \frac{1}{p}\right)^{-z}\right) + O_{\varepsilon}(x^{\varepsilon}) \\
&= |\mathfrak{D}'(x)| \cdot \mathbb{E}(L(1, \mathbb{Y})^z) + O_{\varepsilon}(x^{\varepsilon}),
\end{aligned}$$

where we have used (3.2) together with the binomial series expansions of $\left(1 + \frac{1}{p}\right)^{-z}$ and $\left(1 - \frac{1}{p}\right)^{-z}$. In order to handle the terms $n \neq \square$, we begin by inserting the result of Corollary 3.2 into the right-hand side of (3.3), obtaining

$$\begin{aligned}
\left|\sum_{\substack{n=1 \\ n \neq \square}}^{\infty} \frac{d_z(n)}{n} e^{-n/Z} \sum_{q \in \mathfrak{D}'(x)} \chi_q(n)\right| &\ll x^{1/2} \log x \sum_{n=1}^{\infty} \frac{|d_z(n)|}{n} e^{-n/Z} \log n \\
&\ll x^{1/2} \log x \sum_{n=1}^{\infty} \frac{d_{\lceil |z| \rceil}(n)}{n} e^{-n/Z} \log n,
\end{aligned} \tag{3.4}$$

where we have used $|d_z(n)| \leq d_{|z|}(n)$ and $d_{t_1}(n) \leq d_{t_2}(n)$ for $t_1, t_2 \in \mathbb{R}_{>0}$ and $t_1 \leq t_2$. In [15, (2.4), p. 1001] it is observed that $\sum_{n=1}^{\infty} \frac{d_k(n)}{n} e^{-n/Z} \leq (\log 3Z)^k$ for any positive integer k and real number $Z \geq 2$. One

may adapt that argument to obtain a similar bound for $\sum_{n=1}^{\infty}\frac{d_k(n)}{n}e^{-n/Z}\log n$ by introducing the function

$$\widetilde{\log}(t) := \begin{cases} 2 & \text{if } t < e^2 \\ \log t & \text{if } t \geq e^2. \end{cases}$$

Note that, for any $a_1, a_2, \dots, a_k \in \mathbb{N}$ we have

$$\log(a_1 \cdot a_2 \cdot \dots \cdot a_k) \leq \widetilde{\log}(a_1 \cdot a_2 \cdot \dots \cdot a_k) \leq \widetilde{\log}(a_1) \cdot \widetilde{\log}(a_2) \cdot \dots \cdot \widetilde{\log}(a_k).$$

Furthermore, by estimating a discrete sum by a continuous integral we may see that, for Z large enough,

$$\sum_{a=1}^{\infty}\frac{e^{-a/Z}}{a}\widetilde{\log}(a) \ll (\log(e^2 \cdot Z))^2.$$

Using these facts together with the inequality $d_k(n)e^{-n/Z} \leq e^{k/Z}\sum_{a_1a_2\dots a_k=n} e^{-(a_1+a_2+\dots+a_k)/Z}$, we find that

$$\sum_{n=1}^{\infty}\frac{d_k(n)}{n}e^{-n/Z}\log n \leq \left(e^{1/Z}\sum_{a=1}^{\infty}\frac{e^{-a/Z}}{a}\widetilde{\log}(a)\right)^k \leq (\log(e^2 \cdot Z))^{3k},$$

for Z large enough. Inserting this into (3.4) and taking $Z = \exp\big((\log x)^{10}\big)$, we obtain

$$\begin{aligned}\left|\sum_{\substack{n=1\\ n\neq\square}}^{\infty}\frac{d_z(n)}{n}e^{-n/Z}\sum_{q\in\mathfrak{D}'(x)}\chi_q(n)\right| &\ll x^{1/2}\log x(\log(e^2\cdot Z))^{3\lceil|z|\rceil}, \\ &\ll_{\varepsilon} x^{1/2+\varepsilon}.\end{aligned}$$

This completes the proof of Theorem 3.3. □

3.2. **The proof of Theorem 1.5.** We will largely follow the proof of [40, Theorem 1] with critical modifications in appropriate places; we include the details here for completeness. We make use of the smooth cut-off function

$$\mathfrak{H}_{c,\delta}(x) := \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^s}{s}\left(\frac{(1+\delta)^{s+1}-1}{\delta(s+1)}\right)ds,$$

where the parameters $c, \delta > 0$ will be specified soon. For any $c, \delta > 0$ we have

$$\mathfrak{H}_{c,\delta}(x) = \begin{cases} 1 & \text{if } x \geq 1 \\ (1+\delta-1/x)/\delta & \text{if } (1+\delta)^{-1} \leq x \leq 1 \\ 0 & \text{if } x \leq (1+\delta)^{-1}. \end{cases} \tag{3.5}$$

Just as in [40], by using [15, Theorem 4] , one obtains that

$$\sum_{\substack{h\leq H\\ h \text{ odd}}}\mathcal{F}(h) = \sum_{\substack{q\in\mathfrak{D}'(X)\\ h_q\leq H}} 1 + O_A\left(\frac{H^2}{(\log H)^A}\right) \tag{3.6}$$

for any $A > 0$, where $X := H^2\log H$. By the class number formula, (3.6) and (3.5), it follows that

$$\sum_{\substack{h\leq H\\ h \text{ odd}}}\mathcal{F}(h) \leq \sum_{q\in\mathfrak{D}'(X)}\mathfrak{H}_{c,\delta}\left(\frac{\pi H}{\sqrt{q}L(1,\chi_q)}\right) + O_A\left(\frac{H^2}{(\log H)^A}\right) \leq \sum_{\substack{h\leq H(1+\delta)\\ h \text{ odd}}}\mathcal{F}(h).$$

We will now work with the main term in the middle above, which is

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\sum_{q\in\mathfrak{D}'(X)}\left(\frac{\pi}{\sqrt{q}L(1,\chi_q)}\right)^s\frac{H^s}{s}\left(\frac{(1+\delta)^{s+1}-1}{\delta(s+1)}\right)ds. \tag{3.7}$$

We will put $c := 1/\log H$ and $\delta := 1/(\log H)^{1/2}$. We furthermore set $S := \log X/(10^4(\log\log X)^2)$ and decompose the above interval into

$$\int_{|s|\leq S} + \int_{|s|>S}.$$

The second term is easily seen to be

$$\ll \frac{\mathfrak{D}'(X)}{\delta}\int_{|s|>S}\frac{1}{|s(s+1)|}|ds| \ll \frac{H^2}{(\log H)^{3/2-\varepsilon}}.$$

For the integral over $|s|\leq S$, we will use Theorem 3.3 to re-write the integrand in terms of the appropriate moment of $L(1,\mathbb{Y})$ and then reinterpret $\mathfrak{H}_{c,\delta}$ as a smooth cut-off function as in (3.5). First note that the following equation follows immediately from Theorem 3.3 by partial summation:

$$\sum_{q\in\mathfrak{D}'(X)}(\sqrt{q}L(1,\chi_q))^{-s} = \mathbb{E}(L(1,\mathbb{Y})^{-s})\int_1^X t^{-s/2}d\mathfrak{D}'(t)+O_\varepsilon(X^{1/2+\varepsilon}).$$

Thus, (3.7) is equal to

$$\begin{aligned}&\frac{1}{2\pi i}\int_{|s|\leq S}\mathbb{E}(L(1,\mathbb{Y})^{-s})\int_1^X t^{-s/2}d\mathfrak{D}'(t)\frac{(\pi H)^s}{s}\left(\frac{(1+\delta)^{s+1}-1}{\delta(s+1)}\right)ds+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}\right)\\ =&\mathbb{E}\left(\int_1^X\frac{1}{2\pi i}\int_{|s|\leq S}\left(\frac{\pi H}{\sqrt{t}L(1,\mathbb{Y})}\right)^s\frac{1}{s}\left(\frac{(1+\delta)^{s+1}-1}{\delta(s+1)}\right)ds\,d\mathfrak{D}'(t)\right)+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}\right)\end{aligned} \tag{3.8}$$

Extending the integral to $\int_{c-i\infty}^{c+i\infty}$ and managing the error, we find that

$$\frac{1}{2\pi i}\int_{|s|\leq S}\left(\frac{\pi H}{\sqrt{t}L(1,\mathbb{Y})}\right)^s\frac{1}{s}\left(\frac{(1+\delta)^{s+1}-1}{\delta(s+1)}\right)ds=\mathfrak{H}_{c,\delta}\left(\frac{\pi H}{\sqrt{t}L(1,\mathbb{Y})}\right)+O_\varepsilon\left(\frac{L(1,\mathbb{Y})^{-c}}{(\log H)^{3/2-\varepsilon}}\right).$$

Inserting this into (3.8), we find that (3.7) is equal to

$$\begin{aligned}&\mathbb{E}\left(\int_1^{\min\left(\frac{\pi^2H^2}{L(1,\mathbb{Y})^2},X\right)}d\mathfrak{D}'(t)+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}(1+L(1,\mathbb{Y})^{-c})\right)\right)\\ =&\frac{1}{2}\mathbb{E}\left(\mathrm{Li}\left(\min\left(\frac{\pi^2H^2}{L(1,\mathbb{Y})^2},X\right)\right)\right)+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}\right).\end{aligned} \tag{3.9}$$

Now using [15, Proposition 1], we find that $\min\left(\frac{\pi^2H^2}{L(1,\mathbb{Y})^2},X\right)=\frac{\pi^2H^2}{L(1,\mathbb{Y})^2}+O_A\left(\frac{H^2}{(\log H)^A}\right)$ for any $A>0$, and so we find that (3.9) becomes

$$\frac{1}{2}\mathbb{E}\left(\mathrm{Li}\left(\frac{\pi^2H^2}{L(1,\mathbb{Y})^2}\right)\right)+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}\right).$$

Finally, using the asymptotic $\mathrm{Li}(x)\sim\dfrac{x}{\log x}$ together with the calculation

$$\begin{aligned}\mathbb{E}(L(1,\mathbb{Y})^{-2})=\prod_p\mathbb{E}\left(\left(1-\frac{\mathbb{Y}(p)}{p}\right)^2\right)&=\prod_p\left(\frac{1}{2}\left(1-\frac{1}{p}\right)^2+\frac{1}{2}\left(1+\frac{1}{p}\right)^2\right)\\ &=\prod_p\left(1-\frac{1}{p^4}\right)\left(1-\frac{1}{p^2}\right)^{-1}=\frac{\zeta(2)}{\zeta(4)}=\frac{15}{\pi^2},\end{aligned} \tag{3.10}$$

the proof of Theorem 1.5 is concluded.

Remark 3.4. Our proof shows that in fact

$$\sum_{\substack{h\leq H\\ h\text{ odd}}}\mathcal{F}(h)=\frac{1}{2}\mathbb{E}\left(\mathrm{Li}\left(\frac{\pi^2H^2}{L(1,\mathbb{Y})^2}\right)\right)+O_\varepsilon\left(\frac{H^2}{(\log H)^{3/2-\varepsilon}}\right).$$

We find that the main term in the above expression fits the numerical data much better than the asymptotically equivalent formula given in Theorem 1.5, though it must be stressed that the corrections are of lower order than the error term. In the tables presented in Sections 1 and 6, the number listed under "predicted" refers to the higher order expansion of $\dfrac{\mathfrak{C}}{15}\cdot\mathfrak{c}(h)\cdot h\cdot\mathbb{E}\left(\dfrac{1}{L(1,\mathbb{Y})^2\log(\pi h/L(1,\mathbb{Y}))}\right)$ given in (1.7).

4. Heuristics

4.1. **Heuristics for Conjecture 1.4.** Recall from Remark 3.4 that we have

$$\sum_{\substack{h\leq H\\ h \text{ odd}}} \mathcal{F}(h) \approx \frac{1}{2}\mathbb{E}\left(\mathrm{Li}\left(\frac{\pi^2 H^2}{L(1,\mathbb{Y})^2}\right)\right).$$

Denote the right-hand side by $G(H)$. An average order of $\mathcal{F}$ is given by

$$G(h)-G(h-2)\approx 2\frac{d}{dh}G'(h)=\mathbb{E}\left(\frac{d}{dh}\mathrm{Li}\left(\frac{\pi^2h^2}{L(1,\mathbb{Y})^2}\right)\right)=$$

$$2\pi^2h\mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2\log\left(\pi^2h^2/L(1,\mathbb{Y})^2\right)}\right)=\frac{\pi^2h}{\log(\pi h)}\mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2}\frac{1}{1-\frac{\log L(1,\mathbb{Y})}{\log(\pi h)}}\right).$$

With a high probability, we have $\log L(1,\mathbb{Y})/\log(\pi h)<1$ for large h, so the above can be approximated with

$$\frac{\pi^2h}{\log(\pi h)}\mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2}\left(1+\frac{\log L(1,\mathbb{Y})}{\log(\pi h)}+\frac{\log^2 L(1,\mathbb{Y})}{\log^2(\pi h)}+\cdots\right)\right). \tag{4.1}$$

We will approximate this by keeping the first few terms in the innermost parentheses. In this regard, define $c_0:=\mathbb{E}\left(\frac{1}{L(1,\mathbb{Y})^2}\right)$, $c_1:=\frac{1}{c_0}\mathbb{E}\left(\frac{\log L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)$, $c_2:=\frac{1}{c_0}\mathbb{E}\left(\frac{\log^2 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)$, and $c_3:=\frac{1}{c_0}\mathbb{E}\left(\frac{\log^3 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)$. Recall from (3.10) that $c_0=15/\pi^2$. The constants c_1,c_2 and c_3 may be calculated to arbitrary precision as follows. Write $L_p:=1-\frac{\mathbb{Y}(p)}{p}$. Then $L(1,\mathbb{Y})=\prod_p L_p^{-1}$ and $\log L(1,\mathbb{Y})=-\sum_p\log L_p$. Now

$$\mathbb{E}\left(\frac{\log L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)=\mathbb{E}\left(-\sum_p\log L_p\prod_r L_r^2\right)=-\sum_p\mathbb{E}\left(L_p^2\log L_p\right)\prod_{r\neq p}\mathbb{E}\left(L_r^2\right)=-c_0\sum_p\frac{\mathbb{E}\left(L_p^2\log L_p\right)}{\mathbb{E}\left(L_p^2\right)} \tag{4.2}$$

where $\mathbb{E}\left(L_p^2\right)=1+\frac{1}{p^2}$ and $\mathbb{E}\left(L_p^2\log L_p\right)=\frac{1}{2}\left((1-\frac{1}{p})^2\log(1-\frac{1}{p})+(1+\frac{1}{p})^2\log(1+\frac{1}{p})\right)$. Next

$$\mathbb{E}\left(\frac{\log^2 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)=\mathbb{E}\left(\sum_{p,q}\log L_p\log L_q\prod_r L_r^2\right)=$$

$$\mathbb{E}\left(\sum_{p\neq q}L_p^2L_q^2\log L_p\log L_q\prod_{r\neq p,q}L_r^2+\sum_p L_p^2(\log L_p)^2\prod_{r\neq p}L_r^2\right)=$$

$$\sum_{p\neq q}\mathbb{E}\left(L_p^2\log L_p\right)\mathbb{E}\left(L_q^2\log L_q\right)\prod_{r\neq p,q}\mathbb{E}\left(L_r^2\right)+\sum_p\mathbb{E}\left(L_p^2(\log L_p)^2\right)\prod_{r\neq p}\mathbb{E}\left(L_r^2\right)=$$

$$c_0\sum_{p\neq q}\frac{\mathbb{E}\left(L_p^2\log L_p\right)\mathbb{E}\left(L_q^2\log L_q\right)}{\mathbb{E}\left(L_p^2\right)\mathbb{E}\left(L_q^2\right)}+c_0\sum_p\frac{\mathbb{E}\left(L_p^2(\log L_p)^2\right)}{\mathbb{E}\left(L_p^2\right)}=$$

$$c_0\cdot\left(\left(\sum_p\frac{\mathbb{E}\left(L_p^2\log L_p\right)}{\mathbb{E}\left(L_p^2\right)}\right)^2-\sum_p\left(\frac{\mathbb{E}\left(L_p^2\log L_p\right)}{\mathbb{E}\left(L_p^2\right)}\right)^2+\sum_p\frac{\mathbb{E}\left(L_p^2(\log L_p)^2\right)}{\mathbb{E}\left(L_p^2\right)}\right) \tag{4.3}$$

where $\mathbb{E}\left(L_p^2(\log L_p)^2\right)=\frac{1}{2}\left((1-\frac{1}{p})^2\log^2(1-\frac{1}{p})+(1+\frac{1}{p})^2\log^2(1+\frac{1}{p})\right)$. One may similarly show that

$$-\frac{1}{c_0}\mathbb{E}\left(\frac{\log^3 L(1,\mathbb{Y})}{L(1,\mathbb{Y})^2}\right)=\sum_{\substack{p,q,r\\ \text{distinct}}}\frac{\mathbb{E}\left((\log L_p)L_p^2\right)}{\mathbb{E}\left(L_p^2\right)}\frac{\mathbb{E}\left((\log L_q)L_q^2\right)}{\mathbb{E}\left(L_q^2\right)}\frac{\mathbb{E}\left((\log L_r)L_r^2\right)}{\mathbb{E}\left(L_r^2\right)}+$$

$$3\sum_{p\neq r}\frac{\mathbb{E}\left((\log L_p)^2L_p^2\right)}{\mathbb{E}\left(L_p^2\right)}\frac{\mathbb{E}\left((\log L_r)L_r^2\right)}{\mathbb{E}\left(L_r^2\right)}+\sum_p\frac{\mathbb{E}\left((\log L_p)^3L_p^2\right)}{\mathbb{E}\left(L_p^2\right)}. \tag{4.4}$$

Calculating the expressions (4.2), (4.3) and (4.4) with 10^5 prime terms yields

$$\begin{aligned} c_1 &\approx -0.578071, \\ c_2 &\approx +0.604049, \\ c_3 &\approx -0.526259. \end{aligned} \tag{4.5}$$

Thus, taking the first four terms of (4.1), an approximation of $\mathcal{F}(h)$ for odd h is

$$\frac{\pi^2 h}{\log(\pi h)}\left(c_0 + \frac{c_0c_1}{\log(\pi h)} + \frac{c_0c_2}{\log^2(\pi h)} + \frac{c_0c_3}{\log^3(\pi h)}\right) = \frac{15h}{\log(\pi h)}\left(1 + \frac{c_1}{\log(\pi h)} + \frac{c_2}{\log^2(\pi h)} + \frac{c_3}{\log^3(\pi h)}\right). \tag{4.6}$$

However, this assumes that each number h occurs as $h(d)$ with equal frequency, which is inconsistent with Corollary 2.2. We thus introduce the correction factor

$$\begin{aligned} \tilde{\mathfrak{c}}(h) := \prod_{\substack{p\geq 3 \text{ prime} \\ n\geq 0 \\ p^n \| h}} \frac{\text{Prob}(p^n \parallel h' :\ h' \text{ is an odd class number})}{\text{Prob}(p^n \parallel h' :\ h' \text{ is an odd integer})} &= \prod_{\substack{p\geq 3 \text{ prime} \\ n\geq 0 \\ p^n \| h}} \frac{p^{-n}\prod_{i=n+1}^{\infty}\left(1-\frac{1}{p^i}\right)}{p^{-(n+1)}(p-1)} \\ &= \prod_{\substack{p\geq 3 \text{ prime} \\ n\geq 0 \\ p^n \| h}} \left(1-\frac{1}{p}\right)^{-1} \prod_{i=n+1}^{\infty}\left(1-\frac{1}{p^i}\right) \end{aligned} \tag{4.7}$$

In the above, in addition to using (2.3), we are also using

$$\text{Prob}(h \text{ satisfies } P :\ h \text{ is an odd integer}) := \lim_{x\to\infty} \frac{\#\{h\in\mathbb{N} : h \text{ is odd}, h\leq x, h \text{ satisfies } P\}}{\#\{h\in\mathbb{N} :\ h \text{ is odd}, h\leq x\}}.$$

We emphasize that $n = 0$ is allowed in (4.7), and so the expression defining $\tilde{\mathfrak{c}}(h)$ is an *infinite* product. Note that, heuristically at least, we have

$$\sum_{\substack{h\leq H \\ h \text{ odd}}} \tilde{\mathfrak{c}}(h) \sim \frac{H}{2}, \qquad (H \longrightarrow \infty). \tag{4.8}$$

Indeed, if $\sum_{\substack{h\leq H \\ h \text{ odd}}} \tilde{\mathfrak{c}}(h) \sim B\cdot\frac{H}{2}$, then B has expected value

$$\begin{aligned} B &= \prod_{p \text{ odd}} \sum_{n=0}^{\infty} \text{Prob}(p^n \parallel h :\ h \text{ is an odd integer}) \cdot \frac{\text{Prob}(p^n \parallel h :\ h \text{ is an odd class number})}{\text{Prob}(p^n \parallel h :\ h \text{ is an odd integer})} \\ &= \prod_{p \text{ odd}} \sum_{n=0}^{\infty} \text{Prob}(p^n \parallel h :\ h \text{ is an odd class number}) \\ &= 1. \end{aligned}$$

Noting that

$$\tilde{\mathfrak{c}}(h) = \prod_{\substack{\ell=3 \\ \ell \text{ prime}}}^{\infty} \prod_{i=2}^{\infty}\left(1-\frac{1}{\ell^i}\right)\cdot\mathfrak{c}(h) = \frac{\mathfrak{C}}{15}\cdot\mathfrak{c}(h),$$

we get Conjecture 1.4 by multiplying the average order (4.6) with the local correction factor (4.7).

4.2. **Dampening the three divisibility bias.** Given an odd natural number h, let $k\leq 11$ and $n\leq k-3$ be such that $h\in[3^k, 3^{k+1})$ and $3^n \parallel h$. We define the adjustment $\text{pred}'(h)$ by replacing in $\text{pred}(h)$ the factor

$$\text{Prob}(3^n \parallel h' :\ h' \text{ is an odd class number}) = 3^{-n}\prod_{i=n+1}^{\infty}\left(1-\frac{1}{3^i}\right)$$

in $\tilde{\mathfrak{c}}(h)$ coming from the Cohen-Lenstra heuristic, by the observed value

$$\text{Prob}(3^n \parallel h' : \ h' \text{ is an odd class number} \in [3^k, 3^{k+1})) = \sum_{\substack{h' \in [3^k, 3^{k+1}) \\ h' \text{ odd} \\ 3^n \| h'}} \mathcal{F}(h') \Big/ \sum_{\substack{h' \in [3^k, 3^{k+1}) \\ h' \text{ odd}}} \mathcal{F}(h') \tag{4.9}$$

using our computed values of $\mathcal{F}(h)$ (see Section 6).

As mentioned earlier, this three divisibility bias is connected to other recent work: Belabas [3] noted rather slow convergence in the Davenport-Heilbronn asymptotic average of $H(d)$[3]; Roberts [35] later conjectured that this was due to a negative second order term of size $X^{5/6}$ (here the main term is of order X). Robert's conjecture was recently proved in [4, 42].

4.3. **Heuristics for Conjecture 1.7.** We now give heuristics supporting Conjecture 1.7. Let $\lambda = (n_1, n_2, \dots, n_r)$ be a partition of n, so that

$$n_1 \geq n_2 \geq \cdots \geq n_r \geq 1 \tag{4.10}$$

and $n_1 + n_2 + \cdots + n_r = n$, and let $G_\lambda(p) := \bigoplus_{i=1}^{r} \mathbb{Z}/p^{n_i}\mathbb{Z}$ be the corresponding abelian group. By the assumption (1.2), the expected value of $\mathcal{F}(G_\lambda(p))$ is

$$\mathcal{F}(G_\lambda(p)) \approx P(G_\lambda(p)) \cdot \mathcal{F}(p^n). \tag{4.11}$$

The following proposition evaluates $P(G_\lambda(p))$ explicitly. Let k be the number of distinct parts of λ, and let $m_1, m_2, \dots, m_k$ be the multiplicity of each distinct part. Thus, (4.10) reads

$$n_1 = \cdots = n_{m_1} > n_{m_1+1} = \cdots = n_{m_1+m_2} > \cdots > n_{\sum_{i=1}^{k-1} m_i + 1} = \cdots = n_{\sum_{i=1}^{k} m_i}.$$

Proposition 4.1. *With the notation just given, we have*

$$P(G_\lambda(p)) = p^{\mathrm{c}(\lambda)-n} \cdot \prod_{i=1}^{k} \prod_{j=1}^{m_i} \left(1 - \frac{1}{p^j}\right)^{-1} \prod_{i=1}^{n} \left(1 - \frac{1}{p^i}\right), \tag{4.12}$$

where $\mathrm{c}(\lambda)$ *is given by* (1.11). *In particular, as* $p \longrightarrow \infty$, *we have that*

$$P(G_\lambda(p)) \sim p^{\mathrm{c}(\lambda)-n}. \tag{4.13}$$

Proof. The statement follows immediately by combining Lemma 2.1 with the formula

$$|\operatorname{Aut}(G_\lambda(p))| = p^{2n-\mathrm{c}(\lambda)} \prod_{i=1}^{k} \prod_{j=1}^{m_i} \left(1 - \frac{1}{p^j}\right).$$

This formula is classical, having appeared in a 1907 paper of A. Ranum [34]. For a more modern exposition, see [19] or [27]. □

Inserting (4.13) together with Conjecture 1.4 into the right-hand side of (4.11), and observing that $\mathfrak{c}(p^n) \longrightarrow 1$ as $p \longrightarrow \infty$, we see that Conjecture 1.7 follows. In the case $\mathrm{c}(\lambda) = 0$ we write

$$\sum_{p \leq x} \mathcal{F}(G_\lambda(p)) \sim \sum_{p \leq x} P(G_\lambda(p)) \cdot \mathcal{F}(p^n) \sim \sum_{p \leq x} \mathfrak{C} \cdot \frac{p^{\mathrm{c}(\lambda)}}{\log(p^n)}$$

and use partial summation.

5. Attainable partitions are very rare

We now prove Theorem 1.9. To this end, let $c_{n,r}$ denote the number of attainable partitions of n into r parts. Work of Sellers ([37],[38]) leads to a generating function for the number of partitions of n which satisfy a certain type of linear inequality amongst their parts:

Theorem 5.1 ([37],[38])**.** *The number of partitions $\lambda = (n_1, n_2, \ldots, n_r)$ of n into r non-negative parts satisfying the inequality $n_1 \geq \sum_{i=2}^{r} b_i n_i$, for some non-negative integers b_i with $b_2 > 0$, has generating function*

$$\frac{1}{(1-x)(1-x^{b_2+1})(1-x^{b_2+b_3+2})(1-x^{b_2+b_3+b_4+3})\cdots(1-x^{b_2+b_3+\cdots+b_r+r-1})}.$$

Applying this result to our context requires a slight modification, and leads to the following generating function for the attainable partitions.

Corollary 5.2. *The generating function $C_r(x)$ for the sequence $c_{n,r}$ of length-r attainable partitions of n is given by*

$$C_r(x) = \sum_{n=0}^{\infty} c_{n,r} x^n = \frac{x^{r^2-r}}{(1-x)\prod_{j=2}^{r}(1-x^{j^2-j})}.$$

Proof. First, observe that by definition we require our partitions to be comprised of positive (rather than non-negative) parts. To accommodate this change, we use the easily-verified bijection between partitions of n into r *non-negative* parts satisfying the inequality $b_1 \geq \sum_{i=2}^{r} b_i n_i$ and partitions of $n + \sum_{i=2}^{r} b_i + (r-1)$ into r *positive* parts satisfying the same inequality, given by

$$(n_1, \ldots, n_r) \to (n_1 + \sum_{i=2}^{r} b_i, n_2 + 1, ..., n_r + 1)$$

Thus the analogous generating function to Seller's above for partitions into positive parts is simply a shift of indices away, given by

$$\frac{x^{b_2+b_3+\cdots+b_r+r-1}}{(1-x)(1-x^{b_2+1})(1-x^{b_2+b_3+2})(1-x^{b_2+b_3+b_4+3})\cdots(1-x^{b_2+b_3+\cdots+b_r+r-1})}.$$

Finally, we apply this to attainable partitions, which by definition satisfy an inequality in the form of the theorem, with coefficients $b_i = 2i-3$. The corollary then follows from the observation

$$j - 1 + \sum_{i=2}^{j} b_i = j - 1 + \sum_{i=2}^{j} (2i-3) = j^2 - j$$

for any $2 \leq j \leq r$.

□

Basic results about growth rates about coefficients of rational generating functions leads to an asymptotic count of attainable partitions:

Corollary 5.3. *For fixed r, the proportion of length-r partitions of n which are attainable is asymptotically $\frac{1}{(r-1)!}$.*

Proof. We rewrite our expression for $C_r(x)$ to isolate its singularity on the unit circle with the highest multiplicity ($x = 1$ with multiplicity r) and apply the techniques of singularity analysis. Namely, we write

$$\sum_{n=0}^{\infty} c_{n,r} x^n = \frac{1}{(1-x)^r} \cdot \frac{x^{r^2-r+1}}{\prod_{j=2}^{r}(1 + x + x^2 + \cdots + x^{j^2-j-1})} =: \frac{f_r(x)}{(1-x)^r},$$

where here $f_r(x)$ is analytic at $x = 1$. A partial fraction decomposition shows that the asymptotics for the coefficients are governed by this singularity (see, e.g., [14, p. 256]), and we obtain

$$c_{n,r} \sim \frac{f_r(1) n^{r-1}}{(r-1)!} = \frac{n^{r-1}}{(r-1)! \prod_{j=2}^{r}(j^2 - j)} = \frac{n^{r-1}}{r!(r-1)!^2}.$$

Similarly, by the well-known generating function

$$\sum_{n=0}^{\infty} p_{n,r} x^n = \frac{x^r}{\prod_{j=1}^{r}(1-x^j)} = \frac{1}{(1-x)^r} \cdot \frac{x^r}{\prod_{j=2}^{r}(1 + x + x^2 + \cdots + x^{j-1})},$$

for $p_{n,r}$, the total number of length-r partitions of n, we conclude that

$$p_{n,r} \sim \frac{n^{r-1}}{r!(r-1)!}.$$

Taking the ratio of these gives

$$\lim_{n\to\infty} \frac{c_{n,r}}{p_{n,r}} = \lim_{n\to\infty} \frac{\frac{n^{r-1}}{r!(r-1)!^2}}{\frac{n^{r-1}}{r!(r-1)!}} = \frac{1}{(r-1)!},$$

proving the result. □

Moving from the fixed rank to the fixed order case, we set

$$c_n = \sum_{r=1}^{n} c_{n,r},$$

the total number of partitions of n which are attainable.

Lemma 5.4. *For fixed n, the numbers $c_{n,r}$ satisfy the recurrence relation*

$$c_{n,r+1} = \sum_{i=0}^{\lfloor \frac{n-2r}{r^2+r} \rfloor} c_{n-2r-i(r^2+r),r}.$$

Proof. From Corollary 5.2 we easily deduce the recurrence relation between the successive generating functions:

$$C_{r+1}(x) = \frac{x^{2r}}{1-x^{r^2+r}} C_r(x) = (1 + x^{r^2+r} + x^{2(r^2+r)} + \cdots)(x^{2r} C_r(x)),$$

from which the lemma follows by equating coefficients. □

We prove by induction that for fixed $n \geq 1$ we have $c_{n,r} \leq \frac{n^{r-1}}{(r-1)!^2}$ for all r. This is trivial for $r = 1$ since $c_{n,1} = 1$. For the inductive step, the recurrence relation in Lemma 5.4 gives

$$c_{n,r+1} = \sum_{i=0}^{\lfloor \frac{n-2r}{r^2+r} \rfloor} c_{n-2r-i(r^2+r),r} \leq \frac{1}{(r-1)!^2} \sum_{i=0}^{\lfloor \frac{n-2r}{r^2+r} \rfloor} (n-2r-i(r^2+r))^{r-1}.$$

The terms in this sum are positive and decreasing as a function of i, and so we can compare to the integral:

$$\begin{aligned}\sum_{i=0}^{\lfloor \frac{n-2r}{r^2+r} \rfloor} (n-2r-i(r^2+r))^{r-1} &\leq (n-2r)^{r-1} + \int_0^{\lfloor \frac{n-2r}{r^2+r} \rfloor} (n-2r-i(r^2+r))^{r-1}\, di \\ &= n^{r-1} + \frac{n^r}{r(r^2+r)} - \frac{(n-2r-\lfloor \frac{n-2r}{r^2+r} \rfloor (r^2+r))^r}{r(r^2+r)}\end{aligned}$$

Since the latter term is positive and $r^2 + r \leq n$, we can continue

$$c_{n,r+1} \leq \frac{1}{(r-1)!^2} \sum_{i=0}^{\lfloor \frac{n-2r}{r^2+r} \rfloor} (n-2r-i(r^2+r))^{r-1} \leq \frac{1}{(r-1)!^2} \frac{rn^r + n^r}{r(r^2+r)} = \frac{n^r}{r!^2},$$

completing the induction. Now, summing over r gives

$$c_n = \sum_{r=1}^{n} c_{n,r} \leq \sum_{r=1}^{n} \frac{n^{r-1}}{(r-1)!^2} \leq \sum_{r=1}^{\infty} \frac{n^{r-1}}{(r-1)!^2} = I_0(2\sqrt{n}),$$

where $I_0(x)$ denotes the 0-th modified Bessel function of the first kind. By the asymptotic $I_0(x) \sim \frac{e^x}{\sqrt{2\pi x}}$, we can compare the formula for c_n with the famous asymptotic of Hardy-Ramanujan [16], $p_n \sim \frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}$, for the number of partitions of n. Taking the ratio of the two gives

$$\frac{c_n}{p_n} \ll n^{3/4} e^{(2-\sqrt{\frac{2}{3}}\pi)\sqrt{n}},$$

proving Theorem 1.9.

Remark 5.5. Since $c_{n,1} = 1$ for all $n \geq 1$ (and $c_{0,1} = 0$), Lemma 5.4 provides explicit formulas for the number of attainable partitions of n into a small number of parts. For example,

$$c_{n,2} = \sum_{i=0}^{\lfloor \frac{n-2}{2} \rfloor} c_{n-2-2i,1} = \left\lfloor \frac{n-1}{2} \right\rfloor,$$

agreeing with the easily-checked fact that the partition $[a, b]$ of n is attainable if and only if $a > b$. Less trivially, if we temporarily adopt the simplifying convention that $\lfloor x \rfloor = 0$ for $x < 0$, we have

$$c_{n,3} = \sum_{i=0}^{\lfloor \frac{n-4}{6} \rfloor} c_{n-4-6i,2} = \sum_{i=0}^{\lfloor \frac{n-4}{6} \rfloor} \left\lfloor \frac{n-5-6i}{2} \right\rfloor.$$

This leads to Rademacher-type formulas for computing the exact value of c_n.

6. Numerical computations

With the aid of a supercomputer and assuming GRH, we have computed $\mathcal{F}(h)$ and $\mathcal{F}(G)$ for all odd $h < 10^6$ and all p-groups G of odd size at most 10^6. We have made the computed values available online, see the references [20], [21], [22]. In this section we will describe how this computation was accomplished.

As already noted, by genus theory, if $-q < -8$ is a fundamental discriminant, then $h(-q)$ is odd precisely when q is prime. Corollary 1.3 in [25] states that under GRH,

$$h(-q) \geq \frac{\pi}{12e^\gamma} \sqrt{q} \left(\log\log q - \log 2 + \frac{1}{2} + \frac{1}{\log\log q} + \frac{14 \log\log q}{\log q} \right)^{-1} \tag{6.1}$$

if $-q$ is a fundamental discriminant such that $q \geq 10^{10}$. It is easy to verify that the right-hand side above is monotonic for $q \geq 10^{10}$. This implies that if $q \geq 1.1881 \cdot 10^{15}$ then $h(-q) > 10^6$. Thus it suffices to consider only discriminants in $\mathfrak{D}'(1.1881 \cdot 10^{15})$ (recall that $\mathfrak{D}'(x)$ denotes the set of negative fundamental discriminants $-q$ such that $q \leq x$ is prime.)

We use the procedure `quadclassunit0` in the computer package PARI 2.7.3 to compute the class groups $H(d)$; this procedure guarantees correct results assuming GRH, cf. [31, Section 3.4.70]. In principle, doing this for every $d \in \mathfrak{D}'(1.1881 \cdot 10^{15})$ would suffice, but a number of practical speedups were necessary.

6.1. Brief description of the algorithm.

We give a brief but not complete description of our algorithm. Our computer program iterates over all $d \in \mathfrak{D}'(1.1881 \cdot 10^{15})$ and records for each odd $h < 10^6$ and each noncyclic p-group G, how many times a group of order h or a group isomorphic to G is found, avoiding to compute $h(d)$ or $H(d)$ whenever not necessary (note that if G is a cyclic p-group, then the value of $\mathcal{F}(G)$ can be calculated from the data that we are keeping).

Given a fundamental discriminant $d \in \mathfrak{D}'(1.1881 \cdot 10^{15})$, we begin by calculating an approximation h_{approx} of $h(d)$ together with an explicit error factor E, by setting $h_{\text{approx}} := \frac{\sqrt{|d|}}{\pi} e^{\nu(x_1,d)}$ and $E := e^{\eta(x_1,x_2,d)}$ for suitable x_1, x_2 using Proposition 6.1 below (e.g., towards the end of the discriminant range, it suffices to only consider 7 terms in the truncated Euler product.) If we already at this stage can prove that $h(d) > 10^6$ (that is, if $h_{\text{approx}}/E > 10^6$), then we discard d. This cuts down our search space by roughly a factor of 100, as the lower bound (6.1) is overestimated by roughly this factor in our case.

Otherwise, we compute a candidate h^* for $h(d)$ using Shank's baby-step/giant-step algorithm[4] (specifically, we find an integer h^* near in value to h_{approx} such that g^{h^*} is the identity element for up to three different group elements $g \in H(d)$). We only compute one such candidate, but in practice, this candidate agrees with the true value of $h(d)$ (assuming GRH) with a failure rate of about $1.5 \cdot 10^{-7}$ for d in our range.

Next, we try to find the exponent of the group by determining the smallest divisor e^* of h^* such that g^{e^*} is the identity element for up to 12 different group elements $g \in H(d)$. We have that e^* divides the order of the group, and if moreover the error factor E is small enough such that h^* is the unique multiple of e^* in the interval $h_{\text{approx}} \cdot [\frac{1}{E}, E]$ then we have proven that $h(d) = h^*$. In practice, this step in our program

[4]Using that h is odd we gain a speedup factor of $\sqrt{2}$. In a sense, this speedup is a weak form of Sutherland's Primorial-Steps Algorithm [41].

only catches cyclic groups and groups of the form $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$ ($m \geq 3$), but since the majority of the groups $H(d)$ should be of this form[5], our program stops at this step in 99.7% of all cases it is reached.

If any of the above fails (that is, if g^{h^*} was not the identity element for some g or if E was not small enough) or if $h(d)$ was determined to be an odd prime power, then we proceed to compute the structure of the entire class group $H(d)$ using PARI.

Proposition 6.1. *Assume GRH. Let $d < -8$ be a fundamental discriminant and let $x_1 < x_2$ be two integers such that $x_1 \geq 1$ and $x_2 \geq 10^5$. Then*

$$\frac{1}{e^{\eta(x_1,x_2,d)}} \leq \frac{h(d)}{\frac{\sqrt{|d|}}{\pi} e^{\nu(x_1,d)}} \leq e^{\eta(x_1,x_2,d)} \tag{6.2}$$

where

$$\nu(x_1, d) := \sum_{p \leq x_1} -\log\left(1 - \frac{\left(\frac{d}{p}\right)}{p}\right),$$

$$\eta(x_1, x_2, d) := \frac{1.562 \log|d| + 0.655 \log x_2}{\sqrt{x}_2} + \log\log x_2 + B + \frac{3\log x_2 + 4}{8\pi\sqrt{x}_2} - \sum_{p \leq x_1} \frac{1}{p} + \frac{1}{x_1},$$

and $B := \lim_{x\to\infty} \left(\sum_{p\leq x} \frac{1}{p} - \log\log x\right) \approx 0.2614972128\ldots$ *is the prime reciprocal constant.*

Proof. Let χ be the real-valued character $(\frac{d}{\cdot})$ of modulus $|d| > 1$. Theorem 9.1 combined with Table 4 in [1] states that under GRH,

$$\left|\log L(1,\chi) - \log \prod_{p<x_2} \frac{1}{1 - \frac{\chi(p)}{p}}\right| \leq \frac{1.562\log|d| + 0.655\log x_2}{\sqrt{x}_2} \tag{6.3}$$

for any $x_2 \geq 10^5$. By Taylor expansion, we have

$$\begin{aligned} \log \prod_{p<x_2} \frac{1}{1 - \frac{\chi(p)}{p}} &= \sum_{p<x_2} -\log\left(1 - \frac{\chi(p)}{p}\right) \\ &= \sum_{p\leq x_1} -\log\left(1 - \frac{\chi(p)}{p}\right) + \sum_{x_1<p<x_2} \sum_{m=1}^{\infty} \frac{1}{m}\left(\frac{\chi(p)}{p}\right)^m. \end{aligned} \tag{6.4}$$

We can bound the terms with $m \geq 2$ by

$$\left|\sum_{x_1<p<x_2} \sum_{m=2}^{\infty} \frac{1}{m}\left(\frac{\chi(p)}{p}\right)^m\right| \leq \sum_{x_1<p<x_2} \frac{1}{p^2} \leq \int_{x_1}^{\infty} \frac{dt}{t^2} = \frac{1}{x_1}. \tag{6.5}$$

since $\left|\sum_2^\infty \frac{x^m}{m}\right| \leq \frac{1}{2}\sum_2^\infty |x|^m = \frac{|x|^2}{2(1-|x|)} \leq x^2$ for any $|x| \leq \frac{1}{2}$. For $m = 1$ we have

$$\left|\sum_{x_1<p<x_2} \frac{\chi(p)}{p}\right| \leq \sum_{x_1<p<x_2} \frac{1}{p} < \sum_{p\leq x_2} \frac{1}{p} - \sum_{p\leq x_1} \frac{1}{p}, \tag{6.6}$$

where the first term on the right-hand side can be bounded using inequality (6.21) in [36], which states that under RH,

$$\sum_{p\leq x_2} \frac{1}{p} < \log\log x_2 + B + \frac{3\log x_2 + 4}{8\pi\sqrt{x}_2} \tag{6.7}$$

for any $x_2 \geq 13.5$.

[5]We expect the class group to be cyclic more than 97.7% of the time, and class groups containing $\mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ for prime $q > 5$ are very rare (cf. [10, p. 56].)

Combining the inequalities (6.3), (6.4), (6.5), (6.6), (6.7) we obtain

$$|\log L(1,\chi) - \nu(x_1,d)| \le \eta(x_1,x_2,d), \tag{6.8}$$

and the inequality (6.2) follows from taking exponentials and applying the class number formula $h(d) = \frac{\sqrt{|d|}}{\pi} L(1,\chi)$ for $d < -8$. □

6.2. **Computer resources.** Our program comprises 1500 lines of `C++` code. The total time for the computation was 4.5 CPU years, requiring 1 TB of temporary memory storage.

We used the computer package PARI (cf. [31]) to compute the groups $H(d)$ and we used the computer package `primesieve` [43] to iterate through primes.

References

[1] Eric Bach. Improved approximations for Euler products. In *Number theory (Halifax, NS, 1994)*, volume 15 of *CMS Conf. Proc.*, pages 13–28. Amer. Math. Soc., Providence, RI, 1995.

[2] W. Banks, F. Pappalardi and I. Shparlinski. On group structures realized by elliptic curves over arbitrary finite fields. *Experiment. Math.* **21** (2012), no. 1, 11–25.

[3] K. Belabas. A fast algorithm to compute cubic fields. *Math. Comp.*, 66(219):1213–1237, 1997.

[4] M. Bhargava, A. Shankar, and J. Tsimerman. On the Davenport-Heilbronn theorems and second order terms. *Invent. Math.*, 193(2):439–499, 2013.

[5] A. R. Booker. Quadratic class numbers and character sums. *Math. Comp.*, 75(255):1481–1492 (electronic), 2006.

[6] D. Boyd and H. Kisilevsky. On the exponent of the ideal class groups of complex quadratic fields. *Proc. Amer. Math. Soc.* **31** (1972), 433–436.

[7] D. A. Buell. The last exhaustive computation of class groups of complex quadratic number fields. In *Number theory (Ottawa, ON, 1996)*, volume 19 of *CRM Proc. Lecture Notes*, pages 35–53. Amer. Math. Soc., Providence, RI, 1999.

[8] S. Chowla. An extension of Heilbronn's class number theorem. *Quarterly J. Math.* **5** (1934), 304–307.

[9] L. Claborn. Every abelian group is a class group. *Pacific J. Math.*, 18:219–222, 1966.

[10] H. Cohen and H.W. Lenstra. Heuristics on class groups of number fields. Lecture Notes in Mathematics, **1068**, Springer, 1984, pp. 33–62.

[11] G. Cornell. Abhyankar's Lemma and the class group. In *Number theory, Carbondale 1979* (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), 82–88, Lecture Notes in Math., 751, Springer, Berlin, 1979.

[12] C. David and E. Smith. A Cohen-Lenstra phenomenon for elliptic curves. *J. London Math. Soc.* **89** (2014) 24–44.

[13] J. S. Ellenberg and A. Venkatesh. Reflection principles and bounds for class group torsion. *Int. Math. Res. Not. IMRN*, (1):Art. ID rnm002, 18, 2007.

[14] P. Flajolet and R. Sedgewick. Analytic combinatorics. Cambridge University Press, Cambridge, 2009

[15] A. Granville and K. Soundararajan. The distribution of values of $L(1,\chi_d)$. *Geom. and Funct. Anal.* **13** (2003), 992–1028.

[16] G. Hardy and S. Ramanujan. Asymptotic formulae in combinatory analysis. *Proc. London Math. Soc.* **2** (1918) 75–115.

[17] D.R. Heath-Brown. Imaginary quadratic fields with class group exponent 5. *Forum Math.* **20** (2008), 275–283.

[18] H. A. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.*, 19(3):527–550 (electronic), 2006.

[19] C. Hillar and D. Rhea. Automorphisms of finite abelian groups. *Amer. Math. Monthly* **114** (2007), 917–923.

[20] S. Holmin, P. Kurlberg. List of $\mathcal{F}(h)$ for all odd $h < 10^6$. Available at `https://people.kth.se/~kurlberg/class_group_data/class_group_orders.txt`

[21] S. Holmin, P. Kurlberg. List of $\mathcal{F}(G)$ for all noncyclic p-groups G of odd order $< 10^6$. Available at `https://people.kth.se/~kurlberg/class_group_data/noncyclic_class_groups.txt`

[22] S. Holmin, P. Kurlberg. List of $(d, H(d))$ for all fundamental discriminants $d < 0$ such that $H(d)$ is a noncyclic p-group of odd order $< 10^6$. Available at `https://people.kth.se/~kurlberg/class_group_data/discriminants_of_noncyclic_groups.txt`

[23] M. J. Jacobson, Jr., S. Ramachandran, and H. C. Williams. Numerical results on class groups of imaginary quadratic fields. In *Algorithmic number theory*, volume 4076 of *Lecture Notes in Comput. Sci.*, pages 87–101. Springer, Berlin, 2006.

[24] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem, in A. Frohlich (ed.) *Algebraic Number Fields*, pp. 409–464, Academic Press, 1977.

[25] Youness Lamzouri, Xiannan Li, and Kannan Soundararajan. Conditional bounds for the least quadratic non-residue and related problems. *Math. Comp.*, 84(295):2391–2412, 2015.

[26] C. R. Leedham-Green. The class group of Dedekind domains. *Trans. Amer. Math. Soc.*, 163:493–500, 1972.

[27] J. Lengler. The Cohen-Lenstra Heuristic: Methodology and results. *J. Algebra* **323** (2010), 2960–2976.

[28] J. Lengler. The global Cohen-Lenstra Heuristic. *J. Algebra* **357** (2012), 247–269.

[29] J. Oesterlé. Nombres de classes des corps quadratiques imaginaires. *Astérisque*, (121-122):309–323, 1985. Seminar Bourbaki, Vol. 1983/84.

[30] M. Ozaki. Construction of maximal unramified p-extensions with prescribed Galois groups. *Invent. Math.*, 183(3):649–680, 2011.

[31] The PARI Group, Bordeaux. *PARI/GP version* `2.7.3`, 2015. Available at `http://pari.math.u-bordeaux.fr/pub/pari/unix/pari-2.7.3.tar.gz`.
[32] M. Perret. On the ideal class group problem for global fields. *J. Number Theory*, 77(1):27–35, 1999.
[33] L. B. Pierce. A bound for the 3-part of class numbers of quadratic fields by means of the square sieve. *Forum Math.*, 18(4):677–698, 2006.
[34] A. Ranum. The group of classes of congruent matrices with application to the group of isomorphisms of any abelian group. *Trans. Amer. Math. Soc.* **8** (1907), 71–91.
[35] D. P. Roberts. Density of cubic field discriminants. *Math. Comp.*, 70(236):1699–1705 (electronic), 2001.
[36] Lowell Schoenfeld. Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II. *Math. Comp.*, 30(134):337–360, 1976.
[37] James A. Sellers. Extending a recent result of Santos on partitions into odd parts. *Integers* **3:A4**, 5 pp. (electronic), 2003.
[38] James A. Sellers. Corrigendum to: "Extending a recent result of Santos on partitions into odd parts". *Integers* **4:A8**, 1 pp. (electronic), 2004.
[39] J.-P. Serre. Quelques applications du théorème de densité de Chebotarev. *Publ. Math. I. H. E. S.* **54** (1981), 123–201.
[40] K. Soundararajan. The number of imaginary quadratic fields with a given class number. *Hardy-Ramanujan J.* **30** (2007), 13–18.
[41] A. V. Sutherland. *Order computations in generic groups*. ProQuest LLC, Ann Arbor, MI, 2007. Thesis (Ph.D.)–Massachusetts Institute of Technology.
[42] T. Taniguchi and F. Thorne. Secondary terms in counting functions for cubic fields. *Duke Math. J.*, **162** (2013), 2451–2508.
[43] K. Walisch. `primesieve`. Fast C/C++ prime number generator. Available at `http://primesieve.org/`.
[44] M. Watkins. Class numbers of imaginary quadratic fields. *Math. Comp.* **73** (2003), 907–938.
[45] P. Weinberger. Exponents of the class groups of complex quadratic fields. *Acta Arith.* **22** (1973), 117–124.
[46] O. Yahagi. Construction of number fields with prescribed l-class groups. *Tokyo J. Math.* **1** (1978), no. 2, 275–283.
[47] Y. Yamamoto. On unramified Galois extensions of quadratic number fields, *Osaka J. Math.* **7** (1970), 57–76.

(Samuel Holmin)

- Department of Mathematics, KTH, SE-10044, Stockholm, Sweden.

E-mail address, Samuel Holmin: `holmin@kth.se`

(Nathan Jones, corresponding author)

- Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 322 Science and Engineering Offices (M/C 249), 851 S. Morgan Street, Chicago, IL 60607-7045, USA.
  Phone: (312) 996-3041
  Fax: (312) 996-1491

E-mail address, Nathan Jones: `ncjones@uic.edu`

(Pär Kurlberg)

- Department of Mathematics, KTH, SE-10044, Stockholm, Sweden.

E-mail address, Pär Kurlberg: `kurlberg@math.kth.se`

(Cam McLeman)

- University of Michigan – Flint, Mathematics Department, 402 Murchie Science Building, Flint, MI 48502-1950, USA.

E-mail address, Cam McLeman: `mclemanc@umflint.edu`

(Kathleen L. Petersen)

- Department of Mathematics, Florida State University, 208 Love Building, 1017 Academic Way, Tallahassee, FL 32306-4510, USA.

E-mail address, Kathleen L. Petersen: `petersen@math.fsu.edu`

Paper D

On the free path length distribution for linear motion in an n-dimensional box

Samuel Holmin Pär Kurlberg Daniel Månsson

November 12, 2015

We consider the distribution of free path lengths, or the distance between consecutive bounces of random particles, in an n-dimensional rectangular box. If each particle travels a distance R, then, as $R \to \infty$ the free path lengths coincides with the distribution of the length of the intersection of a random line with the box (for a natural ensemble of random lines) and we determine the mean value of the path lengths. Moreover, we give an explicit formula (piecewise real analytic) for the probability density function in dimension two and three.

In dimension two we also consider a closely related model where each particle is allowed to bounce N times, as $N \to \infty$, and give an explicit (again piecewise real analytic) formula for its probability density function.

Further, in both models we can recover the side lengths of the box from the location of the discontinuities of the probability density functions.

1. Introduction

We consider billiard dynamics on a rectangular domain, i.e., point shaped "balls" moving with linear motion with specular reflections at the boundary, and similarly for rectangular box shaped domains in three dimensions. We wish to determine the distribution of free path lengths of ensembles of trajectories defined by selecting a starting point and direction at random.

The question seems quite natural and interesting on its own, but we mention that it originated from the study of electromagnetic fields in "reverberation chambers" under the assumption of highly directional antennas [7]. Briefly, the connection is as follows (we refer to the forthcoming paper [4] for more details): given an ideal highly directional antenna and a highly transient signal, then the wave pulse dynamics is essentially the same as a point shaped billiard ball traveling inside a chamber, with specular reflection at the boundary. Signal loss is dominated by (linear) "spreading" of the electromagnetic field and by absorption occurring at each interaction ("bounce") with the walls. The first simple model we use in this paper neglects absorption effects, and models signal loss from

spreading by simply terminating the motion of the ball after it has travelled a certain large distance. The second model only takes into account signal loss from absorption, and completely neglects spreading; here the motion is terminated after the ball has bounced a certain number of times.

We remark that the distribution of free path lengths is very well studied in the context of the Lorentz gas — here a point particle interacts with hard spherical obstacles, either placed randomly, or regularly on Euclidean lattices; recently quasicrystal configurations have also been studied (cf. [1–3, 5, 8–10, 12, 13].)

Let $R > 0$ be large and let a rectangular n-dimensional box $K \subseteq \mathbb{R}^n$ be given, where $n \geq 2$. We send off a large number $M > 0$ of particles, each with a random initial position $p^{(i)} \in K$ chosen with respect to a given probability measure μ on K, and each with a uniformly random initial direction $v^{(i)} \in \mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, $i = 1, \ldots, M$, for a total distance R each. Each particle travels along straight lines, changing direction precisely when it hits the boundary of the box, where it reflects specularly. We record the distance travelled between each pair of consecutive bounces for each particle. (Note in particular that we obtain more bounce lengths from some particles than from others.) Let $X_{M,R}$ be the uniformly distributed random variable on this finite set of bounce lengths of all the particles. More precisely, a random sample of $X_{M,R}$ is obtained as follows: first take a random i.i.d. sample of points (with respect to the measure μ) $p^{(1)}, \ldots, p^{(M)} \in K$, and a random sample of directions $v^{(1)}, \ldots, v^{(M)} \in \mathbb{S}^{n-1}$ (with respect to the uniform measure). Each pair $(p^{(i)}, v^{(i)})$ then defines a trajectory T^i of length R, and each such trajectory gives rise to a finite multiset B^i of lengths between consecutive bounces. Finally, with $B = \bigcup_{i=1}^{M} B^i$ denoting the (multiset) union of bounce length multisets $B^1, \ldots, B^M$, we select an element of B with the uniform distribution. (That is, with 1_B denoting the integer valued set indicator function for B, and $B' = \{x : 1_B(x) \geq 1\}$ we select the element $b \in B'$ with probability $1_B(b)/\sum_{x \in B'} 1_B(x)$.)

We are interested in the distribution of $X_{M,R}$ for large M and R, and this turns out to be closely related to a model arising from integral geometry. Namely, let $d\ell$ denote the unique (up to a constant) translation- and rotation-invariant measure on the set of directed lines ℓ in $\mathbb{R}^n$, and consider the restriction of this measure to the set of directed lines ℓ intersecting K, normalized such that it becomes a probability measure. Denote by X the random variable $X := \text{length}(\ell \cap K)$ where ℓ is chosen at random using this measure. Our first result is that $X_{M,R}$ converges in distribution to X as we take $M \to \infty$ and then $R \to \infty$ (or vice versa), and using techniques from integral geometry we find that the mean value of X has a quite simple geometric interpretation.

Theorem 1. *For any dimension $n \geq 2$, and for any distribution μ on the starting points, the random variable $X_{M,R}$ converges in distribution to the random variable X, as we take $R \to \infty$ followed by taking $M \to \infty$, or vice versa. Moreover, the mean value of X is*

$$\mathbb{E}[X] = 2\pi \frac{|\mathbb{S}^{n-1}|}{|\mathbb{S}^n|} \frac{\text{Vol}(K)}{\text{Area}(K)} = 2\sqrt{\pi} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{\text{Vol}(K)}{\text{Area}(K)}$$

where $\text{Area}(K)$ *is the* $(n-1)$*-dimensional surface area of the box* K*,* $\text{Vol}(K)$ *is the volume of the box* K*,* Γ *is the gamma function, and where* $|\mathbb{S}^{n-1}| = 2\pi^{n/2}/\Gamma(n/2)$ *is the*

$(n-1)$-*dimensional surface area of the sphere* $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$.

Throughout the paper, we will write pdf_Z and cdf_Z for the probability density function and the cumulative distribution function of Z, respectively, for random variables Z. We give explicit formulas for the probability density function of X in dimensions two and three.

Theorem 2. *For a box of dimension $n = 2$ with side-lengths $a \leq b$, the probability density function of X is given by*

$$\mathrm{pdf}_X(t) = \frac{1}{a+b} \cdot \begin{cases} 1, & \text{if } t < a, b \\ \dfrac{a^2 b}{t^2\sqrt{t^2-a^2}}, & \text{if } a < t < b \\ -1 + \dfrac{1}{t^2}\left(\dfrac{a^2 b}{\sqrt{t^2-a^2}} + \dfrac{ab^2}{\sqrt{t^2-b^2}}\right), & \text{if } a, b < t. \end{cases}$$

for $0 < t < \sqrt{a^2+b^2}$.

Remark 3. We note that the probability density function in Theorem 2 is analytic on all open subintervals of $(0, \sqrt{a^2+b^2})$ not containing a or b. Moreover, it is constant on the interval $(0, \min(a,b))$ and has singularities of type $(t-a)^{-1/2}$ and $(t-b)^{-1/2}$ just to the right of a and b, respectively. See Figure 1 for more details. For an explanation of these singularities, see Remark 25.

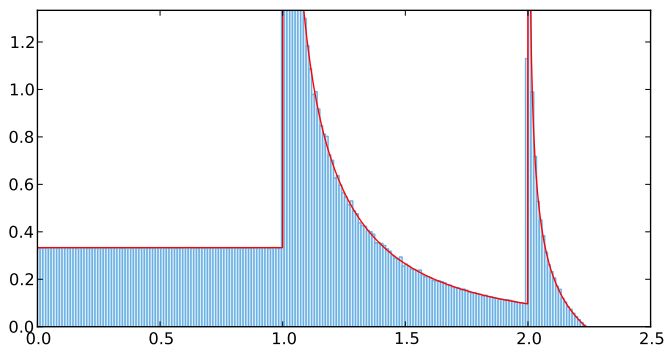


Figure 1: Simulation (blue histogram) vs explicit probability density function (red line) given by Theorem 2 for $(a,b) = (1,2)$. (Simulation used 10^5 particles, each starting at the origin with a uniformly random direction, going for a total distance 1000 each.) The plot is cutoff at $y = 1.3$ since $\mathrm{pdf}_X(t)$ tends to infinity as $t \to 1^+$ and $t \to 2^+$.

Theorem 4. *For a box of dimension $n = 3$ with side-lengths a, b, c, the probability density function of X is given by*

$$\mathrm{pdf}_X(t) = \frac{F(a,b,c,t) + F(b,c,a,t) + F(c,a,b,t)}{3\pi t^3(ab+ac+bc)}$$

where

$$F(a,b,c,t) = t^3(8a - 3t)$$

for $0 < t < a$, and by

$$F(a,b,c,t) = \Big(6t^4 - a^4 + 6\pi a^2 bc\Big) - 4(b+c)\sqrt{|t^2-a^2|}(a^2+2t^2)$$

for $a < t < \sqrt{a^2+b^2}$, and by

$$\begin{aligned} F(a,b,c,t) &= 6\pi a^2 bc + b^4 - 3t^4 - 6a^2b^2 + \\ &\sqrt{|t^2-a^2-b^2|}4c\Big(a^2+b^2+2t^2\Big) + \\ &+4a\sqrt{|t^2-b^2|}(b^2+2t^2) - 12a^2bc\cdot\arctan\left(\frac{\sqrt{|t^2-a^2-b^2|}}{b}\right) + \\ &-4c\sqrt{|t^2-a^2|}(a^2+2t^2) - 12ab^2c\cdot\arctan\left(\frac{\sqrt{|t^2-a^2-b^2|}}{a}\right) \end{aligned}$$

for $\sqrt{a^2+b^2} < t < \sqrt{a^2+b^2+c^2}$.

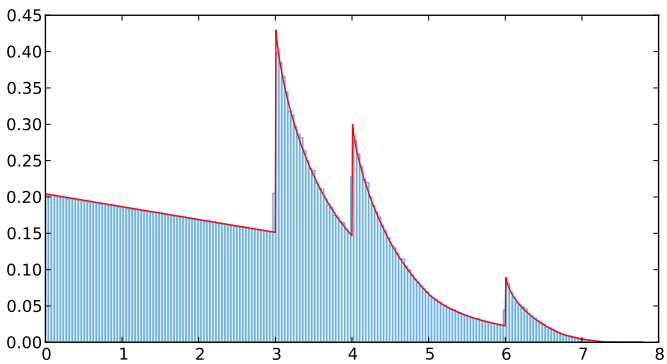


Figure 2: Simulation (blue histogram) vs explicit probability density function (red line) given by Theorem 4 for $(a,b,c) = (3,4,6)$. (Simulation used 10^5 particles, each starting at the origin with a uniformly random direction, going for a total distance 1000 each.) The fact that $\mathrm{pdf}_X(t)$ is not smooth at $t = 5$ is barely noticeable.

Remark 5. We note that the probability density function in Theorem 4 is analytic on all open subintervals of $(0, \sqrt{a^2+b^2+c^2})$ not containing any of the points

$$a, b, c, \sqrt{a^2+b^2}, \sqrt{a^2+c^2}, \sqrt{b^2+c^2}.$$

Moreover, it is linear on the interval $(0, \min(a,b,c))$ and has positive jump discontinuities at the points a, b, c. At the points $\{\sqrt{a^2+b^2}, \sqrt{a^2+c^2}, \sqrt{b^2+c^2}\} \setminus \{a,b,c\}$, it is continuous and differentiable.

Note that the probability distribution $X_{M,R}$ gives a larger "weight" to some particles than others, since some particles get more bounces than others for the same distance R. One could also consider a similar problem where we send off each particle for a certain number $N > 0$ of bounces, and then consider the limit as $M \to \infty$ followed by taking the limit $N \to \infty$, where M is the number of particles. This would give each particle the same "weight". Denote the finite version of this distribution by $Y_{M,N}$ and its limit distribution as $M \to \infty$ and then $N \to \infty$ by Y. With regard to the previous discussion about signal loss, we call the limit distribution X of $X_{M,R}$ the **spreading model** and we call the limit distribution of $Y_{M,N}$ the **absorption model**. Determining the probability density function of the absorption model appears to be the more difficult problem, and we give a formula only in dimension two:

Theorem 6. *For a box of dimension $n = 2$ with side-lengths $a \leq b$, the random variable $Y_{M,N}$ converges in distribution to the random variable Y, as we take $M \to \infty$ followed by taking $N \to \infty$, where the probability density function $\mathrm{pdf}_Y(t)$ is given by*

$$\frac{2}{\pi}\left(\frac{2(a+b)}{(a^2+b^2)} - \frac{2ab}{(a^2+b^2)^{3/2}}\left(\tanh^{-1}\left(\frac{a}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right)\right)\right)$$

for $0 < t < a, b$, and by

$$\frac{2}{\pi}\left(\frac{a\left(b-\sqrt{t^2-a^2}\right)}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} + \frac{2ab+2at-2a\sqrt{t^2-a^2}}{t(a^2+b^2)} + \frac{2ab\left(-\tanh^{-1}\left(\frac{t}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\right) - \tanh^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right)\right)}{(a^2+b^2)^{3/2}}\right)$$

for $a < t < b$, and by

$$\frac{2}{\pi}\left(\frac{a(b-\sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}} + \frac{b(a-\sqrt{t^2-b^2})}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}} + 2\frac{2ab-a\sqrt{t^2-a^2}-b\sqrt{t^2-b^2}}{t(a^2+b^2)} + \frac{2ab\left(-2\tanh^{-1}\left(\frac{t}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\right) + \tanh^{-1}\left(\frac{\sqrt{t^2-b^2}\sqrt{a^2+b^2}}{ta}\right)\right)}{(a^2+b^2)^{3/2}}\right)$$

for $a, b < t < \sqrt{a^2+b^2}$.

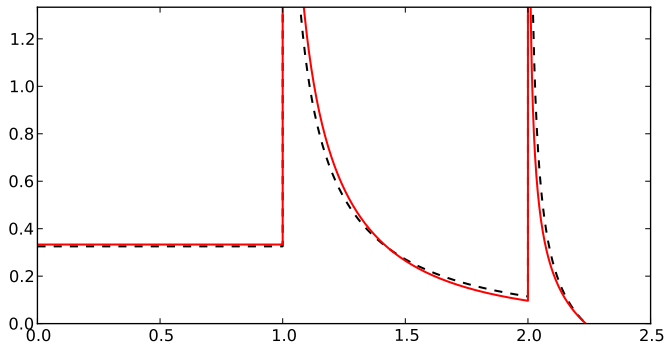


Figure 3: Probability density function for spreading model X (red line) from Theorem 2 vs absorption model (black dashed line) from Theorem 6, for $(a, b) = (1, 2)$.

See Figure 3 for a comparison between the probability density functions for the two different models in dimension 2.

Remark 7. It is not a priori obvious that the two limit distributions should differ, and it is natural to ask how much, if at all, they differ. We start by remarking that the expression for $\mathrm{pdf}_Y(t)$ does not simplify into the expression for $\mathrm{pdf}_X(t)$; indeed, for $(a, b) = (1, 2)$ we have $\mathrm{pdf}_X(t) = 1/3$ but $\mathrm{pdf}_Y(t) \approx 0.32553$ on the interval $(0, 1)$. For very skew boxes, with $a = 1$ and $b \to \infty$, it is straightforward to show that

$$\frac{\mathrm{pdf}_Y(b/2)}{\mathrm{pdf}_X(b/2)} \to \infty$$

as $b \to \infty$.

1.1. Acknowledgements

We would like to thank Zeev Rudnick for some very helpful discussions, especially for suggesting a connection with integral geometry.

S.H. was partially supported by a grant from the Swedish Research Council (621-2011-5498). P.K. was partially supported by grants from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the Swedish Research Council (621-2011-5498).

2. Proof of Theorem 1

In this section, we prove Theorem 1. For notational simplicity, we give the proof in dimension three; the general proof for $n \geq 2$ dimensions is analogous.

Given a particle with initial position p and initial direction v, let $N_{R,p,v}$ be the number of bounce lengths we get from that particle as it has travelled a total distance $R > 0$, and let $N_{R,p,v}(t)$ be the number of such bounce lengths of length at most $t \geq 0$. The uniform probability distribution on the set of bounce lengths of M particles with initial positions $p^{(1)}, \ldots, p^{(M)}$ and initial directions $v^{(1)}, \ldots, v^{(M)}$ has the cumulative distribution function

$$\mathrm{cdf}_{X_{M,R}}(t) = \frac{\sum_{i=1}^{M} N_{R,p^{(i)},v^{(i)}}(t)}{\sum_{i=1}^{M} N_{R,p^{(i)},v^{(i)}}} = \frac{\frac{1}{M}\sum_{i=1}^{M} \frac{N_{R,p^{(i)},v^{(i)}}}{R} \frac{N_{R,p^{(i)},v^{(i)}}(t)}{N_{R,p^{(i)},v^{(i)}}}}{\frac{1}{M}\sum_{i=1}^{M} \frac{N_{R,p^{(i)},v^{(i)}}}{R}}. \tag{8}$$

(Note that the denominator is uniformly bounded from below, which follows from equation (10) below.) By the strong law of large numbers, the function (8) converges almost surely to

$$\frac{\int_K \int_{\mathbb{S}^2} \frac{N_{R,p,v}}{R} \frac{N_{R,p,v}(t)}{N_{R,p,v}} \, dS(v) \, d\mu(p)}{\int_K \int_{\mathbb{S}^2} \frac{N_{R,p,v}}{R} \, dS(v) \, d\mu(p)} \tag{9}$$

as $M \to \infty$, where $d\mu$ is the probability measure with which we choose the starting points, and dS is the surface area measure on the sphere $\mathbb{S}^2$. By symmetry, we may restrict the inner integrals to $\mathbb{S}^2_+ := \{(v_x, v_y, v_z) \in \mathbb{S}^2 : v_x, v_y, v_z > 0\}$. We now look at the limit of (9) as $R \to \infty$, and we note that since the integrands are uniformly bounded, we may move the limit inside the integrals by the Lebesgue dominated convergence theorem. Fix one of the integrands, and denote it by $f(R, p, v, t)$. We will show that its limit $g(p, v, t) := \lim_{R\to\infty} f(R, p, v, t)$ exists for all t and all directions $v \in \mathbb{S}^2$. Moreover, if $p^{(i)}$ and $v^{(i)}$ denote random variables corresponding to an initial position and an initial direction, respectively, as above, then

$$h(p^{(i)}, v^{(i)}, t) := \lim_{R\to\infty} \frac{N_{R,p^{(i)},v^{(i)}}}{R} \frac{N_{R,p^{(i)},v^{(i)}}(t)}{N_{R,p^{(i)},v^{(i)}}}$$

is a random variable with finite variance (and similarly for the terms in the denominator of (8); in particular recall it is uniformly bounded from below), and thus the strong law of large numbers gives that the limit of (8) as $R \to \infty$, and then $M \to \infty$ almost surely equals (9). This shows that $\lim_{M\to\infty} \lim_{R\to\infty} \mathrm{cdf}_{X_{M,R}}(t)$ exists almost surely and is equal to $\lim_{R\to\infty} \lim_{M\to\infty} \mathrm{cdf}_{X_{M,R}}(t)$.

Consider a particle with initial position p and initial direction $v = (v_x, v_y, v_z) \in \mathbb{S}^2_+$. By "unfolding" its motion with specular reflections on the walls of the box to the motion along a straight line in $\mathbb{R}^n$ — see Figure 4 for a 2D illustration — we see that the particle's set of bounce lengths is identical to the set of path lengths between consecutive intersections of the straight line segment $\{p + tv : 0 \leq t \leq R\}$ with any of the planes $x = na, y = nb, z = nc$, $n \in \mathbb{Z}$. Thus we see that

$$N_{R,p,v} = R\frac{v_x}{a} + R\frac{v_y}{b} + R\frac{v_z}{c} + O(1) \tag{10}$$

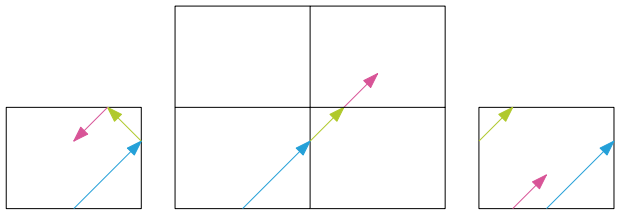


Figure 4: From left to right: Unfolding a motion with specular reflection in a 2D box to a motion the plane and then projecting back to the box.

for large R, and therefore

$$\frac{N_{R,p,v}}{R} \to \frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c} \tag{11}$$

as $R \to \infty$.

Now project the line $\{p+tv : 0 \le t \le R\}$ to the torus $\mathbb{R}^3/\Lambda$ where $\Lambda = \{(n_1 a, n_2 b, n_3 c) : n_1, n_2, n_3 \in \mathbb{Z}\}$ and let us identify the torus with the box K; see Figure 4. Each bounce length corresponds to a line segment which starts in one of the three planes $x = 0$, $y = 0$ or $z = 0$ and runs in the direction v to one of the three planes $x = a, y = b$ or $z = c$. There are $R\frac{v_z}{c} + O(1)$ line segments which start from the plane $z = 0$, and thus the probability that a line segment starts from the plane $z = 0$ is

$$\frac{\frac{v_z}{c}}{\frac{v_x}{a} + \frac{v_y}{b} + \frac{v_z}{c}}$$

as $R \to \infty$. By the ergodicity of the linear flow on tori (for almost all directions), the starting points of these line segments become uniformly distributed on the rectangle $[0, a] \times [0, b] \times \{0\}$ for almost all $v \in \mathbb{S}^2_+$ as $R \to \infty$; from here we will assume that v is such a direction, and we will ignore the measure zero set of directions for which we do not have ergodicity. Consider one of these line segments and denote its length by T and its starting point by $(x_0, y_0, 0)$. For an arbitrary parameter $t \ge 0$, we have $T \le t$ if and only if $tv_x \ge a - x_0$ or $tv_y \ge b - y_0$ or $tv_x \ge c$; the starting points $(x_0, y_0) \in [0, a] \times [0, b]$ which satisfy this are precisely those outside the rectangle $[0, a - tv_x] \times [0, b - tv_y]$ assuming that $tv_z \le c$ and otherwise it is the whole rectangle $[0, a] \times [0, b]$. The area of that region is

$$ab - (a - tv_x)(b - tv_y) \tag{12}$$

if $a \ge tv_x, b \ge tv_y, c \ge tv_z$ and otherwise it is ab. Since the starting points (x_0, y_0) are uniformly distributed in the rectangle $[0, a] \times [0, b]$ as $R \to \infty$, it follows that the probability that $T \le t$ is

$$1 - \frac{(a - tv_x)(b - tv_y)}{ab} \chi(a \ge tv_x, b \ge tv_y, c \ge tv_z),$$

where $\chi(P)$ is the indicator function which is 1 whenever the condition P is true, and 0 otherwise. We get analogous expressions for the case when a line segment starts in the plane $x = 0$ or $y = 0$ instead. Thus the proportion of all line segments with length at most t as $R \to \infty$ is

$$\begin{aligned}\lim_{R\to\infty} \frac{N_{R,p,v}(t)}{N_{R,p,v}} = {} & \frac{\frac{v_x}{a}}{\frac{v_x}{a}+\frac{v_y}{b}+\frac{v_z}{c}}\left(1 - \frac{(b-tv_y)(c-tv_z)}{bc}\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)\right) + \\ & \frac{\frac{v_y}{b}}{\frac{v_x}{a}+\frac{v_y}{b}+\frac{v_z}{c}}\left(1 - \frac{(a-tv_x)(c-tv_z)}{ac}\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)\right) + \\ & \frac{\frac{v_z}{c}}{\frac{v_x}{a}+\frac{v_y}{b}+\frac{v_z}{c}}\left(1 - \frac{(a-tv_x)(b-tv_y)}{ab}\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)\right)\end{aligned}$$

which can be written

$$\begin{aligned}1 - \frac{\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)}{abc(\frac{v_x}{a}+\frac{v_y}{b}+\frac{v_z}{c})} \times \\ \times \Big(v_x(b-tv_y)(c-tv_z) + v_y(a-tv_x)(c-tv_z) + v_z(a-tv_x)(b-tv_y)\Big). \quad (13)\end{aligned}$$

Recognizing that both integrands (11) and (13) are independent of the position p, we see that the limit of (9) as $R \to \infty$ may be written as

$$\begin{aligned}\lim_{R\to\infty}\lim_{M\to\infty} \mathrm{cdf}_{X_{M,R}}(t) = 1 - \frac{1}{\int_{\mathbb{S}^2_+}(v_xbc + av_yc + abv_z)\,dS(v)} \times \\ \times \int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\leq a/t \\ v_y\leq b/t \\ v_z\leq c/t}} ((abv_z + av_yc + v_xbc) - 2t(av_yv_z + v_xbv_z + v_xv_yc) + 3t^2v_xv_yv_z)\,dS(v) \quad (14)\end{aligned}$$

for all $t > 0$. The corresponding formula in n dimensions is given by

$$\lim_{R\to\infty}\lim_{M\to\infty} \mathrm{cdf}_{X_{M,R}}(t) = 1 - \frac{\displaystyle\int_{\substack{v\in\mathbb{S}^{n-1}_+ \\ v_i\leq a_i/t \\ \text{for } i=1,\ldots,n}} \left(\sum_{i=1}^n v_i \prod_{j\neq i}(a_i - tv_j)\right) dS(v)}{\left(\displaystyle\prod_{i=1}^n a_i\right)\displaystyle\int_{\mathbb{S}^{n-1}_+}\left(\sum_{i=1}^n \frac{v_i}{a_i}\right) dS(v)} \quad (15)$$

for all $t > 0$, where the side-lengths of the box K are $a_1, \ldots, a_n$ and dS is the surface area measure on $\mathbb{S}^{n-1}_+ \cap [0,\infty)^n$. (The denominator can be given explicitly by using Lemma 27 below.)

We have thus proved that the random variable $X_{M,R}$ converges in distribution to a random variable with probability density function given by (15) as we take $M \to \infty$ followed by taking $R \to \infty$, or alternatively, first taking $R \to \infty$ followed by taking $M \to \infty$. It remains to prove that this distribution agrees with the distribution of the random variable X defined in the introduction, and to determine the mean value of X.

2.1. Integral geometry

We start by recalling some standard facts from integral geometry (cf. [6, 11].) The set of directed straight lines ℓ in $\mathbb{R}^3$ can be parametrized by pairs (v, q) where $v \in \mathbb{S}^2$ is a unit vector pointing in the same direction as ℓ and $q \in v^\perp$ is the unique point in ℓ which intersects the plane through the origin which is orthogonal to v. The unique translation- and rotation-invariant measure (up to a constant) on the set of directed straight lines in $\mathbb{R}^3$ is $d\ell := dA(q)\, dS(v)$ where dA is the surface measure on the plane through the origin orthogonal to $v \in \mathbb{S}^2$, and dS is the surface area measure on $\mathbb{S}^2$.

Consider the set $L_{a,b,c}$ of directed straight lines in $\mathbb{R}^3$ which intersect the box K. Now, since $abv_z + av_yc + v_xbc$ is the area of the projection of the box K onto the plane $v^\perp$ for $v \in \mathbb{S}^2_+$, it follows that the total measure of $L_{a,b,c}$ with respect to $d\ell$ is

$$C_{a,b,c} := 8 \int_{\mathbb{S}^2_+} (abv_z + av_yc + v_xbc)\, dS(v) = 2\pi(ab + ac + bc)$$

where we used symmetry, and the integral may be evaluated by switching to spherical coordinates. It follows that $d\ell\,/C_{a,b,c}$ is a probability measure on the set of directed lines intersecting the box $L_{a,b,c}$. Let ℓ be a random directed line with respect to this measure, and define the random variable $X := \text{length}(\ell \cap K)$, as in the introduction. Let us determine the probability that $X \leq t$ for an arbitrary parameter $t \geq 0$. By symmetry it suffices to consider only directed lines with $v \in \mathbb{S}^2_+$. The set of all intersection points between the rectangle $[0, a] \times [0, b] \times \{0\}$ and the lines ℓ with $X \leq t$ and direction $v \in \mathbb{S}^2_+$ has area $ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)$, as in (12), and its projection onto the plane $v^\perp$ has area

$$v_z[ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)].$$

By symmetry it follows that the area of the set of directed lines $\ell \in L_{a,b,c}$ with $X \leq t$ and direction $v \in \mathbb{S}^2_+$ projected down to $v^\perp$ is

$$\begin{aligned} U(v, t) := {} & v_x[bc - (b - tv_y)(c - tv_z)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)] + \\ & v_y[ac - (a - tv_x)(c - tv_z)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)] + \\ & v_z[ab - (a - tv_x)(b - tv_y)\chi(a \geq tv_x, b \geq tv_y, c \geq tv_z)], \end{aligned}$$

and it follows that

$$\text{Prob}[X \leq t] = \frac{1}{C_{a,b,c}} \int_{X \leq t} d\ell = \frac{8}{C_{a,b,c}} \int_{\mathbb{S}^2_+} U(v, t)\, dS(v),$$

which we see is identical to (14), and we have thus proved that $X_{M,R}$ converges in distribution to X as we take $M \to \infty$ and then $R \to \infty$.

2.2. Computing the mean value

It remains to determine the mean value of X, and we will do this by exploiting the integral geometry interpretation of the random variable X. By symmetry it suffices to restrict to

directed lines ℓ with $v \in \mathbb{S}^2_+$. For fixed $v \in \mathbb{S}^2_+$, denote by $Q(v) = (K + \mathrm{span}(v)) \cap v^\perp$ the set of $q \in v^\perp$ such that the directed line ℓ parametrized by (v, q) intersects K. We note that $X\, dA(q)$ is a volume element of the box K for any fixed $v \in \mathbb{S}^2_+$, and thus integrating $X\, dA(q)$ over all q yields the volume of the box. Hence the mean value is essentially given by the ratio of the volume of the box to the surface area of the box, or more precisely,

$$\mathbb{E}[X] = \frac{8}{C_{a,b,c}} \int_{\mathbb{S}^2_+} \int_{Q(v)} X\, dA(q)\, dS(v) = \frac{8abc}{C_{a,b,c}} \int_{\mathbb{S}^2_+} dS(v) = \frac{2abc}{ab + ac + bc}.$$

In n dimensions we get a normalizing factor $\frac{\mathrm{Area}(K)}{2} \cdot 2^n \int_{\mathbb{S}^{n-1}_+} v_n\, dS(v)$, so with the aid of the Lemma 27 in the Appendix, it follows that the mean value in n dimensions is

$$\mathbb{E}[X] = \frac{1}{2^n \frac{1}{\pi} \frac{|\mathbb{S}^n|}{2^n} \frac{\mathrm{Area}(K)}{2}} 2^n \,\mathrm{Vol}(K) \frac{|\mathbb{S}^{n-1}|}{2^n} = 2\pi \frac{|\mathbb{S}^{n-1}|}{|\mathbb{S}^n|} \frac{\mathrm{Vol}(K)}{\mathrm{Area}(K)}$$

where $\mathrm{Area}(K)$ is the $(n-1)$-dimensional surface area of the box K, and $\mathrm{Vol}(K)$ is the volume of the box K. This concludes the proof of Theorem 1.

3. Proof of Theorem 2

Using formula (15) in dimension $n = 2$, we get

$$\mathrm{cdf}_X(t) = 1 - \frac{\displaystyle\int_{\substack{v \in \mathbb{S}^1_+ \\ v_x \le a/t \\ v_y \le b/t}} (v_x(b - tv_y) + v_y(a - tv_x))\, dS(v)}{ab \displaystyle\int_{\mathbb{S}^1_+} \left(\frac{v_x}{a} + \frac{v_y}{b} \right) dS(v)}.$$

We use polar coordinates $v_x = \cos\theta, v_y = \sin\theta$ so that $dS(v) = d\theta$. Then the above becomes

$$1 - \frac{\displaystyle\int_{\cos^{-1}(\min(a/t,1))}^{\sin^{-1}(\min(b/t,1))} (b\cos\theta + a\sin\theta - 2t\sin\theta\cos\theta)\, d\theta}{\displaystyle\int_0^{\pi/2} (b\cos\theta + a\sin\theta)\, d\theta} =$$

$$1 - \frac{1}{a+b} \Big[b\sin\theta - a\cos\theta + t\cos^2\theta \Big]_{\cos^{-1}(\min(a/t,1))}^{\sin^{-1}(\min(b/t,1))}. \tag{16}$$

The numerator of the second term may be written

$$\chi(b < t) \left(b \cdot \frac{b}{t} - a\sqrt{1 - \frac{b^2}{t^2}} + t\left(1 - \frac{b^2}{t^2}\right) \right) + \chi(b \ge t)(b - a \cdot 0 + t \cdot 0) +$$

$$-\chi(a < t) \left(b\sqrt{1 - \frac{a^2}{t^2}} - a \cdot \frac{a}{t} + t \cdot \frac{a^2}{t^2} \right) - \chi(a \ge t)(b \cdot 0 - a + t)$$

which can be simplified to

$$\chi(b<t)\left(t-b-a\sqrt{1-\frac{b^2}{t^2}}\right)+\chi(a<t)\left(t-a-b\sqrt{1-\frac{a^2}{t^2}}\right)+(a+b-t).$$

Inserting this into (16) and differentiating yields Theorem 2.

4. Proof of Theorem 4

We will evaluate the cumulative distribution function (14) and then differentiate. The denominator of the second term of (14) is

$$\int_{\mathbb{S}^2_+}(abv_z+av_yc+v_xbc)\,dS(v)=\frac{\pi}{4}(ab+ac+bc),$$

as may be evaluated by switching to spherical coordinates. Define

$$f(a,b,c):=bc\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_x\,dS(v),$$

$$g(a,b,c):=-2tc\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_xv_y\,dS(v),$$

$$h(a,b,c):=3t^2\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_xv_yv_z\,dS(v).$$

By symmetry, we have

$$f(c,a,b)=ab\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le c/t\\ v_y\le a/t\\ v_z\le b/t}} v_x\,dS(v)=ab\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_z\,dS(v),$$

$$f(b,c,a)=ac\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le b/t\\ v_y\le c/t\\ v_z\le a/t}} v_x\,dS(v)=ac\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_y\,dS(v),$$

$$g(c,a,b)=-2tb\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le c/t\\ v_y\le a/t\\ v_z\le b/t}} v_xv_y\,dS(v)=-2tb\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_xv_z\,dS(v),$$

$$g(b,c,a)=-2ta\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le b/t\\ v_y\le c/t\\ v_z\le a/t}} v_xv_y\,dS(v)=-2ta\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} v_yv_z\,dS(v),$$

and thus we can write the numerator in the second term of (14) as

$$f(a,b,c)+f(c,a,b)+f(b,c,a)+g(a,b,c)+g(c,a,b)+g(b,c,a)+h(a,b,c).$$

Exploiting the symmetries, it suffices to evaluate $h(a,b,c), g(a,b,c)$ and $f(b,c,a)$ (note the order of the arguments to f). We will evaluate these integrals by switching to spherical coordinates, but first we need to parametrize the part of the sphere inside the box $0 \le v_x \le a/t, 0 \le v_y \le b/t, 0 \le v_z \le c/t$.

Lemma 17. *Fix $t \in (0, \sqrt{a^2+b^2+c^2})$. We have*

$$\int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\le a/t \\ v_y\le b/t \\ v_z\le c/t}} F(v_x,v_y,v_z)\,dS(v) =$$
$$\left(\int_{\theta_{min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{max}}\int_{\varphi_b}^{\pi/2}\right)\tilde{F}(\theta,\varphi)\sin\theta\,d\varphi\,d\theta$$

for any integrable function $F:\mathbb{S}^2_+\to\mathbb{R}$, where $\tilde{F}(\theta,\varphi) := F(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, where

$$\begin{aligned}
\theta_{min} &:= \cos^{-1}\left\{\frac{c}{t}\right\}_1, \\
\theta_a &:= \max(\theta_{min}, \sin^{-1}\left\{\frac{a}{t}\right\}_1), \\
\theta_b &:= \max(\theta_{min}, \sin^{-1}\left\{\frac{b}{t}\right\}_1), \\
\theta_{max} &:= \sin^{-1}\left\{\frac{\sqrt{a^2+b^2}}{t}\right\}_1, \\
\varphi_a &:= \cos^{-1}\frac{a}{t\sin\theta} \qquad (\textit{whenever } a \le t\sin\theta), \\
\varphi_b &:= \sin^{-1}\frac{b}{t\sin\theta} \qquad (\textit{whenever } b \le t\sin\theta).
\end{aligned}$$

and where we have used the shorthand $\{u\}_1 := \min(u,1)$.

Proof. We will parametrize the set of points $v=(v_x,v_y,v_z)$ on the sphere $\mathbb{S}^2$ such that

$$\begin{aligned}
0 < v_x &\le a/t, \\
0 < v_y &\le b/t, \\
0 < v_z &\le c/t.
\end{aligned} \tag{18}$$

Switch to spherical coordinates $v_x = \sin\theta\cos\varphi, v_y = \sin\theta\sin\varphi, v_z = \cos\theta$. The non-negativity conditions of (18) are equivalent to the condition $\theta,\varphi \in (0,\pi/2)$. For such angles, the condition $v_z \le c/t$ is equivalent to

$$\cos^{-1}\left\{\frac{c}{t}\right\}_1 \le \theta,$$

and the conditions $v_x \le a/t, v_y \le b/t$ are equivalent to

$$\cos^{-1}\left\{\frac{a}{t\sin\theta}\right\}_1 \le \varphi \le \sin^{-1}\left\{\frac{b}{t\sin\theta}\right\}_1. \tag{19}$$

The interval (19) is non-empty for precisely those $\theta \in (0,\pi/2)$ such that $\theta \le \theta_{\max}$ since

$$1 \le \left\{\frac{a}{t\sin\theta}\right\}_1^2 + \left\{\frac{b}{t\sin\theta}\right\}_1^2 \iff 1 \le \left(\frac{a}{t\sin\theta}\right)^2 + \left(\frac{b}{t\sin\theta}\right)^2 \iff$$
$$\sin\theta \le \frac{\sqrt{a^2+b^2}}{t} \iff \theta \le \sin^{-1}\left\{\frac{\sqrt{a^2+b^2}}{t}\right\}_1.$$

Thus we may restrict θ to the interval given by the inequalities

$$\theta_{\min} \le \theta \le \theta_{\max}.$$

Note that we have $\theta_{\min} \le \theta_{\max}$ for all $t \le \sqrt{a^2+b^2+c^2}$ since

$$\theta_{\min} \le \theta_{\max} \iff 1 \le \left\{\frac{c}{t}\right\}_1^2 + \left\{\frac{\sqrt{a^2+b^2}}{t}\right\}_1^2 \iff$$
$$1 \le \left(\frac{c}{t}\right)^2 + \left(\frac{\sqrt{a^2+b^2}}{t}\right)^2 \iff t^2 \le a^2+b^2+c^2.$$

We conclude that we can write

$$\int_{\substack{v\in\mathbb{S}^2_+\\ v_x\le a/t\\ v_y\le b/t\\ v_z\le c/t}} F(v_x,v_y,v_z)\,dS(v) = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1}\left\{\frac{a}{t\sin\theta}\right\}_1}^{\sin^{-1}\left\{\frac{b}{t\sin\theta}\right\}_1} \tilde{F}(\theta,\varphi)\sin\theta\,d\varphi\,d\theta\,. \tag{20}$$

For $\theta \in (0,\pi/2)$, note that $\cos^{-1}\frac{a}{t\sin\theta}$ is defined precisely when $\sin^{-1}\left\{\frac{a}{t}\right\}_1 \le \theta$ and that $\sin^{-1}\frac{b}{t\sin\theta}$ is defined precisely when $\sin^{-1}\left\{\frac{b}{t}\right\}_1 \le \theta$. We have $\theta_{\min} < \theta_a$ if and only if $t < \sqrt{a^2+c^2}$, and we have $\theta_{\min} < \theta_b$ if and only if $t < \sqrt{b^2+c^2}$. Moreover we note that we always have $\theta_a, \theta_b \in [\theta_{\min}, \theta_{\max}]$.

Let us rewrite the integration limits in the right-hand side of (20) in terms of φ_a and φ_b. A priori, we need to distinguish between the two cases $\theta_a \le \theta_b$ and $\theta_b < \theta_a$. If $\theta_a \le \theta_b$ then we get

$$\left(\int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1}\left\{\frac{x}{t\sin\theta}\right\}_1}^{\sin^{-1}\left\{\frac{y}{t\sin\theta}\right\}_1}\right) = \left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_b}\int_{\varphi_a}^{\pi/2} + \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_a}^{\varphi_b}\right) =$$
$$\left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} + \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right) =$$
$$\left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right). \tag{21}$$

If on the other hand $\theta_b < \theta_a$ then

$$\left(\int_{\theta_{\min}}^{\theta_{\max}} \int_{\cos^{-1}\left\{\frac{x}{t\sin\theta}\right\}_1}^{\sin^{-1}\left\{\frac{y}{t\sin\theta}\right\}_1}\right) = \left(\int_{\theta_{\min}}^{\theta_b}\int_0^{\pi/2} + \int_{\theta_b}^{\theta_a}\int_0^{\varphi_b} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\varphi_b}\right) =$$
$$\left(\int_{\theta_{\min}}^{\theta_b}\int_0^{\pi/2} + \int_{\theta_b}^{\theta_a}\int_0^{\pi/2} - \int_{\theta_b}^{\theta_a}\int_{\varphi_b}^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right)$$

which we see is identical to (21). Combining (20) and (21) we get the conclusion of the lemma. □

Applying Lemma 17 we get

$$h(a,b,c) = 3t^2 \int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\le a/t \\ v_y\le b/t \\ v_z\le c/t}} v_x v_y v_z \, dS(v) =$$
$$3t^2\left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right)(\sin^2\theta\cos\theta\cos\varphi\sin\varphi)\sin\theta\, d\varphi\, d\theta\,.$$

An antiderivative of the integrand $\cos\varphi\sin\varphi\cdot\sin^3\theta\cos\theta$ with respect to φ is $-\frac{1}{2}\cos^2\varphi\sin^3\theta\cos\theta$, and thus the above is

$$3t^2\left(\int_{\theta_{\min}}^{\theta_a}\cos^2\varphi\Big|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}}\cos^2\varphi\Big|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}}\cos^2\varphi\Big|_{\varphi=\varphi_b}\right)\frac{1}{2}\sin^3\theta\cos\theta\, d\theta =$$
$$3t^2\left(\int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}}\frac{a^2}{t^2\sin^2\theta} + \int_{\theta_b}^{\theta_{\max}}\left(\frac{b^2}{t^2\sin^2\theta} - 1\right)\right)\frac{1}{2}\sin^3\theta\cos\theta\, d\theta =$$
$$\frac{3}{2}\left(\int_{\theta_{\min}}^{\theta_a} t^2\sin^3\theta\cos\theta\, d\theta + \int_{\theta_a}^{\theta_{\max}} a^2\sin\theta\cos\theta\, d\theta + \int_{\theta_b}^{\theta_{\max}}\left(b^2\sin\theta - t^2\sin^3\theta\right)\cos\theta\, d\theta\right) =$$

$$\frac{3}{2}\left(\left[t^2\frac{1}{4}\sin^4\theta\right]_{\theta_{\min}}^{\theta_a} + \left[a^2\frac{1}{2}\sin^2\theta\right]_{\theta_a}^{\theta_{\max}} + \left[b^2\frac{1}{2}\sin^2\theta - t^2\frac{1}{4}\sin^4\theta\right]_{\theta_b}^{\theta_{\max}}\right). \qquad (22)$$

Next consider

$$g(a,b,c) = -2tc\int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\le a/t \\ v_y\le b/t \\ v_z\le c/t}} v_x v_y \, dS(v) =$$
$$-2tc\left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right)(\sin^2\theta\cos\varphi\sin\varphi)\sin\theta\, d\varphi\, d\theta\,.$$

An antiderivative of the integrand $\cos\varphi\sin\varphi\cdot\sin^3\theta$ with respect to φ is $-\frac{1}{2}\cos^2\varphi\sin^3\theta$,

and thus the above is

$$g(a,b,c) = -2tc \int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\le a/t \\ v_y\le b/t \\ v_z\le c/t}} v_x v_y \, dS(v) =$$

$$-tc\left(\int_{\theta_{\min}}^{\theta_a} \cos^2\varphi\Big|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}} \cos^2\varphi\Big|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}} \cos^2\varphi\Big|_{\varphi=\varphi_b}\right)\sin^3\theta\, d\theta =$$

$$-tc\left(\int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}} \frac{a^2}{t^2\sin^2\theta} + \int_{\theta_b}^{\theta_{\max}} \left(\frac{b^2}{t^2\sin^2\theta} - 1\right)\right)\sin^3\theta\, d\theta =$$

$$-tc\left(\int_{\theta_{\min}}^{\theta_a} \sin^3\theta\, d\theta + \int_{\theta_a}^{\theta_{\max}} \frac{a^2\sin\theta}{t^2}\, d\theta + \int_{\theta_b}^{\theta_{\max}} \left(\frac{b^2\sin\theta}{t^2} - \sin^3\theta\right) d\theta\right) =$$

$$-tc\left(\left[\frac{\cos^3\theta}{3} - \cos\theta\right]_{\theta_{\min}}^{\theta_a} + \frac{a^2}{t^2}[-\cos\theta]_{\theta_a}^{\theta_{\max}} + \left[-\frac{b^2\cos\theta}{t^2} - \frac{\cos^3\theta}{3} + \cos\theta\right]_{\theta_b}^{\theta_{\max}}\right). \quad (23)$$

We obtain $g(b,c,a)$ and $g(c,a,b)$ by switching the roles of a,b,c in (23). We remark that trying to obtain $g(b,c,a)$ and $g(c,a,b)$ directly, by integrating $v_y v_z$ and $v_x v_z$, respectively, by first integrating with respect to φ, taking the limits $\varphi \to \varphi_a$ and $\varphi \to \varphi_b$, and then finding an antiderivative with respect to θ, seem to result in much more complicated expressions.

Finally consider

$$f(b,c,a) = ac \int_{\substack{v\in\mathbb{S}^2_+ \\ v_x\le a/t \\ v_y\le b/t \\ v_z\le c/t}} v_y \, dS(v) =$$

$$ac\left(\int_{\theta_{\min}}^{\theta_a}\int_0^{\pi/2} + \int_{\theta_a}^{\theta_{\max}}\int_{\varphi_a}^{\pi/2} - \int_{\theta_b}^{\theta_{\max}}\int_{\varphi_b}^{\pi/2}\right)(\sin\theta\sin\varphi)\sin\theta\, d\varphi\, d\theta\,.$$

An antiderivative of the integrand $\sin\varphi\cdot\sin^2\theta$ with respect to φ is $-\cos\varphi\cdot\sin^2\theta$, and thus the above is

$$ac\left(\int_{\theta_{\min}}^{\theta_a} \cos\varphi|_{\varphi=0} + \int_{\theta_a}^{\theta_{\max}} \cos\varphi|_{\varphi=\varphi_a} - \int_{\theta_b}^{\theta_{\max}} \cos\varphi|_{\varphi=\varphi_b}\right)\sin^2\theta\, d\theta =$$

$$ac\left(\int_{\theta_{\min}}^{\theta_a} 1 + \int_{\theta_a}^{\theta_{\max}} \frac{a}{t\sin\theta} - \int_{\theta_b}^{\theta_{\max}} \sqrt{1 - \frac{b^2}{t^2\sin^2\theta}}\right)\sin^2\theta\, d\theta =$$

$$ac\left(\int_{\theta_{\min}}^{\theta_a} \sin^2\theta\, d\theta + \int_{\theta_a}^{\theta_{\max}} \frac{a\sin\theta}{t}\, d\theta - \int_{\theta_b}^{\theta_{\max}} \sqrt{\sin^2\theta - \frac{b^2}{t^2}}\sin\theta\, d\theta\right) =$$

$$ac\left(\frac{1}{2}[\theta - \sin\theta\cos\theta]_{\theta_{\min}}^{\theta_a} + \left[\frac{-a\cos\theta}{t}\right]_{\theta_a}^{\theta_{\max}} - \int_{\theta_b}^{\theta_{\max}} \sqrt{1 - \frac{b^2}{t^2} - \cos^2\theta}\sin\theta\, d\theta\right) \quad (24)$$

where the last integral inside the parentheses may be written as

$$\left[-\frac{1}{2}\left(\cos\theta\sqrt{1-\frac{b^2}{t^2}-\cos^2\theta}+\left(1-\frac{b^2}{t^2}\right)\tan^{-1}\left(\frac{\cos\theta}{\sqrt{1-\frac{b^2}{t^2}-\cos^2\theta}}\right)\right)\right]_{\theta_b}^{\theta_{\max}} =$$

$$\left[-\frac{1}{2}\left(\cos\theta\sqrt{\sin^2\theta-\frac{b^2}{t^2}}+\left(1-\frac{b^2}{t^2}\right)\tan^{-1}\left(\frac{\cos\theta}{\sqrt{\sin^2\theta-\frac{b^2}{t^2}}}\right)\right)\right]_{\theta_b}^{\theta_{\max}}$$

whenever $\theta_b < \pi/2$, by using the fact that $\frac{1}{2}\left(x\sqrt{c-x^2}+c\tan^{-1}\left(\frac{x}{\sqrt{c-x^2}}\right)\right)$ is an antiderivative of $\sqrt{c-x^2}$ with respect to x when c is a constant. We obtain $f(b,c,a)$ and $f(c,a,b)$ by switching the roles of a,b,c in (24).

It remains to insert the limits $\theta_{\min},\theta_a,\theta_b,\theta_{\max}$ into the antiderivatives (22), (23) and (24) above. Noting that $\theta_{\min},\theta_a,\theta_b,\theta_{\max}$ are expressed in terms of piecewise-defined functions, the following manipulations will be useful. For any function ψ, we have

$$\begin{aligned}\psi(\theta_{\min}) &= \psi\left(\cos^{-1}\frac{c}{t}\right)\chi_c+\psi(\cos^{-1}1)(1-\chi_c)\\ &= \left(\psi\left(\cos^{-1}\frac{c}{t}\right)-\psi(0)\right)\chi_c+\psi(0)\end{aligned}$$

where $\chi_c := \chi(t>c)$. Similarly,

$$\psi(\theta_{\max}) = \left(\psi\left(\sin^{-1}\frac{\sqrt{a^2+b^2}}{t}\right)-\psi(\pi/2)\right)\chi_{a,b}+\psi(\pi/2)$$

where $\chi_{a,b} := \chi(\sqrt{a^2+b^2}>t)$, and

$$\begin{aligned}\psi(\theta_a) &= (1-\chi_a)\psi(\pi/2)+(\chi_a-\chi_{a,c})\psi\left(\sin^{-1}\frac{a}{t}\right)+\chi_{a,c}\psi\left(\cos^{-1}\frac{c}{t}\right)\\ &= \chi_{a,c}\cdot\left(\psi\left(\cos^{-1}\frac{c}{t}\right)-\psi\left(\sin^{-1}\frac{a}{t}\right)\right)+\chi_a\cdot\left(\psi\left(\sin^{-1}\frac{a}{t}\right)-\psi(\pi/2)\right)+\psi(\pi/2)\end{aligned}$$

and similarly, $\psi(\theta_b)$ can be written as

$$\chi_{b,c}\cdot\left(\psi\left(\cos^{-1}\frac{c}{t}\right)-\psi\left(\sin^{-1}\frac{b}{t}\right)\right)+\chi_b\cdot\left(\psi\left(\sin^{-1}\frac{b}{t}\right)-\psi(\pi/2)\right)+\psi(\pi/2).$$

With this we can evaluate $[\psi]_{\theta_{\min}}^{\theta_a}, [\psi]_{\theta_a}^{\theta_{\max}}, [\psi]_{\theta_b}^{\theta_{\max}}$. But since we know that we will get a function symmetric with respect to the values a,b,c, it suffices to keep only those terms with χ_a and $\chi_{a,b}$, say, and then the other terms may be evaluated by just switching the

order of a, b, c. Upon inserting the limits and differentiating, one obtains (after tedious calculations) that

$$\mathrm{pdf}_X(t) = \frac{F(a,b,c,t) + F(b,c,a,t) + F(c,a,b,t)}{3\pi t^3(ab+ac+bc)}$$

where

$$\begin{aligned}
F(a,b,c,t) &:= (8at^3 - 3t^4) + \\
\chi(t \geq a)&\left(\left(6t^4 - a^4 + 6\pi a^2 bc\right) - (8at^3 - 3t^4) - 4(b+c)\sqrt{|t^2-a^2|}(a^2+2t^2) \right) + \\
\chi(t \geq \sqrt{a^2+b^2})&\Bigg[a^4 + b^4 - 9t^4 - 6a^2b^2 + \sqrt{|t^2-a^2-b^2|}4c\left(a^2+b^2+2t^2\right) + \\
&\quad 4a\sqrt{|t^2-b^2|}(b^2+2t^2) - 12a^2bc \cdot \arctan\left(\frac{\sqrt{|t^2-a^2-b^2|}}{b}\right) + \\
&\quad 4b\sqrt{|t^2-a^2|}(a^2+2t^2) - 12ab^2c \cdot \arctan\left(\frac{\sqrt{|t^2-a^2-b^2|}}{a}\right) \Bigg].
\end{aligned}$$

Rewriting F as a piecewise function, we get Theorem (4).

5. Proof of Theorem 6

Consider the distribution of the random variable $Y_{M,N}$. Since we record the *same* number of bounces for each choice of angle φ we may replace the M-particle system with a one particle system Y_N as follows: randomly select, with uniform distribution, the angle φ and generate N bounce lengths and randomly select one of these bounce lengths (with uniform distribution); by the strong law of large numbers, $Y_{M,N}$ converges in distribution to Y_N as $M \to \infty$.

We now determine the limit distribution of Y_N. As before, we first unfold the motion, and replace motion in a box with specular reflections on the walls with motion in $\mathbb{R}^2$; see Figure 4. The path lengths between bounces is then the same as the lengths between the intersections with horizontal or vertical grid lines. To understand the spatial distribution, we project the dynamics to the torus $\mathbb{R}^2/\Lambda$ where Λ is the lattice

$$\Lambda = \{(n_1 a, n_2 b) : n_1, n_2 \in \mathbb{Z}\},$$

and we may identify the torus with the rectangle $[0,a] \times [0,b]$.

Let us first consider the motion of a single particle with an arbitrary initial position, and direction of motion given by an angle φ. Taking symmetries into account, we may assume that $\varphi \in [0, \pi/2]$. (Note that $\frac{d\varphi}{\pi/2}$ gives a probability measure on these angles.) If the particle travels a large distance $R > 0$, the number of intersections with horizontal, respectively vertical, grid lines is $\frac{R\sin\varphi}{b} + O(1)$, respectively $\frac{R\cos\varphi}{a} + O(1)$. Thus, in the limit $R \to \infty$, the probability of a line segment beginning at a horizontal (respectively

vertical) grid line is given by P_h, respectively P_v (here we suppress the dependence on φ) where

$$P_h := \frac{\frac{\sin\varphi}{b}}{\frac{\sin\varphi}{b} + \frac{\cos\varphi}{a}}, \quad P_v := \frac{\frac{\cos\varphi}{a}}{\frac{\sin\varphi}{b} + \frac{\cos\varphi}{a}}.$$

The unfolded flow on the torus is ergodic for almost all φ, and thus the starting points of the line segments becomes uniformly distributed as $R \to \infty$ for almost all φ.

Let

$$T = T(\varphi) := a/\cos\varphi.$$

Since $\sin\varphi = \sqrt{T^2 - a^2}/T$, we obtain that

$$P_h = \frac{\sqrt{T^2 - a^2}}{b + \sqrt{T^2 - a^2}}, \quad P_v = \frac{b}{b + \sqrt{T^2 - a^2}}.$$

Let $\theta = \arctan b/a$ denote the angle of the diagonal in the box, and assume that $0 \le \varphi \le \theta$. We then observe the following regarding the line segment lengths.

First, if the segment begins at a horizontal line, it must end at a vertical line, and the possible lengths of these segment lie between 0 and T. We find that these lengths are uniformly distributed in $[0, T]$ since the starting points of the segments are uniformly distributed.

On the other hand, if the line segment begins at a vertical line, it can either end at a vertical or horizontal line. Since the starting points are uniformly distributed, the former happens with probability

$$\frac{a\tan\varphi}{b} = \frac{a\frac{\sqrt{T^2-a^2}}{a}}{b} = \frac{\sqrt{T^2 - a^2}}{b}$$

and the length of the segment is again uniformly distributed in $[0, T]$, whereas the latter happens with probability

$$\frac{b - a\tan\varphi}{b} = 1 - \frac{\sqrt{T^2 - a^2}}{b}$$

in which case the segment is always of length T.

Now, $\varphi \in [0, \theta]$ implies that $T \in [a, \sqrt{a^2 + b^2}]$, and noting that

$$\frac{d\varphi}{dT} = \frac{a}{T\sqrt{T^2 - a^2}}$$

we find that the probability of observing a line segment of length t is the sum of a "singular part" (the segment begins and ends on vertical lines; note that all such segments have the *same* lengths) and a "smooth part" (the segment does not begin and end on vertical lines). Moreover, the smooth part contribution equals

$$\frac{1}{\pi/2}\int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{1}{T}\left(P_h + P_v\frac{a\tan\varphi}{b}\right)\frac{d\varphi}{dT}\,dT$$

which, on inserting (5), equals

$$\frac{1}{\pi/2}\int_{\max(a,t)}^{\sqrt{a^2+b^2}}\frac{1}{T}\cdot\left(\frac{\sqrt{T^2-a^2}}{b+\sqrt{T^2-a^2}}+\frac{b}{b+\sqrt{T^2-a^2}}\frac{a\tan\varphi}{b}\right)\cdot\frac{a}{T\sqrt{T^2-a^2}}\,dT=$$
$$\frac{1}{\pi/2}\int_{\max(a,t)}^{\sqrt{a^2+b^2}}\frac{1}{T}\cdot\left(\frac{\sqrt{T^2-a^2}}{b+\sqrt{T^2-a^2}}+\frac{b}{b+\sqrt{T^2-a^2}}\frac{\sqrt{T^2-a^2}}{b}\right)\cdot\frac{a}{T\sqrt{T^2-a^2}}\,dT=$$
$$\frac{1}{\pi/2}\int_{\max(a,t)}^{\sqrt{a^2+b^2}}\frac{2a}{b+\sqrt{T^2-a^2}}\cdot\frac{dT}{T^2}.$$

On the other hand, the "singular part contribution", provided $t\geq a$, to the probability of a segment having length t equals

$$\frac{P_v}{\pi/2}\cdot\frac{b-a\tan\varphi}{b}\cdot\frac{d\varphi}{dt}=\frac{1}{\pi/2}\cdot\frac{b}{b+\sqrt{t^2-a^2}}\cdot\left(1-\frac{\sqrt{t^2-a^2}}{b}\right)\cdot\frac{a}{t\sqrt{t^2-a^2}}=$$
$$\frac{1}{\pi/2}\cdot\frac{a}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}}\cdot\Big(b-\sqrt{t^2-a^2}\Big).$$

In case $\theta\leq\varphi\leq\pi/2$, a similar argument (we simple reverse the roles of a and b) shows that the smooth contribution equals

$$\frac{1}{\pi/2}\int_{\max(b,t)}^{\sqrt{a^2+b^2}}\frac{2b}{a+\sqrt{T^2-b^2}}\cdot\frac{dT}{T^2}$$

and that the singular contribution (if $t\geq b$) equals

$$\frac{1}{\pi/2}\cdot\frac{b}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}}\cdot\Big(a-\sqrt{t^2-b^2}\Big).$$

Thus, if we let $P_{\text{sing}}(t)$ denote the "singular contribution" to the probability density function we find the following: if $t<a$, then

$$P_{\text{sing}}(t)=0$$

if $t\in[a,b]$, then

$$P_{\text{sing}}(t)=\frac{1}{\pi/2}\cdot\frac{a\Big(b-\sqrt{t^2-a^2}\Big)}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}}$$

and if $t\in[b,\sqrt{a^2+b^2}]$, then

$$P_{\text{sing}}(t)=\frac{1}{\pi/2}\cdot\left(\frac{a(b-\sqrt{t^2-a^2})}{t(b+\sqrt{t^2-a^2})\sqrt{t^2-a^2}}+\frac{b(a-\sqrt{t^2-b^2})}{t(a+\sqrt{t^2-b^2})\sqrt{t^2-b^2}}\right).$$

Remark 25. Note that P_{sing} has a singularity of type $(t-a)^{-1/2}$ just to the right of $t=a$ (and similarly just to the right of $t=b$). In a sense this singularity arises from the singularity in the change of variables $\varphi\mapsto T$ since $\frac{d\varphi}{dT}=\frac{a}{T\sqrt{T^2-a^2}}$. The reason for the singularities in the spreading model for $n=2$ is similar, as the spreading model can be obtained from the absorption model by a smooth change of the angular measure.

Similarly, the "smooth part" of the contribution is (for $t \in [0, \sqrt{a^2+b^2}]$) given by

$$P_{\text{smooth}}(t) = \frac{1}{\pi/2}\left(\int_{\max(a,t)}^{\sqrt{a^2+b^2}} \frac{2a}{b+\sqrt{T^2-a^2}} \cdot \frac{dT}{T^2} + \int_{\max(b,t)}^{\sqrt{a^2+b^2}} \frac{2b}{a+\sqrt{T^2-b^2}} \cdot \frac{dT}{T^2}\right)$$

Hence the probability density function of the distribution of the segment length t is given by

$$\text{pdf}_Y(t) = P_{\text{sing}}(t) + P_{\text{smooth}}(t).$$

We will now evaluate $P_{\text{smooth}}(t)$. An antiderivative of $\frac{2a}{b+\sqrt{T^2-a^2}} \cdot \frac{1}{T^2}$ with respect to T for $T \in (a, \sqrt{a^2+b^2})$ is

$$\frac{2a(\sqrt{T^2-a^2}-b)}{T(a^2+b^2)} + \frac{2ab\Big(\tanh^{-1}\Big(\frac{T}{\sqrt{a^2+b^2}}\Big) - \tanh^{-1}\Big(\frac{\sqrt{T^2-a^2}\sqrt{a^2+b^2}}{Tb}\Big)\Big)}{(a^2+b^2)^{3/2}} \tag{26}$$

where $\tanh^{-1}(z) = \frac{1}{2}\log\frac{1+z}{1-z}$ for $|z|<1$. (A quick calculation shows that $\frac{\sqrt{T^2-a^2}\sqrt{a^2+b^2}}{Tb} < 1$ whenever $a < T < \sqrt{a^2+b^2}$.) We can rewrite (26) as

$$\frac{2a(\sqrt{T^2-a^2}-b)}{T(a^2+b^2)} + \frac{ab\log\left(\frac{\left(\sqrt{a^2+b^2}+T\right)\left(Tb-\sqrt{T^2-a^2}\sqrt{a^2+b^2}\right)}{\left(\sqrt{a^2+b^2}-T\right)\left(Tb+\sqrt{T^2-a^2}\sqrt{a^2+b^2}\right)}\right)}{(a^2+b^2)^{3/2}}$$

By l'Hôpital's rule we have

$$\lim_{T\to\sqrt{a^2+b^2}+} \frac{Tb-\sqrt{T^2-a^2}\sqrt{a^2+b^2}}{\sqrt{a^2+b^2}-T} = \lim_{T\to\sqrt{a^2+b^2}+} \frac{b-\frac{T}{\sqrt{T^2-a^2}}\sqrt{a^2+b^2}}{-1} = \frac{a^2}{b}$$

so the limit of (26) as $T \to \sqrt{a^2+b^2}+$ is

$$\frac{ab\log\left(\left(\frac{a^2}{b}\right)\cdot\frac{\left(\sqrt{a^2+b^2}+\sqrt{a^2+b^2}\right)}{\left(b\sqrt{a^2+b^2}+b\sqrt{a^2+b^2}\right)}\right)}{(a^2+b^2)^{3/2}} = \frac{2ab\log\left(\frac{a}{b}\right)}{(a^2+b^2)^{3/2}}.$$

The limit of (26) as $T \to a+$ is

$$\frac{-2b}{(a^2+b^2)} + \frac{2ab\tanh^{-1}\Big(\frac{a}{\sqrt{a^2+b^2}}\Big)}{(a^2+b^2)^{3/2}}.$$

Thus, assuming $a < b$, we can write $\frac{\pi}{2}P_{\text{smooth}}(t)$ as

$$\frac{2(a+b)}{(a^2+b^2)} - \frac{2ab}{(a^2+b^2)^{3/2}}\left(\tanh^{-1}\left(\frac{a}{\sqrt{a^2+b^2}}\right) + \tanh^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right)\right)$$

if $t < a, b$, or as

$$\frac{2ab + 2at - 2a\sqrt{t^2 - a^2}}{t(a^2 + b^2)} + \frac{2ab\Big(-\tanh^{-1}\Big(\frac{t}{\sqrt{a^2+b^2}}\Big) + \tanh^{-1}\Big(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\Big) - \tanh^{-1}\Big(\frac{b}{\sqrt{a^2+b^2}}\Big)\Big)}{(a^2 + b^2)^{3/2}}$$

if $a < t < b$ or as

$$2\frac{2ab - a\sqrt{t^2 - a^2} - b\sqrt{t^2 - b^2}}{t(a^2 + b^2)} + \frac{2ab\Big(-2\tanh^{-1}\Big(\frac{t}{\sqrt{a^2+b^2}}\Big) + \tanh^{-1}\Big(\frac{\sqrt{t^2-a^2}\sqrt{a^2+b^2}}{tb}\Big) + \tanh^{-1}\Big(\frac{\sqrt{t^2-b^2}\sqrt{a^2+b^2}}{ta}\Big)\Big)}{(a^2 + b^2)^{3/2}}$$

if $a, b < t$. Adding $P_{\text{sing}}(t)$ to this, we get Theorem 6.

A. Calculation of an integral

Lemma 27. *Write $|\mathbb{S}^{n-1}|$ for the $(n-1)$-dimensional surface area of the sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$. Then we have*

$$\int_{\mathbb{S}^{n-1}_+} v_n \, dS(v) = \frac{1}{\pi} \frac{|\mathbb{S}^n|}{2^n}.$$

where $\mathbb{S}^{n-1}_+ := \mathbb{S}^{n-1} \cap (0, \infty)^n$ is the part of the sphere $\mathbb{S}^{n-1}$ with positive coordinates.

Proof. We may parametrize $v = (v_1, \ldots, v_n) \in \mathbb{S}^{n-1}_+$ with

$$\begin{aligned}
v_1 &= \cos\theta_1 \\
v_2 &= \sin\theta_1 \cos\theta_2 \\
v_3 &= \sin\theta_1 \sin\theta_2 \cos\theta_3 \\
&\vdots \\
v_{n-1} &= \sin\theta_1 \cdots \sin\theta_{n-2} \cos\theta_{n-1} \\
v_n &= \sin\theta_1 \cdots \sin\theta_{n-2} \sin\theta_{n-1}
\end{aligned}$$

for $\theta_1, \ldots, \theta_{n-1} \in (0, \pi/2)$. We have the spherical area element

$$dS(v) = \sin^{n-2}\theta_1 \sin^{n-3}\theta_2 \cdots \sin\theta_{n-2} \, d\theta_1 \cdots d\theta_{n-1} \, .$$

Thus we get

$$\int_{\mathbb{S}^{n-1}_+} v_n \, dS(v) = \prod_{i=1}^{n-1} \int_0^{\pi/2} \sin^{n-1-i}\theta_i \, d\theta_i \, .$$

Introducing an additional integration variable θ_n, we recognize the integrand as the spherical area element in $n+1$ dimensions, and thus the above is

$$\frac{1}{\int_0^{\pi/2} d\theta_n} \prod_{i=1}^{n} \int_0^{\pi/2} \sin^{n-1-i} \theta_i \, d\theta_i = \frac{1}{\pi/2} \frac{|\mathbb{S}^n|}{2^{n+1}}.$$

since $\int_{\mathbb{S}^n_+} dS(v) = |\mathbb{S}^n|/2^{n+1}$. □

References

[1] C. Boldrighini, L. A. Bunimovich, and Y. G. Sinaĭ. On the Boltzmann equation for the Lorentz gas. *J. Statist. Phys.*, 32(3):477–501, 1983.

[2] J. Bourgain, F. Golse, and B. Wennberg. On the distribution of free path lengths for the periodic Lorentz gas. *Comm. Math. Phys.*, 190(3):491–508, 1998.

[3] L. A. Bunimovich and Y. G. Sinaĭ. Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.*, 78(4):479–497, 1980/81.

[4] M. Bäckström, S. Holmin, P. Kurlberg, D. Månsson. *Randomized Ray Tracing for Modeling UWB Transients in a Reverberation Chamber.* In preparation.

[5] F. Golse and B. Wennberg. On the distribution of free path lengths for the periodic Lorentz gas. II. *M2AN Math. Model. Numer. Anal.*, 34(6):1151–1163, 2000.

[6] D. A. Klain and G.-C. Rota. *Introduction to geometric probability.* Lezioni Lincee. [Lincei Lectures]. Cambridge University Press, Cambridge, 1997.

[7] D. Månsson, personal communication.

[8] J. Marklof and A. Strömbergsson. The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems. *Ann. of Math. (2)*, 172(3):1949–2033, 2010.

[9] J. Marklof and A. Strömbergsson. The Boltzmann-Grad limit of the periodic Lorentz gas. *Ann. of Math. (2)*, 174(1):225–298, 2011.

[10] J. Marklof and A. Strömbergsson. Free path lengths in quasicrystals. *Comm. Math. Phys.*, 330(2):723–755, 2014.

[11] L. A. Santaló. *Integral geometry and geometric probability.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2004. With a foreword by Mark Kac.

[12] H. Spohn. The Lorentz process converges to a random flight process. *Comm. Math. Phys.*, 60(3):277–290, 1978.

[13] B. Wennberg. Free path lengths in quasi crystals. *J. Stat. Phys.*, 147(5):981–990, 2012.