# Lattice points in bodies of revolution II

Fernando Chamizo        Carlos Pastor

**Abstract**

In [3] it was shown that when a three-dimensional smooth convex body has rotational symmetry around a coordinate axis one can find better bounds for the lattice point discrepancy than what is known for more general convex bodies. To accomplish this, however, it was necessary to assume a non-vanishing condition on the third derivative of the generatrix. In this article we drop this condition, showing that the aforementioned bound holds for a wider family of revolution bodies, which includes those with analytic boundary. A novelty in our approach is that, besides the usual analytic methods, it requires studying some Diophantine properties of the Taylor coefficients of the phase on the Fourier transform side.

## 1    Introduction

Let $\mathbb{E} \subset \mathbb{R}^d$, $d > 1$, be a compact convex body with non-empty interior, and whose boundary $\partial\mathbb{E}$ is a smooth $(d-1)$-submanifold with positive Gaussian curvature (for short, a *smooth convex body*). Denote by $\mathcal{N}(R)$ the number of points of $\mathbb{Z}^d$ lying in $\mathbb{E}$ after being dilated by a factor $R > 1$, *i.e.*,

$$\mathcal{N}(R) = \#\{\vec{n} \in \mathbb{Z}^d \,:\, \vec{n}/R \in \mathbb{E}\}.$$

A central question in lattice point theory consists in estimating how big the discrepancy

$$\mathcal{E}(R) = \mathcal{N}(R) - R^d|\mathbb{E}|$$

can be for a particular choice of $\mathbb{E}$, where $|\mathbb{E}|$ stands for the volume of $\mathbb{E}$. We denote by $\alpha_\mathbb{E}$ the minimal exponent (using Landau's notation)

$$\alpha_\mathbb{E} = \inf\big\{\alpha \,:\, \mathcal{N}(R) = R^d|\mathbb{E}| + O(R^\alpha)\big\}.$$

A geometrical observation originally due to Gauss for the circle [5], shows that $\alpha_\mathbb{E} \leq d-1$. This elementary upper bound has been improved by numerous authors. The state of the art is the following: Huxley [11] has proved $\alpha_\mathbb{E} \leq 131/208$ for $d = 2$, and Guo [7] $\alpha_\mathbb{E} \leq 73/158$ for $d = 3$, and $\alpha_\mathbb{E} \leq (d^2 + 3d + 8)/(d^3 + d^2 + 5d + 4)$ for $d \geq 4$.

For some particular choices of $\mathbb{E}$ it is possible to do better. For example, when $\mathbb{E}$ is the $d$-dimensional unit ball, the problem is essentially solved for $d \geq 4$: the equality $\alpha_\mathbb{E} = d-2$ follows from classical results on representations of integers by quadratic forms. The best known upper bound for the three-dimensional unit ball is $\alpha_\mathbb{E} \leq 21/16$ shown by Heath-Brown [8], while for the circle a recent preprint of Bourgain and Watt [2] improves Huxley's result to $\alpha_\mathbb{E} \leq 517/824$.
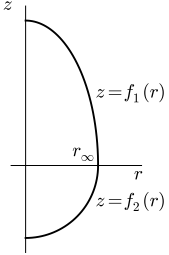
The aim of the article [3] was to show that, albeit the substantial differences in lattice point problems for general smooth convex bodies and balls that the

previous results evidentiate, if one assumes rotational symmetry around a coordinate axis then one can obtain intermediate results even from the simplest van der Corput's estimate. We consider three-dimensional smooth convex bodies of the form

$$\mathbb{E} = \left\{ (x,y,z) \in \mathbb{R}^3 \; : \; f_2(r) \leq z \leq f_1(r), \; 0 \leq r \leq r_\infty \right\} \tag{1}$$

where $r = \sqrt{x^2 + y^2}$.

In other words, $\mathbb{E}$ is the solid generated by the rotation around the $z$-axis of the curve

$$\gamma(t) = \begin{cases} \big(t, 0, f_1(t)\big) & 0 \leq t \leq r_\infty \\[2mm] \big(2r_\infty - t, 0, f_2(2r_\infty - t)\big) & r_\infty \leq t \leq 2r_\infty \end{cases}$$

Theorem 1.1 of [3] reads:

**Theorem.** *Let $\mathbb{E} \subset \mathbb{R}^3$ be a body of revolution as before and suppose that the functions $\frac{1}{r} f_i'''(r)$ (extended by continuity to $r = 0$) do not vanish for $0 \leq r < r_\infty$, where $i = 1, 2$. Then the inequality $\alpha_\mathbb{E} \leq 11/8$ holds.*

This result is not entirely satisfactory, as the extra hypothesis concerning the non-vanishing of the function $\frac{1}{r} f_i'''(r)$ lacks geometrical meaning. This kind of technical hypothesis, which often appear when applying van der Corput's estimates, are usually very difficult to deal with. The problem worsened when multiple exponent pairs were introduced. For instance, until the arrival of the discrete Hardy-Littlewood method [10], stating and checking non-vanishing conditions was a substatial task even for the circle and divisor problems (see [9] and [13]). The problem still persists when $d > 2$. In this article, however, we prove

**Theorem 1.** *Let $\mathbb{E} \subset \mathbb{R}^3$ be a body of revolution as before and suppose that the functions $f_i$ are real analytic for $0 \leq r < r_\infty$ and $i = 1, 2$. Then the inequality $\alpha_\mathbb{E} \leq 11/8$ holds.*

In fact much less is needed as evident from the proof below. Theorem 1 holds as long as every zero of the function $f_i'''$ is of finite order for both $i = 1, 2$, i.e. $f_i'''(r) = 0$ implies we can find an integer $n > 3$ such that $f_i^{(n)}(r) \neq 0$. As remarked in the next section, the result also holds if in the definition (1) we take $r = \sqrt{Q(x + \alpha, y + \beta)}$ with $Q$ a positive definite rational quadratic form and $\alpha, \beta \in \mathbb{R}$. In other words, Theorem 1 extends to the case in which the horizontal sections are rational ellipses with a common center when projected onto the $xy$-plane.

The idea of the proof is the following: we transform the problem via Poisson summation into estimating an exponential sum, as it is customary; and then slice the sum diadically in pieces corresponding to the zeros of $f_i'''(r)$. For

2

the pieces where van der Corput's estimation falls short the phase is almost linear, and we are in position to apply the first derivative test (Kuzmin-Landau inequality [6]). This, by itself, is not good enough, as the derivative of the phase function might happen to be close to an integer way too often. Showing that this cannot be the case requires –in some ranges– studying certain diophantine properties of the Taylor coefficient in question. This goes beyond the utterly analytic treatment in the classical (van der Corput's) theory of exponential sums and vaguely resembles to the situation in [1] (the seminal paper for the discrete Hardy-Littlewood method) in which the arithmetic properties of the Taylor coefficients play a fundamental role.

It is worth mentioning that while looking for examples where the non-vanishing condition is blatantly violated, the authors were led to the case of revolution paraboloids (and more generally elliptic paraboloids) for which it turns out that one can explicitly determine $\alpha_{\mathbb{E}}$ by relating the associated exponential sum to some analytic functions enjoying automorphic properties [4]. In some sense a related phenomenon is happening here, as very close to a zero of $\frac{1}{r}f_i'''(r)$ the function $f_i(r)$ essentially looks like a parabola, and some of the arithmetic leaks in in the form of the aforementioned diophantine properties of the Taylor coefficient.

Throughout this article we use Vinogradov's notation $f \ll g$ with the same meaning as Landau's $f = O(g)$, i.e., $|f| \leq C|g|$ for an unspecified constant $C$. We employ $f \gg g$ to denote $g = O(f)$, and $f \asymp g$ when both of these conditions hold. We also abbreviate $e^{2\pi it}$ by $e(t)$.

## 2  The exponential sum

Our starting point will be the truncated Hardy-Voronoï formula given by Proposition 2.1 of [3]. We restate it here for convenience. Fix a smooth even function $\eta \in C_0^\infty\big((-1,1)\big)$ with $\eta(0) = 1$ and satisfying that the Fourier transform of $\eta(\|\vec{x}\|)$ is positive (this latter condition can be easily fulfilled by considering the convolution of a radial function with itself).

**Proposition 1.** *Let $\mathbb{E} \subset \mathbb{R}^3$ be a smooth convex body, $\eta$ as before, and fix $\epsilon > 0$ and $0 < c < 1$. Then for any given $R > 2$ there exists $R' \in \big(R-2, R+2\big)$ such that*

$$\mathcal{E}(R) = -\frac{R'}{\pi} \sum_{\vec{n} \in \mathbb{Z}^3 - \{0\}} \eta\big(\delta\|\vec{n}\|\big) \frac{\cos\big(2\pi R' g(\vec{n})\big)}{\|\vec{n}\|^2 \sqrt{\kappa(\vec{n})}} + O\big(R^{2+\epsilon}\delta\big),$$

*where $\delta = R^{-c}$, $g$ is the* support function *$g(\vec{n}) = \sup\{\vec{x} \cdot \vec{n} \, : \, \vec{x} \in \mathbb{E}\}$ and $\kappa(\vec{n})$ stands for the Gaussian curvature of $\partial\mathbb{E}$ at the point whose unit outer normal is $\vec{n}/\|\vec{n}\|$.*

To obtain an error term of the form $O\big(R^{11/8+\epsilon}\big)$ we must choose $c \geq 5/8$, and the larger we pick $c$ the longer the exponential sum we have to bound becomes. The natural choice is therefore $c = 5/8$ and $\delta = R^{-5/8}$.

All the functions of $\vec{n}$ involved in the expression for $\mathcal{E}(R)$ given by Proposition 1 are invariant under rotations on the first two variables. Grouping the

corresponding terms and applying summation by parts,

$$\mathcal{E}(R) \ll \sup_{N,M^2 \le \delta^{-2}} \frac{R^{1+\epsilon}}{N+M^2} \left| \sum_{1 \le n \le N} \sum_{1 \le |m| \le M} r_2(n)e\big(R'h(n,m)\big) \right| + R^{11/8+\epsilon}, \quad (2)$$

where $h(n,m) = g\big(\sqrt{n},0,m\big)$ and $r_2(n)$ stands for the number of representations of $n$ as sum of two squares (the contribution corresponding to the terms with $n = 0$ or $m = 0$ is negligible). The summation by parts step is justified as long as we can guarantee

$$\int_1^{\delta^{-2}} \int_1^{\delta^{-1}} (u+v^2) \left| \frac{\partial^2}{\partial u \partial v} \frac{\eta\big(\delta\sqrt{u+v^2}\big)}{(u+v^2)\sqrt{\kappa(\sqrt{u},0,v)}} \right| du\,dv \ll R^\epsilon,$$

and a similar bound with $v$ replaced with $-v$. This becomes evident after performing the change of variables $u \mapsto u^2$ and changing to polar coordinates (note $\kappa(u,0,v)$ depends smoothly on the angle $\theta = \arctan v/u$).

When $r$ in (1) is replaced by $r = \sqrt{Q(x+\alpha, y+\beta)}$ with $Q$ a positive definite rational quadratic form, as mentioned in the introduction, rescaling $f_i$ we can assume that the *dual form* $Q^*$ (the one having the inverse matrix) is an integral quadratic form. By the properties of the Fourier transform, (2) still holds but replacing $r_2(n)$ by
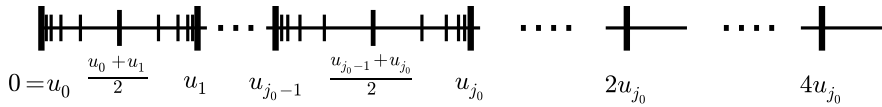
$$r_2^*(n) = \frac{1}{\det Q} \sum_{Q^*(x,y)=n} e(\alpha x + \beta y) \qquad \text{where } x,y \in \mathbb{Z}.$$

See §6 of [3] for more details. In what follows the only fact that will be used about the arithmetic function $r_2$ is the bound $r_2(n) \ll n^\epsilon$ for any $\epsilon > 0$, also satisfied by $r_2^*(n)$. Therefore all the forthcoming arguments may be readily applied in the context of rational elliptic sections.

If instead of the convex body $\mathbb{E}$ we consider its specular reflection over the plane $z = 0$, we arrive at exactly the same expression (2) but with the sign of $m$ reversed (because the invariance of the Fourier transform by symmetries). This means that we can restrict our attention to the half of the sum consisting of those terms with $m > 0$, and then apply the same argument to the specular reflection of $\mathbb{E}$ to bound the other half in the same way. In what follows, therefore, we restrict $m$ to be positive, and rename $f_1$ to $f$ to avoid the excessive use of subindices.

Let $0 = r_0 < r_1 < \cdots < r_{j_0-1}$ be the zeros of $f'''$ in $[0, r_\infty)$ and fix any $r_{j_0}$ satisfying $r_{j_0-1} < r_{j_0} < r_\infty$. Denote $u_j = (f'(r_j))^2$ for $0 \le j \le j_0$. We are going to split the summation domain of (2) dyadically in $m$ as $m \to +\infty$, and in $n/m^2$ as it approaches either some $u_j$ or $+\infty$ (see figure below); decomposing the sum in an at most a constant times $\log R$ number of pieces of the form

$$S(U_1, U_2, M) = \sum_{\substack{U_1 \le n/m^2 < U_2 \\ M \le m < 2M}} \sum r_2(n)e\big(Rh(n,m)\big). \qquad (3)$$



$$0 = u_0 \quad \frac{u_0+u_1}{2} \quad u_1 \qquad u_{j_0-1} \quad \frac{u_{j_0-1}+u_{j_0}}{2} \quad u_{j_0} \qquad 2u_{j_0} \qquad 4u_{j_0}$$

4

It is clear that after this dyadic subdivision, we can deduce Theorem 1 from the following theorems. The part $0 \le n/m^2 < u_{j_0}$ of the double sum in (2) is covered by Theorem 2, except for the terms with $n/m^2 = u_j$ that can be estimated trivially, and the part $n/m^2 \ge u_{j_0}$ is covered by Theorem 3.

**Theorem 2.** *Given $\epsilon > 0$ and $0 \le j < j_0$, for any $R > 1$, $2M \le R^{5/8}$ and $0 < U \le (u_{j+1} - u_j)/4$ we have*

$$\left| S(u_{j+1} - 2U, u_{j+1} - U, M) \right| + \left| S(u_j + U, u_j + 2U, M) \right| \ll M^2 R^{3/8+\epsilon}.$$

**Theorem 3.** *Given $\epsilon > 0$, for any $R > 1$, $2M \le R^{5/8}$ and $u_{j_0} \le U \le R^{5/4}M^{-2}$ we have*

$$S(U, 2U, M) \ll UM^2 R^{3/8+\epsilon}.$$

# 3 Weyl step

In order to be able to estimate the sum $S$ given by (3) using the van der Corput method we must first get rid of the arithmetic function $r_2$. We do this by performing a so-called *Weyl step* (*cf.* §8.2 of [12]).

**Proposition 2.** *Let $S$ as before and fix $\epsilon > 0$. For any $1 \le M \le R^{5/8}$, $0 < U_1 < U_2 \le R^{5/4}$ and $1 \le L \le M$, satisfying $U_2 - U_1 = U$ and $U_2 L + 1 \ll UM$, we have*

$$\left| S(U_1, U_2, M) \right|^2 \ll R^\epsilon (U^2 M^6 L^{-1} + UM^3 T)$$

*where $T = T(U_1, U_2, M, L)$ is given by*

$$T = \frac{1}{L} \sum_{1 \le \ell \le L} \left| \sum_{\substack{U_1 \le n/(m+\ell)^2, n/m^2 < U_2 \\ M \le m, m+\ell < 2M}} e\big(R(h(n, m + \ell) - h(n, m))\big) \right|. \qquad (4)$$

*Proof.* Consider

$$\psi_{n,m} = \begin{cases} e\big(Rh(n,m)\big) & \text{if } U_1 \le n/m^2 < U_2 \text{ and } M \le m < 2M, \\ 0 & \text{otherwise.} \end{cases}$$

It suffices to prove the inequality when $L$ is an integer. We may therefore write

$$LS = \sum_{M-L \le m < 2M} \sum_{U_1 m^2 \le n < U_2(m+L)^2} r_2(n) \sum_{1 \le \ell \le L} \psi_{n,m+\ell}.$$

The length of the first sum is $\ll M$ and the length of the second one $\ll UM^2$, hence squaring and applying Cauchy-Schwarz,

$$L^2 S^2 \ll R^\epsilon UM^3 \sum_{M-L \le m < 2M} \sum_{U_1 m^2 \le n < U_2(m+L)^2} \sum_{1 \le \ell_1, \ell_2 \le L} \psi_{n,m+\ell_1} \overline{\psi_{n,m+\ell_2}}.$$

Separating the diagonal contribution $\ell_1 = \ell_2$ and interchanging the summation order, which can be done because $\psi_{n,m}$ keeps track of the summation domain,

$$L^2 S^2 \ll R^\epsilon LU^2 M^6 + R^\epsilon UM^3 \Re \sum_{1 \le \ell_2 < \ell_1 \le L} \sum_n \sum_m \psi_{n,m+\ell_1} \overline{\psi_{n,m+\ell_2}}.$$

To obtain the desired inequality perform the change of variables $m \mapsto m - \ell_2$ and group the terms corresponding to each value of $\ell = \ell_1 - \ell_2$. $\qquad \square$

# 4 The function $h$

In this section we prove the estimates we need about the function $h$. Note that the convexity of $-f$ implies that $-f' : [0, r_\infty) \to \mathbb{R}^+$ is one-to-one, and therefore its inverse function $\phi$ is well-defined. By Lemma 4.1 of [3] we know that

$$\frac{\partial}{\partial m} h(n, m) = F(n/m^2) \quad \text{where} \quad F(u) = f(\phi(\sqrt{u})). \tag{5}$$

The estimates for $h$ near the "bad" points $u_j$ will depend on the order of vanishing of $f'''(r)$. By definition, each $u_j$ is the preimage by the function $\phi(\sqrt{u})$ of a zero $r_j$ of $f'''(r)$, except the last one which is added for convenience. If $r_j \neq 0$ we define $d_j$ as the unique nonnegative integer satisfying $f'''(r) \asymp (r - r_j)^{d_j}$ as $r \to r_j$. For $r_0 = 0$ we define $d_0$ as the unique nonnegative integer satisfying $f'''(r) \asymp r^{2d_0+1}$ as $r \to 0^+$. We also set $d_\infty = -5/2$.

**Lemma 1.** *We have $F'(u) \asymp (1 + u)^{-3/2}$ for $0 \leq u < \infty$. We also have $F''(u) \neq 0$ for $u \neq u_j$, $0 \leq j \leq j_0$, and*

$$F''(u) \asymp (u - u_j)^{d_j} \quad as \quad u \to u_j \quad and \quad F''(u) \asymp u^{d_\infty} \quad as \quad u \to \infty.$$

*Proof.* Let $k(r)$ denote the curvature of $r \mapsto (r, f(r))$, which admits the explicit formula (see p. 11 of [14])

$$f''(r) = k(r)(1 + |f'(r)|^2)^{3/2}, \tag{6}$$

and set $c(u) = k(\phi(\sqrt{u}))$. Differentiating $F$ and recalling that $\phi$ is the inverse function of $-f'$ we obtain (*cf.* the proof of Lemma 4.2 of [3])

$$F'(u) = \frac{1}{2c(u)(1 + u)^{3/2}},$$

$$F''(u) = \frac{f'''(\phi(\sqrt{u}))}{4(c(u))^3(1 + u)^{9/2}u^{1/2}}.$$

Now all but the last claim of the lemma is clear as $c_1 < c(u) < c_2$ for some constants $c_1, c_2 > 0$ and $\phi(\sqrt{u})$ is a regular function for $u > 0$, and behaves like $C\sqrt{u}$ for some $C \neq 0$ as $u \to 0^+$. To establish the last claim, we note that by (6) and L'Hôpital's rule,

$$k(r_\infty) = \lim_{r \to r_\infty^-} \frac{f''(r)}{(f'(r))^3} = \lim_{r \to r_\infty^-} \frac{f'''(r)}{3(f'(r))^2 f''(r)} = \lim_{u \to \infty} \frac{f'''(\phi(\sqrt{u}))}{3c(u)(1 + u)^{3/2}u}.$$

Therefore $f'''(\phi(\sqrt{u})) \asymp u^{5/2}$ when $u \to \infty$, and $F''(u) \asymp u^{-5/2}$. $\qquad\square$

With Lemma 1 we can estimate higher derivatives of $h$.

**Proposition 3.** *Let $(n, m) \in (\mathbb{R}^+)^2$ with $m \asymp M$. If $n/m^2 < u_{j_0}$ let $U$ be distance of $n/m^2$ to the closest $u_i$, say $u_j$. If $n/m^2 \geq u_{j_0}$ take $U = n/m^2$ and $j = \infty$. Then*

$$\frac{\partial^3 h}{\partial n^2 \partial m}(n, m) \asymp \frac{U^{d_j}}{M^4}.$$

6

*Proof.* By (5) the partial derivative is $m^{-4}F''(n/m^2)$ and the result follows from Lemma 1. $\qquad\square$

**Proposition 4.** *Let* $(n,m) \in \left(\mathbb{R}^+\right)^2$ *with* $m \asymp M$ *and fix* $j$ *with* $d_j > 0$. *If* $U = |n/m^2 - u_j|$ *is small enough, then*

$$\frac{\partial^3 h}{\partial n \partial m^2}(n,m) \asymp \frac{1}{M^3}.$$

*Proof.* The partial derivative here is $-2m^{-3}\big(F''(n/m^2)n/m^2 + F'(n/m^2)\big)$. By Lemma 1 the function $F'$ remains positive and bounded in bounded subintervals of $\mathbb{R}^+$, while $F''(n/m^2)n/m^2 \to 0$ when $U \to 0$. $\qquad\square$

**Proposition 5.** *Let* $(n,m) \in \left(\mathbb{R}^+\right)^2$ *with* $m \asymp M$ *and fix* $j$ *with* $d_j > 0$. *If* $U = |n/m^2 - u_j|$ *is small enough and* $1 \leq \ell \leq UM$,

$$\frac{\partial h}{\partial n}(n,m+\ell) - \frac{\partial h}{\partial n}(n,m) = C_j \frac{\ell}{m(m+\ell)} + O\left(\frac{\ell U^{d_j+1}}{M^2}\right)$$

*for some constant* $C_j \neq 0$.

*Proof.* We express the left hand side as

$$
\begin{aligned}
\int_0^\ell \frac{\partial^2 h}{\partial n \partial m}(n,m+t)\, dt &= \int_0^\ell F'\left(\frac{n}{(m+t)^2}\right) \frac{dt}{(m+t)^2} \\
&= \int_0^\ell \left[ F'(u_j) + \int_{u_j}^{n/(m+t)^2} F''(v)\, dv \right] \frac{dt}{(m+t)^2} \\
&= F'(u_j) \frac{\ell}{m(m+\ell)} + O\left(\frac{\ell U^{d_j+1}}{M^2}\right).
\end{aligned}
$$

To bound the error term we have applied Lemma 1 noting that $n/(m+t)^2 - u_j = O(U)$ for $0 \leq t \leq UM$. $\qquad\square$

## 5 The van der Corput estimate

In this section we generalize the argument in [3] to prove Theorem 2 in some ranges and Theorem 3 using the simplest van der Corput bound for $T$. The first named author wants to take the opportunity to point out that the case $L = M$ was neglected there (although it does not affect the result). To simplify the proofs, we will assume from now on that $UM \geq R^{3/8}$, as otherwise the trivial estimate $S \ll R^\epsilon UM^3$ suffices to prove the desired inequalities. We will also refer to the arguments of $S$ in the statements of these theorems as $U_1$ and $U_2$ for the sake of convenience.

**Proposition 6.** *Let* $R$, $M$, $U$, $U_1$ *and* $U_2$ *be as in the hypothesis of either Theorem 2 or 3, setting* $j = \infty$ *in the second case. Then*

$$\sum_n e\big(R(h(n,m+\ell) - h(n,m))\big)$$

$$\ll R^{1/2}\ell^{1/2}U^{(d_j+2)/2} + R^{-1/2}\ell^{-1/2}U^{-d_j/2}M^2,$$

*where the range of the summation is* $U_1(m+\ell)^2 \leq n < U_2 m^2$.

7

*Proof.* By the mean value theorem and Proposition 3 we have

$$\frac{\partial^2}{\partial n^2}\big(h(n, m+\ell) - h(n,m)\big) = \ell\frac{\partial^3 h}{\partial n^2 \partial m}(n, \tilde{m}) \asymp \ell\frac{U^{d_j}}{M^4}.$$

Applying now the simplest van der Corput well-known estimate (see, for instance, Theorem 2.2 of [6]),

$$\sum_n e\big(R(h(n,m+\ell) - h(n,m))\big) \ll UM^2\big(R\ell U^{d_j}M^{-4}\big)^{1/2} + \big(R\ell U^{d_j}M^{-4}\big)^{-1/2}.$$

This concludes the proof. $\qquad\square$

**Proposition 7.** *Theorem 2 holds when $d_j = 0$, 1, or when $d_j \geq 2$ and $U \gg R^{-5/(8d_j-8)}$.*

*Proof.* Note that since $U_2 \ll 1$ we are in position to apply Proposition 2 as long as we take $L \leq UM$. Using Proposition 6 to bound $T(U_1, U_2, M, L)$ we obtain

$$R^{-\epsilon}M^{-4}|S|^2 \ll L^{-1}U^2M^2 + R^{1/2}L^{1/2}U^{(d_j+4)/2} + R^{-1/2}L^{-1/2}U^{(2-d_j)/2}M^2. \tag{7}$$

We choose $L = \min\{R^{1/2}U^{-d_j}, UM\}$. If $L = R^{1/2}U^{-d_j}$ then using $M \leq R^{5/8}$ we obtain $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$, as desired. Hence assume $L = UM$ and $U^{d_j+1} < R^{1/2}M^{-1}$. We have

$$R^{-\epsilon}M^{-4}|S|^2 \ll UM + R^{1/2}U^{(d_j+5)/2}M^{1/2} + R^{-1/2}U^{(1-d_j)/2}M^{3/4}.$$

Using the inequality $U^{d_j+1} < R^{1/2}M^{-1}$ on the second summand and the hypothesis of this proposition we conclude again $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$. $\qquad\square$

*Proof of Theorem 3.* We proceed similarly as in the previous proof. Note that now $U_2 \asymp U$ and we may take $1 \leq L \leq M$ in Proposition 2. Using Proposition 6 to bound $T(U_1, U_2, M, L)$ we obtain exactly the same bound (7) with $d_\infty = -5/2$:

$$R^{-\epsilon}M^{-4}|S|^2 \ll L^{-1}U^2M^2 + R^{1/2}L^{1/2}U^{3/4} + R^{-1/2}L^{-1/2}U^{9/4}M^2.$$

The choice $L = \min\{R^{1/2}, M\}$ also works in exactly the same way, using $U \leq R^{5/4}M^{-2}$ and $M \leq R^{5/8}$, to show $M^{-4}|S|^2 \ll U^2R^{3/4+\epsilon}$. $\qquad\square$

# 6   Diophantine approximation of the phase

As $U$ gets smaller than $R^{-5/(8d_j-8)}$ the van der Corput estimate is not good enough to prove Theorem 2 anymore. The reason is that the phase of the exponential sum in (4) is almost linear in $n$, as Proposition 5 shows, and the oscillation is not captured by a second derivative test.

Throughout this section we will assume that $R$, $M$, $U$, $U_1$, $U_2$ and $j$ are as in the statement of Theorem 2, $UM \geq R^{5/8}$ (see comments in §5) and $M \leq m < 2M$. Let $I_{m,\ell} = [U_1(m+\ell)^2, U_2m^2]$, which we may assume non-empty by restricting the possible values of $m$, and define the quantities

$$\phi_\ell(n,m) = R\left(\frac{\partial h}{\partial n}(n, m+\ell) - \frac{\partial h}{\partial n}(n,m)\right),$$
$$\Phi_\ell(m) = \min_{x \in I_{m,\ell}} \text{dist}\big(\phi_\ell(x,m), \mathbb{Z}\big).$$

The function $\phi_\ell$ is the derivative of the phase of the exponential sum in $n$ appearing in (4). Since by Proposition 3 it is monotone in $n$, we can apply Kuzmin-Landau's lemma (Theorem 2.1 of [6]) to obtain the bound

$$\left| \sum_{n \in I_{m,\ell}} e\big(R(h(n, m+\ell) - h(n,m))\big) \right| \ll \big(\Phi_\ell(m)\big)^{-1}. \qquad (8)$$

Suppose we can find another bound $H_\ell$ for the same exponential sum, this time uniform in $m$, to apply in those cases when $\Phi_\ell \approx 0$. Then knowing very little about the distribution of the values $\Phi_\ell(m)$ we can find a good bound for $T$. The underlying idea here is to gain from some control of the spacing. In [1] and [11] this is accomplished via large sieve inequalities, while we introduce the spacing through the following simple result:

**Lemma 2.** *Assume we have a finite sequence of points $a_m \in [0, \frac{1}{2}]$ satisfying the following condition:*

$$\#\{m : a_m \leq x\} \leq A + Bx \qquad \text{for every} \quad 0 \leq x \leq 1/2.$$

*Then for any $H > 0$ we have*

$$\sum_m \min\{H, a_m^{-1}\} \leq AH + B(1 + |\log H|).$$

Assuming, in our setting, that $A_\ell$, $B_\ell$ and $H_\ell$ are functions such that $\ell A_\ell$, $\ell B_\ell$ and $\ell H_\ell$ are non-decreasing in $\ell$, and $H_\ell$ is bounded above and below by powers of $R$, it is immediate that for any fixed $\epsilon > 0$,

$$T(U_1, U_2, M, L) \ll R^\epsilon \big(A_L H_L + B_L\big). \qquad (9)$$

*Proof of Lemma 2.* Let us say that the finite sequence is $0 \leq a_1 \leq a_2 \leq \cdots \leq a_N \leq 1/2$. Note that, by hypothesis, $m \leq A + B a_m$. Let $f : [0, \frac{1}{2}] \to \mathbb{R}$ be a non-increasing function and extend it to the negative real numbers as the constant function $f(0)$. Then

$$\sum_m f(a_m) \leq \sum_m f\left(\frac{m-A}{B}\right) \leq B \int_{-A/B}^{1/2} f(x)\, dx = A f(0) + B \int_0^{1/2} f(x)\, dx.$$

The result follows applying this inequality with $f(x) = \min\{H, x^{-1}\}$. $\qquad \square$

The upper bound $H_\ell$ will be either the trivial estimate $UM^2$, or the second term in the van der Corput estimate given by Proposition 6 (the first one may be neglected in the range $U \ll R^{-5/(8d_j - 8)}$, $UM \geq R^{3/8}$). The pair $(A_\ell, B_\ell)$ will be given by one of the following two propositions.

**Proposition 8.** *Assume $U^{d_j + 1} M$ is small enough and $1 \leq \ell \leq UM$. Then*

$$\#\{m : \Phi_\ell(m) \leq x\} \ll 1 + \frac{R\ell}{M^2} + M\left(1 + \frac{M^2}{R\ell}\right)x \qquad \big(0 \leq x \leq 1/2\big).$$

*Proof.* Choose for each pair $(m, \ell)$ a point $x_m \in I_{m,\ell}$ (depending implicitly in $\ell$) satisfying

$$\Phi_\ell(m) = \text{dist}\big(\phi_\ell(x_m, m), \mathbb{Z}\big).$$

By the mean value theorem, $\phi_\ell(x_{m+1}, m+1) - \phi_\ell(x_m, m)$ equals

$$R\ell \frac{\partial^3 h}{\partial n \partial m^2}(x_1, y_1) + R\ell(x_{m+1} - x_m)\frac{\partial^3 h}{\partial n^2 \partial m}(x_2, y_2),$$

for some points $(x_1, y_1)$, $(x_2, y_2)$ lying in the rectangle

$$[U_1(m+\ell)^2, U_2(m+1)^2] \times [m, m+\ell+1].$$

The function $x/y^2$ over this rectangle satisfies

$$U_1(1 - 4M^{-1}) \leq x/y^2 \leq U_2(1 + 4M^{-1}),$$

and since $UM \geq R^{3/8}$ we have $|u_j - x_i/y_i^2| \asymp U$ for $i = 1, 2$. Using the estimates given by Propositions 3 and 4,

$$\phi_\ell(x_{m+1}, m+1) - \phi_\ell(x_m, m) \asymp \frac{R\ell}{M^3} + O\left(R\ell \cdot UM^2 \cdot \frac{U^{d_j}}{M^4}\right) \asymp \frac{R\ell}{M^3}, \quad (10)$$

the sign of the left hand side being always the same.

Since $M \leq m < 2M$, we deduce from (10) that the number of integers $k$ satisfying $|\phi_\ell(x_m, m) - k| \leq 1/2$ for some $m$ is at most a constant times $1 + R\ell M^{-2}$. On the other hand we deduce again from (10) that for each of those $k$ and any $x \geq 0$

$$\#\{m \,:\, |\phi_\ell(x_m, m) - k| \leq x\} \ll 1 + R^{-1}\ell^{-1}M^3 x.$$

Therefore,

$$\#\{m : \Phi_\ell(m) \leq x\} \ll \left(1 + \frac{R\ell}{M^2}\right)\left(1 + \frac{M^3}{R\ell}x\right)$$

for every $0 \leq x \leq 1/2$. $\qquad\square$

**Proposition 9.** *Fix $\epsilon > 0$. For any $0 < U \leq 1$ we have*

$$\#\{m \,:\, \Phi_\ell(m) \leq x\} \ll R^\epsilon\left(1 + R\ell U^{d_j+1} + M^2 x\right) \qquad (0 \leq x \leq 1/2).$$

*Proof.* Let $C_j$ the constant involved in Proposition 5, and assume that we have

$$\text{dist}\left(C_j \frac{R\ell}{m(m+\ell)}, \mathbb{Z}\right) \leq x \qquad \text{for some} \quad x \geq 0.$$

This means that there exists an integer $k = k(m, \ell)$ satisfying

$$|C_j R\ell - km(m+\ell)| \leq m(m+\ell)x \leq 2M^2 x.$$

In particular, $m$ must divide a certain integer $km(m+\ell)$ lying in the interval centered at $C_j R\ell$ of half-length $2M^2 x$. Since there are at most $1 + 4M^2 x$ of these integers, and each has at most $O(R^\epsilon)$ divisors, we conclude

$$\#\{m : \text{dist}\left(C_j R\ell/(m(m+\ell)), \mathbb{Z}\right) \leq x\} \ll R^\epsilon\left(1 + M^2 x\right).$$

Replacing $x$ by $x + O\left(R\ell U^{d_j+1}M^{-2}\right)$ the result follows from Proposition 5. $\qquad\square$

The following two propositions, together with Proposition 7 in §5, complete the proof of Theorem 2, and hence also the proof of Theorem 1.

**Proposition 10.** *If $U \ll R^{-5/(8d_j+8)}$ for a sufficiently small constant then Theorem 2 holds when $d_j \leq 4$, or when $d_j \geq 5$ and $U \gg R^{-5/(4d_j-16)}$.*

*Proof.* We apply Proposition 2 to bound $S$ with $L = R^{-3/4}U^2M^2$, which always lies in the interval $[1, UM]$. Using (9) with $(A_L, B_L)$ given by Proposition 8 (note the hypothesis imply $U^{d_j+1}M$ is small enough) we obtain

$$R^{-\epsilon}M^{-4}|S|^2 \ll R^{3/4} + \frac{UH_L}{M}\left(1 + \frac{RL}{M^2}\right) + U\left(1 + \frac{M^2}{RL}\right). \tag{11}$$

We choose either $H_L = UM^2$ or $H_L = R^{-1/8}U^{-(d_j+2)/2}M$ (second term in Proposition 6) depending on whether $RL/M^2 \leq 1$ or not. In the first case, the right hand side of (11) may be majored by $R^{3/4} + U^2M + R^{-1/4}U^{-1}$, and using $M \leq R^{5/8}$ and $U \geq R^{-1/4}$ (from $UM \geq R^{3/8}$) we conclude $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$. In the second case, the right hand side of (11) may be majored by $R^{3/4} + R^{1/8}U^{-(d_j-4)/2} + U$, which also leads to $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$ under the hypothesis of this proposition. $\square$

**Proposition 11.** *Theorem 2 holds when $U \ll R^{-5/(4d_j+24)}$.*

*Proof.* We proceed similarly as in the proof of Proposition 10. We apply Proposition 2 to bound $S$ with $L = R^{-3/4}U^2M^2$, and use (9) with $(A_L, B_L)$ given by Proposition 9 to obtain

$$R^{-\epsilon}M^{-4}|S|^2 \ll R^{3/4} + \frac{UH_L}{M}\left(1 + RLU^{d_j+1}\right) + UM. \tag{12}$$

We choose either $H_L = UM^2$ or $H_L = R^{-1/8}U^{-(d_j+2)/2}M$ depending on whether $RLU^{d_j+1} \leq 1$ or not. In the first case (12) shows that $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$ is satisfied trivially, while in the second case the right hand side of (12) may be majored by $R^{3/4} + R^{1/8}U^{(d_j+6)/2}M^2 + UM$, which also leads to $M^{-4}|S|^2 \ll R^{3/4+\epsilon}$ under the hypothesis of this proposition. $\square$

# Acknowledgements

# References

[1] E. Bombieri and H. Iwaniec. On the order of $\zeta(\frac{1}{2} + it)$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 13(3):449–472, 1986.

[2] J. Bourgain and N. Watt. Mean square of zeta function, circle problem and divisor problem revisited. arXiv:1709.04340 [math.AP], 2017.

[3] F. Chamizo. Lattice points in bodies of revolution. *Acta Arith.*, 85(3):265–277, 1998.

[4] F. Chamizo and C. Pastor. Lattice points in elliptic paraboloids. arXiv: 1611.04498 [math.NT], 2016.

[5] K. F. Gauss. De nexu inter multitudinem classium, in quas formae binariae secundi gradus distribuntur, earumque determinantem. In *Werke*, volume 2, pages 269–291. Georg Olms Verlag, Hildesheim, New York, 1981.

[6] S. W. Graham and G. Kolesnik. *van der Corput's method of exponential sums*, volume 126 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1991.

[7] J. Guo. On lattice points in large convex bodies. *Acta Arith.*, 151(1):83–108, 2012.

[8] 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.

[9] L.-K. Hua. The lattice-points in a circle. *Quart. J. Math., Oxford Ser.*, 13:18–29, 1942.

[10] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.

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

[12] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[13] G. Kolesnik. On the order of $\zeta(\frac{1}{2} + it)$ and $\Delta(R)$. *Pacific J. Math.*, 98(1):107–122, 1982.

[14] M. Spivak. *A comprehensive introduction to differential geometry. Vol. II.* Published by M. Spivak, Brandeis Univ., Waltham, Mass., 1970.