MAT 419 Final Project
Analytic Theory of Quadratic Forms and Kloosterman's Circle Method

Aleksa Milojević

Abstract

We discuss extensions of Hasse-Minkowski local-global principle to the case of integral quadratic forms. In particular, we present Kloosterman's refinement of the Hardy-Littlewood circle method and show that every large enough integer which is primitively locally represented by a quadratic form Q of rank at least 4 is also represented by Q over $\mathbb{Z}$.

I pledge my honor that I have not violated the University guidelines during the work on this project.

1 Introduction

The fundamental question in the theory of quadratic forms asks to determine when the equation $Q(x_1, \dots, x_r) = \sum_{1 \leq i,j \leq r} a_{ij} x_i x_j = n$ has a solution $x \in \mathbb{Z}^r$, for arbitrary integers a_{ij} and a positive integer n. When this equation has a solution, we say that the quadratic form Q *represents* n. In general, this question is very hard to answer, and throughout the years it inspired the growth of a very deep theory. However, much more is known about this question is we restrict to certain special cases, for example to positive definite quadratic forms.

A particularly elegant answer to the above question can be found in case we allow for rational solutions instead of integer ones. In this case, the local-global principle due to Hasse and Minkowski answers the question completely.

Theorem 1. *(Local-global principle) Let $Q(x_1, \dots x_r) = \sum_{1 \leq i,j \leq r} a_{ij} x_i x_j$ be a positive definite quadratic form of rank r and let $n \in \mathbb{Z}$. The equation $Q(x) = n$ has a solution $x \in \mathbb{Q}^r$ if and only if it has a solution in every p-adic completion of the rationals $\mathbb{Q}_p$, including $\mathbb{R} = \mathbb{Q}_{-1}$.*

If the equation $Q(x) = n$ has a solution for over $\mathbb{Q}_p$ for all $p \geq -1$, we say that n is *locally represented* by Q. Therefore, the local-global principle reduces the question of solving a complicated quadratic equation in rationals to discussing the behavior of the quadratic form Q modulo various primes. A refined analysis also shows that it suffices to consider only the primes dividing the determinant $\det(Q)$, which effectively reduces solving $Q(x) = n$ to a finite problem which can be solved by checking a bounded number of prime moduli.

However, trying to determine when $Q(x) = n$ has integral solutions $x \in \mathbb{Z}^r$ is considerably harder, because the local-global principle does not hold anymore. For example, the form $Q(x, y, z) = 2x^2 + 2y^2 + 5z^2 + 2yz$ locally represents all positive integers which are not of the form $9k \pm 3$ or $2^{\text{odd}}(8k + 7)$. Still, Q does not represent 1, as all coefficients of Q are larger than 1. Hence, to correct the statement of the local-global principle for integral quadratic forms, one needs to require the represented integer n to be large enough. And indeed, the following theorem of Tartakowski shows that the local-global principle still holds in this altered form for quadratic forms of rank at least 5.

Theorem 2. *([13]) Let Q be a positive-definite quadratic form of rank $r \geq 5$ and let n be a large enough integer. Then, n is represented by Q if and only if n is locally represented by Q.*

Although it may seem slightly unnatural at first, the requirement that Q has rank at least 5 is very important - the local-global as stated above does not hold for quaternary quadratic forms. For example, [11] and [15] give an example of the form $Q(x, y, z, t) = x^2 + y^2 + 7z^2 + 7t^2$ which locally

represents the integers of the form $3 \cdot 7^{2k} = (7^k)^2 + (\sqrt{2} \cdot 7^k)^2 + 0^2 + 0^2$. However, Q does not represent $3 \cdot 7^{2k}$ over $\mathbb{Z}$ as any solution $x^2 + y^2 + 7z^2 + 7t^2 = 3 \cdot 7^{2k}$ would need to have $7^k | x, y, z, t$, which would reduce the above equation to $\left(\frac{x}{7^k}\right)^2 + \left(\frac{y}{7^k}\right)^2 + 7\left(\frac{z}{7^k}\right)^2 + 7\left(\frac{t}{7^k}\right)^2 = 3$, which obviously has no solutions. Hence, we need to impose a stronger technical condition in order to extend Tartakowski's theorem to quaternary quadratic forms.

Namely, for an integer n, we say that n is *primitively locally represented* by Q if, for all primes p and $k \geq 1$, the equation $Q(x) \equiv n (\mathrm{mod}\ p^k)$ has a solution x with not all entries divisible by p. With this additional requirement, we are able to show the version of the local-global principle for quaternary forms. The main step towards this theorem was made by Kloosterman in 1926, and the missing steps were filled in shortly after.

Theorem 3. *([9]) Let Q be a positive-definite quaternary quadratic form and let n be a large enough integer. Then, n is represented by Q if n is primitively locally represented by Q.*

It is possible to extend the above theorem further, by weakening the conditions of primitive local representability. Namely, it suffices that no *anisotropic* prime p divides the integer n to high power. Here, a prime p is called *isotropic* if $Q(x) = 0$ has a nonzero solution in $\mathbb{Q}_p$ and *anisotropic* otherwise. There is no fundamental difference in either of the proofs, and therefore we will present Theorem 3 in a slightly weaker form. For more detail, one can refer to the survey [12]. A further extension of this theorem is possible in the case of ternary forms [1]. However, the techniques used to prove this result are based on the modular forms and differ significantly from the approach we take here. Hence, we will not go in depth any further. An interested reader can consult books [10], [7] for more information about modular forms.

The proof of Theorem 3 relies heavily on the circle method, which initially developed by Hardy, Littlewood and Ramanujan in 1918 in order to estimate the partition number $p(n)$, [4], and on their work on Waring's problem, [5]. In 1926, Kloosterman [9] refined some of the estimates using a different partition on the circle, which ultimately allowed him to improve the error term and essentially prove Theorem 3. Since then, a number of further applications of the circle method have been developed. For reference, see [6], [14].

2 Overview of the proof

We begin this section by setting up the notation which will be used and proceed to highlight some of the key steps of the proof.

Throughout the paper, we will fix a positive-definite quadratic form Q of rank $r \geq 4$, with integral coefficients. To the form Q, we can also associate the matrix A which satisfies $Q(x) = \frac{1}{2}A[x] = \frac{1}{2}x^T A x$, for $x \in \mathbb{Z}^r$. The dependence of bounds on Q will sometimes be suppressed in the notation, although all derived bounds will implicitly depend on Q.

Let $r(n)$ be the number of integer solutions $x \in \mathbb{Z}^r$ to the equation $Q(x) = n$. The main goal of the proof will be to derive a asymptotic expression for $r(n)$, which will give

$$r(n) = \frac{(2\pi)^{r/2}}{|A|^{1/2}\Gamma\left(\frac{r}{2}\right)} n^{\frac{r}{2}-1}\sigma(n) + O_\varepsilon(n^{\frac{r}{4}-\frac{1}{4}+\varepsilon}), \tag{1}$$

for all $\varepsilon > 0$, where the singular series $\sigma(n)$ is given by $\sigma(n) = \sum_{c \geq 0} c^{-r} T_0(c, n)$ and

$$T_0(c, n) = \sum_{d \bmod c}^{*} \sum_{h \bmod c} e\left(\frac{d}{c}(Q(h) - n)\right).$$

The key analytic object that will allow us to understand $r(n)$ is the *theta function* associated to the quadratic form Q. The theta function is defined as the generating function of the sequence $\{r(n)\}_{n=0}^{\infty}$,

and it can be expressed as

$$\Theta(z) = \sum_{n \geq 0} r(n) e(nz) = \sum_{m \in \mathbb{Z}^r} e\left(Q(m)z\right),$$

for z in the upper half-plane $\mathbb{H}$. Here, we adopt the standard notation $e(z) = e^{2i\pi z}$, for $z \in \mathbb{C}$.

To arrive at the equation 1, the first step is to reconstruct values of $r(n)$ from the associated theta function. In particular, for $z = x + iy$ where $y = n^{-1}$ is fixed, one can write

$$\int_0^1 \Theta(z) e(-nz) dx = \int_0^1 \sum_{n' \geq 0} r(n') e\left((n' - n)z\right) dx = r(n). \tag{2}$$

The next step expression 1 is exploit the rich symmetries possessed by theta functions in order to estimate the mean of the theta function given in 2. Although we only take a glance at these symmetries in Section 4, we note that a much broader theory of understanding theta functions as modular forms can be developed (for reference, see [7], [10]). In particular, we will show the Jacobi inversion formula, which gives in a special case $\Theta_Q(z) = |A|^{-1/2} \left(\frac{i}{z}\right)^{r/2} \Theta_{Q^*}(z)$, where $\Theta_{Q^*}(z)$ is the theta function associated to the adjoint quadratic form $Q^*(m) = \frac{1}{2} A^{-1}[m]$. Using Jacobi inversion, one can transform the expression 2 to get:

$$r(n) = |A|^{-1/2} \int_0^1 \left(\frac{i}{z}\right)^{r/2} e(-nz) \sum_{m \in \mathbb{Z}^r} e\left(-\frac{Q^*(m)}{z}\right) dx. \tag{3}$$

The leading term in the above sum occurs at $m = 0$, and hence evaluating the asymptotics of $r(n)$ amounts to finding the integral of the type $\int_0^1 \left(\frac{i}{z}\right)^{r/2} e(-nz) dx$, which we do in Section 6. A rough heuristic for this is that $m = 0$ is the only term with constant phase, while the oscillations in the other terms induce significant cancellation. However, in order to formalize this intuition and estimate the terms $m \neq 0$ precisely, one needs to decompose the interval $[0, 1)$ into smaller, carefully chosen pieces, on which cancellations occur due to specific behavior of Kloosterman sums. The specific decomposition we apply is described in Section 3, while the cancellations in Kloosterman sums are discussed in Section 5.

Finally, to make any concrete inference about the asymptotics of $r(n)$ from 1, we need to describe the behavior of $\sigma(n)$. Therefore, in Section 7 we discuss the behavior of $\sigma(n)$. If n satisfies the appropriate local conditions, we derive a lower bound $\sigma(n) \geq \Omega(1)$ in case $r \geq 5$ and $\sigma(n) \geq \Omega\left(\frac{1}{\log n}\right)$ when $r = 4$.

The proof presented in this paper is based on the approach taken in [7] and [8]. For the discussion of local densities in Section 7, we referred to [3] and [11].

3 The circle method and the Farey decomposition

In this section, we derive a formula which expresses the mean of a fucntion $f : \mathbb{R} \to \mathbb{C}$ of period 1 by decomposing the interval $[0, 1)$ into a set of smaller intervals, centered at rational points of small denominator.

Lemma 4. *Let $f : \mathbb{R} \to \mathbb{C}$ be a function satisfying $f(x) = f(x+1)$ and $f(-x) = \overline{f(x)}$. Then, for any integer $K \geq 1$ we have the following expression:*

$$\int_0^1 f(x) dx = 2\mathrm{Re} \sum_{c=1}^{K} \sum_{\substack{K < d \leq c+K \\ (c,d)=1}} \int_0^{1/cd} f\left(x - \frac{\overline{d}}{c}\right) dx, \tag{4}$$

where $\overline{d}$ denotes the multiplicative inverse of d modulo c.

Proof. To derive expression 4, we use the Farey sequences. Recall that the Farey sequence of order K, denoted by F_K, is the sequence of rational numbers $0 \le \frac{a}{c} < 1$ with $(a,c) = 1$ and $c \le K$, in increasing order. Before continuing the proof, we will recall a couple of properties of Farey sequences.

If $\frac{a}{c} < \frac{b}{d}$ are adjacent elements of F_K, we must have $c+d > K$. Otherwise, the element $\frac{a+b}{c+d}$ appears in F_K and $\frac{a}{c} < \frac{a+b}{c+d} < \frac{b}{d}$ implies that $\frac{a}{c}$ and $\frac{b}{d}$ are not adjacent. Similarly, we must have $bc - ad = 1$, which we show using induction. Suppose that this statement holds in $F_{K'}$ for some K'. For a pair of adjacent elements $\frac{a}{c} < \frac{b}{d}$ of $F_{K'}$, the rational number $\frac{x}{y}$ with the smallest denominator lying between them is precisely $\frac{a+b}{c+d}$. To see why this holds, note that $\frac{1}{cd} = \frac{b}{d} - \frac{a}{c} = (\frac{b}{d} - \frac{x}{y}) + (\frac{x}{y} - \frac{a}{c}) \ge \frac{1}{dy} + \frac{1}{cy} = \frac{c+d}{cdy}$. Hence, we must have $y \ge c + d$, with the equality only if $by - dx = cx - ay = 1$, i.e. if $\frac{x}{y} = \frac{a+b}{c+d}$. This discussion implies that the first element to be inserted between $\frac{a}{c}$ and $\frac{b}{d}$ will be $\frac{a+b}{c+d}$. However, even after inserting this element, the above property is maintained because $b(c+d) - d(a+b) = bc - ad = 1$ and $(a+b)c - (c+d)a = bc - ad = 1$.

Once we have established these basic properties, let us describe a decomposition of the interval $[0,1]$ naturally arising from F_K. If $\frac{a_1}{c_1} < \frac{a}{c} < \frac{a_2}{c_2}$ are the elements adjacent to $\frac{a}{c}$ in F_K, we have that $ac_1 - a_1c = 1$ and $c + c_1 > K$. Hence, c_1 is the unique integer satisfying $K - c < c_1 \le K$ for which $c_1 a \equiv_c 1$. Similarly, c_2 is the unique integer satisfying $K - c < c_2 \le K$ with $c_2 a \equiv_c -1$. Then, a_1 and a_2 are determined by relations $ac_1 - a_1c = 1, a_2c - ac_2 = 1$. Note that these definitions still make sense for $\frac{a}{c} = \frac{0}{1}$ and $\frac{a}{c} = \frac{K-1}{K}$, giving $-\frac{1}{K+1} < \frac{0}{1} < \frac{1}{K}$ in the first case and $\frac{K-2}{K-1} < \frac{K-1}{K} < \frac{K}{K+1}$ in the second case.

To an element $\frac{a}{c} \in F_K$ we associate the interval $I(\frac{a}{c}) = \left(\frac{a+a_1}{c+c_1}, \frac{a+a_2}{c+c_2}\right]$. Then, the intervals $I(\frac{a}{c})$ partition $\left(-\frac{1}{K+1}, \frac{K}{K+1}\right]$, and hence, using the periodicity of f, one can write

$$\begin{aligned}
\int_0^1 f(x)dx &= \sum_{\frac{a}{c}\in F_K} \int_{I(\frac{a}{c})} f(x)dx = \sum_{\frac{a}{c}\in F_K} \int_{\frac{a}{c}-\frac{1}{c(c+c_1)}}^{\frac{a}{c}+\frac{1}{c(c+c_2)}} f(x)dx = \sum_{\frac{a}{c}\in F_K} \int_{-\frac{1}{c(c+c_1)}}^{\frac{1}{c(c+c_2)}} f\left(x+\frac{a}{c}\right)dx \\
&= \sum_{\frac{a}{c}\in F_K} \int_0^{\frac{1}{c(c+c_1)}} f\left(\frac{a}{c}-x\right) + \int_0^{\frac{1}{c(c+c_2)}} f\left(x+\frac{a}{c}\right)dx \\
&= \sum_{c=1}^{K} \sum_{d=K+1}^{K+c}{}^{*} \int_0^{\frac{1}{cd}} f\left(\frac{\overline{d}}{c}-x\right) + \int_0^{\frac{1}{cd}} f\left(x-\frac{\overline{d}}{c}\right)dx.
\end{aligned}$$

Now, one can use the additional symmetry of f given by $f(-x) = \overline{f(x)}$ to arrive at the final expression:

$$\begin{aligned}
\int_0^1 f(x)dx &= \sum_{c=1}^{K} \sum_{d=K+1}^{K+c}{}^{*} \int_0^{\frac{1}{cd}} f\left(\frac{\overline{d}}{c}-x\right) + \int_0^{\frac{1}{cd}} f\left(x-\frac{\overline{d}}{c}\right)dx \\
&= \sum_{c=1}^{K} \sum_{d=K+1}^{K+c}{}^{*} \int_0^{\frac{1}{cd}} \overline{f\left(x-\frac{\overline{d}}{c}\right)} + \int_0^{\frac{1}{cd}} f\left(x-\frac{\overline{d}}{c}\right)dx \\
&= 2\mathrm{Re} \sum_{c=1}^{K} \sum_{d=K+1}^{K+c}{}^{*} \int_0^{1/cd} f\left(x-\frac{\overline{d}}{c}\right)dx.
\end{aligned}$$

This completes the proof of the decomposition lemma. □

Remark. In the above lemma, the assumption that f possesses an additional symmetry $f(-x) = \overline{f(x)}$ is purely a technical one, and it can be removed if one is willing to work with a slightly more cumbersome expressions.

Lemma 4 can now be applied to derive an expression for $r(n)$, as indicated in the Section 2. Setting $f(x) = \Theta(z)e(-nz)$, where $z = x + iy$, one can take the equation 2 further and get:

$$r(n) = \int_0^1 \Theta(z)e(-nz)dx = 2\mathrm{Re}\sum_{c=1}^{K} \sum_{\substack{K<d\leq c+K \\ (c,d)=1}} \int_0^{1/cd} \Theta\left(z - \frac{\overline{d}}{c}\right) e\left(-n\left(z - \frac{\overline{d}}{c}\right)\right) dx$$

Hence, we arrive at an analytic expression for $r(n)$, given by

$$r(n) = 2\mathrm{Re}\sum_{c=1}^{K} \int_0^{1/cK} T(c, n; x)e(-nz)dx, \tag{5}$$

where $T(c, n; x)$ corresponds to the following incomplete Kloosterman sum

$$T(c, n; x) = \sum_{\substack{d=K+1 \\ cdx<1}}^{c+K\,*} e\left(n\frac{\overline{d}}{c}\right) \Theta\left(z - \frac{\overline{d}}{c}\right). \tag{6}$$

4 Theta functions and the Jacobi inversion

This section is dedicated to exploring the symmetries possessed by a theta function associated to a quadratic form. Namely, the theta functions can be viewed as a modular form, which means that they posses symmetries with respect to transformations in a certain congruence subgroup of $SL_2(\mathbb{Z})$. A nice treatment of this can be found in Chapter 10 of [7].

However, in this paper, we will not explore the theta function in greater depth than what is needed for the proof of the main theorem. Hence, we only prove the most basic version of Jacobi inversion formula.

Whenever we explore the symmetries of theta functions, the key ingredient is the Poisson summation, which relates the sum of a function and its Fourier transform over a lattice. The precise statement follows.

Lemma 5. *(Poisson summation) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a rapidly decaying Schwartz function, and let $\hat{f}(\xi) = \int_{x\in\mathbb{R}^n} f(x)e(-x \cdot \xi)dx$ be its Fourier transform. Then, we have*

$$\sum_{x\in\mathbb{Z}^n} f(x) = \sum_{\xi\in\mathbb{Z}^n} \hat{f}(\xi).$$

Equipped with the Possion summation, we proceed to show the Jacobi inversion formula.

Lemma 6. *(Jacobi inversion) For any $t \in \mathbb{R}^r$ and $z \in \mathbb{H}$, we have:*

$$\sum_{m\in\mathbb{Z}^r} e\left(\frac{1}{2}A[m + t]z\right) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} \sum_{m\in\mathbb{Z}^r} e\left(-\frac{A^{-1}[m]}{2z} + m \cdot t\right). \tag{7}$$

Proof. We begin by computing the Fourier transform of the function $f(x) = e\left(\frac{1}{2}A[x]z\right)$. As A is a real symmetric positive definite matrix, it is diagonalizable over $\mathbb{R}$ (in the sense of quadratic forms). In other words, we can represent $A = B^T IB = B^T B$. Then, $A[x] = y \cdot y$ for $y = Bx$. Then, we can compute the Fourier transform of f:

$$\begin{aligned}
\hat{f}(\xi) &= \int_{x\in\mathbb{R}^r} f(x)e(-x\cdot\xi)dx = \int_{y\in\mathbb{R}^r} e\left(\frac{1}{2}y\cdot yz - y\cdot(B^T)^{-1}\xi\right)|B|^{-1}dy = \\
&= |B|^{-1}\prod_{j=1}^{r}\int_{y_j\in\mathbb{R}} e\left(\frac{1}{2}y_j^2 z - y_j\xi_j'\right)dy_j \\
&= |A|^{-1/2}\prod_{j=1}^{r}\int_{y_j\in\mathbb{R}} \exp\left(i\pi z\left(y_j - \frac{\xi_j'}{z}\right)^2 - 2i\pi\frac{\xi_j'^2}{2z}\right)dy_j \\
&= |A|^{-1/2}e\left(-\frac{\xi'\cdot\xi'}{2z}\right)\prod_{j=1}^{r}\int_{y_j\in\mathbb{R}}\left(\frac{i}{z}\right)^{1/2}\exp\left(-\pi y_j^2\right)dy_j \\
&= |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2}e\left(-\frac{\xi'\cdot\xi'}{2z}\right),
\end{aligned}$$

where $\xi' = (B^T)^{-1}\xi$ implies $\xi'\cdot\xi' = A^{-1}[\xi]$. Before applying Possion summation, note that we need to use $g(x) = e\left(\frac{1}{2}A[x+t]z\right) = f(x+t)$ in the Poisson summation. Still, from the above computation, it is easy to see that the Fourier transform of g is $\hat{g}(\xi) = \hat{f}(\xi)e(\xi\cdot t)$. Therefore, Lemma 5 implies

$$\sum_{m\in\mathbb{Z}^r} g(m) = \sum_{m\in\mathbb{Z}^r}\hat{f}(m)e(m\cdot t) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2}\sum_{m\in\mathbb{Z}^r} e\left(-\frac{A^{-1}[m]}{2z} + m\cdot t\right),$$

which completes the proof. □

Remark. Even though the Jacobi inversion formula does not directly reference the Theta function, note that one can rewrite 7 to give $\Theta(z) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2}\Theta^*(-1/z)$, where $\Theta^*(z)$ is the theta function associated to the adjoint quadratic form $Q^*(m) = \frac{1}{2}A^{-1}[m]$.

Lemma 6 can be used to rewrite the expression 5 in the following way. We begin by splitting $\Theta\left(z - \frac{\overline{d}}{c}\right)$ into the congruence theta functions $\Theta(z;h)$ and applying Jacobi inversion to them:

$$\begin{aligned}
\Theta\left(z - \frac{\overline{d}}{c}\right) &= \sum_{m\in\mathbb{Z}^r} e\left(-Q(m)\frac{\overline{d}}{c}\right)e\left(Q(m)z\right) = \\
&= \sum_{h \bmod c} e\left(-Q(h)\frac{\overline{d}}{c}\right)\Theta(z;h),
\end{aligned}$$

where $\Theta(z;h) = \sum_{m\equiv_c h} e(Q(m)z)$. Applying Lemma 6 now implies:

$$\begin{aligned}
\Theta(z;h) = \sum_{m\in\mathbb{Z}^r} e\left(\frac{1}{2}A[cm+h]z\right) &= \sum_{m\in\mathbb{Z}^r} e\left(\frac{1}{2}A\left[m + \frac{h}{c}\right](c^2 z)\right) \\
&= |A|^{-1/2}\left(\frac{i}{c^2 z}\right)^{r/2}\sum_{m\in\mathbb{Z}^r} e\left(-\frac{A^{-1}[m]}{2c^2 z} + m\cdot\frac{h}{c}\right) \\
&= |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2}c^{-r}\sum_{m\in\mathbb{Z}^r} e\left(-\frac{A^{-1}[m]}{2c^2 z} - m\cdot\frac{h}{c}\right).
\end{aligned}$$

Plugging this back into the expression for $\Theta(z-\frac{\overline{d}}{c})$, we obtain

$$\Theta\left(z-\frac{\overline{d}}{c}\right) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} c^{-r} \sum_{h \bmod c} \sum_{m\in\mathbb{Z}^r} e\left(-Q(h)\frac{\overline{d}}{c} - \frac{A^{-1}[m]}{2c^2 z} - m\cdot\frac{h}{c}\right)$$
$$= |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} c^{-r} \sum_{m\in\mathbb{Z}^r} e\left(-\frac{Q^*(m)}{c^2 z}\right) \sum_{h \bmod c} e\left(-\frac{\overline{d}}{c}Q(h) - \frac{m\cdot h}{c}\right),$$

where Q^* is the adjoint of Q defined by $Q^*(m) = \frac{1}{2}A^{-1}[m]$. Changing the variables $h \to hd$, we get

$$\Theta\left(z-\frac{\overline{d}}{c}\right) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} c^{-r} \sum_{m\in\mathbb{Z}^r} e\left(-\frac{Q^*(m)}{c^2 z}\right) \sum_{h \bmod c} e\left(-\frac{d}{c}Q(h) - \frac{d}{c}m\cdot h\right) =$$
$$= |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} c^{-r} \sum_{m\in\mathbb{Z}^r} e\left(-\frac{Q^*(m)}{c^2 z}\right) G_m\left(-\frac{d}{c}\right),$$

where $G_m\left(\frac{d}{c}\right)$ is the Gauss sum given by

$$G_m\left(-\frac{d}{c}\right) = \sum_{h \bmod c} e\left(-\frac{d}{c}(Q(h) + h\cdot m)\right).$$

Looking back at the definition of $T(c,n;x)$ in 6

$$T(c,n;x) = |A|^{-1/2}\left(\frac{i}{z}\right)^{r/2} c^{-r} \sum_{m\in\mathbb{Z}^r} e\left(-\frac{Q^*(m)}{c^2 z}\right) \sideset{}{^*}\sum_{\substack{d=K+1\\ cdx<1}}^{K+c} e\left(n\frac{\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right) \tag{8}$$

From the above equation, it is clear that the quantity of interest is

$$T_m(c,n;x) = \sideset{}{^*}\sum_{\substack{d=K+1\\ cdx<1}}^{K+c} e\left(n\frac{\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right). \tag{9}$$

As we will see in the next section, $T_m(c,n;x)$ is closely related to Kloosterman sums (or Salié sums if r is odd). Hence, the key step of the proof is obtaining a good upper bound on $T_m(c,n;x)$

5 Upper bounds on Kloosterman sums

The goal for this section is to prove the following upper bound on $T_m(c,n;x)$.

Lemma 7. *Let $T_m(c,n;x)$ be defined by expression 9. Then, for every $\varepsilon > 0$ one has*

$$T_m(c,n;x) \le O_\varepsilon\left((n,c_1)^{1/2} c_0^{1/2} c^{\frac{r}{2}+\frac{1}{2}+\varepsilon}\right), \tag{10}$$

where c_1 is the largest factor of c with $(c_1, 2|A|) = 1$ and $c_0 = c/c_1$.

Proof. The proof has four main steps. We begin by completing the sum $T_m(c,n;x)$ to eliminate dependence on x, which will allow us to focus on the complete sums $K(l,n,m;c)$. Next, we show that the complete sums satisfy a multiplicativity property, which allows us to break the analysis into two moduli - one that it coprime to $2|A|$ and one that is not. To analyze the arising sums, we show how to estimate and compute some of the Gauss sums involved. Finally, this reduces the problem of estimating the $K(l,n,m;c)$ to the problem of estimating Kloosterman sums, which can be done either using the Weil bound, or by bounding higher moments of the Kloosterman sums.

As explained, the first step of the proof separates the dependence on x from the sum 9, hence making the sums easier to understand.

Lemma 8. *Let* $\gamma(l) = \frac{1}{c}\sum_{b=K+1}^{\min\{K+c,\frac{1}{cx}\}} e\left(-\frac{bl}{c}\right)$, $K(l,n,m;c) = \sum^*_{d \bmod c} e\left(\frac{ld+n\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right)$. *Then,*

$$T_m(c,n;x) = \sum_{l \bmod c} \gamma(l)K(l,n,m;c). \tag{11}$$

Proof of Lemma 8. Expression 11 can be derived from the following simple computation.

$$\begin{aligned}
\sum_{l \bmod c} \gamma(l)K(l,n,m;c) &= \frac{1}{c}\sum_{l \bmod c}\sum_{b=K+1}^{\min\{K+c,1/cx\}}\sum_{d \bmod c}^{*} e\left(-\frac{bl}{c}\right) e\left(\frac{ld+n\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right) \\
&= \sum_{b=K+1}^{\min\{K+c,1/cx\}}\sum_{d \bmod c}^{*} e\left(\frac{n\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right)\frac{1}{c}\sum_{l \bmod c} e\left(\frac{l}{c}(b-d)\right) \\
&= \sum_{b=K+1}^{\min\{K+c,1/cx\}}\sum_{d \bmod c}^{*} e\left(\frac{n\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right)\mathbf{1}_{b=d} \\
&= \sum_{d=K+1}^{\min\{K+c,1/cx\}}{}^{*} e\left(\frac{n\overline{d}}{c}\right) G_m\left(-\frac{d}{c}\right) = T_m(c,n;x)
\end{aligned}$$

□

Lemma 8 reduces the task of bounding $T_m(c,n;x)$ to that of bounding $\gamma(l)$ and $K(l,n,m;c)$. If we choose l such that $|l| \le c/2$, the term $\gamma(l)$ can be bounded by $|\gamma(l)| \le O\left(\frac{1}{l+1}\right)$. This is trivial for $l = 0$, and for $l > 0$ we have

$$|\gamma(l)| = \frac{1}{c}\left|\sum_{b=K+1}^{\min\{K+c,1/cx\}} e\left(-\frac{bl}{c}\right)\right| = \frac{1}{c}\left|e\left(-\frac{(K+1)b}{c}\right)\frac{e\left(-\frac{lS}{c}\right)-1}{e\left(-\frac{l}{c}\right)-1}\right|$$

Note that $\left|e(-\frac{l}{c})-1\right| \ge \sin(\frac{l}{c}) \ge \Omega(\frac{l}{c})$ and therefore $|\gamma(l)| \le \frac{1}{c}O\left(\frac{1}{l/c}\right) = O\left(\frac{1}{l+1}\right)$. Now, we direct our attention to a more involved task of bounding $K(l,n,m;c)$. In order to do that, we begin by discussing a multiplicativity property obeyed by these sums.

Lemma 9. *Let* c_1 *be the largest divisor of* c *with* $(c_1, 2|A|) = 1$ *and let* $c_0 = c/c_1$. *Then, the following multiplicativity property holds for the sums* $K(l,n,m;c)$*:*

$$K(l,n,m;c) = K^{(c_0)}(l,n,m;c_1)K^{(c_1)}(l,n,m;c_0),$$

where the superscipts $(c_0),(c_1)$ *indicate that* $e(z)$ *is replaced by* $e(\overline{c_0}z), e(\overline{c_1}z)$*, respectively. Here* $\overline{c_0}, \overline{c_1}$ *denote the multiplicative inverses of* c_0, c_1 *modulo* c_1, c_0*, respectively.*

Proof. The proof of the lemma is based on the Chinese Remainder theorem, which states that for coprime integers c_0, c_1 there is a ring homomorphism $\phi : \mathbb{Z}/c_0\mathbb{Z} \times \mathbb{Z}/c_1\mathbb{Z} \to \mathbb{Z}/c\mathbb{Z}$ given by $\phi(d_0,d_1) = \overline{c_1}c_1d_0 + \overline{c_0}c_0d_1$, which commutes with the canonical projections. We have

$$\begin{aligned}
&K^{(c_0)}(l,n,m;c_1)K^{(c_1)}(l,n,m;c_0) = \\
&\quad= \sum_{d_0 \bmod c_0}^{*}\sum_{d_1 \bmod c_1}^{*} e\left(\frac{\overline{c_1}ld_0 + \overline{c_1}n\overline{d_0}}{c_0}\right) G_m^{(c_1)}\left(-\frac{d_0}{c_0}\right) e\left(\frac{\overline{c_0}ld_1 + \overline{c_0}n\overline{d_1}}{c}\right) G_m^{(c_0)}\left(-\frac{d_1}{c_1}\right) \\
&\quad= \sum_{d_0 \bmod c_0}^{*}\sum_{d_1 \bmod c_1}^{*} e\left(\frac{ld+n\overline{d}}{c}\right) G_m^{(c_1)}\left(-\frac{d_0}{c_0}\right) G_m^{(c_1)}\left(-\frac{d_1}{c_1}\right),
\end{aligned}$$

where $d = \phi(d_0, d_1)$. The Gauss sums $G_m^{(c_1)}, G_m^{(c_0)}$ also satisfy the same multiplicativity property, as the following straightforward calculation shows

$$G_m^{(c_1)}\left(-\frac{d_0}{c_0}\right) G_m^{(c_1)}\left(-\frac{d_1}{c_1}\right) = \sum_{h_0 \bmod c_0} \sum_{h_1 \bmod c_1} e\left(-\frac{\overline{c_1} d_0}{c_0}(Q(h_0) + m \cdot h_0) - \frac{\overline{c_0} d_1}{c_1}(Q(h_1) + m \cdot h_1)\right)$$
$$= \sum_{h_0 \bmod c_0} \sum_{h_1 \bmod c_1} e\left(-\frac{d}{c}(Q(h) + m \cdot h)\right) = G_m\left(-\frac{d}{c}\right),$$

where $h = \phi(h_0, h_1)$. Combining the previous two conclusions completes the proof. □

We will now estimate the sums $K^{(c_1)}(l, n, m; c_0)$ and $K^{(c_0)}(l, n, m; c_1)$ separately, by investigating the behavior of Gauss sums modulo c_0, c_1. We begin by an ad hoc bound for Gauss sums modulo c_0.

Lemma 10. *If $(c, d) = 1$, we have $G_m\left(\frac{d}{c}\right) \leq O(c^{r/2})$.*

Proof. We have

$$\left|G_m\left(\frac{d}{c}\right)\right|^2 = \sum_{h_1, h_2 \bmod c} e\left(\frac{d}{c}(Q(h_1) - Q(h_2)) + \frac{d}{c}(h_1 - h_2) \cdot m\right).$$

It is not hard to see that

$$Q(h_1) - Q(h_2) = \frac{1}{2} h_1^T A h_1 - \frac{1}{2} h_2^T A h_2 = \frac{1}{2}(h_1 + h_2)^T A (h_1 - h_2).$$

Hence, substituting $x = h_1 + h_2, y = h_1 - h_2$ gives

$$\left|G_m\left(\frac{d}{c}\right)\right|^2 = \sum_{x, y \bmod c} e\left(y \cdot \left(\frac{d}{2c} A x + \frac{d}{c} m\right)\right) = c^r \left|\left\{x \bmod c : \frac{1}{2} A x + d m \equiv_c 0\right\}\right|.$$

Hence, $\left|G_m\left(\frac{d}{c}\right)\right|^2 \leq c^r |\ker_{\mathbb{Z}/c\mathbb{Z}} A|$, where A is though of as a linear map from $(\mathbb{Z}/c\mathbb{Z})^r$ to $(\mathbb{Z}/c\mathbb{Z})^r$. We will now show that $|\ker_{\mathbb{Z}/c\mathbb{Z}} A| \leq |A|^r$, which immediately implies the statement of the lemma. Note that $|\operatorname{im}_{\mathbb{Z}/c\mathbb{Z}} A| \cdot |\ker_{\mathbb{Z}/c\mathbb{Z}} A| = c^r$. One can show that $|\operatorname{im}_{\mathbb{Z}/c\mathbb{Z}} A| \geq c^r/(c, |A|)^r$ by seeing that $|A| x = A(|A| A^{-1} x) \in \operatorname{im}_{\mathbb{Z}/c\mathbb{Z}} A$ for all $x \in (\mathbb{Z}/c\mathbb{Z})^r$. Therefore, we have $|\ker_{\mathbb{Z}/c\mathbb{Z}} A| \leq (c, |A|)^r \leq |A|^r$, which completes the proof. □

Using Lemma 10, one trivially gets $|K^{(c_1)}(l, n, m; c_0)| \leq O(c_0^{r/2+1})$. Although this bound is clearly not the best possible, it will suffice for our purposes. On the other hand, we can exploit the fact $(c_1, 2|A|) = 1$ to evaluate the Gauss sums appearing in $K^{(c_0)}(l, n, m; c_1)$ precisely.

Lemma 11. *Assume that $(c, 2|A|d) = 1$. Then,*

$$G_m\left(\frac{d}{c}\right) = \left(\frac{|A|}{c}\right) \left(\varepsilon_c \left(\frac{2d}{c}\right) \sqrt{c}\right)^r e\left(-\frac{d}{c} Q^*(m)\right),$$

where $\varepsilon_c = 1$ if $c \equiv_4 1$ and $\varepsilon_c = i$ if $c \equiv_4 -1$.

Proof of Lemma 11. In this proof, we will use the fact that the Gauss sum is easy to compute in the one dimensional case, giving $\sum_{h \in \mathbb{Z}/c\mathbb{Z}} e(\frac{d}{c} h^2) = \left(\frac{d}{c}\right) \varepsilon_c \sqrt{c}$. To compute the Gauss sum in the general case with linear terms, we begin by diagonalizing the matrix A, and completing the square. More precisely, as $(c, 2|A|) = 1$, the matrix A is diagonalizable modulo c, i.e. there exists a matrix B with

$B^TAB = D$, where B is an invertible matrix and D is diagonal. Then, $Q(h) = \frac{1}{2}h^TAh = \frac{1}{2}g^TDg$, for $g = B^{-1}h$. Moreover, if we let $m' = B^Tm$, we get:

$$Q(h) + h \cdot m \equiv_c \frac{1}{2}D[g] + g \cdot m' \equiv_c \frac{1}{2}\sum_i D_{ii}\left(g_i + \frac{m_i'}{D_{ii}}\right)^2 - \frac{m_i'^2}{D_{ii}}$$

It is not hard to see that $\sum_i \frac{m_i'^2}{D_{ii}^2} \equiv_c D^{-1}[B^Tm] \equiv_c m^TBB^{-1}A^{-1}(B^T)^{-1}B^Tm \equiv_c A^{-1}[m]$. Hence,

$$\begin{aligned}
G_m\left(\frac{d}{c}\right) &= \sum_{h \bmod c} e\left(\frac{d}{c}(Q(h) + h \cdot m)\right) \\
&= \sum_{g \bmod c} e\left(\frac{d}{c}\sum_i \frac{D_{ii}}{2}\left(g_i + m_i'\overline{D_{ii}}\right)^2\right) e\left(-\frac{d}{2c}A^{-1}[m]\right) \\
&= e\left(-\frac{d}{c}Q^*(m)\right)\prod_{i=1}^{r}\sum_{g_i \bmod c} e\left(\frac{d}{c}\frac{D_{ii}}{2}\left(g_i + m_i'\overline{D_{ii}}\right)^2\right) \\
&= e\left(-\frac{d}{c}Q^*(m)\right)\prod_{i=1}^{r}\varepsilon_c\left(\frac{dD_{ii}/2}{c}\right)\sqrt{c} \\
&= e\left(-\frac{d}{c}Q^*(m)\right)\left(\frac{|A|}{c}\right)\left(\varepsilon_c\left(\frac{2d}{c}\right)\sqrt{c}\right)^r,
\end{aligned}$$

where we have used $\prod_i D_{ii} = |D| \equiv_c |A||B|^2$. This completes the proof. □

Using the expression given by Lemma 11, we can rewrite $K^{(c_0)}(l, n, m; c_1)$ in the following way

$$\begin{aligned}
K^{(c_0)}(l, n, m; c_1) &= \sideset{}{^*}\sum_{d_1 \bmod c_1} e\left(\overline{c_0}\frac{ld_1 + n\overline{d_1}}{c_1}\right) G_m\left(-\frac{\overline{c_0}d_1}{c_1}\right) \\
&= \left(\frac{|A|}{c_1}\right)\varepsilon_{c_1}^r c_1^{r/2}\left(\frac{-2\overline{c_0}}{c_1}\right)^r \sideset{}{^*}\sum_{d_1 \bmod c_1}\left(\frac{d_1}{c_1}\right)^r e\left(\overline{c_0}\frac{ld_1 + n\overline{d_1}}{c_1}\right) e\left(\frac{d_1\overline{c_0}}{c_1}Q^*(m)\right).
\end{aligned}$$

Depending on the parity of r, the arising sum is either a classical Kloosterman sum or a Salié sum. In both cases, we have the following bound for these:

$$S_r(l', n'; c_1) = \sideset{}{^*}\sum_{d_1 \bmod c_1}\left(\frac{d_1}{c_1}\right)^r e\left(\frac{l'd_1 + n'\overline{d_1}}{c_1}\right) \leq (l', n', c_1)^{\frac{1}{2}} c_1^{\frac{1}{2}}\tau(c_1). \tag{12}$$

This bound follows from Weil's proof of Riemann hypothesis over finite fields. Unfortunately, it is out of the scope of this paper to prove it. At the same time, this bound is essentially the best possible.

It is worth noting that Kloosterman's initial proof did not use the Weil bound. Kloosterman came up with an elementary argument which shows that $|S_r(l, n; p)| \leq p^{\frac{3}{4}}$, by evaluating the higher moments of the $V_k(p) = \sum_a S_r(1, a; p)^k$ using arithmetic arguments. A good reference on this topic is Chapter 4 of [7].

In any case, by using the bound 12, we immediately get $|K^{(c_0)}(l, n, m; c_1)| \leq (n, c_1)^{\frac{1}{2}} c_1^{\frac{r+1}{2}}\tau(c_1)$. Bringing the bounds for $K^{(c_0)}(l, n, m; c_1)$ and $K^{(c_1)}(l, n, m; c_0)$ together, we have

$$|K(l, n, m; c)| \leq O\left((n, c_1)^{\frac{1}{2}} c_0^{\frac{1}{2}} c^{\frac{r+1}{2}}\tau(c_1)\right).$$

If we sum over all l, using the bound $\gamma(l) \leq O(\frac{1}{l+1})$, we get the wanted bound for $T_m(c, n; x)$:

$$T_m(c, n; x) \leq O((n, c_1)^{\frac{1}{2}} c_0^{\frac{1}{2}} c^{\frac{r+1}{2}}\tau(c_1)\log c)$$

□

6 Deriving the asymptotics of $r(n)$

Let us recall the formula determining $r(n)$, and let us expand it in terms of $T_m(c,n;x)$, by combining expressions 5, 8 and 9.

$$r(n) = 2|A|^{-1/2}\mathrm{Re}\sum_{c=1}^{K} c^{-r}\int_0^{1/cK}\left(\frac{i}{z}\right)^{r/2} e(-nz)\sum_{m\in\mathbb{Z}^r} e\left(-\frac{Q^*(m)}{c^2 z}\right)T_m(c,n;x)dx$$

We will show that the leading term in the above expression comes from $m=0$ when $0<x<\frac{1}{c(c+K)}$, while all other terms are of lower order. In what follows, the notation $\sum'$ implies that the term with $m=0$, $0<x<\frac{1}{c(c+K)}$ is not included in the sum. The goal of the following computation is to bound the error term arising in the final formula for $r(n)$. If we denote the complete sum $T_0(c,n;0)$ by $T_0(c,n)$, we have

$$T(c,n;x) = |A|^{-1/2}c^{-r}\left(\frac{i}{z}\right)^{r/2} T_0(c,n)\mathbf{1}_{0<x<\frac{1}{c(c+K)}} + \underbrace{+O\left(|z|^{-r/2}c^{-r}\sum_{m\in\mathbb{Z}^r}{}'\left|e\left(-\frac{Q^*(m)}{c^2 z}\right)\right|(n,c_1)^{1/2}c_0^{1/2}c^{\frac{r+1}{2}+\varepsilon}\right)}_{E} \tag{13}$$

Using the fact that $Q^*(m)\geq\alpha|m|^2$ for some $\alpha>0$, which follows from positive definiteness, we have

$$E\leq O\left(|z|^{-\frac{r}{2}}c^{-\frac{r}{2}+\frac{1}{2}+\varepsilon}(n,c_1)^{1/2}c_0^{1/2}\sum_{m\in\mathbb{Z}^r}{}'\exp\left(-\frac{2\pi\alpha y}{c^2|z|^2}|m|^2\right)\right),$$

where $z=x+iy$. If we set $y=1/n$, $K=\sqrt{n}$ it is easy to see that $\frac{y}{c^2|z|^2}\geq\frac{1}{2}$. This is because $c^2|z|^2/y = c^2(x^2+y^2)/y\leq c^2(\frac{n}{K^2c^2}+n^{-1})\leq 2$.

Note that there exists an absolute constant C such that for all $t\geq\pi\alpha$, $m\neq 0$, one has $\exp(-t|m|^2)\leq Ct^{-r/4}\exp(-|m|^2)$. Hence, for $x<\frac{1}{c(c+K)}$, i.e. when $m=0$ is not included in the sum we set $t=\frac{2\pi\alpha y}{c^2|z|^2}$ to get

$$|z|^{-\frac{r}{2}}c^{-\frac{r}{2}}\sum_{m\in\mathbb{Z}^r}{}'\exp\left(-\frac{2\pi\alpha y}{c^2|z|^2}|m|^2\right)\leq O\left(|z|^{-\frac{r}{2}}c^{-\frac{r}{2}}\sum_{m\in\mathbb{Z}^r}{}'\left(\frac{2\pi\alpha y}{c^2|z|^2}\right)^{-r/4}\exp(-|m^2|)\right)\leq O\left(y^{-r/4}\right).$$

If $x>\frac{1}{c(c+K)}$, we also have $\frac{y}{c^2|z|^2}\leq 4$. This is because $c^2|z|^2/y\geq c^2\frac{1}{c^2(c+K)^2y}\geq\frac{n}{4K^2}\geq\frac{1}{4}$. As the sum has only one term with $m=0$ one also derives

$$|z|^{-\frac{r}{2}}c^{-\frac{r}{2}}\sum_{m\in\mathbb{Z}^r}{}'\exp\left(-\frac{2\pi\alpha y}{c^2|z|^2}|m|^2\right)\leq O(y^{-r/4}).$$

Hence, in either case we have $E\leq O(n^{r/4+\varepsilon}(n,c_1)^{1/2}c_0^{1/2}c^{1/2})$. Thus, the expression for $r(n)$ becomes

$$r(n) = 2|A|^{-1/2}\mathrm{Re}\int_0^{1/c(K+c)}\left(\frac{i}{z}\right)^{r/2}e(-nz)dx + O\left(n^{r/4+\varepsilon}\sum_{c\leq K}\frac{(n,c_1)^{1/2}c_0^{1/2}c^{1/2}}{cK}\right).$$

Again, it is not hard to bound the error term by noting that

$$\sum_{c\leq K}\frac{(n,c_1)^{1/2}c_0^{1/2}}{c^{1/2}} = \sum_{\substack{\delta=1\\ \delta|n}}^{K}\sum_{\substack{c=1\\ (c_1,n)=\delta}}^{K}\frac{\delta^{1/2}}{c_1^{1/2}}\leq\sum_{\substack{\delta=1\\ \delta|n}}^{K}\delta^{1/2}\sum_{\substack{c_0=1\\ \mathrm{rad}(c_0)|2|A|}}^{K}\sum_{l=1}^{K/\delta}\frac{1}{l^{1/2}}\leq\sum_{\substack{\delta=1\\ \delta|n}}^{K}\delta^{1/2}K^{\varepsilon}\frac{K^{1/2}}{\delta^{1/2}}\leq O(K^{1/2+\varepsilon}\tau(n)).$$

Therefore, we arrive at the final bound for the error term, giving

$$r(n) = 2|A|^{-1/2} \sum_{c \le K} c^{-r} T_0(c,n) \mathrm{Re} \int_0^{1/c(K+c)} \left(\frac{i}{z}\right)^{r/2} e(-nz)dx + O\left(n^{\frac{r}{4}-\frac{1}{4}+\varepsilon}\right). \tag{14}$$

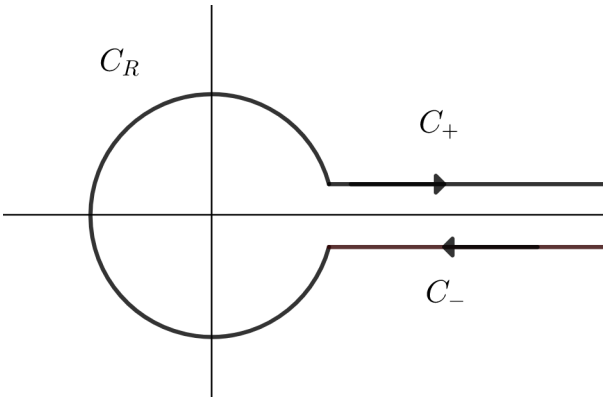
The final step of the computation amounts to evaluating the remaining integral, which is done using the following lemma.

Lemma 12. *For $y > 0$ we have*

$$\int_{-\infty+iy}^{\infty+iy} \left(\frac{i}{z}\right)^{r/2} e(-nz)dx = \frac{(2\pi)^{r/2}}{\Gamma(\frac{r}{2})} n^{\frac{r}{2}-1}.$$

Proof. The proof of this lemma is based on the standard contour integration techniques. The cases of odd and even r differ slightly, as the integrand is a meromorphic function when r is even. When r is odd, the integrand has a branch cut along the positive real axis.

In any case, using Cauchy's theorem, the contour $-\infty + iy \to \infty + iy$ can be transformed into the contour $C_- \cup C_R \cup C_+$, as drawn on the diagram.



In the drawn contour, C_R is an arc of the circle of radius R, while C_+ and C_- can be made arbitrarily close to the real axis. We will prove the lemma by induction on r. Let us begin by discussing cases $r = 1, 2$.

When $r = 1$, as $R \to 0$ we have $\int_{C_R} f(z)dz \le 2\pi R \left|\frac{1}{\sqrt{R}}\right| \to 0$. Moreover, it is easy to see that f differs by a sign above and below the real axis and therefore

$$\int_{C_-} f(z)dz + \int_{C_+} f(z)dz = 2\int_0^\infty \frac{\sqrt{i}}{\sqrt{x}} e^{-2i\pi nx} dx = 2\int_0^\infty \frac{\sqrt{i}}{\sqrt{x}} e^{-2i\pi nx} dx.$$

Introducing $y = 2i\pi nx$, we arrive at

$$2\int_0^{i\infty} \frac{\sqrt{i}}{\sqrt{x}} e^{-2i\pi nx} dx = \sqrt{\frac{2}{n\pi}} \int_0^{i\infty} \frac{e^{-y}}{\sqrt{y}} dy = \sqrt{\frac{2}{n}},$$

where we have used once more Cauchy's theorem and evaluated $\Gamma(\frac{1}{2}) = \int_0^\infty \frac{e^{-y}}{\sqrt{y}} dy = \sqrt{\pi}$.

In case $r = 2$, the integral over C_R is 2π, while the contours C_+ and C_- cancel out due to the fact that the integrand is continuous in this case. Hence, the above formula holds in cases $r = 1, 2$.

For $r > 2$, we apply partial integration on $u = e(-nz), dv = \left(\frac{i}{z}\right)^{r/2}$, giving

$$\int_C \left(\frac{i}{z}\right)^{r/2} e(-nz)dz = -\int_C \frac{2i^2\pi n}{-\frac{r}{2}+1} \left(\frac{i}{z}\right)^{r/2-1} e(-nz)dz = \frac{2\pi n}{\frac{r}{2}-1} \frac{(2\pi)^{\frac{r}{2}-1}}{\Gamma\left(\frac{r}{2}-1\right)} n^{\frac{r}{2}-2} = \frac{(2\pi)^{r/2}}{\Gamma(\frac{r}{2})} n^{\frac{r}{2}-1}.$$

This completes the proof of the lemma. □

Deriving the final formula for $r(n)$ is now simple. First, use the fact that

$$2\mathrm{Re}\int_0^{\frac{1}{c(c+K)}}\left(\frac{i}{z}\right)^{r/2}e(-nz)dx=\int_{-\infty}^{\infty}\left(\frac{i}{z}\right)^{r/2}e(-nz)dx+O\left((cK)^{\frac{r}{2}-1}\right).$$

It suffices to note that the error term introduced by changing the integration bounds in 14 is bounded by $O\left((cK)^{\frac{r}{2}-1}\right)$ in every integral, and therefore the total error term can be bounded by

$$\begin{aligned}\sum_{c\le K}c^{-r}T_0(c,n)(cK)^{\frac{r}{2}-1}&\le O\left(\sum_{c\le K}c^{-r}(n,c_1)^{\frac{1}{2}}c_0^{1/2}c^{\frac{r}{2}+\frac{1}{2}+\varepsilon}(cK)^{\frac{r}{2}-1}\right)\\&\le O\left(K^{r/2-1+\varepsilon}\sum_{c\le K}\frac{(n,c_1)^{\frac{1}{2}}}{c_1^{1/2}}\right)\\&\le O\left(K^{r/2-1/2+2\varepsilon}\right),\end{aligned}$$

which is still less than the previous error term acquired. Note that the tail of the series $\sum_{c\ge K}c^{-r}T_0(c,n)$ can be bounded using Lemma 7, and hence we can extend the sum to an infinite one. Therefore, we arrive at the asymptotic formula for $r(n)$:

$$r(n)=|A|^{-1/2}\frac{(2\pi)^{r/2}}{\Gamma(\frac{r}{2})}n^{\frac{r}{2}-1}\sum_{c\ge 1}c^{-r}T_0(c,n)+O\left(n^{\frac{r}{4}-\frac{1}{4}+\varepsilon}\right)$$

The term $\sigma(n)=\sum_{c\ge1}c^{-r}T_0(c,n)$ converges absolutely by virtue of bound 7. Still, the term $\sigma(n)$ is called the *singular series* because of its relation to rational singularities.

This is the term which is characterized by local representability of n. In the next section, we will show that, if n satisfies the appropriate local conditions, we must have $\sigma(n)\ge\Omega\left(\frac{1}{\log n}\right)$. Then, we see that for $r\ge4$ one has $r(n)\sim\Theta(n^{\frac{r}{2}-1}\sigma(n))$, implying that n can be represented by a quadratic form Q.

7 Singular series and local densities

The goal of this section is to investigate in more detail the singular series $\sigma(n)$, which is needed in order to derive Theorem 3 from the asymptotic expression 1. More specifically, we will show the following lemma.

Lemma 13. *If Q has rank $r\ge5$ and n is a positive integer which is locally represented by Q, then $\sigma(n)=\Theta(1)$. Similarly, if rank of Q is $r=4$, we have $\sigma(n)\ge\Omega\left(\frac{1}{\log n}\right)$ for all n that are primitively represented by Q.*

Note that Lemma 13 implies Theorem 3 almost immediately - it suffices to notice that the error term in 1 is smaller than the leading term whenever n satisfies the appropriate local conditions. Hence, we must have $r(n)>0$ for all such n.

The proof of Lemma 13 will be based on the relation between $\sigma(n)$ and the local densities $\delta_p(n)$. Therefore, before passing to the proof of Lemma 13, we need to introduce the notion of the local density, which captures how often an integer is represented in $\mathbb{Z}_p$. If $p\ge2$ is a prime, we define the *local density* $\delta_p(n)$ as

$$\delta_p(n)=\lim_{k\to\infty}\frac{\left|\{h\in(\mathbb{Z}/p^k\mathbb{Z})^r:Q(h)\equiv_{p^k}n\}\right|}{p^{k(r-1)}}.$$

It is not hard to see that, if n is a positive integer, the above ratio stabilizes for large enough k and hence the limit is well-defined. The intuition behind this definition is that there are p^{rk} possible values for h, and generically we expect about p^{-k} proportion of them to have $Q(h) \equiv_{p^k} n$. Hence, $\delta_p(n)$ is the ratio of the real number of solutions to $Q(h) \equiv_{p^k} n$ with the generically expected one. Understanding or computing local densities in general is not very easy, and [3] provides a nice discussion of known results.

The following lemma shows the tight connection between the local densities and the singular series, which will be the key to understanding the behavior of $\sigma(n)$.

Lemma 14. *Let P be a set of primes. If $\pi(c)$ denotes the set of prime divisors of c, we have:*

$$\sum_{c:\pi(c)\subseteq P} c^{-r} T_0(c,n) = \prod_{p\in P} \delta_p(n). \tag{15}$$

In particular, if P is the set of all primes we have $\sigma(n) = \prod_{p\geq 2} \delta_p(n)$.

Proof. In this proof, we apply the well-known Möbius inversion formula, which states the following: if $f, g : \mathbb{Z}_{>0} \to \mathbb{C}$ are multiplicative functions with $f(n) = \sum_{m|n} g(m)$, then we have $g(m) = \sum_{n|m} \mu\left(\frac{m}{n}\right) f(n)$. Here, the Möbius function $\mu : \mathbb{Z}_{>0} \to \mathbb{C}$ is a multiplicative function given by $\mu(p_1 \dots p_k) = (-1)^k$ for primes $p_1, \dots, p_k$ and $\mu(n) = 0$ for all n which are not squarefree.

To prove equation 15, we recall the definition of $T_0(c,n)$:

$$T_0(c,n) = \sum_{h \bmod c} \sideset{}{^*}\sum_{d \bmod c} e\left(\frac{d}{c}(Q(h)-n)\right).$$

Note that the inner sum in $T_0(c,n)$ is a simple Ramanujan sum which can be evaluated using the Möbius inversion formula.

Lemma 15. *For an arbitrary integer b, we have*

$$\sideset{}{^*}\sum_{d \bmod c} e\left(\frac{bd}{c}\right) = \sum_{q|(b,c)} \mu\left(\frac{c}{q}\right) q. \tag{16}$$

Proof. Let us define $f(c) = \sideset{}{^*}\sum_{d \bmod c} e\left(\frac{bd}{c}\right)$ and $g(q) = \mathbf{1}_{b|q} q$, where $\mathbf{1}_{b|q}$ is the indicator which evaluates to 1 if and only if $b|q$. With these definitions, we have

$$\sum_{c|q} f(c) = \sum_{c|q} \sideset{}{^*}\sum_{d \bmod c} e\left(\frac{bd}{c}\right) = \sum_{c|q} \sum_{\substack{d \bmod q \\ (d,q)=\frac{q}{c}}} e\left(\frac{bd}{\frac{q}{c}c}\right) = \sum_{d \bmod q} e\left(\frac{bd}{q}\right) = \mathbf{1}_{b|q} q = g(q).$$

Applying Möbius inversion hence gives:

$$\sideset{}{^*}\sum_{d \bmod c} e\left(\frac{bd}{c}\right) = \sum_{q|c} \mu\left(\frac{c}{q}\right) \mathbf{1}_{q|b} q = \sum_{q|(b,c)} \mu\left(\frac{c}{q}\right) q,$$

completing the proof of Lemma 15. □

In light of expression 16, one has the expression for $T_0(c,n)$ and a partial sum of $\sigma(n)$:

$$\begin{aligned} T_0(c,n) = \sum_{h \bmod c} \sum_{q|(c,Q(h)-n)} \mu\left(\frac{c}{q}\right) q &= \sum_{q|c} \mu\left(\frac{c}{q}\right) \left|\{h \bmod c : Q(h) \equiv_q n\}\right| \\ &= c^r \sum_{q|c} \mu\left(\frac{c}{q}\right) q^{1-r} \left|\{h \bmod q : Q(h) \equiv_q n\}\right|. \end{aligned}$$

If P is finite, letting $s = (\prod_{p\in P} p)^k$ and using Möbius inversion once more, we get the following conclusion:

$$\sum_{c|s} c^{-r} T_0(c,n) = s^{1-r} \left|\{h \bmod s : Q(h) \equiv_s n\}\right|$$

As we let k go to infinity, the absolute convergence of the singular series gives that:

$$\sum_{\pi(c)\subseteq P} c^{-r} T_0(c,n) = \prod_{p\in P} \delta_p(n). \tag{17}$$

In case P is an infinite set of primes, $P = \{p_1, p_2, \dots\}$, defining $s_k = \prod_{i=1}^{k} p_i^k$ and letting $k \to \infty$ gives the same result. In any case, we conclude that the singular series $\sigma(n)$ can be expressed as the product of $\delta_p(n)$ over all primes p, $\sigma(n) = \prod_p \delta_p(n)$. Thus, the proof of Lemma 14 is complete. □

Remark. In fact, one can interpret the leading term $\frac{(2\pi)^{r/2} n^{\frac{r}{2}-1}}{|A|^{1/2}\Gamma\left(\frac{r}{2}\right)}$ in the formula 1 as the local density with respect to $\mathbb{R}$. From this perspective, we can say that $r(n) \sim \prod_{p\geq -1} \delta_p(n)$.

Equipped with the relation between $\delta_p(n)$ and $\sigma(n)$, we are ready to prove Lemma 13.

Proof of Lemma 13. Once the formula $\sigma(n) = \prod_p \delta_p(n)$ is established, the proof is based on analyzing local densities at various primes, depending on whether p divides $2|A|$ or not. We begin by discussing the primes not dividing $2|A|$. In case $r \geq 5$, it is easy to bound their contribution by a constant, while $r = 4$ requires a little more work and has a slightly worse bound.

Lemma 16. *If $r = 4$, we have:*

$$\prod_{p\nmid 2|A|} \delta_p(n) = \Theta\left(\prod_{p|n,(p,2|A|)=1} (1 - \chi_D(p)p^{-1})\right) \geq \Omega\left(\frac{1}{\log n}\right),$$

where $D = (-1)^{r/2}|A|$ and $\chi_D(p) = \left(\frac{D}{p}\right)$.

Proof. As we are working with odd primes not dividing $|A|$, Lemma 11 gives an efficient way to evaluate the Gauss sums, which proves the statement quickly. We begin by fixing a prime p with $(p, 2|A|) = 1$ and computing the local density $\delta_p(n)$. After than, a series of simple estimates will yield the desired statement.

To isolate the local density modulo p, we plug $s = p^{v_p(n)+1} = p^{\alpha_p}$ into equation 15 and get

$$\delta_p(n) = \sum_{c=p^k} c^{-r} \sideset{}{^*}\sum_{d \bmod c} G_0\left(\frac{d}{c}\right) e\left(-n\frac{d}{c}\right).$$

As every term c appearing in the sum is coprime to $2|A|$, Lemma 11 applies to give:

$$\delta_p(n) = \sum_{c=p^k} c^{-r} \sideset{}{^*}\sum_{d \bmod c} \left(\frac{|A|}{c}\right)\left(\varepsilon_c\left(\frac{2d}{c}\right)\sqrt{c}\right)^r e\left(-n\frac{d}{c}\right) = \sum_{c=p^k} \left(\frac{D}{c}\right) c^{-r/2} \sideset{}{^*}\sum_{d \bmod c} e\left(-n\frac{d}{c}\right),$$

We recognize the expression from the Lemma 15 in the above equation, and by using multiplicativity of μ and χ_D, we get

$$\begin{aligned}
\delta_p(n) &= \sum_{c=p^k} \chi_D(c) c^{-r/2} \sideset{}{^*}\sum_{d \bmod c} e\left(-n\frac{d}{c}\right) \\
&= \sum_{c=p^k} \chi_D(c) c^{-r/2} \sum_{q|(c,n)} \mu\left(\frac{c}{q}\right) q \\
&= \sum_{q|(s,n)} q \sum_{q|c|s} \mu\left(\frac{c}{q}\right) \chi_D(c) c^{-r/2} \\
&= \sum_{q|(s,n)} q^{1-\frac{r}{2}} \chi_D(q) \sum_{c|\frac{s}{q}} \mu(c)\chi_D(c) c^{-r/2} \\
&= \sum_{q|(s,n)} q^{1-\frac{r}{2}} \chi_D(q) \prod_{p|\frac{s}{q}} (1-\chi_D(p)p^{-r/2}).
\end{aligned}$$

Using the fact that $v_p(s) = \alpha_p = v_p(n) + 1$ we get $(s,n) = p^{\alpha_p - 1}$ implying

$$\delta_p(n) = \frac{1-\chi_D(p^{\alpha_p})p^{\alpha_p\left(1-\frac{r}{2}\right)}}{1-\chi_D(p)p^{1-\frac{r}{2}}}(1-\chi_D(p)p^{-\frac{r}{2}}).$$

Having derived an expression for $\delta_p(n)$, we proceed to estimate the product of local densities at primes p coprime to $2|A|$:

$$\prod_{(p,2|A|)=1} \delta_p(n) = \underbrace{\prod_{\substack{(p,2|A|)=1 \\ p|n}} \left(1-\chi_D(p^{\alpha_p})p^{\alpha_p\left(1-\frac{r}{2}\right)}\right)}_{P_1} \underbrace{\prod_{\substack{(p,2|A|)=1 \\ p|n}} \frac{1}{1-\chi_D(p)p^{1-\frac{r}{2}}}}_{P_2} \underbrace{\prod_{(p,2|A|)=1} \left(1-\chi_D(p)p^{-\frac{r}{2}}\right)}_{P_3}.$$

The products P_1, P_3 are simple to bound by absolute constants, as the appearing exponents are all at least 2. Namely, we have:

$$\prod_p \left(1-\frac{1}{p^2}\right) \le \min\{P_1, P_3\} \le \max\{P_1, P_3\} \le \prod_p \left(1+\frac{1}{p^2}\right).$$

Therefore, we conclude that $P_1, P_3 = \Theta(1)$. As $r = 4$, the product P_2 can be bounded by

$$\prod_{p|n} \left(1+\frac{1}{p}\right) \ge P_2 \ge \prod_{p|n} \left(1-\frac{1}{p}\right).$$

After taking logarithms, this transforms to

$$\sum_{p|n} \frac{1}{p} + \Theta(1) \ge \log P_2 \ge -\sum_{p|n} \frac{1}{p} + \Theta(1).$$

Using the fact that the sum of reciprocals of all primes up to n is $O(\log\log n)$, we have $O(\log n) \ge P_2 \ge \Omega\left(\frac{1}{\log n}\right)$, which completes the proof. □

Remark. For $r = 4$, the estimate from the proof can be improved to deduce $\prod_{(p,2|A|)=1} \delta_p(n) \ge O((\log\log n)^{-1})$, through a more precise estimate of the sum $\sum_{p|n} \frac{1}{p}$, which can in fact be bounded by $O(\log\log\log n)$. However, this has no effect on the proof itself and hence, for the sake of brevity, we do not present this argument.

Now, we deal with the case $r \geq 5$ and $(p, 2|A|) = 1$.

Lemma 17. *If $r \geq 5$, we have:*

$$\prod_{p \nmid 2|A|} \delta_p(n) = \Theta(1).$$

Proof. The proof will be somewhat similar to the proof of Lemma 16 in that we will use Lemma 11 to evaluate the arising Gauss sums. However, the fact $r \geq 5$ ensures fast convergence of the singular series using the simplest bounds, which will allows to bound the product of local densities by absolute constants.

We begin by using Lemma 11 to deduce the trivial bound $|T_0(c, n)| \leq c^{\frac{r}{2}+1}$ when $(c, 2|A|) = 1$. Therefore we have the upper bound,

$$\sum_{(c,2|A|)=1} c^{-r} T_0(c, n) \leq \sum_{(c,2)=1} c^{1-\frac{r}{2}} \leq (1 - 2^{-3/2})\zeta\left(\frac{3}{2}\right) \approx 1.69.$$

This also translates to give the lower bound by:

$$\sum_{(c,2|A|)=1} c^{-r} T_0(c, n) \geq 1 - \sum_{\substack{(c,2)=1 \\ c>1}} c^{1-\frac{r}{2}} \geq 2 - (1 - 2^{-3/2})\zeta\left(\frac{3}{2}\right) \approx 0.31.$$

Combined with Lemma 14, this completes the proof. □

Lemmas 16 and 17 complete the discussion of primes coprime to $2|A|$. Now, we have two lemmas estimating local densities at the remaining finitely many primes.

Lemma 18. *If p divides $2|A|$ and n is primitively locally represented by Q, then $\delta_p(n) \geq \Omega(1)$.*

Proof. We will present the proof in case p is odd. The case $p = 2$ can be dealt with in a similar fashion, with a little more care.

We begin by diagonalizing Q and writing $Q = \sum_i p^{a_i} c_i x_i^2$, for some integers $a_1, \ldots, a_r \geq 0$ and some units $c_1, \ldots, c_r \in (\mathbb{Z}/p\mathbb{Z})^*$. Moreover, we let $k \geq 2\max_i\{a_i\} + 1 = M$. We will show that the number of primitive solutions $x = (x_1, \ldots, x_r) \in S_k = \prod_{i=1}^r \mathbb{Z}/p^{k-a_i}\mathbb{Z}$ is at least $p^{(r-1)(k-M)}$, where a primitive solution is one where not all entries are divisible by p. In order to show this, we describe a lifting process which starts from a primitive solution to the equation $Q(x) \equiv_{p^k} n$, living in S_k and lifts it to p^{r-1} different solutions to the equation $Q(x) \equiv_{p^{k+1}} n$ in S_{k+1}.

To describe this lifting process, suppose we are given a primitive solution $x \in S_k$ and consider a potential lift $x' \in S_{k+1}$ given by $x_i' = x_i + \alpha_i p^{k-a_i}$, for $(\alpha_1, \ldots, \alpha_r) \in (\mathbb{Z}/p\mathbb{Z})^r$. This lift satisfies the equation $Q(x') \equiv_{p^{k+1}} n$ if and only if

$$Q(x_1, \ldots, x_n) + p^k \sum_{i=1}^r 2x_i c_i \alpha_i \equiv_{p^{k+1}} n, \tag{18}$$

where we has used the fact $k \geq M$ to get rid of higher powers of p. As $Q(x) \equiv_{p^k} n$, we can write $Q(x) \equiv_{p^{k+1}} n + ap^k$, and the equation 18 becomes

$$\sum_{i=1}^r 2x_i c_i \alpha_i \equiv_p a \tag{19}$$

The assumption that x is a primitive solution ensures that some α_i has a non-zero coefficient on the left hand side, meaning that the equation 19 has at least p^{r-1} solutions. Note that throughout this

lifting process, all possible lifts of x retain the same projection onto S_k. Therefore, lifts arising from different solutions $x_1 \neq x_2$ cannot be the same. Also, note that all lifts of a primitive solution remain primitive, by the same argument.

The above argument inductively implies that, starting from a single primitive solution in S_M, one can generate at least $p^{(r-1)(k-M)}$ solutions in S_k, which gives a lower bound on the local density at p

$$\delta_p(n) \geq p^{-(r-1)M}.$$

As p is bounded in size by $|A|$, this completes the proof of the lemma. □

Lemma 19. *If $r \geq 5$ and n is locally represented by Q, then $\delta_p(n) \geq \Omega(1)$ for all p dividing $2|A|$.*

Proof. The proof strategy will be similar to Lemma 18, in that it suffices to take a primitive solution to $Q(x) \equiv_{p^M} n$ and lift it to produce more solutions. However, as we only have the assumption that n is locally represented modulo p, we also need to construct the primitive solution. Again, we focus on the case when p is odd, and leave $p = 2$ to the reader. We also assume Q is diagonalized, i.e. $Q = \sum_i p^{a_i} c_i x_i^2$ and $2\max_i\{a_i\} + 1 = M$.

We have to consider two cases: $v_p(n) \geq M$ and $v_p(n) < M$. The first case is simple, but it is the case where we use the fact that the rank of the quadratic form is at least 5.

Namely, a quadratic form of rank $r \geq 5$ is isotropic in $\mathbb{Q}_p$ in the sense there exists a nonzero vector $x \in \mathbb{Q}_p^r$ such that $Q(x) = 0$. To see why this is true, we recall that all forms in at least 3 variables are isotropic in $\mathbb{Z}/p\mathbb{Z}$. Then, we can write the form Q as $Q(y, z) = \sum_i c_i y_i^2 + p\sum_j d_j z_j^2 = Q_1 \oplus pQ_2$ over $\mathbb{Q}_p$. At least one of the forms Q_1, Q_2 has rank ≥ 3, implying it is isotropic over $\mathbb{Z}/p\mathbb{Z}$.

Hence, there exists a non-trivial solution $(y, z) \in (\mathbb{Z}/p\mathbb{Z})^r$ to $Q(y, z) \equiv_p 0$. Additionally, we may rearrange the variables and assume $y_1 \not\equiv_p 0$ or $z_1 \not\equiv_p 0$. If $y_1 \not\equiv_p 0$, we fix $y_2, \ldots, y_n$ and look at the equation $Q_1(y_1, \ldots, y_n) \equiv_p 0$ as an equation in y_1. As $\frac{\partial}{\partial y_1} Q_1(y_1, \ldots, y_n) = c_1 y_1 \not\equiv_p 0$, Hensel's lemma implies that the root y_1 of this equation can be lifted to give $Q_1(y_1, \ldots, y_r) = 0$ in $\mathbb{Q}_p$. Plugging in a zero vector for z gives $Q(y, z) = 0$ in $\mathbb{Q}_p$, as needed. The case $z_1 \not\equiv_p 0$ can be dealt with in a similar fashion after dividing by p once.

We conclude that there exists a nontrivial solution $x \in \mathbb{Q}_p^r$ to $Q(x) = 0$, which can be scaled by a power of p to become a primitive solution in $\mathbb{Z}_p$. As $v_p(n) \geq M$, this x satisfies $Q(x) \equiv_{p^M} 0 \equiv_{p^M} n$, and therefore the solution can be lifted to show the local density is bounded below by $\Omega(1)$.

In case $v_p(n) \leq M$, we can take an arbitrary solution x to $Q(x) \equiv_{p^{2M}} n$ and consider $\alpha = \min\{v_p(x_i)\}$. Clearly, $\alpha \leq M/2$ as $v_p(n) \leq M$. Then, we can define $x_i' = x_i/p^\alpha$ to obtain a primitive solution x' to $Q(x') \equiv_{p^{2M-2\alpha}} n/p^{2\alpha}$. As $2M - 2\alpha > M$, this solution can be lifted to give at least p^{l-M} solutions to $Q(x') \equiv_{p^l} n/p^{2\alpha}$. Rescaling $x' \mapsto p^\alpha x'$ now provides at least $p^{l-M-2\alpha}$ solutions to $Q(x') \equiv_{p^l} n$, which shows that $\delta_p(n) \geq \Omega(1)$, thus completing the proof. □

Having dealt with all cases in the previous four lemmas, it is easy to see that $\prod_p \delta_p(n) = \Theta(1)$ if $r \geq 5$ and $\prod_p \delta_p(n) \geq \Omega\left(\frac{1}{\log n}\right)$ if $r = 4$, which completes the proof of Lemma 13. □

Remark. Note that the proof of Lemma 13 actually gives an asymptotic result on the size of $\sigma(n)$, which can be extended to get the asymptotics of $r(n)$, by the means of formula 1. In this way, one can conclude:

$$r(n) = \Theta\left(n^{\frac{r}{2}-1}\right), \text{ if } r \geq 5, \text{ and}$$

$$r(n) = \Theta\left(n^{\frac{r}{2}-1} \prod_{p|n} \left(1 - \chi_D(p)p^{-1}\right)\right), \text{ if } r = 4.$$

I would like to thank Shilin Lai, Weibo Fu and Nina Zubrilina for helpful discussions about this project.

References

[1] William Duke and Rainer Schulze-Pillot, Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids, *Invent. Math.* , 99 (1990): 49–57.

[2] Jonathan Hanke, Quadratic Forms and Automorphic Forms, https://arxiv.org/pdf/1105.5759

[3] Jonathan Hanke, Local densities and explicit bounds for representability by a quadratric form, *Duke Math. J.* 124 (2004), no. 2, 351–388.

[4] Godfrey H. Hardy and Srinivasa Ramanujan, Asymptotic formulae in combinatorial analysis, *Proc. London Math. Soc.* 17 (1918): 75 - 115.

[5] Godfrey H. Hardy and John E. Littlewood, A new solution of Waring's Problem, *Q. J. Math.* 48 (1919): 272-293.

[6] Roger Heath-Brown. A new form of the circle method, and its application to quadratic forms. *Journal für die reine und angewandte Mathematik* 481 (1996): 149-206.

[7] Henryk Iwaniec. *Topics in classical automorphic forms*, volume 17 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1997.

[8] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[9] Hendryk D. Kloosterman, On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* 49 (1926), 407-464.

[10] Toshitsune Miyake, *Modular forms*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, english edition, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[11] Jeremy Rouse, Integers represented by positive-definite quadratic forms and Petersson inner products, *Acta Arithmetica* 187 (2019), 81–100.

[12] Rainer Schulze-Pillot. Representation by integral quadratic forms—a survey. In *Algebraic and arithmetic theory of quadratic forms*, volume 344 of *Contemp. Math.*, pages 303–321. Amer. Math. Soc., Providence, RI, 2004.

[13] W. Tartakowsky. Die gesamtheit der zahlen, die durch eine positive quadratische form $f(x_1, x_2, \ldots, x_s)$ $(s \geq 4)$ darstellbar sind. i, ii. *Bull. Ac. Sc. Leningrad,* 2 (1929): 111–122; 165–196.

[14] Robert C. Vaughan, *The Hardy–Littlewood Method*, Cambridge Tracts in Mathematics, vol. 125 (2nd ed.), Cambridge University Press, 1997.

[15] George L. Watson, *Integral quadratic forms*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 51, Cambridge University Press, New York, 1960.