THE CIRCLE PROBLEM
OPEN PROBLEMS IN NUMBER THEORY
SPRING 2018, TEL AVIV UNIVERSITY

ZEÉV RUDNICK

Contents

1. Lattice point problems 1
1.1. The circle problem 1
1.2. The Dirichlet divisor problem 3
2. A better remainder term 5
2.1. A smooth counting function 6
2.2. A lower bound 9
3. Higher dimension 11
3.1. An Omega result 11
3.2. Sums of d squares - a survey 12
Appendix A. Background on Fourier analysis 13

1. Lattice point problems

1.1. The circle problem. Let $N(R)$ denote the number of lattice points in a circle of radius R

$$N(R) := \#\mathbb{Z}^2 \cap B(0,R)$$

where $B(R) = \{x \in \mathbb{R}^2 : |x| \leq R\}$, $|(x_1,x_2)|^2 = x_1^2 + x_2^2$. We will want to know the asymptotic behaviour of $N(R)$ as $R \to \infty$. More generally, we will consider a nice bounded domain $\Omega \subset \mathbb{R}^2$, say containing the origin, and ask for the number of lattice points in the homogeneously expanding domain $R\Omega$ for $R \to \infty$:

$$N_\Omega(R) := \#\mathbb{Z}^2 \cap R\Omega$$

The natural guess would be the area of $R\Omega$, which is $\operatorname{area}(R\Omega) = R^2 \operatorname{area}(\Omega)$. Our first result would be to confirm this guess; we do this only for the circle $\Omega = B(0,1)$, which has area π.

Proposition 1.1.

$$N(R) = \pi R^2 + O(R)$$

We will see that the result is really in terms of the geometry of the problem:

$$N(R) = \operatorname{area} B(0,R) + O\Big(\operatorname{length}\left(\partial B(0,R)\right) \Big)$$

We will give two proofs, essentially similar.

Date: May 2, 2018.

1.1.1. *First proof.*

Proof. We consider two polygons inscribing and circumscribing the disk $B(0,R)$:

$$P_- \subseteq B(0,R) \subseteq P_+$$

where

- P_- is the union of all squares with unit side $\Box_p$ centered at lattice points $p \in \mathbb{Z}^2 \cap B(0,R)$ so that $\Box_p \subseteq B(0,R)$ is wholly contained in the disk. Hence $P_- \subseteq B(0,R)$ and $\operatorname{area} P_- \leq \pi R^2$. Moreover

$$L_- := \left\{p \in \mathbb{Z}^2 : \Box_p \subseteq P_-\right\} \subseteq \mathbb{Z}^2 \cap B(0,R)$$

- P_+ is the unit of unit squares $\Box_p$ centered at lattice points $p \in \mathbb{Z}^2$, so that $\Box_p \cap B(0,R) \neq \emptyset$ intersects the disk.

$$\mathbb{Z}^2 \cap B(0,R) \subseteq L_+ := \left\{p \in \mathbb{Z}^2 : \Box_p \subseteq P_+\right\}$$

Hence

$$\#L_- \leq N(R) := \#\mathbb{Z}^2 \cap B(0,R) \leq \#L_+$$

Now because $P_\pm$ are unions of *unit* squares, their area is just the number of squares they contain, that is $\#L_\pm = \operatorname{area} P_\pm$. Hence

$$\operatorname{area} P_- \leq N(R) \leq \operatorname{area} P_+$$

The important observation is that every point sufficiently far from the boundary of $B(0,R)$ is inside a unit square $\Box_p$ wholly contained in $B(0,R)$, so that

$$P_- \supseteq B(0, R - \sqrt{2})$$

and every unit square $\Box_p$ that intersects $B(0,R)$ is at distance at most $R + \sqrt{2}$ from the origin, so that

$$P_+ \subset B(0, R + \sqrt{2})$$

Hence

$$\pi(R - \sqrt{2})^2 \leq \operatorname{area} P_- \leq N(R) \leq \operatorname{area} P_+ \leq \pi(R + \sqrt{2})^2$$

Expanding out gives

$$|N(r) - \pi R^2| \leq 2\sqrt{2}\pi R + 2\pi = O(R).$$

□

Important note: The argument works when we replace the circle $B(0,R)$ by the dilate $R\Omega$ of any fixed *convex* region with (say) smooth boundary.

1.1.2. *Second proof.*

Proof. We slice the ball $B(0,R)$ by vertical lines (n,y), and count the number of lattice points in each such line segment. The segment L_n with x-coordinate set to be n has y running between $-\sqrt{R^2 - n^2} \leq y \leq \sqrt{R^2 - n^2}$. Now the number of lattice points in a line segment satisfies

$$\#\{n \in \mathbb{Z} : a \leq n \leq b\} = (b - a) + O(1)$$

and therefore the line segment L_n contains $2\sqrt{R^2-n^2}+O(1)$ lattice points. Summing over all admissible n's, namely $-R \le n \le R$ we get

$$\begin{aligned} N(R) &= \sum_{-R\le n\le R}\Big(2\sqrt{R^2-n^2}+O(1)\Big) \\ &= 2\sum_{-R\le n\le R}\sqrt{R^2-n^2}+O(R) \\ &= r\sum_{1\le n\le R}\sqrt{R^2-n^2}+O(R) \end{aligned}$$

To evaluate the sum, we use summation by parts, with $a_n = 1$, $f(t) = \sqrt{R^2-t^2}$

$$\sum_{1\le n\le R}\sqrt{R^2-n^2} = \lfloor t\rfloor\sqrt{R^2-t^2}\Big|_0^R + \int_0^R \lfloor t\rfloor \frac{t}{\sqrt{R^2-t^2}}dt$$

Writing $\lfloor t\rfloor = t+O(1)$, we obtain

$$\begin{aligned} \int_0^R \lfloor t\rfloor \frac{t}{\sqrt{R^2-t^2}}dt &= \int_0^R \frac{t^2}{\sqrt{R^2-t^2}}dt + O\Big(\int_0^R \frac{t}{\sqrt{R^2-t^2}}dt\Big) \\ &= R^2\int_0^1 \frac{x^2}{\sqrt{1-x^2}}dx + O(R) = \frac{\pi}{4}R^2+O(R) \end{aligned}$$

Hence

$$\sum_{1\le n\le R}\sqrt{R^2-n^2} = \frac{\pi}{4}R^2+O(R)$$

which gives $N(R) = \pi R^2 + O(R)$. □

The goal is to understand the remainder term in the lattice point problem

$$P(R) := N(R) - \pi R^2$$

We saw that $P(R) = O(R)$.

Open Problem 1. *Show that for all $\varepsilon > 0$,*

$$P(R) = O(R^{1/2+\varepsilon}) = O\Big(\Big(\operatorname{length}\partial B(0,R)\Big)^{1/2+\varepsilon}\Big)$$

1.2. **The Dirichlet divisor problem.** Let $d(n) = \#\{(a,b) : a,b \ge 1, ab = n\}$ be the number of divisors of n. We have $d(1) = 1$, $d(p) = 2$ for p prime, and more generally $d(p^k) = k+1$. It is a multiplicative function: $d(mn) = d(m)d(n)$ if m,n are coprime. Hence we get a formula in terms of the prime decomposition of n:

$$d(\prod_j p_j^{k_j}) = \prod_j (k_j+1)$$

if p_j are distinct primes.

We can compute the *average value* of $d(n)$ by solving a lattice point problem:

$$\frac{1}{N}\sum_{n=1}^{N} d(n) = \frac{1}{N}\#\{(a,b)\in\mathbb{Z}^2 : a,b\ge 1, a\cdot b\le N\}$$

so we want the number of lattice points under the hyperbola $xy = n$ and in the positive quadrant $x,y \ge 1$. Thus let

$$D(N) := \#\{(a,b)\in\mathbb{Z}^2 : a,b\ge 1, a\cdot b\le N\}$$

Theorem 1.2.
$$D(N) = N\log N + (2C-1)N + O(\sqrt{N})$$
where $C = \lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n} - \log N\right) = 0.57721\ldots$ *is Euler's constant.*

Proof. As a first step, we try to reproduce Proof 1.1.2, by slicing the hyperbola with vertical segments
$$L_n = \{(n,y) : 1 \le y \le N\}$$
each of length $\text{length}(L_n) = N/n$ and then
$$\begin{aligned} D(N) &= \sum_{1\le n\le N} \#L_n = \sum_{1\le n\le N}\Big(\text{length}(L_n) + O(1)\Big) \\ &= \sum_{1\le n\le N}\frac{N}{n} + O(N) = N\Big(\log N + C + O(\frac{1}{N})\Big) + O(N) \end{aligned}$$
Thus we obtain
$$D(N) = N\log N + O(N)$$
Note that the area of the hyperbolic region $\{(x,y) : x,y \ge 1, x\cdot y \le N\}$ is $N\log N$, so again the main term is an area. However, the remainder term $O(N)$ falls far short of the $O(\sqrt{\text{area}(B(0,R)}) = O(R)$ remainder term that we obtained for the circle problem.

1.2.1. *Dirichlet's hyperbola method:* To overcome this, observe that the estimate
$$\#L_n = \text{length}(L_n) + O(1) = N/n + O(1)$$
is not good if n is large. So instead, divide the hyperbolic region into three parts (see Figure 1): A square $\square = \{1 \le x, y \le \sqrt{N}\}$, and two symmetric hyperbolic regions
$$H_1 = \{(x,y) : 1 \le x \le \sqrt{N}, \sqrt{N} < y \le \frac{N}{x}\}, \quad H_2 = \{(x,y) : 1 \le y \le \sqrt{N}, \sqrt{N} < x \le \frac{N}{x}\}$$
which contain the same number of lattice points. Hence
$$D(N) = \#\square + 2\#H_1$$
It is easy to compute $\#\square$:
$$\#\square = \Big(\lfloor\sqrt{N}\rfloor\Big)^2 = \Big(\sqrt{N} + O(1)\Big)^2 = N + O(\sqrt{N})$$
To compute H_1, use the slicing method again to obtain
$$\begin{aligned} \#H_1 &= \sum_{1\le n\le\sqrt{N}} \#\{\sqrt{N} < m \le \frac{N}{n}\} = \sum_{1\le n\le\sqrt{N}}\Big(\frac{N}{n} - \sqrt{N} + O(1)\Big) \\ &= N\Big(\log\Big(\sqrt{N} + O(1)\Big) + C + O\left(\frac{1}{\sqrt{N}}\right)\Big) - \Big(\sqrt{N} + O(1)\Big)\sqrt{N} + O(\sqrt{N}) \\ &= \frac{1}{2}N\log N + CN - N + O(\sqrt{N}) \end{aligned}$$
Thus
$$\begin{aligned} D(N) &= \#\square + 2\#H_1 = N + O(\sqrt{N}) + 2\Big(\frac{1}{2}N\log N + CN - N + O(\sqrt{N})\Big) \\ &= N\log N + (2C-1)N + O(\sqrt{N}) \end{aligned}$$
□

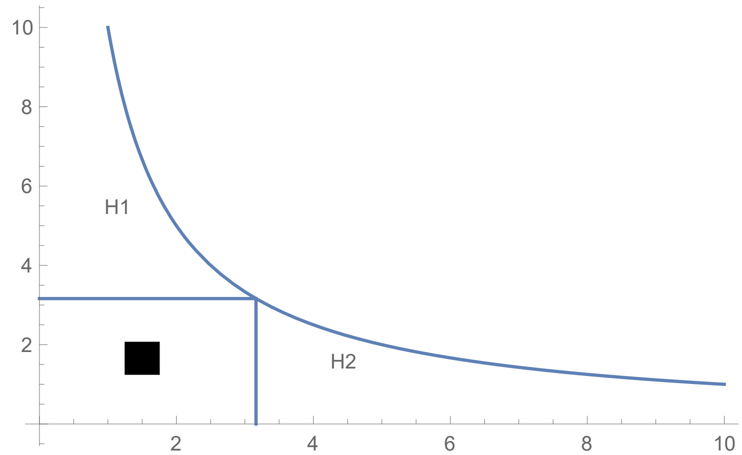

FIGURE 1. Dirichlet's hyperbola method.

Open Problem 2. *Show that* $\Delta(N) = O(N^{1/4+\varepsilon})$ *for all* $\varepsilon > 0$.

2. A better remainder term

We want to show that

Theorem 2.1. $P(R) := N(R) - \pi R^2$ *satisfies*

$$P(R) \ll R^{2/3}$$

History: The exponent $2/3$ was proved by different methods by Voronoi (1903), Sierpinski (1906), van der Corput (1923). The first improvement on $2/3 = 0.6666\dots$ was to $33/50 = 0.66000$ by van der Corput (1922). The have been various improvements over the past 100 years, the current record due to Bourgain (2017): $517/824 = 0.627427\dots$. The conjecture is that the remainder is $P(R) = O(R^{1/2+o(1)})$.

We will need some basics of Fourier analysis, and estimates on oscillatory integrals (van der Corput).

2.0.1. *Approximate identity.* Let $0 \le \Psi \le 1$ be a bump function on $\mathbb{R}^2$, smooth, rotationally symmetric and supported in the ball $B(0,1)$ of radius 1, and normalized so that $\int_{\mathbb{R}^2} \Psi(\vec{x}) d\vec{x} = 1$.

We can create such a function by taking a one-dimensional *even* bump function $\psi \in C^\infty(-1,1)$ and taking $\Psi(\vec{x}) := \psi(|\vec{x}|)$, and normalizing appropriately so that $\int_{\mathbb{R}^2} \Psi(\vec{x}) d\vec{x} = 2\pi \int_0^\infty \psi(r) r dr = 1$. (why is it possible?).

For $\varepsilon > 0$, set

$$\Psi_\varepsilon(\vec{x}) := \frac{1}{\varepsilon^2} \Psi(\frac{\vec{x}}{\varepsilon})$$

which is now supported in the ball $B(0,\varepsilon)$ and still has total mass 1.

Note: Such a family Ψ_ε is an *approximate identity*: For various classes of function spaces, we have

$$f * \Psi_\varepsilon \to f, \quad \text{as } \varepsilon \to 0$$

in the appropriate topology.

Let χ be the indicator function of the unit ball $B(0,1)$, and set

$$\chi_\varepsilon := \chi * \Psi_\varepsilon$$

Lemma 2.2. *χ_ε is supported in $B(0,1+\varepsilon)$ and coincides with χ in the smaller ball $B(0,1-\varepsilon)$:*

$$\chi_\varepsilon(\vec{x}) = \begin{cases} 1, & |\vec{x}| < 1-\varepsilon \\ 0, & |\vec{x}| > 1+\varepsilon \end{cases}$$

Moreover, $0\chi_\epsilon \leq 1$.

Proof. By definition

$$\chi_\varepsilon(x) = \int_{B(0,\varepsilon)} \frac{1}{\varepsilon^2}\Psi(\frac{z}{\varepsilon})\chi(x-z)dz$$

Now if $|x| \leq 1-\varepsilon$ and $z \in B(0,\varepsilon)$ then

$$|x-z| \leq |x| + |z| \leq 1-\varepsilon+\varepsilon = 1$$

so that $\chi(x-z) = 1$, and then

$$\chi_\varepsilon(x) = \int_{B(0,\varepsilon)} \Psi_\varepsilon(z)\chi(x-z)dz = \int_{B(0,\varepsilon)} \Psi_\varepsilon(z)dz = 1.$$

If $|x| > 1+\varepsilon$, and $z \in B(0,\varepsilon)$ then

$$|x-z| \geq \Big||x|-|z|\Big| = |x|-|z| > 1+\varepsilon-\varepsilon = 1$$

and then $\chi(x-z) = 0$, so that

$$\chi_\varepsilon(x) = \int_{B(0,\varepsilon)} \Psi_\varepsilon(z)\chi(x-z)dz = 0$$

for all $x \notin B(0,1+\varepsilon)$.

To see that $0 \leq \chi_\epsilon \leq 1$, just observe that since both χ and Ψ are non-negative, so is their convolution, and since $\chi \leq 1$ we have

$$\chi_\varepsilon(x) = \int \Psi_\varepsilon(z)\chi(x-z)dz \leq \int \Psi_\varepsilon(z)\cdot 1\, dz = 1$$

□

2.1. **A smooth counting function.** Define

$$N_\varepsilon(R) = \sum_{n\in\mathbb{Z}^2} \chi_\varepsilon(\frac{n}{R})$$

which counts lattice points with the smooth weight χ_ε. We claim that

Lemma 2.3. *For $0 < \varepsilon \ll 1$*

$$N_\varepsilon(\frac{R}{1+\varepsilon}) \leq N(R) \leq N_\varepsilon(\frac{R}{1-\varepsilon})$$

Proof. We first show

$$N\Big(R(1-\varepsilon)\Big) \leq N_\varepsilon(R) \leq N\Big(R(1+\varepsilon)\Big)$$

Indeed, to be counted in the sum for $N_\varepsilon(R)$, we must have $n/R \in \operatorname{supp}\chi_\varepsilon \subseteq B(0,1+\varepsilon)$, so that $|n/R| \leq 1+\varepsilon$. Since $\chi_\varepsilon \leq 1$, we obtain

$$N_\varepsilon(R) = \sum_{n\in\mathbb{Z}^2} \chi_\varepsilon(\frac{n}{R}) \leq \sum_{|n|\leq R(1+\varepsilon)} 1 = N\Big(R(1+\varepsilon)\Big)$$

Likewise, if $|n/R| < 1-\varepsilon$ then $\chi_\varepsilon(n/R) = 1$, so that

$$N_\varepsilon(R) \geq \sum_{|n|<R(1-\varepsilon)} 1 = N\Big(R(1-\varepsilon)\Big)$$

Changing variables $R \mapsto R/(1\pm\varepsilon)$ we deduce our claim. □

2.1.1. *Evaluating the smooth counting function.*

Lemma 2.4.

$$N_\varepsilon(R) = \pi R^2 + O\Big(\frac{1}{\varepsilon^{1/2}}\Big)$$

Proof. We use Poisson summation to transform N_ε:

$$N_\varepsilon(R) = \sum_{n\in\mathbb{Z}^2} \chi_\varepsilon(\frac{n}{R}) = \sum_{m\in\mathbb{Z}^2} R^2 \widehat{\chi}_\varepsilon(Rm)$$

since the Fourier transform of a dilated function $f(x/R)$ is $R^2\widehat{f}(Ry)$.

Now the Fourier transform of the convolution χ_ε is

$$\widehat{\chi}_\varepsilon = \widehat{\chi * \Psi_\varepsilon} = \widehat{\chi}\ \cdot\ \widehat{\Psi}_\varepsilon$$

and $\widehat{\Psi}_\varepsilon(y) = \widehat{\Psi}(\varepsilon y)$, so that

$$\widehat{\chi}_\varepsilon(Rm) = \widehat{\chi}(Rm)\widehat{\Psi}(R\varepsilon m)$$

and hence

$$\begin{aligned} N_\varepsilon(R) &= \sum_{m\in\mathbb{Z}^2} R^2\widehat{\chi}(Rm)\widehat{\Psi}(R\varepsilon m) \\ &= \widehat{\chi}(0)R^2 + R^2\sum_{m\neq 0}\widehat{\chi}(Rm)\widehat{\Psi}(R\varepsilon m) \end{aligned}$$

We have

$$\widehat{\chi}(0) = \int_{\mathbb{R}^2}\chi(y)dy = \operatorname{area} B(0,1) = \pi.$$

It does no great harm to pretend to that $\widehat{\Psi}$ is compactly supported (rather than just rapidly decaying), so that the sum is truncated at $R\varepsilon|m| \ll 1$, or $|m| < 1/(R\varepsilon)$. Thus up to an error which we will estimate later ???

$$N_\varepsilon(R) = \pi R^2 + O\Big(\sum_{0<|m|<(R\varepsilon)^{-1}} R^2\widehat{\chi}(Rm)\Big)$$

Now we use van der Corput's bound **??**

$$\widehat{\chi}(Rm) \ll (R|m|)^{-3/2}, \quad |m| \geq 1$$

to obtain

$$\sum_{0<|m|<(R\varepsilon)^{-1}} R^2\widehat{\chi}(Rm) \ll R^{1/2}\sum_{0<|m|<(R\varepsilon)^{-1}}\frac{1}{|m|^{3/2}}$$

We estimate the lattice sum (using partial summation) by the integral (exercise 1)

$$\sum_{0<|m|<M}\frac{1}{|m|^{3/2}} \ll \int_{1<|x|<M}\frac{dx}{|x|^{3/2}} \ll \int_1^M \frac{rdr}{r^{3/2}} \ll M^{1/2}$$

Thus

$$R^{1/2}\sum_{0<|m|<(R\varepsilon)^{-1}}\frac{1}{|m|^{3/2}} \ll R^{1/2}(R\varepsilon)^{-1/2} = \varepsilon^{-1/2}$$

which gives $N_\varepsilon(R) = \pi R^2 + O(\varepsilon^{-1/2})$. □

Exercise 1.

$$\sum_{0<|m|<M} \frac{1}{|m|^{3/2}} \ll M^{1/2}$$

We can now prove Theorem 2.1: We use Lemma 2.3 and Lemma 2.4 to deduce that

$$\pi(\frac{R}{1+\varepsilon})^2 + O(\varepsilon^{-1/2}) \le N(R) \le \pi(\frac{R}{1-\varepsilon})^2 + O(\varepsilon^{-1/2})$$

Now

$$(\frac{R}{1\pm\varepsilon})^2 = R^2(1+O(\varepsilon)) = R^2 + O(R^2\varepsilon)$$

and so

$$N(R) = \pi R^2 + O\Big(R^2\varepsilon + \varepsilon^{-1/2}\Big)$$

Choosing $\varepsilon^{-1/2} = R^2\varepsilon$, that is $\varepsilon = R^{-4/3}$, gives

$$N(R) = \pi R^2 + O(R^{2/3})$$

as claimed.

2.2. **A lower bound.** We next show that the conjectured exponent of $P(R) = O(R^{1/2+\varepsilon})$ cannot be improved, by showing

Theorem 2.5. *There is some $c > 0$ so that there are arbitrarily large R for which $|P(R)| > cR^{1/2}$.*

Let $S(R)$ be the normalized remainder term $P(R)/R^{1/2}$:

$$S(t) = \frac{N(t) - \pi t^2}{\sqrt{t}} = t^{-1/2} P(t)$$

We invoke, without providing a proof[1], a series representation of $S(t)$:

Proposition 2.6. *For any $T \gg 1$, uniformly for $t \in [T/(10), 10T]$*

$$S(t) = -\frac{1}{\pi} \sum_{\substack{0<|\vec{m}|\le T^{3/4} \\ 0\neq \vec{m}\in\mathbb{Z}^2}} \frac{\cos(2\pi|\vec{m}| \cdot t + \frac{\pi}{4})}{|\vec{m}|^{3/2}} + O(T^{-1/4+o(1)})$$

Motivation: We saw that the Fourier transform of the unit disk played a role in the formula for the smooth counting function. We can pretend that we can apply Poisson summation to the sharp counting function $N(R)$, and try to write

$$N(R) - \pi R^2 \text{``} = \text{''} \sum_{0\neq \vec{m}\in\mathbb{Z}^2} R^2 \widehat{\chi}(R\vec{m})$$

We expressed $\widehat{\chi}$ as an oscillatory integral

$$\widehat{\chi}(\vec{y}) = \frac{i}{2\pi|\vec{y}|} \int_0^{2\pi} \langle \dot{\gamma}(t), \frac{y^\perp}{|y|} \rangle e^{i2\pi|\vec{y}|\langle \gamma(t), \frac{\vec{y}}{|\vec{m}|}\rangle} dt$$

Now recall the stationary phase asymptotics of Theorem **??** (not just the van der Corput bound),

$$\int A(x) e^{i\lambda(\phi(x)} dx \sim e^{i\frac{\pi}{4}\mathrm{sign}(\phi''(x_0))} A(x_0) \sqrt{\frac{2\pi}{|\phi''(x_0)|}} \cdot \frac{e^{i\lambda\phi(x_0)}}{\sqrt{\lambda}}, \quad \text{as } \lambda \to +\infty,$$

which give

$$R^2 \widehat{\chi}(R\vec{m}) \sim *R^{1/2} \frac{\cos(2\pi|\vec{m}|R + \frac{\pi}{4})}{|\vec{m}|^{3/2}}$$

which is the form that appears in Proposition 2.6.

2.2.1. *Proof of Theorem 2.5.* . To get a lower bound on $|P(R)|$, it suffices to show that there is some $c > 0$ so that for arbitrarily large t, we have $|S(t)| > c$. To do so, we consider the integral

$$J(T) := e^{i\pi/4} \int_T^{2T} S(t) e(t) w(\frac{t}{T}) \frac{dt}{T}$$

where $w(x) \in C_c^\infty[1, 2]$ is a smooth weight function, supported in $[1, 2]$, and of total mass unity: $\int w(x)dx = 1$. It suffices to show that

$$\lim_{T\to\infty} J(T) = -\frac{2}{\pi} \neq 0$$

since if we had $S(t) = o(1)$ then the integral $J(T) = o(1)$ would also tend to zero.

[1]See (12.4.4) in E.C. Titchmarsh The Theory of the Riemann Zeta-function, 2nd ed., Oxford Univ. Press, Oxford 1986.

Plugging in Proposition 2.6, we see that

$$J(T) = -\frac{1}{\pi} \sum_{\substack{0<|\vec{m}|\leq T^{3/4} \\ 0\neq \vec{m}\in\mathbb{Z}^2}} \frac{1}{|\vec{m}|^{3/2}} e^{i\pi/4} \int \cos(2\pi|\vec{m}|\cdot t + \frac{\pi}{4}) e(t) w(\frac{t}{T}) \frac{dt}{T} + o(1).$$

The integral is essentially a Fourier transform of the dilate of w:

$$e^{i\pi/4} \int \cos(2\pi|\vec{m}|\cdot t + \frac{\pi}{4}) e(t) w(\frac{t}{T}) \frac{dt}{T} = \frac{1}{2}\widehat{w}\Big(T(|\vec{m}|-1)\Big) + \frac{i}{2}\widehat{w}\Big(T(|\vec{m}|+1)\Big).$$

There are 4 vectors of norm one $|\vec{m}| = 1$, which contribute the term

$$-\frac{1}{\pi} 4 \frac{1}{2} \widehat{w}(0) = -\frac{2}{\pi}\int_{-\infty}^{\infty} w(x)dx = -\frac{2}{\pi}.$$

We now use the rapid decay of the Fourier transform of the weight function w, say $|\widehat{w}(y)| < y^{-10}$ for $|y| \geq 1$, to find that for any nonzero $\vec{m}$,

$$\widehat{w}\Big(T(|\vec{m}|+1)\Big) \ll \frac{1}{(T|\vec{m}|)^{10}}$$

and if $|\vec{m}| \neq 1, 0$ then $|\vec{m}| - 1 \geq \sqrt{2} - 1 > \min(\sqrt{2}-1, |\vec{m}|/2)$,

$$\widehat{w}\Big(T(|\vec{m}|-1)\Big) \ll \frac{1}{(T|\vec{m}|)^{10}}, \qquad |\vec{m}| \neq 1, 0.$$

Hence

$$J(T) = -\frac{2}{\pi} + O\Big(\sum_{\vec{m}\neq 0} \frac{1}{|\vec{m}|^{3/2}} \frac{1}{(T|\vec{m}|)^{10}}\Big).$$

Since the sum $\sum_{\vec{m}\neq 0} \frac{1}{|\vec{m}|^{3/2+10}} < \infty$ is convergent, we find

$$J(T) = -\frac{2}{\pi} + O(\frac{1}{T^{10}})$$

as claimed. □

3. Higher dimension

3.1. **An Omega result.** Let $N_d(R)$ be the number of lattice points in the d-dimensional ball of radius R:

$$N_d(R) = \#\mathbb{Z}^d \cap B(0,R)$$

Arguing as in the two-dimensional case shows that

$$N_d(R) = \omega_d R^d + O(R^{d-1})$$

where $\omega_d = \operatorname{vol} B(0,1)$. Let $P_d(r) = N_d(R) - \omega_d R^d$ be the remainder term. We want to note that in dimension $d \geq 4$, it is not the case that we get square root cancellation, that is it is not true that $P_d(R)$ is $O(R^{(d-1)/2})$. To see this, we will show that $P_d(R) = \Omega(R^{d-2})$, that is there is some $c > 0$ and arbitrarily large R's so that $|P(R)| > cR^{d-2}$. Thus if $d-2 > (d-1)/2$, i.e. $d > 3$ (so $d \geq 4$), we cannot get square root cancellation.

The reason will be that there will be arbitrarily large R's so that on the boundary of the sphere $\{|x| = R\}$ there are $\gg R^{d-2}$ lattice points. Once we establish this, we pick such a sequence of R's, and note that

$$\begin{aligned} R^{d-2} &\ll \#\{x \in \mathbb{Z}^d : |x| = R\} \leq N_d(R + \frac{1}{R^2}) - N_d(R - \frac{1}{R^2}) \\ &= \omega_d \Big((R + \frac{1}{R^2})^d - (R - \frac{1}{R^2})^d \Big) + P_d(R + \frac{1}{R^2}) - P_d(R - \frac{1}{R^2}) \\ &= O(R^{d-3}) + P_d(R + \frac{1}{R^2}) - P_d(R - \frac{1}{R^2}) \end{aligned}$$

If we assume that $|P_d(R)| \ll R^\theta$ then we obtain

$$R^{d-2} \ll R^{d-3} + R^\theta$$

which forces $\theta \geq d-2$. Thus $P_d(R) = \Omega(R^{d-2})$.

Now to see that there are arbitrarily large R's for which $\mathbb{Z}^d \cap \{|x| = R\} \gg R^{d-2}$: Let $d \geq 2$, and for $n \geq 0$ an integer let

$$r_d(n) = \#\{x \in \mathbb{Z}^d : \sum_{j=1}^d x_j^2 = n\}$$

be the number of representations of an integer n as a sum of d squares. We show that $r_d(n) = \Omega(n^{(d-2)/2})$ which is our claim.

Now if $r_d(n) = o(n^{(d-2)/2})$ then we would get

$$\sum_{n=1}^N r_d(n) = o(\sum_{n=1}^N n^{\frac{d}{2}-1}) = o(N^{d/2})$$

But

$$\sum_{n=1}^N r_d(n) = N_d(\sqrt{N}) \sim \omega_d N^{d/2}$$

which gives a contradiction.

3.2. **Sums of d squares - a survey.** The problem of understanding which integers are sums of d squares, and if so in how many ways, is a very old topic. We will later discuss the two dimensional case.

It is an old result that every positive integer is a sum of 4 squares (Lagrange's four-square theorem), so that $r_4(n) \neq 0$ for all $n \geq 0$. For prime p, we have (Jacobi)

$$r_4(p) = 8(p+1)$$

and $r_4(n)/8$ is multiplicative, with

$$r_4(n) = 8 \sum_{\substack{d|n \\ 4\nmid d}} d$$

For n odd we have $r_4(n) = n^{1+o(1)}$. (note that this is the exponent $(d-2)/2 = 1$ here).

For $d \geq 5$, we certainly have $r_d(n) \geq r_4(n) > 0$. The "circle method" shows that

$$r_d(n) \sim \mathfrak{S}_d(n) n^{(d-2)/2}$$

where the "singular series" is bounded away from zero and infinity:

$$0 < c_d < \mathfrak{S}_d(n) < C_d < \infty$$

The three-dimensional case is quite subtle. A celebrated result of Legendre/Gauss asserts that n is a sum of three squares if and only if $n \neq 4a(8b+7)$. If $n = 4^a$ then $r_3(4^a) = 6$. It is known that $r_3(n) = O(n^{1/2+o(1)})$. If there are primitive lattice points, that is $x = (x_1, x_2, x_3)$ with $\gcd(x_1, x_2, x_3) = 1$ such that $x_1^2 + x_2^2 + x_3^2 = n$ (which happens if and only if $n \neq 0, 4, 7 \bmod 8$) then there is a lower bound of $r_3(n) > n^{1/2-o(1)}$ (Siegel's theorem).

Exercise 2. $r_3(4^a) = 6$.

Appendix A. Background on Fourier analysis

The Fourier transform of an L^1 function on the real line (or more generally on $\mathbb{R}^d$) is defined as

$$\mathcal{F}(f) = \widehat{f}(y) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot y}dx$$

It is clearly a linear map (but we haven't specified the domain and range; we will see below that it preserves the "Schwartz space" $\mathcal{S}$).

An example: In dimension one, let $\mathbf{1}_{[-1/2,1/2]}$ to be the indicator function of a unit interval (clearly not in $\mathcal{S}(\mathbb{R})$). Then

$$\widehat{\mathbf{1}}_{[-1/2,1/2]}(x) = \frac{\sin(\pi x)}{\pi x}$$

Exercise 3. *In dimension 3, take f to be the indicator function of the unit ball $B(0,1) \subset \mathbb{R}^3$. Compute $\widehat{f}$.*

Answer: $\widehat{f}(\xi) = -\frac{\cos(2\pi|\xi|)}{\pi|\xi|^2} + \frac{\sin(2\pi|\xi|)}{2\pi^2|\xi|^3}$.

Definition. *The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ consisting of smooth functions f so that f and all its derivatives decay rapidly:*

$$\mathcal{S}(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) : \forall \alpha, \beta \in \mathbb{N}^d, \sup_{x\in\mathbb{R}^d} |x^\alpha \partial^\beta f(x)| < \infty\}$$

where

$$x^\alpha := \prod_{j=1}^d x_j^{\alpha_j}, \quad \partial^\beta f := \frac{\partial^{\beta_1+\cdots+\beta_d} f}{\partial^{\beta_1}x_1 \ldots \partial^{\beta_d}x_d}.$$

Clearly $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ for all $p \geq 1$.

Exercise 4. *The Gaussian $g(x) = e^{-\pi x^2}$ lies in $\mathcal{S}(\mathbb{R})$. Show that $\widehat{g} = g$.*

Here are some simple and easily checked properties of the Fourier transform: For $f \in \mathcal{S}$,

- The Fourier transform exchanges differentiation and translation: If $T_z f(x) = f(x+z)$, then
$$\widehat{T_z f}(y) = e^{2\pi i z\cdot x}\widehat{f}(x)$$
and consequently converts differentiation to multiplication by $2\pi i x$:
$$\widehat{\frac{df}{dx}} = 2\pi i x \cdot \widehat{f}(x)$$
- Convolution:
$$(f * g)(x) := \int_{\mathbb{R}^d} f(y)g(x-y)dy \quad \Rightarrow \quad \widehat{f * g}(y) = \widehat{f}(y)\cdot\widehat{g}(y)$$
- The Fourier transform intertwines dilation operators: If $\lambda > 0$, and $(D_\lambda f)(x) := f(x/\lambda)$, then
$$\widehat{(D_\lambda f)}(y) = \lambda^d \widehat{f}(\lambda y)$$

Lemma A.1. *If $f \in \mathcal{S}$ then so is $\widehat{f}$.*

Proof. We just treat the one-dimensional case. We need to show that $\widehat{f}$ and all its derivatives decay faster than $1/|x|^N$ for all $N \geq 1$. Since $\partial^n \widehat{f}(x) = (-2\pi i x)^n \widehat{f}$, it suffices to just show that $\widehat{f}$ is rapidly decaying. Indeed, again using the relation $\widehat{\partial^n f}(x) = (2\pi i x)^n \widehat{f}$ gives

$$\widehat{f}(x) = \frac{1}{(2\pi i x)^n} \widehat{\partial^n f}(x)$$

so that

$$|\widehat{f}(x)| \leq \frac{1}{(2\pi|x|)^n} \int_{-\infty}^{\infty} |\partial^n f(y)| dy \ll \frac{||\partial^n f||_\infty}{|x|^n}$$

where we note that if $F \in \mathcal{S}$ then so are all its derivatives $\partial^n f$, so in particular $\partial^n f \in L^1(\mathbb{R})$. □

The main properties of the Fourier transform:

- For functions in $\mathcal{S}$ we have Plancherel's formula

$$||\widehat{f}||_{L^2(\mathbb{R}^d)} = ||f||_{L^2(\mathbb{R}^d)}$$

  and since $\mathcal{S}$ is dense in L^2, the Fourier transform extends to an isometry $\mathcal{F} = \widehat{\ }: L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$.
- Fourier inversion: For $f \in \mathcal{S}$,

$$\widehat{(\widehat{f})}(x) = f(-x)$$

  so that

$$f(x) = \int_{\mathbb{R}^d} \widehat{f}(y) e^{2\pi i x \cdot y} dy$$

- We saw that if $f \in \mathcal{S}$ then so is its Fourier transform, so is in particular rapidly decreasing. We also saw from the example of $f = \mathbf{1}_{[-1/2,1/2]}$ that its Fourier transform $\sin(\pi x)/\pi x$ does decay at infinity, but not rapidly. The decay at infinity is shared by all L^1 functions:

  Theorem (The Riemann-Lebesgue Lemma)**.** *If $f \in L^1(\mathbb{R}^d)$ then $\widehat{f}(y) \to 0$ as $|y| \to \infty$.*

- The Poisson summation formula:

  Theorem A.2. *For $f \in \mathcal{S}(\mathbb{R}^d)$,*

$$\sum_{n \in \mathbb{Z}^d} f(n) = \sum_{m \in \mathbb{Z}^d} \widehat{f}(m)$$