Mathematische Annalen (2023) 386:1237–1265
https://doi.org/10.1007/s00208-022-02404-w

Mathematische Annalen

On the squarefree values of $a^4 + b^3$

Gian Cordana Sanjaya[1] · Xiaoheng Wang[1]

Received: 21 July 2021 / Revised: 23 March 2022 / Accepted: 30 March 2022 /
Published online: 18 July 2022

Abstract
In this article, we prove that the density of integers a, b such that $a^4 + b^3$ is squarefree, when ordered by $\max\{|a|^{1/3}, |b|^{1/4}\}$, equals the conjectured product of the local densities. We show that the same is true for polynomials of the form $\beta a^4 + \alpha b^3$ for any fixed integers α and β. We give an exact count for the number of pairs (a, b) of integers with $\max\{|a|^{1/3}, |b|^{1/4}\} < X$ such that $\beta a^4 + \alpha b^3$ is squarefree, with a power-saving error term.

Mathematics Subject Classification 11N32 · 11N36 · 11N45

1 Introduction

A classical question in analytic number theory is to determine the probability that a given polynomial F with integer coefficients takes squarefree values when evaluated at random integers. The simplest case of one-variable and degree-one asks for the probability that a random integer is squarefree, which is well-known to be $6/\pi^2$. In general, one conjectures that the desired probability equals the product over all primes p of the probabilities that the values of F are not divisible by p^2.

The one-variable degree-two case can also be solved by elementary methods. The one-variable degree-three case was solved by Hooley [10]. For homogeneous polynomials of two variables, the question is known up to degree 6 due to Greaves [8]. For non-homogeneous polynomials of two variables that factor completely into a product of linear factors over some extension of $\mathbb{Q}$, the question is known also up to degree 6 due to Hooley [11]. Very recently, Kowalski [12] proved the case where F is a sum of at least 3 cubic polynomials in different variables. The cases when F is the discrimi-

✉ Xiaoheng Wang
x46wang@uwaterloo.ca

Gian Cordana Sanjaya
gcsanjaya@uwaterloo.ca

[1] Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada

nant of monic polynomials or when F is the discriminant of general polynomials were proven to equal the conjectured probability by Bhargava–Shankar–Wang [5, 6].

Conditional on the abc-conjecture, Granville [7] proved the one-variable case in general and a bound on the error term was later obtained by Murty–Pasten [13]. Also conditional on the abc-conjecture, Poonen [14] proved the multi-variable case where the variables are growing to infinity one by one. Unconditionally, very little is known otherwise. In most cases, it is even unknown whether the polynomial takes squarefree values infinitely often—the most famous example being $a^4 + 2$.

In this paper, we consider for the first time the polynomial $a^4 + b^3$. Our method in fact allows us to consider all polynomials of the form $\beta a^4 + \alpha b^3$ for any fixed integers α and β. We prove:

Theorem 1 *Let α and β be fixed nonzero integers such that $\gcd(\alpha, \beta)$ is squarefree. Let*

$$N(X; \alpha, \beta) = \#\{(a, b) \in \mathbb{Z}^2 : \max\{|a|^{1/3}, |b|^{1/4}\} < X, \beta a^4 + \alpha b^3 \text{ is squarefree}\}.$$

For any positive integer m, let $\rho_{\alpha,\beta}(m) = \#\{(a, b) \bmod m : m \mid \beta a^4 + \alpha b^3\}$ and let

$$C(\alpha, \beta) = \prod_p (1 - \rho_{\alpha,\beta}(p^2)p^{-4}).$$

Then

$$N(X; \alpha, \beta) = C(\alpha, \beta) \cdot 4X^7 + O_\epsilon(X^{6.992+\epsilon}).$$

The implied constant depends on α and β.

The case $\alpha = 256$ and $\beta = -27$ is of special importance since $256b^3 - 27a^4$ is the discriminant of the quartic polynomial $x^4 + ax + b$. An elementary calculation shows that $\rho_{256,-27}(p^2)$ equals p^3 for $p = 2, 3$; and equals $2p^2 - p$ for $p \geq 5$. Therefore, we have:

Theorem 2 *When pairs (a, b) of integers are ordered by $H(a, b) = \max\{|a|^{1/3}, |b|^{1/4}\}$, the density of quartic polynomials of the form $x^4 + ax + b$ having squarefree discriminant exists and is equal to*

$$\frac{1}{3} \prod_{p \geq 5} \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right)$$

which is approximately 28.03%.

It is also of interest to determine the density of irreducible quartic polynomials $f(x) = x^4 + ax + b$ such that $\mathbb{Z}[x]/(f(x))$ is the ring of integers of $\mathbb{Q}[x]/(f(x))$. This density is proved to be $\zeta(2)^{-1}$ for the case of general monic polynomials of any degree in [5]. It is then not surprising that the same density holds for the case of trinomial quartics.

Theorem 3 *When pairs (a, b) of integers are ordered by $H(a, b) = \max\{|a|^{1/3}, |b|^{1/4}\}$, the density of quartic polynomials $f(x)$ of the form $x^4 + ax + b$ that are irreducible and such that $\mathbb{Z}[x]/(f(x))$ is the ring of integers of $\mathbb{Q}[x]/(f(x))$ exists and is equal to $\zeta(2)^{-1}$.*

It is easy to see that the Euler product $C(\alpha, \beta)$ gives an upper bound for the desired density, if it exists, by applying the Chinese Remainder Theorem to more and more primes. As is standard in sieve theory, to demonstrate the lower bound, a "tail estimate" is required to show that there are not too many pairs (a, b) of integers such that $\beta a^4 + \alpha b^3$ is divisible by m^2 for some squarefree integer m. More precisely, we prove:

Theorem 4 *Let α and β be fixed nonzero integers such that $\gcd(\alpha, \beta)$ is squarefree. For any squarefree integer m, let*

$$N_m(X; \alpha, \beta) = \#\{(a, b) \in \mathbb{Z}^2 : |a| \le X^3, |b| \le X^4, m^2 \mid \beta a^4 + \alpha b^3\}.$$

Then for any positive real number M and $\epsilon > 0$,

$$\sum_{\substack{m > M \\ m \text{ squarefree}}} N_m(X; \alpha, \beta) = O_\epsilon\left(\frac{X^{7+\epsilon}}{\sqrt{M}}\right) + O_\epsilon(X^{6.992+\epsilon}) \tag{1}$$

The implied constants depend on α and β.

We note that since the exponents 3 and 4 are coprime, it is enough to prove Theorem 4 for one choice of α, β. Indeed, we have

$$-256 \cdot 27 \cdot \alpha^8 \beta^3 (\beta a^4 + \alpha b^3) = 256(-3\alpha^3 \beta b)^3 - 27(4\alpha^2 \beta a)^4,$$

which implies that,

$$N_m(X; \alpha, \beta) \le N_m(c_{\alpha,\beta} X;\ 256, -27)$$

for some constant $c_{\alpha,\beta}$ depending only on α, β. Hence the power saving bound (1) for $\alpha = 256$ and $\beta = -27$ implies it for all other α and β. We simplify notation by writing $\Delta(a, b)$ for $256b^3 - 27a^4$.

For any prime p and pair (a, b) of integers such that $p^2 \mid \Delta(a, b)$, we say p^2 *strongly divides* $\Delta(a, b)$ if $p^2 \mid \Delta(a', b')$ for any integers $a' \equiv a \pmod{p}$ and $b' \equiv b \pmod{p}$; otherwise, we say p^2 *weakly divides* $\Delta(a, b)$. Note in this case, for $p \ge 5$, p^2 strongly divides $\Delta(a, b)$ if and only if $p \mid a$ and $p \mid b$. For any squarefree integer m, let $\mathcal{W}_m^{(1)}$ (respectively $\mathcal{W}_m^{(2)}$) denote the set of pairs (a, b) of integers such that p^2 strongly divides (respectively weakly divides) $\Delta(a, b)$ for every prime $p \mid m$. Then we prove:

Theorem 5 *For any positive real number M and $\epsilon > 0$,*

$$\text{(a)} \quad \# \bigcup_{\substack{m>M \\ m \text{ squarefree}}} \{(a,b) \in \mathcal{W}_m^{(1)} : H(a,b) < X\} = O\left(\frac{X^7}{M}\right) + O\left(X^4\right); \tag{2}$$

$$\text{(b)} \quad \# \bigcup_{\substack{m>M \\ m \text{ squarefree}}} \{(a,b) \in \mathcal{W}_m^{(2)} : H(a,b) < X\} = O_\epsilon\left(X^{6.992+\epsilon}\right) + O_\epsilon\left(\frac{X^{7+\epsilon}}{M}\right), \tag{3}$$

where the implied constants are independent of M and X.

We now briefly describe our methods. Theorem 5(a) is immediate with the first term counting the contribution from $a \neq 0$ and the second term counting the contribution from $a = 0$. We devote the rest of the paper to proving Theorem 5(b). We follow the strategy of [5] to embed $\mathcal{W}_m^{(2)}$ into the space W of 4×4 symmetric matrices. More precisely, let A_0 denote the 4×4 matrix with 1's on the anti-diagonal and 0's elsewhere. The group $G = \mathrm{PSO}(A_0) = \mathrm{SO}(A_0)/\langle \pm I \rangle$ acts on W via the action $g \cdot B = gBg^t$ for $g \in G$ and $B \in W$. Define the *invariant polynomial* of an element $B \in W$ by

$$f_B(x) = \det(A_0 x - B).$$

Then f_B is a monic quartic polynomial. We extend the definition $H(a,b)$ to arbitrary monic quartic polynomials by

$$H(x^4 + c_1x^3 + c_2x^2 + c_3x + c_4) = \max\{|c_1|, |c_2|^{1/2}, |c_3|^{1/3}, |c_4|^{1/4}\}.$$

Define the discriminant and height of an element $B \in W$ by the discriminant and height of f_B, respectively. We then construct a map

$$\sigma_m : \mathcal{W}_m^{(2)} \to \frac{1}{4}W(\mathbb{Z})$$

with $f_{\sigma_m(a,b)} = x^4 + ax + b$ as in [5], where $\frac{1}{4}W(\mathbb{Z})$ is the lattice of elements B whose coefficients have denominators dividing 4. We note that the existence of the map σ_m (and the formula for $f_{\sigma_m(a,b)}$) is not immediate and is crucial to our method. It thus remains to count $G(\mathbb{Z})$-orbits in $\frac{1}{4}W(\mathbb{Z})$ that intersect the image of σ_m for some squarefree $m > M$, and have height bounded by X.

The space W has several subspaces: W_{00} consisting of $B \in W$ whose $(1,1)$- and $(1,2)$-entries are 0; W_{01} consisting of $B \in W$ whose $(1,1)$- and $(1,3)$-entries are 0; and W_0 consisting of $B \in W$ whose $(1,1)$-entry is 0. From our construction of σ_m in Section 2.2, we see that $\sigma_m(a,b)$ in fact lands in W_{00} for any $(a,b) \in \mathcal{W}_m^{(2)}$, and so is guaranteed to be *distinguished* in the sense of [9]. We obtain a bound of $O_\epsilon(X^{7+\epsilon}/M)$ for the distinguished cusps W_{00} and W_{01} and a bound of $O_\epsilon(X^{6+\epsilon})$ for the "thick" cusp $W_0 \backslash (W_{00} \cup W_{01})$.

The main novelty of this paper is on counting orbits of distinguished elements in the main body $W \backslash W_0$. We use the circle method to handle the condition that the invariant polynomials have vanishing x^2-coefficients, combined with the Selberg sieve to impose the distinguished condition to obtain the desired power saving.

We remark that Heath–Brown's result [15] on the density of integers n such that $n^d + c$ is k-free specializes to the case of squarefree values of the cubic polynomial $n^3 + c$ where c is a constant. A major observation of [15] is that counting triples (n, s, t) with $n^3 + c = s^2 t$ when n and s are large and c is fixed, is akin to counting points close to the projective curve $N^3 = S^2 T$. The bigger c is, which in our case can be as big as n^3, the worse the estimate gets. As such, we cannot patch the results of [15] together to prove Theorem 1.

This paper is organized as follows. In Sect. 2, we set up the embedding into W and collect some results on the invariant theory for the action of G on W, which allows us to reduce Theorem 5(b) to a result on counting $G(\mathbb{Z})$-orbits in $\frac{1}{4}W(\mathbb{Z})$. In Sect. 3, we apply Bhargava's averaging trick and count in the thick cusp and the distinguished cusps. In Sect. 4, we use the circle method and the Selberg sieve to count in the main body. Finally, in Sect. 5, we prove Theorem 1, Theorem 3 and Theorem 4.

2 Embedding into the space of 4 × 4 symmetric matrices

Let A_0 be the 4×4 matrix with 1's on the anti-diagonal and 0's elsewhere. The group $G = \mathrm{PSO}(A_0) = \mathrm{SO}(A_0)/\langle \pm I \rangle$ acts on the space W of symmetric 4×4 matrices via the action $g \cdot B = gBg^t$ for $g \in G$ and $B \in W$. The ring of polynomial invariants over $\mathbb{C}$ is freely generated by the coefficients of the invariant polynomial $f_B(x) = \det(A_0 x - B)$, which is a monic quartic polynomial. Define G-invariant discriminant $\Delta(B)$ and height $H(B)$ of an element $B \in W$ by $\Delta(B) = \Delta(f_B)$ and $H(B) = H(f_B)$. We recall some of the arithmetic invariant theory for this representation. See [5, 9] for more detail.

2.1 Invariant theory for the representation *W* of *G*

Let k be a field of characteristic not 2. For any monic quartic polynomial $f(x) \in k[x]$ such that $\Delta(f) \neq 0$, let C_f denote the smooth hyperelliptic curve $y^2 = f(x)$ of genus 1, let J_f denote its Jacobian (which is an elliptic curve), and let $J_f[2]$ denote the 2-torsion subgroup scheme of J_f. The stabilizer in $G(k)$ of an element $B \in W(k)$ with $f_B(x) = f(x)$ is naturally isomorphic to $J_f[2](k)$, which in turn is in bijection with the set of *even factorization* of $f(x)$ over k. An even factorization of $f(x)$ over k is an unordered pair $(g(x), h(x))$ of quadratic polynomials with $g(x)h(x) = f(x)$ such that either (i) g and h are both defined over k; or (ii) they are (defined and) conjugate over a quadratic extension of k.

An element $B \in W(k)$ or its $G(k)$-orbit is said to be: *k-soluble* if $\Delta(B) \neq 0$ and there exists a nonzero vector $v \in k^4$ such that

$$v^t A_0 v = 0 = v^t B v; \tag{4}$$

k-distinguished if $\Delta(B) \neq 0$ and there exist linearly independent vectors $v, w \in k^4$ such that

$$v^t A_0 v = v^t B v = w^t A_0 w = v^t A_0 w = v^t B w = 0. \tag{5}$$

Moreover, the set of k-lines $\mathrm{Span}(v)$ satisfying (4), if nonempty, is in bijection with $J_{f_B}(k)$; and the set of k-flags $\mathrm{Span}(v) \subset \mathrm{Span}(v, w)$ satisfying (5), if nonempty, is in bijection with $J_{f_B}[2](k)$. The set of k-soluble orbits with $f_B(x) = f(x)$ is in bijection with $J_f(k)/2J_f(k)$. The number of k-distinguished orbits with $f_B(x) = f(x)$ is 1 if $f(x)$ has a linear factor over k or if $f(x)$ admits a factorization of the form $g(x)h(x)$ where g and h are not rational over k but are conjugate over a quadratic extension of k; and is 2 otherwise.

Let W_{00} denote the subspace of W consisting of matrices B whose $(1, 1)$- and $(1, 2)$-entries are 0. Let W_{01} denote the subspace of W consisting of matrices B whose $(1, 1)$- and $(1, 3)$-entries are 0. Let W_0 denote the subspace of W consisting of matrices B whose $(1, 1)$-entry is 0. Let $\{e_1, e_2, e_3, e_4\}$ denote the standard basis for k^4. Then we see that the elements in $W_0(k)$ with nonzero discriminants are k-soluble with $v = e_1$ in (4); the elements in $W_{00}(k)$ with nonzero discriminants are k-distinguished with $v = e_1$ and $w = e_2$ in (5); and the elements in $W_{01}(k)$ with nonzero discriminants are k-distinguished with $v = e_1$ and $w = e_3$ in (5). A further polynomial invariant, called the Q-invariant, is defined on W_{00} in [5, Sect. 3.1]. For the case of 4×4 matrices B, this is simply the $(1, 3)$-entry b_{13}. The Q-invariant has the following important property:

Proposition 1 *Let $B \in W_{00}(\mathbb{Q})$ be an element whose invariant polynomial $f_B(x)$ has no even factorizations over $\mathbb{Q}$. If $B' \in W_{00}(\mathbb{Q})$ is any element that is $G(\mathbb{Z})$-equivalent to B, then the $(1, 3)$-entries of B' and B are equal up to sign. If $B' \in W_{01}(\mathbb{Q})$ is any element that is $G(\mathbb{Z})$-equivalent to B, then the $(1, 2)$-entry of B' equals the $(1, 3)$-entry of B up to sign.*

Proof We prove the statement for $B' \in W_{01}(\mathbb{Q})$. The statement for $W_{00}(\mathbb{Q})$ follows by a similar argument (see also [5, Proposition 3.1]).

Let γ_0 be the element of $\mathrm{SO}(A_0)(\mathbb{Z}[i])$ defined by

$$\gamma_0(e_1) = ie_1, \quad \gamma_0(e_2) = ie_2, \quad \gamma_0(e_3) = -ie_3, \quad \gamma_0(e_4) = -ie_4,$$

where $i = \sqrt{-1}$ is a root to $x^2 + 1 = 0$. Then any $\gamma \in \mathrm{PSO}(A_0)(\mathbb{Z})$ can either be lifted to some $\widetilde{\gamma} \in \mathrm{SO}(A_0)(\mathbb{Z})$ or to $\gamma_0\widetilde{\gamma} \in \mathrm{SO}(A_0)(\mathbb{Z}[i])$ for some $\widetilde{\gamma} \in \mathrm{SO}(A_0)(\mathbb{Z})$.

Suppose $B' = \gamma B \gamma^t$ for some $\gamma \in \mathrm{PSO}(A_0)(\mathbb{Z})$. Then either $B' = \widetilde{\gamma} B \widetilde{\gamma}^t$ or $B' = \gamma_0 \widetilde{\gamma} B \widetilde{\gamma}^t \gamma_0^t$ for some $\widetilde{\gamma} \in \mathrm{SO}(A_0)(\mathbb{Z})$. Since B' satisfies (5) with $v = e_1$ and $w = e_3$, we see that in either case, B satisfies (5) with $v = \widetilde{\gamma}^t e_1$ and $w = \widetilde{\gamma}^t e_3$. Since $B \in W_{00}(\mathbb{Q})$, we also see what B satisfies (5) with $v = e_1$ and $w = e_2$. The assumption that f_B has no even factorizations over $\mathbb{Q}$ then implies that $\mathrm{Span}_{\mathbb{Q}}(e_1) = \mathrm{Span}_{\mathbb{Q}}(\widetilde{\gamma}^t e_1)$ and $\mathrm{Span}_{\mathbb{Q}}(e_1, e_2) = \mathrm{Span}_{\mathbb{Q}}(\widetilde{\gamma}^t e_1, \widetilde{\gamma}^t e_3)$. Since $\widetilde{\gamma}^t$ is a matrix with integer entries, we see that there are integers $\alpha_1, \alpha_2, \alpha_3$ such that

$$\widetilde{\gamma}^t e_1 = \alpha_1 e_1,$$
$$\widetilde{\gamma}^t e_3 = \alpha_2 e_2 + \alpha_3 e_1.$$

Since $\widetilde{\gamma} \in \mathrm{SO}(A_0)(\mathbb{Z})$, we must then have

$$\widetilde{\gamma}^t e_4 = \alpha_1^{-1} e_4 - \alpha_3 \alpha_1^{-1} \alpha_2^{-1} e_3,$$
$$\widetilde{\gamma}^t e_2 = \alpha_2^{-1} e_3,$$

with $\alpha_1 = \pm 1$ and $\alpha_2 = \pm 1$. The $(1, 2)$-entry b'_{12} of B' is then either $(\widetilde{\gamma}^t e_1)^t B(\widetilde{\gamma}^t e_2)$ or $(i\widetilde{\gamma}^t e_1)^t B(i\widetilde{\gamma}^t e_2)$. In both cases, we have $b'_{12} = \pm\alpha_1\alpha_2^{-1} e_1^t B e_3 = \pm b_{13}$. □

Let $U \simeq \mathbb{A}^2 \backslash \{\Delta = 0\}$ be the space of monic quartic polynomials of the form $x^4 + ax + b$ with nonzero discriminant. Note if $f \in U(\mathbb{Z})$ has an even factorization over $\mathbb{Q}$, then it is either reducible over $\mathbb{Q}$ or factors as $g(x)h(x)$ where g and h are conjugate over some quadratic extension of $\mathbb{Q}$. The next result then shows that the number of elements of $U(\mathbb{Z})$ failing the condition of Proposition 1 is negligible.

Proposition 2 *The number of elements $f \in U(\mathbb{Z})$ with $H(f) < X$ such that $f(x)$ is either reducible over $\mathbb{Q}$ or factors as $g(x)h(x)$ where g and h are conjugate over some quadratic extension of $\mathbb{Q}$ is $O(X^4 \log X)$.*

Proof Throughout this proof, we use repeatedly the classical result that the sum $\sum_{|n|<X} d(n)$ of the divisor function is $O(X \log X)$ and that the sum $\sum_{|n|<X} \tau_3(n)$ of the triple-divisor function is $O(X \log^2 X)$. See for example [1, Sect. 3.5].

Suppose first $f(x) = x^4 + ax + b$ has a linear factor $x - r$ over $\mathbb{Q}$. When $b = 0$, one can choose a freely. When $b \neq 0$, then since $r \mid b$, we get $O(X^4 \log X)$ choices for the pair (r, b), which then uniquely determines a since $a = -r^3 - b/r$. Hence, there are $O(X^4 \log X)$ such $f(x)$ with a linear factor.

Next we consider the case where $f(x) = x^4 + ax + b$ does not have a linear factor but factors as $(x^2 + cx + d)(x^2 - cx + e)$ over $\mathbb{Q}$. Since $f(x)$ does not have a linear factor, we see that $b \neq 0$. Then from $de = b$, we get $O(X^4 \log X)$ choices for the triple (d, e, b). Comparing the x^2-coefficients gives $c^2 = d + e$, and so c is determined given d and e. Comparing the x-coefficients then uniquely determines a. Hence, there are $O(X^4 \log X)$ such $f(x)$ that factors as a product of two irreducible quadratic polynomials.

Finally, we consider the case where $f(x) = x^4 + ax + b$ is irreducible over $\mathbb{Q}$ but factors as

$$\left(x^2 + e_1\sqrt{d}x + \frac{c_2 + e_2\sqrt{d}}{2}\right)\left(x^2 - e_1\sqrt{d}x + \frac{c_2 - e_2\sqrt{d}}{2}\right)$$

over the ring of integers in $\mathbb{Q}(\sqrt{d})$ for some d. If $a = 0$, then we have $O(X^4)$ choices for b. Suppose now $a \neq 0$. Comparing the x-coefficients gives $e_1 e_2 d = a$. Hence there are $O(X^3 \log^2 X)$ choices for the tuple (e_1, e_2, d, a). Comparing the x^2-coefficients gives $c_2 - e_1^2 d = 0$, and so c_2 is determined given e_1 and d. Comparing the constant terms then uniquely determines b. Hence, there are $O(X^4)$ such $f(x)$ that factors into conjugate quadratic polynomials over some quadratic extension of $\mathbb{Q}$. □

We end this section with a bound on distinguished elements over finite fields, which will be used in the Selberg sieve in Sect. 4.

Proposition 3 *Let $p \geq 7$ be a prime. Then the number d_p of elements $B \in W(\mathbb{F}_p)$ with $f_B \in U(\mathbb{F}_p)$ and is not $\mathbb{F}_p$-distinguished satisfies*

$$\frac{1}{16}p^8 + O(p^7) \leq d_p \leq \frac{3}{4}p^8 + O(p^7).$$

Proof Over the finite field $\mathbb{F}_p$, every orbit with nonzero discriminant is $\mathbb{F}_p$-soluble. Moreover, for any monic quartic polynomial $f(x) \in \mathbb{F}_p[x]$ with nonzero discriminant, the number $\#J_f(\mathbb{F}_p)/2J_f(\mathbb{F}_p)$ of $\mathbb{F}_p$-orbits with invariant polynomial f equals the size $\#J_f[2](\mathbb{F}_p)$ of any stabilizer with invariant polynomial f. Hence, the number of $B \in W(\mathbb{F}_p)$ with $f_B = f$ equals $\#G(\mathbb{F}_p) = p^2(p^2-1)^2$. There are $p^2 + O(p)$ polynomials $f \in U(\mathbb{F}_p)$ and so a total of $p^8 + O(p^7)$ elements $B \in W(\mathbb{F}_p)$ with $f_B \in U(\mathbb{F}_p)$. Moreover, for any $f \in U(\mathbb{F}_p)$, there is at least one $\mathbb{F}_p$-distinguished orbit with stabilizer having size at most 4. Hence, we have the upper bound $d_p \leq \frac{3}{4}p^8 + O(p^7)$.

Consider next quartic polynomials of the form

$$g_{a,b}(x) := (x-a)(x-b)(x^2 + (a+b)x + (a^2+ab+b^2)) \in U(\mathbb{F}_p).$$

Since $g_{a,b}(x)$ has a linear factor, there is only one distinguished orbit with invariant $g_{a,b}$. Moreover, we have $2 \leq \#J_{g_{a,b}}[2](\mathbb{F}_p) \leq 4$. Hence, there is at least one non-distinguished orbit of size at least $|G(\mathbb{F}_p)|/4$. It remains to count the number of such $g_{a,b}(x)$ with nonzero discriminant, which is equivalent to requiring that $a \neq b$, that a is not a root of the quadratic factor, and that the quadratic factor has nonzero discriminant. In other words, we have $a \neq b$, $3(a+b/3)^2 + (2/3)b^2 \neq 0$ and $3(a+b/3)^2 + (8/3)b^2 \neq 0$. Given any b, there are at least $p-5$ choices for a. Finally, given any $g_{a,b}$ with nonzero discriminant, we see that $g_{a,b} = g_{a',b'}$ if and only if $x^2+(a+b)x+(a^2+ab+b^2) = (x-a')(x-b')$ or $(x-a)(x-b) = (x-a')(x-b')$, as any other possibility contradicts $\Delta(g_{a,b}) \neq 0$. Hence, there are at least $p(p-5)/4$ quartic polynomials with nonzero discriminant of the form $g_{a,b}$ for some $a, b \in \mathbb{F}_p$. Therefore, we have at least $\frac{1}{16}p^8 + O(p^7)$ non-distinguished elements B in $W(\mathbb{F}_p)$ with $f_B \in U(\mathbb{F}_p)$. □

2.2 Embedding $\mathcal{W}_m^{(2)}$ into $\frac{1}{4}W(\mathbb{Z})$

In light of Proposition 2, it is sufficient to prove Theorem 5 with $\mathcal{W}_m^{(2)}$ replaced by the set of pairs $(a,b) \in \mathcal{W}_m^{(2)}$ such that $f_{a,b}(x) := x^4 + ax + b$ is irreducible and does not factor into a product of quadratic polynomials conjugate over some quadratic extension of $\mathbb{Q}$. We prove some preliminary results in order to use the map σ_m defined in [5, Sect. 3.2].

Fix $(a,b) \in \mathcal{W}_m^{(2)}$ and fix any prime $p \mid m$. For any $(a',b') \in \mathbb{Z}^2$ with $a' \equiv a \pmod{p}$ and $b' \equiv b \pmod{p}$, we have $f_{a',b'}(x) \equiv f_{a,b}(x) \pmod{p}$. Since p^2 weakly divides $\Delta(a,b)$, we see that $f_{a,b}(x)$ has a unique double root mod p. Let $r \in \mathbb{Z}$ be an integer such that $f_{a,b}(x+r) = x^4 + b_1x^3 + b_2x^2 + b_3x + b_4$ with $p \mid b_3$ and $p \mid b_4$. We claim that $p^2 \mid b_4$. Note the discriminant of a quartic polynomial is of

the form

$$\Delta(x^4 + b_1x^3 + b_2x^2 + b_3x + b_4) = b_4\Delta'(b_1, b_2, b_3, b_4) + b_3^2\Delta(x^3 + b_1x^2 + b_2x + b_3)$$

where Δ' is some polynomial with integer coefficients. Suppose for a contradiction that $p^2 \nmid b_4$. Then, since $\Delta(f_{a,b}(x + r)) = \Delta(f_{a,b}(x)) = \Delta(a, b)$. we have $p^2 \mid \Delta(f_{a,b}(x + r))$ and so $p \mid \Delta'(b_1, b_2, b_3, b_4)$. Hence, $p^2 \mid \Delta(g(x))$ for any monic quartic polynomial $g(x)$ congruent to $f_{a,b}(x + r)$. Now for any $(a', b') \in \mathbb{Z}^2$ with $a' \equiv a \pmod{p}$ and $b' \equiv b \pmod{p}$, we have $f_{a',b'}(x+r) \equiv f_{a,b}(x+r) \pmod{p}$ and so $p^2 \mid \Delta(f_{a',b'}(x + r))$. Since $\Delta(f_{a',b'}(x + r)) = \Delta(a', b')$, this contradicts the assumption that p^2 weakly divides $\Delta(a, b)$.

By the Chinese Remainder Theorem, there exists an integer r such that $f_{a,b}(x+r) = x^4+c_1x^3+c_2x^2+mc_3x+m^2c_4$ for some integers c_1, c_2, c_3, c_4. Consider the following matrix:

$$B(c_1, c_2, c_3, c_4) = \begin{pmatrix} 0 & 0 & m & 0 \\ 0 & 1 & -c_1/2 & 0 \\ m & -c_1/2 & c_1^2/4 - c_2 & -c_3/2 \\ 0 & 0 & -c_3/2 & -c_4 \end{pmatrix}.$$

A direct computation shows that $f_{B(c_1,c_2,c_3,c_4)}(x) = x^4 + c_1x^3 + c_2x^2 + mc_3x + m^2c_4$. We now set $\sigma_m(a, b) = B(c_1, c_2, c_3, c_4) + rA_0 \in \frac{1}{4}W(\mathbb{Z})$. Then

$$f_{\sigma_m(a,b)}(x) = \det(xA_0 - (B(c_1, c_2, c_3, c_4) + rA_0)) = f_{B(c_1,c_2,c_3,c_4)}(x - r) = f_{a,b}(x).$$

Note in fact that the image of σ_m lies inside $W_{00}(\mathbb{Q})$ and the $(1, 3)$-entry of any element in the image of σ_m is m. We combine Proposition 1 and the above in the following theorem.

Theorem 6 *Let m be any squarefree integer. There is a map $\sigma_m : \mathcal{W}_m^{(2)} \to \frac{1}{4}W(\mathbb{Z})$ such that the following two conditions are satisfied:*

(a) $f_{\sigma_m(a,b)} = f_{a,b}(x)$ for any $(a, b) \in \mathcal{W}_m^{(2)}$;
(b) the $(1, 3)$-entry (respectively the $(1, 2)$-entry) of any element in $W_{00}(\mathbb{Q})$ (respectively $W_{01}(\mathbb{Q})$) that is $G(\mathbb{Z})$-equivalent to some element in $\sigma_m(\mathcal{W}_m^{(2)})$ equals m in absolute value.

3 Averaging and counting in the cusp

Fix any positive real number M. Let $\mathcal{L}_M$ denote the set of elements in $\frac{1}{4}W(\mathbb{Z})$ that are $G(\mathbb{Z})$-equivalent to some elements in $\sigma_m(\mathcal{W}_m^{(2)})$ for some squarefree integer $m > M$. Write $N(\mathcal{L}_M, X)$ for the number of $G(\mathbb{Z})$-orbits in $\mathcal{L}_M$ having height at most X. Since $G(\mathbb{Z})$-equivalent elements have the same invariant polynomials, we see by Theorem

6(a) that

$$N(\mathcal{L}_M, X) \geq \# \bigcup_{\substack{m > M \\ m \text{ squarefree}}} \{f \in \mathcal{W}_m^{(2)} : H(f) < X\}.$$

Therefore, Theorem 5(b) follows from the following result.

Theorem 7 *For any positive real number M and any $\epsilon > 0$, we have*

$$N(\mathcal{L}_M, X) = O_\epsilon\left(X^{6.992+\epsilon}\right) + O_\epsilon\left(\frac{X^{7+\epsilon}}{M}\right). \tag{6}$$

In Sect. 3.1, we recall the set up in [9] for counting $G(\mathbb{Z})$-orbits in $\frac{1}{4}W(\mathbb{Z})$ and divide up a fundamental domain $\mathcal{F}$ for the left-multiplication action of $G(\mathbb{Z})$ on $G(\mathbb{R})$ into the main body, the thick cusp, and the distinguished cusps. In Sect. 3.2, we obtain bounds for the contribution from the thick cusp and the distinguished cusps. Finally in Sect. 4, we obtain bounds for the contribution from the main body and complete the proof of Theorem 7.

3.1 Counting $G(\mathbb{Z})$-orbits in $\frac{1}{4}W(\mathbb{Z})$

The counting problem for the representation W of G is studied in [9]. In this section, we recall some of the set up and results of [9].

Let R be a fundamental domain for the action of $G(\mathbb{R})$ on the elements of $W(\mathbb{R})$ having nonzero discriminant and height bounded by 1 as constructed in [9, Sect. 4.1]. Let $\mathcal{F}$ be a fundamental set for the left-multiplication action of $G(\mathbb{Z})$ on $G(\mathbb{R})$ obtained using the Iwasawa decomposition of $G(\mathbb{R})$. More explicitly, we have

$$G(\mathbb{R}) = N(\mathbb{R})TK,$$

where N is a unipotent group consisting of lower triangular matrices, K is compact, and T is the split torus of G given by

$$T = \left\{ \begin{pmatrix} t_1^{-1} & & & \\ & t_2^{-1} & & \\ & & t_2 & \\ & & & t_1 \end{pmatrix} \right\}.$$

We also make the following change of variables: set

$$s_1 = t_1/t_2, \quad s_2 = t_1 t_2.$$

We denote an element of T with coordinates t_i (resp. s_i) by (t) (resp. (s)). We may take $\mathcal{F}$ to be contained in a *Siegel set*, i.e., contained in $N'T'K$, where N' consists

of elements in $N(\mathbb{R})$ whose entries are absolutely bounded and $T' \subset T$ consists of elements in $(s) \in T$ with $s_1 \geq c$ and $s_2 \geq c$ for some positive constant c.

For any $h \in G(\mathbb{R})$, since $\mathcal{F}h$ remains a fundamental domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$, the set $(\mathcal{F}h) \cdot (XR)$ (when viewed as a multiset) is a finite cover of a fundamental domain for the action of $G(\mathbb{Z})$ on the elements in $W(\mathbb{R})$ with nonzero discriminant and height bounded by X. The degree of the cover depends only on the size of stabilizer in $G(\mathbb{R})$ and is thus absolutely bounded by 4. The presence of these stabilizers is in fact the reason we consider $(\mathcal{F}h) \cdot (XR)$ as a multiset. Hence, we have

$$N(\mathcal{L}_M, X) \ll \#\big\{\big((\mathcal{F}h) \cdot (XR)\big) \cap \mathcal{L}_M\big\}. \tag{7}$$

Let $\mathcal{G}_1$ be a compact left K-invariant set in $G(\mathbb{R})$ which is the closure of a nonempty open set. Averaging (7) over $h \in \mathcal{G}_1$ and exchanging the order of integration as in [4, Theorem 2.5], we obtain

$$N(\mathcal{L}_M, X) \ll \int_{\gamma \in \mathcal{F}} \#\big\{\big((\gamma\mathcal{G}_1) \cdot (XR)\big) \cap \mathcal{L}_M\big\} d\gamma, \tag{8}$$

where the implied constant depends only on $\mathcal{G}_1$ and R, and where $d\gamma$ is a Haar measure on $G(\mathbb{R})$ given by

$$d\gamma = dn\, s_1^{-1} s_2^{-1} d^\times s\, dk,$$

where dn is a Haar measure on the unipotent group $N(\mathbb{R})$, dk is a Haar measure on the compact group K, and $d^\times s = s_1^{-1} ds_1\, s_2^{-1} ds_2$ is the standard Haar measure on $\mathbb{G}_m^2$ (see [9, (20)]).

Since $s_i \geq c$ for every i, there exists a compact subset N'' of $N(\mathbb{R})$ containing $(t)^{-1} N'(t)$ for all $t \in T'$. Since N'', K, $\mathcal{G}_1$ are compact and R is bounded, the set $E = N'' K \mathcal{G}_1 R$ is bounded. Then we have

$$N(\mathcal{L}_M, X) \ll \int_{s_i \gg 1} \#\big\{\big((s) \cdot XE\big) \cap \mathcal{L}_M\big\} s_1^{-1} s_2^{-1} d^\times s. \tag{9}$$

The (i, j)-entry of any $B \in XE$ is bounded by $c_0 X$, where $c_0 > 0$ is a constant depending only on $\mathcal{G}_1$ and R. The action of the torus T then scales each entry of B. We denote the coordinates of W by b_{ij} for $1 \leq i \leq j \leq 4$ and define

$$\begin{aligned} w(b_{11}) = s_1^{-1} s_2^{-1},\ w(b_{12}) = s_2^{-1},\quad & w(b_{13}) = s_1^{-1},\quad && w(b_{14}) = 1, \\ w(b_{22}) = s_1 s_2^{-1},\ w(b_{23}) = 1,\quad & w(b_{24}) = s_1, \\ & w(b_{33}) = s_1^{-1} s_2,\ && w(b_{34}) = s_2, \\ & && w(b_{44}) = s_1 s_2. \end{aligned}$$

Then the (i, j)-entry of any $B \in (s) \cdot XE$ is bounded by $c_0 X w(b_{ij})$.

We define two distinguished cusps: $T_{00} \subset T'$ consisting of elements (s) such that $c_0 X w(b_{11}) < 1/4$ and $c_0 X w(b_{12}) < 1/4$; and $T_{01} \subset T'$ consisting of elements (s) such that $c_0 X w(b_{11}) < 1/4$ and $c_0 X w(b_{13}) < 1/4$. We define the thick cusp T_0 to be

the subset of T' consisting of elements (s) such that $c_0 X w(b_{11}) < 1/4$, $c_0 X w(b_{12}) \geq 1/4$, and $c_0 X w(b_{13}) \geq 1/4$. We define the main body T'' to be the complement $T' \backslash (T_{00} \cup T_{01} \cup T_0)$. Then for any $(s) \in T_{00}$, we have $\big((s) \cdot XE\big) \cap \frac{1}{4} W(\mathbb{Z}) \subset W_{00}(\mathbb{Q})$; for any $(s) \in T_{01}$, we have $\big((s) \cdot XE\big) \cap \frac{1}{4} W(\mathbb{Z}) \subset W_{01}(\mathbb{Q})$; and for any $(s) \in T_0$. we have $\big((s) \cdot XE\big) \cap \frac{1}{4} W(\mathbb{Z}) \subset W_0(\mathbb{Q})$.

Since the invariant polynomials of elements in $\mathcal{L}_M$ have the form $x^4 + ax + b$, we now express the conditions of the invariant polynomial having vanishing x^3- and x^2-coefficients in terms of the coordinates b_{ij}. The x^3-coefficient is the anti-trace, and so we have

$$b_{23} = -b_{14}.$$

After replacing b_{23} by $-b_{14}$, we see that the x^2-coefficient is the following quadratic form:

$$q(b_{ij}) := -b_{11}b_{44} - b_{22}b_{33} - 2b_{12}b_{34} - 2b_{13}b_{24} - 2b_{14}^2.$$

3.2 Counting in the cusps

In this section, we compute the contribution to (9) for $(s) \in T_{00}$, $(s) \in T_{01}$ and for $(s) \in T_0$.

Proposition 4 *For any positive real number M and $\epsilon > 0$, we have*

$$\int_{(s)\in T_{00}} \#\big\{\big((s) \cdot XE\big) \cap \mathcal{L}_M\big\}\, s_1^{-1} s_2^{-1} d^\times s = O_\epsilon\Big(\frac{X^{7+\epsilon}}{M}\Big), \tag{10}$$

$$\int_{(s)\in T_{01}} \#\big\{\big((s) \cdot XE\big) \cap \mathcal{L}_M\big\}\, s_1^{-1} s_2^{-1} d^\times s = O_\epsilon\Big(\frac{X^{7+\epsilon}}{M}\Big), \tag{11}$$

$$\int_{(s)\in T_0} \#\big\{\big((s) \cdot XE\big) \cap \mathcal{L}_M\big\}\, s_1^{-1} s_2^{-1} d^\times s = O_\epsilon\Big(X^{6+\epsilon}\Big). \tag{12}$$

Proof Consider first the distinguished cusp T_{00}. In this case, any element in $\big((s) \cdot XE\big) \cap \mathcal{L}_M$ is an element in $W_{00}(\mathbb{Q})$ that is $G(\mathbb{Z})$-equivalent to some element in $\sigma_m(\mathcal{W}_m^{(2)})$ for some $m > M$. Hence, by Theorem 6, we have $|b_{13}| > M$ for any element $B \in \big((s) \cdot XE\big) \cap \mathcal{L}_M$. In other words, we have $X s_1^{-1} \gg M$. Moreover, note that if $b_{11} = b_{12} = b_{22} = 0$, then $\det(xA_0 - B) = ((x - b_{14})(x - b_{23}) - b_{13}b_{24})^2$ which implies that $\Delta(B) = 0$. Hence we may assume that $X s_1 s_2^{-1} \gg 1$. Let T'_{00} denote the subset of T_{00} consisting of elements (s) with $X s_1^{-1} \gg M$ and $X s_1 s_2^{-1} \gg 1$. Note we also have $s_1 \ll X$ and $s_2 \ll X^2$ for $(s) \in T'_{00}$.

The quadratic form $q(b_{ij})$ when restricted to W_{00} simplifies to $q_1(b_{ij}) = -b_{22}b_{33} - 2b_{13}b_{24} - 2b_{14}^2$. Hence, we have

$$\begin{aligned}\#\big((s) \cdot XE \cap \mathcal{L}_M\big) \ll_\epsilon \big((Xw(b_{14}))^{1+\epsilon}(Xw(b_{22}) + Xw(b_{33})) \\ + (Xw(b_{14})Xw(b_{13})Xw(b_{24}))^{1+\epsilon}\big) Xw(b_{34})Xw(b_{44})\end{aligned}$$

$$\begin{aligned} &\ll X^{4+\epsilon} s_1^2 s_2 + X^{4+\epsilon} s_2^3 + X^{5+\epsilon} s_1 s_2^2 \\ &\ll X^{4+\epsilon} s_1^2 s_2 + X^{5+\epsilon} s_1 s_2^2. \end{aligned}$$

Integrating these two terms separately gives

$$\begin{aligned} \int_{(s)\in T'_{00}} X^{4+\epsilon} s_1^2 s_2 s_1^{-1} s_2^{-1} d^\times s &= \int_{(s)\in T'_{00}} X^{4+\epsilon} s_1 d^\times s \ \ll \ \frac{X^{5+\epsilon}\log X}{M}, \\ \int_{(s)\in T'_{00}} X^{5+\epsilon} s_1 s_2^2 s_1^{-1} s_2^{-1} d^\times s &= \int_{(s)\in T'_{00}} X^{5+\epsilon} s_2 d^\times s \\ &\ll \int_{(s)\in T'_{00}} X^{6+\epsilon} s_1 d^\times s \ll \frac{X^{7+\epsilon}\log X}{M}. \end{aligned}$$

The integral over the other distinguished cusp T_{01} has the same bound via the same analysis with s_1 and s_2 switched.

Finally, we consider the thick cusp T_0. In this case, we have $Xs_1^{-1} \gg 1$ and $Xs_2^{-1} \gg 1$. The quadratic form $q(b_{ij})$ when restricted to W_0 simplifies to $q_2(b_{ij}) = -b_{22}b_{33} - 2b_{12}b_{34} - 2b_{13}b_{24} - 2b_{14}^2$. The above analysis shows that the number of choices for $(b_{22}, b_{33}, b_{13}, b_{24}, b_{14})$ such that $q_1(b_{ij}) = 0$ is $O_\epsilon(X^{2+\epsilon}s_1 s_2^{-1} + X^{2+\epsilon}s_1^{-1}s_2 + X^{3+\epsilon})$. Multiplying it by $(Xw(b_{12}) + Xw(b_{34}))Xw(b_{44})$ gives a bound of $O_\epsilon(X^{4+\epsilon}s_1^2 s_2 + X^{4+\epsilon}s_2^3 + X^{5+\epsilon}s_1 s_2^2)$ for the number of $B \in \big((s)\cdot XE\big)\cap \mathcal{L}_M$ with $q_1(b_{ij}) = 0$. The contribution from $q_1(b_{ij}) \neq 0$ is

$$O_\epsilon\Big((Xw(b_{22})Xw(b_{33})Xw(b_{13})Xw(b_{24})Xw(b_{14}))^{1+\epsilon}Xw(b_{44})\Big) = O_\epsilon(X^{6+\epsilon}s_1 s_2).$$

Using the bound $s_1 \ll X$ and $s_2 \ll X$, we have

$$\#\big\{\big((s)\cdot XE\big)\cap \mathcal{L}_M\big\} \ll_\epsilon X^{4+\epsilon}s_1^2 s_2 + X^{4+\epsilon}s_2^3 + X^{5+\epsilon}s_1 s_2^2 + X^{6+\epsilon}s_1 s_2 \ll X^{6+\epsilon}s_1 s_2.$$

Multiplying by $s_1^{-1}s_2^{-1}$ and integrating then give the desired bound (12). □

For the main body T'', we have $Xs_1^{-1}s_2^{-1} \gg 1$. Since both s_1 and s_2 are bounded below by some absolute constant, we still have the bound $s_1 \ll X$ and $s_2 \ll X$. The above analysis gives a bound of

$$\begin{aligned} &O_\epsilon\Big(\big((X^{2+\epsilon}s_1 s_2^{-1} + X^{2+\epsilon}s_1^{-1}s_2 + X^{3+\epsilon})(Xw(b_{12}) + Xw(b_{34})) \\ &\quad + X^{5+\epsilon}\big)(Xw(b_{11}) + Xw(b_{44}))\Big) \\ &= O_\epsilon\Big((X^{3+\epsilon}s_1^{-1}s_2^2 + X^{4+\epsilon}s_2 + X^{5+\epsilon})Xs_1 s_2\Big) \\ &= O_\epsilon\Big(X^{6+\epsilon}s_1 s_2\Big) \end{aligned}$$

for the number of $B \in \big((s)\cdot XE\big)\cap \mathcal{L}_M$ with $q_2(b_{ij}) = 0$. Multiplying by $s_1^{-1}s_2^{-1}$ and integrating give a bound of $O_\epsilon(X^{6+\epsilon})$.

It remains to consider the contribution to the main body integral from the number of $B \in \big((s) \cdot XE\big) \cap \mathcal{L}_M$ with $q_2(b_{ij}) \neq 0$. We have a trivial bound of $O_\epsilon(X^{7+\epsilon})$ for the number of such B. For any positive real number δ, let T''_δ denote the subset of T'' where $s_1 \gg X^\delta$ or $s_2 \gg X^\delta$. Then we have

$$\int_{(s)\in T''_\delta} \#\big\{\big((s)\cdot XE\big)\cap \mathcal{L}_M\big\}\, s_1^{-1}s_2^{-1}d^\times s = O_\epsilon(X^{7-\delta+\epsilon}). \tag{13}$$

Therefore, it remains to consider the main body integral under the additional assumption that $s_1 \ll X^\delta$ and $s_2 \ll X^\delta$ where δ is some small enough positive real number.

4 Counting in the main body using the circle method

In this section, we consider the contribution to (9) from the main body under the additional assumption that $s_1 \ll X^\delta$ and $s_2 \ll X^\delta$. Let $V \simeq \mathbb{A}^9$ denote the subspace of W cut out by $b_{14} = -b_{23}$. For any $B \in V(\mathbb{Q})$, let

$$q(B) = -b_{11}b_{44} - b_{22}b_{33} - 2b_{12}b_{34} - 2b_{13}b_{24} - 2b_{14}^2.$$

Since scaling an element $B \in V(\mathbb{Q})$ by 4 does not affect the vanishing of $q(B)$ or whether it is $\mathbb{Q}$-distinguished, it is enough to count points in a box in $V(\mathbb{Z})$ defined by $|b_{ij}| \leq 4c_0 X w(b_{ij})$. The assumption on s_1 and s_2 implies that $4c_0 X w(b_{ij}) = O(X^{1+2\delta})$ for all i, j. The goal of this section is to prove the following theorem:

Theorem 8 *Let $\delta < 0.01$ be a positive real number. Let $\mathcal{B}$ be a box in $V(\mathbb{R})$ defined by $|b_{ij}| \leq X_{ij}$ for $(i,j) = (1,1), (1,2), (1,3), (1,4), (2,2), (2,4), (3,3), (3,4), (4,4)$ where X_{ij} are real numbers satisfying $c_1^{-1}X^{1-2\delta} \leq X_{ij} \leq c_1 X^{1+2\delta}$, $X_{14} = c_2 X$ and*

$$X_{11}X_{44} = X_{22}X_{33} = X_{12}X_{34} = X_{13}X_{24} = c_2^2 X^2,$$

for some positive constants c_1, c_2. Let $N_q^{\mathrm{dist}}(\mathcal{B})$ denote the number of $\mathbb{Q}$-distinguished elements $B \in \mathcal{B} \cap V(\mathbb{Z})$ with $q(B) = 0$. Then

$$N_q^{\mathrm{dist}}(\mathcal{B}) = O_\epsilon\left(X^{\frac{209}{30}+\frac{137}{45}\delta+\epsilon}\right). \tag{14}$$

Multiplying the bound (14) by $s_1^{-1}s_2^{-1}$ and integrating over $1 \ll s_1, s_2 \ll X^\delta$, combining with (13), (9) and Proposition 4, and setting $\delta = 3/364$ then completes the proof of Theorem 7.

We will prove Theorem 8 by applying a Selberg sieve. To do so, we need to count elements in $\mathcal{B} \cap V(\mathbb{Z})$ satisfying congruence conditions. For any $B, B' \in V(\mathbb{Z})$ and any integer r, we write $B \equiv B' \pmod r$ if and only if $B - B' \in rV(\mathbb{Z})$. We prove:

Theorem 9 *Let m be an odd squarefree positive integer with $m \ll X^{1/3}$. Let $\mathcal{B}$ and δ be as in Theorem 8. Let $B_0 \in V(\mathbb{Z})$ be an element such that $m \mid q(B_0)$ and B_0 is*

nonzero modulo p for each prime factor p of m. Let $N_q(\mathcal{B}; m, B_0)$ denote the number of $B \in \mathcal{B} \cap V(\mathbb{Z})$ such that $B \equiv B_0 \pmod m$ and $q(B) = 0$.

For each $r \geq 1$, set

$$C_q(r) = \frac{1}{r^9} \sum_{\substack{0 \leq a < r \\ \gcd(a,r)=1}} \sum_{B \bmod r} e\left(\frac{a}{r} q(B)\right) \tag{15}$$

Define the singular series

$$\mathfrak{S}(q) = \sum_{r \geq 1} C_q(r), \tag{16}$$

and for each prime p, the series

$$\mathfrak{S}(q; p) = \sum_{\ell \geq 0} C_q(p^\ell). \tag{17}$$

Define the singular integral

$$\mathfrak{S}_\infty(\mathcal{B}; q) = \int_{\mathbb{R}} \int_{\mathcal{B}} e(\theta q(B))\, dB\, d\theta, \tag{18}$$

where $e(x) = e^{2\pi i x}$ and dB denotes the Euclidean measure on $V(\mathbb{R})$. Then,

$$N_q(\mathcal{B}; m, B_0) = \frac{1}{m^8} \left(\prod_{p \mid m} \mathfrak{S}(q; p)^{-1} \right) \mathfrak{S}(q) \mathfrak{S}_\infty(\mathcal{B}; q) + O\left(\frac{X^{6.85(1+2\delta)}}{m^{5.5}} \log X \right), \tag{19}$$

with the implied constant being absolute. All the series defined above converge absolutely and they have positive value.

We note that the conditions $m \ll X^{1/3}, \delta < 0.01$ and eventually picking $\delta = 3/364$ are not optimal. They are only to make sure that the term $O_\epsilon(X^{6.992+\epsilon})$ in Theorem 7 beats $O(X^7)$.

4.1 Proof of Theorem 9 using the circle method

Fix an odd squarefree m. For any $\alpha \in [0, 1]$, let

$$S_{\mathcal{B}}(\alpha; m, B_0) = \sum_{\substack{B \in \mathcal{B} \cap V(\mathbb{Z}) \\ B \equiv B_0 \bmod m}} e\left(\frac{\alpha}{m} q(B)\right).$$

Then,

$$N_q(\mathcal{B}; m, B_0) = \int_0^1 S_{\mathcal{B}}(\alpha; m, B_0) d\alpha.$$

Let r_1 and r_2 be positive real numbers, to be picked later, with $r_1 \ll \frac{X^{1-2\delta}}{m}$ and $r_2 \gg X^{1+2\delta}$. Split the interval $[0, 1]$ into the major arcs $\mathfrak{M}$ and the minor arcs $\mathfrak{m} = [0, 1] \setminus \mathfrak{M}$, where

$$\mathfrak{M} = \left\{\alpha : \left|\alpha - \frac{a}{r}\right| \le \frac{1}{rr_2}, \gcd(a, r) = 1, 0 \le a < r \le r_1\right\}.$$

4.1.1 Major arc estimate

We estimate first the major arc integral

$$\int_{\mathfrak{M}} S_{\mathcal{B}}(\alpha; m, B_0)d\alpha = \sum_{r\le r_1} \sum_{\substack{0\le a<r\\ \gcd(a,r)=1}} \int_{|\theta|\le \frac{1}{rr_2}} S_{\mathcal{B}}\left(\frac{a}{r} + \theta; m, B_0\right) d\theta.$$

Fix some $\alpha = \frac{a}{r} + \theta \in \mathfrak{M}$, where $|\theta| \le \frac{1}{rr_2}$. We have

$$\begin{aligned} S_{\mathcal{B}}(\alpha; m, B_0) &= \sum_{\substack{B_1 \bmod rm\\ B_1\equiv B_0 \bmod m}} e\left(\frac{a}{rm}q(B_1)\right) \sum_{\substack{B\in\mathcal{B}\cap V(\mathbb{Z})\\ B\equiv B_1 \bmod rm}} e\left(\frac{\theta}{m}q(B)\right) \\ &= \sum_{\substack{B_1 \bmod rm\\ B_1\equiv B_0 \bmod m}} e\left(\frac{a}{rm}q(B_1)\right) \sum_{B'\in\mathcal{B}'\cap V(\mathbb{Z})} e\left(\frac{\theta}{m}q(rmB' + B_1)\right), \end{aligned}$$

where $\mathcal{B}' = \{B' \in V(\mathbb{R}) : rmB' + B_1 \in \mathcal{B}\}$ is another box. To compute the exponential sum over a box, we use the following result from [16, Proposition 8.7].

Lemma 1 *Let $f(x)$ be a real function on an interval $[a, b]$ such that $|f'(x)| \le \frac{1}{2}$ for all $x \in (a, b)$. Suppose further that $f''(x) \ge 0$ on (a, b) or that $f''(x) \le 0$ on (a, b). Then,*

$$\sum_{a<n<b} e(f(n)) = \int_a^b e(f(x))\,dx + O(1),$$

with the implied constant being absolute.

We note that [16, Proposition 8.7] requires that $f''(x) > 0$, but the same proof applies when $f''(x) \ge 0$ or when $f''(x) \le 0$. The following multivariable version also follows immediately.

Lemma 2 *Let $f(x_1, \ldots, x_\ell)$ be a real function on a box $\mathcal{R} = \prod_i [a_i, b_i]$ such that $|\frac{\partial f}{\partial x_i}(x)| \le \frac{1}{2}$ on $\mathcal{R}$ for all $i = 1, \ldots, \ell$. Suppose for any $i = 1, \ldots, \ell$ and for any fixed $x_j \in (a_j, b_j)$ for all $j \ne i$, the second partial derivative $\frac{\partial^2 f}{\partial x_i^2}(x)$ as a function of*

x_i is either non-negative on (a_i, b_i) or non-positive on (a_i, b_i). Then

$$\sum_{n \in \mathcal{R} \cap \mathbb{Z}^\ell} e(f(n)) = \int_{\mathcal{R}} e(f(x))\, dx + O(\max\{\mathrm{Vol}(\bar{\mathcal{R}}), 1\}),$$

where $\mathrm{Vol}(\bar{\mathcal{R}})$ *denotes the greatest d-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace obtained by equating $\ell - d$ coordinates to zero, where d takes all values from* 1 *to* $\ell - 1$. *The implied constant depends only on ℓ.*

We apply Lemma 2 to the box $\mathcal{B}'$ and the quadratic polynomial $f(b_{ij}) = \frac{\theta}{m} q(rmB' + B_1)$ viewed as a function in the coordinates of B'. The partial derivative of f with respect to b_{ij} equals $\theta r \frac{\partial q}{\partial b_{ij}}(rmB' + B_1)$ which is bounded by $c_3 c_1 \theta r X^{1+2\delta}$ where c_3 is a constant depending only on q (and equals 2 in this case). Hence, we can bound the first order partial derivatives by $\frac{1}{2}$ by taking $r_2 \geq 2c_3 c_1 X^{1+2\delta}$. The second partial derivative of f with respect to any b_{ij} is a constant since f is quadratic. Finally, the side lengths of $\mathcal{B}'$ are of the form $2X_{ij}/(rm) \gg X^{1-2\delta}/(r_1 m) \gg 1$ by the assumption on r_1. Hence, we have

$$\begin{aligned}
&\sum_{B' \in \mathcal{B}' \cap V(\mathbb{Z})} e\left(\frac{\theta}{p} q(rmB' + B_1)\right) \\
&= \int_{\mathcal{B}'} e\left(\frac{\theta}{m} q(rmB' + B_1)\right) dB' + O\left(\left(\frac{X^{1+2\delta}}{rm}\right)^8\right) \\
&= \frac{1}{r^9 m^9} \int_{\mathcal{B}} e\left(\frac{\theta}{m} q(B)\right) dB + O\left(\left(\frac{X^{1+2\delta}}{rm}\right)^8\right).
\end{aligned}$$

Summing over the r^9 possible B_1's then gives

$$S_{\mathcal{B}}(\alpha; m, B_0) = c_q(a; r, m, B_0) \int_{\mathcal{B}} e\left(\frac{\theta}{m} q(B)\right) dB + O\left(\frac{r X^{8(1+2\delta)}}{m^8}\right), \tag{20}$$

where

$$c_q(a; r, m, B_0) = \frac{1}{r^9 m^9} \sum_{\substack{B_1 \bmod rm \\ B_1 \equiv B_0 \bmod m}} e\left(\frac{a}{rm} q(B_1)\right).$$

In the light of (15), we define for any integer $r \geq 1$ and any integer a coprime to r,

$$c_q(a; r) = \frac{1}{r^9} \sum_{B \bmod r} e\left(\frac{a}{r} q(B)\right).$$

Lemma 3 *If* $\gcd(r, m) = 1$, *then* $c_q(a; r, m, B_0) = \frac{1}{m^9} c_q(a; r)$. *Otherwise* $c_q(a; r, m, B_0) = 0$.

Proof We consider the case $\gcd(r, m) = 1$ first. Let $\bar{m}$ be any integer such that $m\bar{m} \equiv 1 \pmod{r}$. For any integer n divisible by m, we have $\frac{a}{rm}n \equiv \frac{a\bar{m}}{r}n \pmod{1}$. Suppose now $B_1, B' \in V(\mathbb{Z})$ with $B_1 \equiv B_0 \pmod{m}$ and $B_1 \equiv B' \pmod{r}$. Then $q(B_1) \equiv q(B_0) \equiv 0 \pmod{m}$ and $q(B_1) \equiv q(B') \pmod{r}$ and so

$$e\left(\frac{a}{rm}q(B_1)\right) = e\left(\frac{a\bar{m}}{r}q(B_1)\right) = e\left(\frac{a\bar{m}}{r}q(B')\right).$$

Since m and r are coprime, we have by the Chinese Remainder Theorem,

$$\sum_{\substack{B_1 \bmod rm \\ B_1 \equiv B_0 \bmod m}} e\left(\frac{a}{rm}q(B_1)\right) = \sum_{B' \bmod r} e\left(\frac{a}{r}q(B')\right).$$

Dividing by r^9m^9 gives us $c_q(a; r, m, B_0) = \dfrac{1}{m^9}c_q(a; r)$.

Now, we consider the case $\gcd(r, m) > 1$. Suppose that p is a prime dividing $\gcd(r, m)$. We rewrite the sum as

$$\sum_{\substack{B_1 \bmod rm \\ B_1 \equiv B_0 \bmod m}} e\left(\frac{a}{rm}q(B_1)\right) = \sum_{\substack{B' \bmod rm/p \\ B' \equiv B_0 \bmod m}} \sum_{B_0' \bmod p} e\left(\frac{a}{rm}q\left(\frac{rm}{p}B_0' + B'\right)\right).$$

Given $v, w \in V$, we write $\langle v, w\rangle = q(v + w) - q(v) - q(w)$ for the associated bilinear form. Hence, we have

$$q\left(\frac{rm}{p}B_0' + B'\right) = q(B') + \frac{rm}{p}\langle B', B_0'\rangle + \frac{r^2m^2}{p^2}q(B_0').$$

Since $rm \mid \frac{r^2m^2}{p^2}$, the inner sum equals

$$\sum_{B_0' \bmod p} e\left(\frac{a}{rm}\left(q(B') + \frac{rm}{p}\langle B', B_0'\rangle\right)\right) = e\left(\frac{a}{rm}q(B')\right) \sum_{B_0' \bmod p} e\left(\frac{a}{p}\langle B', B_0'\rangle\right).$$

Since $B' \equiv B_0$ is nonzero modulo p and q is non-degenerate modulo p for $p \geq 3$, the linear form $\langle B', *\rangle : V(\mathbb{F}_p) \to \mathbb{F}_p$ is nonzero. Moreover, a is coprime to p since $p \mid r$ and a is coprime to r. Therefore, the above exponential sum vanishes and as a result, $c_q(a; r, m, B_0) = 0$. □

Integrating (20) over the arc $|\theta| \leq \frac{1}{rr_2}$ and summing over a and r now give

$$\begin{aligned}&\int_{\mathfrak{M}} S(\alpha; m, B_0)d\alpha \\ &= \sum_{\substack{r \leq r_1 \\ \gcd(r,m)=1}} \sum_{\substack{0 \leq a < r \\ \gcd(a,r)=1}} \frac{1}{m^9}c_q(a; r) \int_{|\theta| \leq \frac{1}{rr_2}} \int_{\mathcal{B}} e\left(\frac{\theta}{m}q(B)\right) dB\, d\theta + O\left(\frac{r_1^2X^{8(1+2\delta)}}{r_2m^8}\right)\end{aligned}$$

$$= \sum_{\substack{r \le r_1 \\ \gcd(r,m)=1}} \frac{1}{m^8} C_q(r) \int_{|\theta| \le \frac{1}{mrr_2}} \int_{\mathcal{B}} e(\theta q(B))\, dB\, d\theta + O\left(\frac{r_1^2 X^{8(1+2\delta)}}{r_2 m^8}\right), \tag{21}$$

Our aim is to replace the above truncated sum by the singular series

$$\mathfrak{S}_m(q) = \sum_{\gcd(r,m)=1} \frac{1}{m^8} C_q(r) \tag{22}$$

and the above integral by the singular integral $\mathfrak{S}_\infty(\mathcal{B}; q)$. To this end, we prove the following bounds:

Lemma 4 *With notations as above, we have:*

(a) for all $r \ge 1$,

$$|C_q(r)| \le 4r^{-7/2}; \tag{23}$$

(b) for all $\theta \ne 0$,

$$\int_{\mathcal{B}} e(\theta q(B))\, dB \ll \min\{X^9, |\theta|^{-9/2}\}; \tag{24}$$

(c) the singular integral

$$\mathfrak{S}_\infty(\mathcal{B}; q) = \int_{\mathbb{R}} \int_{\mathcal{B}} e(\theta q(B))\, dB d\theta \ll X^7. \tag{25}$$

The above implied constants depend only on q (which is fixed).

Proof We prove first the bound

$$\sum_{B \bmod r} e\left(\frac{a}{r} q(B)\right) \le 8r^{9/2}. \tag{26}$$

Also, we prove the bound with the constant 8 replaced by $\sqrt{2}$ for r odd. Recall that $q(b_{ij}) = -b_{11}b_{44} - b_{22}b_{33} - 2b_{12}b_{34} - 2b_{13}b_{24} - 2b_{14}^2$. Hence (26) follows from

$$\sum_{x,y \bmod r} e\left(\frac{a}{r} xy\right) = \gcd(a, r)r, \quad \left| \sum_{x \bmod r} e\left(\frac{a}{r} x^2\right) \right| \le (2 \gcd(a, r)r)^{1/2},$$

where $\gcd(a, r) \mid 2$. Note that the second sum is a standard quadratic Gauss sum and the bound follows, for example, from [3, Sects. 1.3–1.6]. Thus, for r odd, $|C_q(r)| \le \sqrt{2}r^{-9/2}\phi(r) \le \sqrt{2}r^{-7/2}$, where $\phi(r)$ is the Euler's totient function. Meanwhile, for r even, we have $|C_q(r)| \le 8r^{-9/2}\phi(r) \le 4r^{-7/2}$. This proves (23).

Next we prove the bound (24). The X^9 bound is trivial since $\mathrm{Vol}(\mathcal{B}) \ll X^9$. We may also assume that $\theta > 0$ as the case $\theta < 0$ follows by complex conjugation. Setting

$B' = \theta^{1/2} B$, we see that it suffices to prove the following general statement: for any box $\mathcal{B}'$ centered at the origin,

$$\int_{\mathcal{B}'} e(q(B'))\, dB' \ll 1.$$

Again, using the explicit formula of q, it reduces to proving that for any $X, Y > 0$,

$$\int_{-X}^{X} \int_{-Y}^{Y} e(xy)\, dy dx \ll 1, \quad \int_{-X}^{X} e(x^2) dx \ll 1.$$

The first integral can be computed as follows:

$$\int_{-X}^{X} \int_{-Y}^{Y} e(xy) dx\, dy = \int_{-X}^{X} \frac{\sin(2\pi x Y)}{\pi x} dx = \int_{-XY}^{XY} \frac{\sin(2\pi x)}{\pi x} dx \ll 1.$$

Now, we bound the second integral. Since $e(x^2)$ is an even function, we can write

$$\int_{-X}^{X} e(x^2) dx = 2 \int_{0}^{X} e(x^2) dx.$$

For $0 < X < 1$, we can use the trivial estimate. For $X \geq 1$, we use the trivial estimate for $x \in [0, 1]$ and partial integration for $x \in [1, X]$:

$$\int_{0}^{X} e(x^2) dx \ll 1 + \left[\frac{e(x^2)}{2x}\right]_1^X + \int_1^X \frac{e(x^2)}{2x^2} dx \ll 1 + 1 + \left[\frac{1}{2x}\right]_1^X \ll 1.$$

Finally, by using (24), we have

$$\mathfrak{S}_\infty(\mathcal{B}; q) \ll \int_{|\theta| \leq X^{-2}} X^9 d\theta + \int_{|\theta| \geq X^{-2}} |\theta|^{-9/2} d\theta \ll X^7,$$

which is the desired bound (25). □

Note the bound (23) on $C_q(r)$ implies that

$$|\mathfrak{S}(q) - 1| \leq 4(\zeta(7/2) - 1) < 1$$

and that for each prime p,

$$|\mathfrak{S}(q; p) - 1| \leq 4 \sum_{\ell \geq 1} p^{-(7/2)\ell} = \frac{4}{p^{7/2} - 1} < 1.$$

Hence, the series defined by (16) and (17) have positive values.

Combining the bounds (23), (24) and (25) with (21), we have

$$\begin{aligned}
&\int_{\mathfrak{M}} S(\alpha; m, B_0)d\alpha \\
&= \sum_{\substack{r\le r_1\\ \gcd(r,m)=1}} \left(\frac{1}{m^8}C_q(r)(\mathfrak{S}_\infty(\mathcal{B}; q) + O((mrr_2)^{7/2})\right) + O\left(\frac{r_1^2X^{8(1+2\delta)}}{r_2m^8}\right) \\
&= \left(\mathfrak{S}_m(q) + \sum_{r>r_1} O(r^{-7/2}m^{-8})\right)\mathfrak{S}_\infty(\mathcal{B}; q) + \sum_{r\le r_1} O(r^{-7/2}m^{-8}(mrr_2)^{7/2}) \\
&\quad + O\left(\frac{r_1^2X^{8(1+2\delta)}}{r_2m^8}\right) \\
&= \mathfrak{S}_m(q)\mathfrak{S}_\infty(\mathcal{B}; q) + O\left(\frac{X^7}{r_1^{5/2}m^8} + \frac{r_1r_2^{7/2}}{m^{9/2}} + \frac{r_1^2X^{8(1+2\delta)}}{r_2m^8}\right), \qquad (27)
\end{aligned}$$

where $\mathfrak{S}_m(q)$ is defined in (22).

4.1.2 Minor arc estimate

We now estimate the minor arc integral. Fix some $\alpha = \frac{a}{r} + \theta \in \mathfrak{m}$, where $r_1 < r \le r_2$ and $|\theta| \le \frac{1}{rr_2}$. Then

$$\begin{aligned}
|S_{\mathcal{B}}(\alpha; m, B_0)|^2 &= \sum_{\substack{B', B''\in\mathcal{B}\cap V(\mathbb{Z})\\ B', B''\equiv B_0 \bmod m}} e\left(\frac{\alpha}{m}(q(B'') - q(B'))\right) \\
&= \sum_{B\in V(\mathbb{Z})} \sum_{\substack{B'\in\mathcal{B}\cap V(\mathbb{Z})\\ B'\equiv B_0 \bmod m\\ B'+mB\in\mathcal{B}\cap V(\mathbb{Z})}} e\left(\frac{\alpha}{m}(m^2q(B) + m\langle B', B\rangle)\right),
\end{aligned}$$

where the second equality follows by setting $B'' = B' + mB$. The set of $B \in V(\mathbb{Z})$ for which the inner sum is non-empty is contained in the box $\mathcal{B}'' = \frac{1}{m}(\mathcal{B} - \mathcal{B}) = \{\frac{1}{m}(B'' - B') : B', B'' \in \mathcal{B}\}$. Taking absolute values now give

$$|S_{\mathcal{B}}(\alpha; m, B_0)|^2 \le \sum_{B\in\mathcal{B}''\cap V(\mathbb{Z})} \left| \sum_{\substack{B'\in\mathcal{B}\cap V(\mathbb{Z})\\ B'\equiv B_0 \bmod m\\ B'+mB\in\mathcal{B}\cap V(\mathbb{Z})}} e\left(\alpha\langle B', B\rangle\right)\right|$$

$$= \sum_{B \in \mathcal{B}'' \cap V(\mathbb{Z})} \left| \sum_{\substack{B_1 \in V(\mathbb{Z}) \\ B_0 + mB_1 \in \mathcal{B} \\ B_0 + mB_1 + mB \in \mathcal{B}}} e\left(\alpha m \langle B_1, B \rangle\right) \right|. \tag{28}$$

Let b_{ij} denote the entries of B and let x_{ij} denote the entries of B_1. Then from the explicit formula for q, we have

$$\begin{aligned} \langle B_1, B \rangle = &-b_{11}x_{44} - b_{44}x_{11} - b_{22}x_{33} - b_{33}x_{22} \\ &-2b_{12}x_{34} - 2b_{34}x_{12} - 2b_{13}x_{24} - 2b_{24}x_{13} - 4b_{14}x_{14}. \end{aligned} \tag{29}$$

Each x_{ij} takes all integer values within an interval, depending only on b_{ij}, of length at most $2X_{ij}/m$. Hence, the inner sum in (28) factors into a product of geometric sums. For each (i, j), let c_{ij} denote the integer coefficient in front of each b_{ij} in (29) and let I_{ij} denote the closed interval $[-2X_{ij}/m, 2X_{ij}/m]$. Then we have

$$|S_{\mathcal{B}}(\alpha; m, B_0)|^2 \ll \prod_{(i,j)} \sum_{b_{ij} \in I_{ij} \cap \mathbb{Z}} \min\left\{ \frac{X_{ij}}{m}, ||\alpha c_{ij} m b_{ij}||^{-1} \right\},$$

where $||\cdot||$ is the distance to the nearest integer function.
Recalling that $\alpha = \frac{a}{r} + \theta$ with $|\theta| \leq \frac{1}{rr_2}$, we have

$$|\theta c_{ij} m b_{ij}| \leq \frac{4mb_{ij}}{rr_2} \leq \frac{8X_{ij}}{rr_2} \leq \frac{8c_1 X^{1+2\delta}}{rr_2} \leq \frac{1}{2r}$$

by taking $r_2 \geq 16c_1 X^{1+2\delta}$. So we have the lower bound

$$||\alpha c_{ij} m b_{ij}|| \geq \frac{1}{2} ||\frac{a}{r} m c_{ij} b_{ij}||.$$

Write $r_{ij} = r/\gcd(r, mc_{ij}) \geq r/(4m)$ and $a_{ij} = amc_{ij}/\gcd(r, mc_{ij})$. We have

$$|S_{\mathcal{B}}(\alpha; m, B_0)|^2 \ll \prod_{(i,j)} \sum_{b_{ij} \in I_{ij} \cap \mathbb{Z}} \min\left\{ \frac{X_{ij}}{m}, ||\frac{a_{ij}}{r_{ij}} b_{ij}||^{-1} \right\}.$$

Now if $r_{ij} > 4X_{ij}/m$, then

$$\sum_{b_{ij} \in I_{ij} \cap \mathbb{Z}} \min\left\{ \frac{X_{ij}}{m}, ||\frac{a_{ij}}{r_{ij}} b_{ij}||^{-1} \right\} \leq \frac{X_{ij}}{m} + 2 \sum_{\ell=1}^{\lceil 2X_{ij}/m \rceil} \frac{r_{ij}}{\ell} \ll \frac{X_{ij}}{m} + r_{ij} \log X_{ij}. \tag{30}$$

If $r_{ij} \le 4X_{ij}/m$, then

$$\sum_{b_{ij}\in I_{ij}\cap\mathbb{Z}} \min\left\{\frac{X_{ij}}{m}, ||\frac{a_{ij}}{r_{ij}}b_{ij}||^{-1}\right\} \ll \frac{X_{ij}}{m}\frac{4X_{ij}/m}{r_{ij}} + \frac{4X_{ij}/m}{r_{ij}}\sum_{\ell=1}^{\lfloor r_{ij}/2\rfloor}\frac{r_{ij}}{\ell}$$
$$\ll \frac{X_{ij}^2}{rm} + \frac{X_{ij}}{m}\log X_{ij}. \tag{31}$$

Combining (30) and (31) then gives

$$\sum_{b_{ij}\in I_{ij}\cap\mathbb{Z}} \min\left\{\frac{X_{ij}}{m}, ||\frac{a_{ij}}{r_{ij}}b_{ij}||^{-1}\right\} \ll \frac{X^{2(1+2\delta)}}{rm} + r_2\log X.$$

Raising it to the power 9 and taking square root give

$$S_{\mathcal{B}}(\alpha; m, B_0) \ll \frac{X^{9(1+2\delta)}}{r^{9/2}m^{9/2}} + r_2^{9/2}\log^{9/2} X.$$

Finally, integrating over the minor arc gives

$$\int_{\mathfrak{m}} S_{\mathcal{B}}(\alpha; m, B_0) \ll \sum_{r_1<r\le r_2}\sum_{\substack{0\le a<r\\(a,r)=1}}\int_{|\theta|\le\frac{1}{rr_2}}\left(\frac{X^{9(1+2\delta)}}{r^{9/2}m^{9/2}} + r_2^{9/2}\log^{9/2} X\right)d\theta$$
$$\ll \sum_{r_1<r\le r_2}\left(\frac{X^{9(1+2\delta)}}{r_2 r^{9/2}m^{9/2}} + r_2^{7/2}\log^{9/2} X\right)$$
$$\ll \frac{X^{9(1+2\delta)}}{r_2 r_1^{7/2}m^{9/2}} + r_2^{9/2}\log^{9/2} X, \tag{32}$$

where the last bound follows from $r_2 \gg X^{1+2\delta}$.

4.1.3 Proof of Theorem 4.2

We are ready to prove Theorem 9. By (27) and (32), we have

$$N_q(\mathcal{B}; m, B_0) = \mathfrak{S}_m(q)\mathfrak{S}_\infty(\mathcal{B}; q)$$
$$+O\left(\frac{X^7}{r_1^{5/2}m^8} + \frac{r_1 r_2^{7/2}}{m^{9/2}} + \frac{r_1^2X^{8(1+2\delta)}}{r_2m^8} + \frac{X^{9(1+2\delta)}}{r_2r_1^{7/2}m^{9/2}} + r_2^{9/2}\log^{9/2} X\right),$$

where $\mathfrak{S}_m(q)$ is defined in (22). Take

$$r_1 = X^{\frac{2}{11}(1+2\delta)}m^{\frac{7}{11}}, \qquad r_2 = \frac{X^{1.522(1+2\delta)}}{m^{1.223}\log X}.$$

Since $m \ll X^{1/3}$ and $\delta < 0.01$, we see that $r_1 m \ll X^{1-2\delta}$ and $X^{1+2\delta} \ll r_2 \ll X^2$. With this choice of r_1 and r_2, we have

$$N_q(\mathcal{B}; m, B_0) = \mathfrak{S}_m(q)\mathfrak{S}_\infty(\mathcal{B}; q) + O\left(\frac{X^{6.85(1+2\delta)}}{m^{5.5}} \log X\right).$$

Finally, since C_q is multiplicative (as easily verified), the singular series $\mathfrak{S}_m(q)$ defined in (22) equals

$$\mathfrak{S}_m(q) = \frac{1}{m^8}\left(\prod_{p|m} \mathfrak{S}(q; p)^{-1}\right)\mathfrak{S}(q).$$

This completes the proof of Theorem 9.

4.2 Proof of Theorem 4.1 using the Selberg sieve

For any prime p, we say an element $B \in V(\mathbb{F}_p)$ (or $W(\mathbb{F}_p)$) is $\mathbb{F}_p$-*reducible* if either $\Delta(B) = 0 \in \mathbb{F}_p$ or $\Delta(B) \neq 0$ and B is $\mathbb{F}_p$-distinguished in the sense of Section 2.1. We begin by proving that any $B \in V(\mathbb{Z})$ that is $\mathbb{Q}$-distinguished is $\mathbb{F}_p$-reducible for every prime p. For any prime p, let $\alpha_p : V(\mathbb{Z}) \to V(\mathbb{F}_p)$ and $\beta_p : \mathbb{Z}^4 \to \mathbb{F}_p^4$ denote the reduction-mod-p maps.

Lemma 5 *Suppose $B \in V(\mathbb{Z})$ is $\mathbb{Q}$-distinguished. Let p be a prime such that $p \nmid \Delta(B)$. Then $\alpha_p(B)$ is $\mathbb{F}_p$-distinguished.*

Proof Since B is $\mathbb{Q}$-distinguished, there exist linearly independent vectors $v, w \in \mathbb{Q}^4$ satisfying (5); namely

$$v^t A_0 v = v^t B v = w^t A_0 w = v^t A_0 w = v^t B w = 0.$$

By scaling v and w, we may assume that $v, w \in \mathbb{Z}^4$ and $\beta_p(v), \beta_p(w) \neq 0$. If $\beta_p(v), \beta_p(w)$ are linearly independent over $\mathbb{F}_p$, then $\alpha_p(B)$ is $\mathbb{F}_p$-distinguished since the vectors $\beta_p(v), \beta_p(w)$ satisfy (5) and $\Delta(\alpha_p(B)) \neq 0$ since $p \nmid \Delta(B)$. If $\beta_p(v), \beta_p(w)$ are linearly dependent over $\mathbb{F}_p$, then there exists $a_1 \in \{0, 1, \ldots, p-1\}$ and $w_1 \in \mathbb{Z}^4$ such that $w = a_1 v + p w_1$. Note that v, w_1 also satisfy (5). If $\beta_p(v), \beta_p(w_1)$ are linearly independent over $\mathbb{F}_p$, then we are done. Otherwise, there exists $a_2 \in \{0, 1, \ldots, p-1\}$ and $w_2 \in \mathbb{Z}^4$ such that $w_1 = a_2 v + p w_2$. Note now $w = (a_1 + p a_2)v + p^2 w_2$. We may now repeat this process. If it terminates at some $v, w_n \in \mathbb{Z}^4$ with $\beta_p(v), \beta_p(w_n)$ linearly independent over $\mathbb{F}_p$, then we are done. If it does not terminate, then there exists a sequence $a_1, a_2, \ldots \in \{0, 1, \ldots, p-1\}$ such that for any $n \geq 1$, $w - (a_1 + p a_2 + \cdots + p^{n-1} a_n)v \in p^n \mathbb{Z}^4$. This implies that v and w are linearly dependent over $\mathbb{Z}_p$ and so also over $\mathbb{Q}_p$, which contradicts the assumption that they are linearly independent over $\mathbb{Q}$ since $v, w \in \mathbb{Q}^4$. □

We now apply the Selberg sieve ([16, Theorem 6.4]) to prove Theorem 8. We follow the setup as in [17, Sect. 3]. Let z be a number less than $X^{1/3}$. Let P be the

product of all primes p with $N \le p < z$ where N is some large absolute constant to be determined later. For each $m \mid P$, let a_m be the number of elements $B \in \mathcal{B} \cap V(\mathbb{Z})$ such that:

- $q(B) = 0$;
- for any prime $p \mid \frac{P}{m}$, B is $\mathbb{F}_p$-reducible;
- for any prime $p \mid m$, B is not $\mathbb{F}_p$-reducible.

For $m \nmid P$, we set $a_m = 0$. Then, applying the Selberg sieve will give us the count for

$$a_1 = \sum_{\gcd(n,P)=1} a_n,$$

which is the number of elements $B \in \mathcal{B} \cap V(\mathbb{Z})$ with $q(B) = 0$ and is $\mathbb{F}_p$-reducible for all primes $p \mid P$.

For any squarefree $m \mid P$, the expression

$$\sum_{n \equiv 0 \bmod m} a_n$$

counts the number of elements $B \in \mathcal{B} \cap V(\mathbb{Z})$ such that $q(B) = 0$ and B is not $\mathbb{F}_p$-reducible for any $p \mid m$. Recall that for any prime p, we defined d_p in Proposition 3 for the number of $B_0 \in W(\mathbb{F}_p)$ with $f_{B_0} \in U(\mathbb{F}_p)$ and are not $\mathbb{F}_p$-distinguished, which is the same as the number of $B_0 \in V(\mathbb{F}_p)$ with $q(B_0) = 0$ and are not $\mathbb{F}_p$-reducible. The condition that $\Delta(B_0) \neq 0$ in $\mathbb{F}_p$ also implies that B_0 is nonzero modulo p. Thus, by Proposition 3 and Theorem 9, we have

$$\begin{aligned}\sum_{n \equiv 0 \bmod m} a_n &= \frac{1}{m^8} \prod_{p \mid m} \left(d_p \mathfrak{S}(q; p)^{-1}\right) \mathfrak{S}(q) \mathfrak{S}_\infty(\mathcal{B}; q) + O\left(X^{6.85(1+2\delta)} m^{2.5} \log X\right) \\ &= \left(\prod_{p \mid m} \frac{d_p \mathfrak{S}(q; p)^{-1}}{p^8}\right) \mathfrak{S}(q) \mathfrak{S}_\infty(\mathcal{B}; q) + O\left(X^{6.85(1+2\delta)} m^{2.5} \log X\right).\end{aligned}$$

We set $g(m) = \prod_{p \mid m} g(p)$ and $u_m = O\left(X^{6.85(1+2\delta)} m^{2.5} \log X\right)$ for each squarefree $m \mid P$, where

$$g(p) = \frac{d_p \mathfrak{S}(q; p)^{-1}}{p^8}$$

for each prime $p \mid P$. By (23), we have

$$\mathfrak{S}(q; p) = 1 + O\Big(\sum_{\ell \ge 1} p^{-7\ell/2}\Big) = 1 + O(p^{-7/2}).$$

Recall from Proposition 3, we have the bound

$$\frac{1}{16} + O(p^{-1}) \leq \frac{d_p}{p^8} \leq \frac{3}{4} + O(p^{-1}).$$

Hence, by taking N large enough, we have the bound $\frac{1}{32} \leq g(p) \leq \frac{7}{8}$ for $p \geq N$.

Now, set $h(m) = \prod_{p|m} \frac{g(p)}{1-g(p)}$ for all squarefree $m \mid P$. Let $D > 1$ with $D < z$ be a real number to be picked later and set

$$H = \sum_{\substack{m < \sqrt{D} \\ m|P}} h(m).$$

Then, by [16, Theorem 6.4], we have

$$a_1 = \sum_{\gcd(n,P)=1} a_n \leq H^{-1} \mathfrak{S}(q) \mathfrak{S}_\infty(\mathcal{B}; q) + R,$$

where

$$\begin{aligned} |R| &\leq \sum_{\substack{m < \sqrt{D} \\ m|P}} \tau_3(m) u_m \ll_\epsilon X^{6.85(1+2\delta)} \log X \sum_{m < \sqrt{D}} m^{2.5+\epsilon} \\ &\ll_\epsilon X^{6.85(1+2\delta)} D^{1.75+\epsilon} \log X \end{aligned}$$

for any $\epsilon > 0$.

Meanwhile, for p prime, we have $\frac{1}{31} \leq h(p) \leq 8$ and so for any $\epsilon > 0$,

$$H \gg \pi(\sqrt{D}) \gg_\epsilon D^{0.5-\epsilon}.$$

Thus, we get

$$N_q^{\text{dist}}(\mathcal{B}) \leq a_1 \ll_\epsilon X^7 D^{-0.5+\epsilon} + X^{6.85(1+2\delta)} D^{1.75+\epsilon} \log X.$$

Taking $D = X^{(1/15)-(54.8/9)\delta}$ gives the desired bound (14).

5 Proof of Theorem 1, Theorem 3 and Theorem 4

We prove Theorem 4 first. By the paragraph following Theorem 4, it is enough to consider the case $\alpha = 256$ and $\beta = -27$. We note that for $(a, b) \in \mathbb{Z}^2$ with $H(a, b) < X$, there are at most X^ϵ integers m whose square divides $\Delta(a, b)$. Hence, it is enough

to prove:

$$X^{\epsilon} \cdot \# \bigcup_{\substack{m>M \\ m \text{ squarefree}}} \{(a,b) \in \mathbb{Z}^2 \colon H(a,b) < X, m^2 \mid \Delta(a,b)\} \ll_{\epsilon} \frac{X^{7+\epsilon}}{\sqrt{M}} + X^{6.992+\epsilon}.$$

Moreover, if $m^2 \mid \Delta(a,b)$, then we can factor $m = m_1 m_2$ where m_1 is the product of all prime factors p of m such that p^2 strongly divides $\Delta(a,b)$, and m_2 is the product of all prime factors p of m such that p^2 weakly divides $\Delta(a,b)$. Since at least one of m_1 or m_2 is at least m' for some squarefree integer $m' \geq \sqrt{m}$, we have

$$\bigcup_{\substack{m>M \\ m \text{ squarefree}}} \{(a,b) \in \mathbb{Z}^2 \colon H(a,b) < X, m^2 \mid \Delta(a,b)\} \subset \bigcup_{\substack{m'>\sqrt{M} \\ m' \text{ squarefree}}} \mathcal{W}_{m'}^{(1)} \cup \bigcup_{\substack{m'>\sqrt{M} \\ m' \text{ squarefree}}} \mathcal{W}_{m'}^{(2)}.$$

Theorem 4 now follows from Theorem 5.

Next, we prove Theorem 1 using an inclusion-exclusion sieve. We have

$$N(X;\alpha,\beta) = \sum_{m} \mu(m) N_m(X;\alpha,\beta).$$

By covering the box $(-X^3, X^3) \times (-X^4, X^4)$ by $(2X^3 m^{-2} + O(1))(2X^4 m^{-2} + O(1))$ boxes of size $m^2 \times m^2$, each of which contains $\rho_{\alpha,\beta}(m^2)$ integral points (a,b) such that $m^2 \mid \beta a^4 + \alpha b^3$, we have the following individual count

$$N_m(X;\alpha,\beta) = 4X^7 m^{-4} \rho_{\alpha,\beta}(m^2) + O(X^4 m^{-2} \rho_{\alpha,\beta}(m^2)) + O(\rho_{\alpha,\beta}(m^2)).$$

Since $\rho_{\alpha,\beta}(m^2) = O(m^2)$, we sum over $m < X^{\eta}$ for some $\eta > 0$ to get

$$\begin{aligned} &\sum_{m<X^{\eta}} \mu(m) N_m(X;\alpha,\beta) \\ &= 4X^7 \sum_{m<X^{\eta}} \mu(m) \frac{\rho_{\alpha,\beta}(m^2)}{m^4} + O(X^{4+\eta}) + O(X^{1+3\eta}) \\ &= C(\alpha,\beta) \cdot 4X^7 + O(X^{7-\eta}) + O(X^{4+\eta}) + O(X^{1+3\eta}). \end{aligned} \tag{33}$$

We take $\eta = 0.1$ and apply Theorem 4 with $M = X^{0.1}$ to get

$$N(X;\alpha,\beta) = C(\alpha,\beta) \cdot 4X^7 + O(X^{6.9}) + O_{\epsilon}(X^{6.9+\epsilon} + X^{6.992+\epsilon}).$$

The proof of Theorem 1 is now complete.

Finally, we prove Theorem 3. By Proposition 2, we see that 100% of the quartics of the form $x^4 + ax + b$ are irreducible. For any prime p and any irreducible monic

polynomial $f(x) \in \mathbb{Z}[x]$, if $\mathbb{Z}[x]/(f(x))$ is not maximal at p, then $p^2 \mid \Delta(f)$. Hence the tail estimate for the number of monic quartics of the form $x^4 + ax + b$ whose discriminant is divisible by the square of a large prime implies the tail estimate for the number of monic quartics $f(x) = x^4 + ax + b$ such that $\mathbb{Z}[x]/(f(x))$ is not maximal at a large prime. Therefore, it remains to compute the p-adic density for monic quartics $f(x) = x^4 + ax + b$ such that $\mathbb{Z}[x]/(f(x))$ is maximal at p.

Fix a prime p. For any $g \in \mathbb{Z}[x]$, let $\bar{g}$ denote its reduction in $\mathbb{F}_p[x]$. By [2, Corollary 3.2], $\mathbb{Z}[x]/(f(x))$ is not maximal at p if and only if there exists a monic polynomial $u \in \mathbb{Z}[x]$ such that $\bar{u} \in \mathbb{F}_p[x]$ is irreducible and $f \in (p^2, pu, u^2) \subset \mathbb{Z}[x]$. Suppose $f(x) = x^4 + ax + b \in \mathbb{Z}[x]$ and $u(x)$ is monic with $f \in (p^2, pu, u^2)$. Then $\bar{u}^2 \mid \bar{f}$. Hence $\deg(u) \leq 2$. Suppose $u(x) = x^2 + cx + d \in \mathbb{Z}[x]$ has degree 2. Then

$$f(x) - u^2 = 2cx^3 + (2d + c^2)x^2 + (2cd - a)x + (d^2 - b) \in p\mathbb{Z}[x].$$

If $p \neq 2$, then we have $p \mid c$ and $p \mid d$, in which case $\bar{u} = x^2$ is not irreducible. If $p = 2$, then from the x^2-coefficient, we have $2 \mid c$, in which case $\bar{u} = x^2 + \bar{d}$ is also not irreducible. Hence $u(x) = x - r$, for some $r \in \mathbb{Z}$, is linear. We now have $f(x + r) \in (p^2, px, x^2)$, which is equivalent to $p \mid f'(r)$ and $p^2 \mid f(r)$. Note this implies $p^2 \mid f(r')$ for any $r' \equiv r \pmod{p}$. We may then take $r \in \{0, 1, \ldots, p - 1\}$. From $p \mid f'(r)$, we get $a \equiv -4r^3 \pmod{p}$. Once we fix one of the p choices of $a \in \{0, 1, \ldots, p^2 - 1\}$, from $p^2 \mid f(r)$, we get $b \equiv -r^4 - ar \pmod{p^2}$. Thus there are p pairs (a, b) associated to each $r \in \{0, 1, \ldots, p - 1\}$.

Suppose now a pair (a, b) arises from two distinct $r_1, r_2 \in \{0, 1, \ldots, p - 1\}$. Then r_1, r_2 are both double roots of $\bar{f}(x)$ and so we have $\bar{f}(x) = (x - r_1)^2(x - r_2)^2 = (x^2 - (r_1 + r_2)x + r_1 r_2)^2$. Since $f(x)$ has vanishing x^3- and x^2-coefficients, the same is true for $\bar{f}(x)$. When $p \neq 2$, this is possible only if $(x - r_1)(x - r_2) = x^2$ which implies that $r_1 \equiv r_2 \equiv 0 \pmod{p}$ and so $r_1 = r_2$. When $p = 2$, this implies $r_1 + r_2 \equiv 0 \pmod{2}$ and thus $r_1 = r_2$. Both cases yield a contradiction.

As a result, there are p^2 pairs $(a, b) \in \{0, 1, \ldots, p^2 - 1\}^2$ such that $\mathbb{Z}[x]/(x^4 + ax + b)$ is not maximal at p. We therefore obtain the desired $1 - p^{-2}$ for the p-adic density of monic quartics $f(x) = x^4 + ax + b$ such that $\mathbb{Z}[x]/(f(x))$ is maximal at p.

Acknowledgements It is a pleasure to thank Manjul Bhargava for many helpful comments. The first named author is supported by the University of Waterloo through an MURA project. The second named author is supported by an NSERC Discovery Grant.

Author Contributions Not applicable.

Funding The first named author is supported by the University of Waterloo through an MURA project. The second named author is supported by an NSERC Discovery Grant.

Availability of data and material Not applicable.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Code availability Not applicable.

References

1. Apostal, T.: Introduction to Analytic Number Theory. Undergraduate Texts in Mathematics, Springer, New York (1976)
2. Ash, A., Brakenhoff, J., Zarrabi, T.: Equality of polynomial and field discriminants. Exp. Math. **16**, 367–374 (2007)
3. Berndt, B., Evans, R., Williams, K.: Gauss and Jacobi Sums. Wiley, New York (1998)
4. Bhargava, M., Shankar, A.: Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves. Ann. Math. (2) **181**, 191–242 (2015)
5. Bhargava, M., Shankar, A., Wang, X.: Squarefree values of polynomial discriminants I. Invent. Math. (2022). https://doi.org/10.1007/s00222-022-01098-w
6. Bhargava, M., Shankar, A., Wang, X.: Squarefree values of polynomial discriminants II. Preprint
7. Granville, A.: ABC allows us to count squarefrees. Int. Math. Res. Not. **19**, 991–1009 (1998)
8. Greaves, G.: Power-free values of binary forms. Q. J. Math. Oxford Ser. (2) **43**, 45–65 (1992)
9. Shankar, A., Wang, X.: Rational points on hyperelliptic curves having a marked non-Weierstrass point. Compos. Math. **154**(1), 188–222 (2018)
10. Hooley, C.: On the square-free values of cubic polynomials. J. Reine Angew. Math. **229**, 147–154 (1968)
11. Hooley, C.: On the power-free values of polynomials in two variables: II. J. Number Theory **129**(6), 1443–1455 (2009)
12. Kowalski, J.: On the proportion of squarefree numbers among sums of cubic polynomials. Ramanujan J. **54**(2), 343–354 (2021)
13. Murty, R., Paston, H.: Counting squarefree values of polynomials with error term. Int. J. Number Theory **10**(7), 1743–1760 (2014)
14. Poonen, B.: Squarefree values of multivariable polynomials. Duke Math. J. **118**(2), 353–373 (2003)
15. Heath-Brown, D.: Power-free values of polynomials. Q. J. Math. **64**, 177–188 (2013)
16. Iwaniec, H., Kowalski, E.: Analytic Number Theory. American Mathematical Society Colloquium Publications 53, Providence, RI (2004)
17. Shankar, A., Tsimerman, J.: Counting S_5-fields with a power saving error term. Forum Math. Sigma **2**, e13 (2014)