Forum of Mathematics, Pi (2020), Vol. 8, e3, 128 pages
doi:10.1017/fmp.2019.8

PRIMES REPRESENTED BY INCOMPLETE NORM FORMS

JAMES MAYNARD
Magdalen College, Oxford, England OX1 4AU, UK;
email: james.alexander.maynard@gmail.com

Received 6 January 2016; accepted 17 November 2019

Abstract

Let $K = \mathbb{Q}(\omega)$ with ω the root of a degree n monic irreducible polynomial $f \in \mathbb{Z}[X]$. We show that the degree n polynomial $N(\sum_{i=1}^{n-k} x_i \omega^{i-1})$ in $n-k$ variables takes the expected asymptotic number of prime values if $n \geqslant 4k$. In the special case $K = \mathbb{Q}(\sqrt[n]{\theta})$, we show that $N(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})$ takes infinitely many prime values, provided $n \geqslant 22k/7$.

Our proof relies on using suitable 'Type I' and 'Type II' estimates in Harman's sieve, which are established in a similar overall manner to the previous work of Friedlander and Iwaniec on prime values of $X^2 + Y^4$ and of Heath-Brown on $X^3 + 2Y^3$. Our proof ultimately relies on employing explicit elementary estimates from the geometry of numbers and algebraic geometry to control the number of highly skewed lattices appearing in our final estimates.

2010 Mathematics Subject Classification: 11N05 (primary); 11N35, 11N32 (secondary)

1. Introduction

It is believed that any integer polynomial satisfying some simple necessary conditions should represent infinitely many primes. Specifically, we have the following quantitative strengthening of Bunyakovsky's conjecture, which is the Bateman–Horn conjecture [1] in the special case of one polynomial.

CONJECTURE. *Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree d with positive leading coefficient and no fixed prime divisor. Then we have*

$$\#\{1 \leqslant a \leqslant x : f(a) \text{ prime}\} = \mathfrak{S}_f \frac{x}{d \log x} + o_f\left(\frac{x}{\log x}\right),$$

© The Author 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

where

$$\mathfrak{S}_f = \prod_p \left(1 - \frac{\nu_f(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1},$$
$$\nu_f(p) = \#\{1 \leqslant a \leqslant p : f(a) \equiv 0 \pmod p\}.$$

It follows from a classical result of Kronecker (or the later Frobenius or Chebotarev density theorems) that the infinite product $\mathfrak{S}_f$ converges to a positive constant.

Unfortunately, no case of the above conjecture is known other than when f is linear, and the problem seems to be well beyond the current techniques. A nonlinear polynomial f represents $O(x^{1/2})$ integers less than x, and there are essentially no examples of sets containing $O(x^{1/2})$ integers less than x which contain infinitely many primes (beyond artificial examples). (The seemingly simpler problem of showing the existence of a prime in the short interval $[x, x + x^{1/2}]$, for example, is not known even under the assumption of the Riemann hypothesis.) Thus, the sparsity of the set of values of f presents a major obstacle.

As an approximation to the conjecture, one can look at polynomials $f \in \mathbb{Z}[X_1, \ldots, X_n]$ in multiple variables, so the resulting sets are less sparse. If the number of variables is sufficiently large (relative to other measures of the complexity of f), then, in principle, the Hardy–Littlewood circle method can be used to show that every *integer* satisfying necessary local conditions is represented by f. It follows from the seminal work of Birch [2], for example, that any homogeneous nonsingular $f \in \mathbb{Z}[X_1, \ldots, X_n]$ of degree d with no fixed prime divisor represents infinitely many prime values, provided $n > (d-1)2^d$.

When the number of variables is not larger than the degree, only a few polynomials are known to represent infinitely many primes, and these tend to have extra algebraic structure. Iwaniec [17] has shown that any suitable binary quadratic polynomial represents infinitely many primes. If $K/\mathbb{Q}$ is a number field with an $\mathbb{Z}$-basis $\{\beta_1, \ldots, \beta_n\}$ of $\mathcal{O}_K$, then the norm form $N_{K/\mathbb{Q}}(X_1\beta_1 + \cdots + X_n\beta_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ is a degree n polynomial in n variables, which represents infinitely many primes, since every degree 1 principal prime ideal of K gives rise to a prime value of $N_{K/\mathbb{Q}}$.

The groundbreaking work of Friedlander–Iwaniec [8] shows that the polynomial $X_1^2 + X_2^4$ takes the expected number of prime values. Along with the work of Heath-Brown [13] on $X_1^3 + 2X_2^3$ (and its generalizations due to Heath-Brown and Moroz [16], [15] and the recent work of Heath-Brown–Li [14] on $X^2 + p^4$), these are the only known examples of a set of polynomial values containing $O(x^c)$ elements less than x (for some constant $c < 1$) which contain infinitely many prime values. A key feature in the proofs are the fact that these

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

polynomials are closely related to norm forms; $N_{\mathbb{Q}(i)/\mathbb{Q}}(X_1 + X_2^2 i) = X_1^2 + X_2^4$ and $N_{\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}}(X_1 + X_2\sqrt[3]{2}) = X_1^3 + 2X_2^3$. This allows the structure of the prime factorization in the number field to be combined with bilinear techniques to count primes in these cases.

The paper of Heath-Brown [13] suggested that one might hope to utilize similar techniques when considering higher degree norm forms with appropriate variables set equal to zero. We address this problem in this paper, thereby giving further examples of thin polynomials which represent infinitely many primes.

THEOREM 1.1. *Let n, k be positive integers. Let $f \in \mathbb{Z}[X]$ be a monic irreducible polynomial of degree n with root $\omega \in \mathbb{C}$. Let $K = \mathbb{Q}(\omega)$ be the corresponding number field of degree n, and let $N_K \in \mathbb{Z}[X_1, \dots, X_{n-k}]$ be the 'incomplete norm form'*

$$N_K(\mathbf{a}) = N_K(a_1, \dots, a_{n-k}) = N_{K/\mathbb{Q}}\left(\sum_{i=1}^{n-k} a_i \omega^{i-1}\right).$$

If $n \geqslant 4k$, then as $X \to \infty$, we have

$$\#\{\mathbf{a} \in [1, X]^{n-k} : N_K(\mathbf{a}) \text{ prime}\} = (\mathfrak{S} + o(1))\frac{X^{n-k}}{n \log X},$$

where

$$\mathfrak{S} = \prod_p \left(1 - \frac{\nu(p)}{p^{n-k}}\right)\left(1 - \frac{1}{p}\right)^{-1},$$

$$\nu(p) = \#\{1 \leqslant a_1, \dots, a_{n-k} \leqslant p : N_K(\mathbf{a}) \equiv 0 \pmod p\}.$$

All implied constants depend only on ω and are effectively computable.

THEOREM 1.2. *Let n, k be positive integers. Let $f(X) = X^n - \theta \in \mathbb{Z}[X]$ be irreducible, $K = \mathbb{Q}(\sqrt[n]{\theta})$ and $N_K(\mathbf{a}) = N_{K/\mathbb{Q}}(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})$, as in Theorem 1.1 in the case $f(X) = X^n - \theta$.*

If $n \geqslant 22k/7$ and X is sufficiently large, then

$$\#\{\mathbf{a} \in [1, X]^{n-k} : N_K(\mathbf{a}) \text{ prime}\} \gg \mathfrak{S}\frac{X^{n-k}}{\log X}.$$

All implied constants depend only on θ and are effectively computable, and $\mathfrak{S}$ is the constant defined in Theorem 1.1.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

A sieve upper bound shows $\#\{\mathbf{a} \in [1, X]^{n-k} : N_K(\mathbf{a}) \text{ prime}\} \ll \mathfrak{S} X^{n-k}/\log X$, and so the lower bound in Theorem 1.2 is of the correct order of magnitude. We note $22/7 = 3.14\ldots < 4$.

Theorems 1.1 and 1.2 give examples of sets of polynomial values containing roughly $x^{1-k/n}$ elements less than x which contain many primes. We obtain an asymptotic for the number of primes in the sets of Theorem 1.1 which contain $\gg x^{3/4}$ values less than x, and a lower bound of the correct order of magnitude for the sets of Theorem 1.2 which contain $\gg x^{15/22}$ elements. By way of comparison, the Friedlander–Iwaniec polynomial $X_1^2 + X_2^4$ takes roughly $x^{3/4}$ values less than x, which is at the limit of the range for asymptotic estimates in Theorem 1.1, whilst Heath-Brown's polynomial $X_1^3 + 2X_2^3$ takes roughly $x^{2/3}$ values less than x, which is thinner than the sets considered in Theorem 1.1 or Theorem 1.2.

By virtue of being homogeneous, the algebraic structure of the polynomials considered in Theorems 1.1 and 1.2 are simpler in some key aspects to the Friedlander–Iwaniec polynomial $X_1^2 + X_2^4$; much of the paper [**8**] is spent employing sophisticated techniques to handle sums twisted by a quadratic character caused by the nonhomogeneity. In our situation, the key multiplicative machinery is instead just a Siegel–Walfisz-type estimate for Hecke L-functions. (The fact that $n > 3k$ means that characters of large conductor do not play a role, and so we do not even require a large sieve type estimate as in [**13**].) On the other hand, the fact that we consider polynomials in an arbitrary number of variables and with multiple coordinates of the norm form set to 0 introduces different complications of a geometric nature. It is handling such issues, which is the key innovation of this paper. In particular, if just one coefficient were set equal to zero, then the result would follow from an adaption of the paper of Heath-Brown. Thus, the fact that we are able to take a moderately large positive proportion of the coefficients to be equal to zero should be viewed as the key feature of Theorem 1.1.

Unlike the previous estimates, the implied constants in Theorems 1.1 and 1.2 are effectively computable. This is a by-product of the fact that we explicitly treat the contribution of a possible exceptional quadratic character in order to be able to utilize a Siegel–Walfisz-type estimate in a slightly wider range of uniformity of conductor. This extra range of uniformity enables us to restrict ourselves to simpler algebraic estimates.

In view of the results of Friedlander–Iwaniec and Heath-Brown, the restrictions of $n \geqslant 4k$ and $n \geqslant 22k/7$ in Theorems 1.1 and 1.2 might seem unnatural at first sight, but it turns out that these are natural barriers to any simple argument used to establish 'Type I' and 'Type II' estimates. If one simply bounds the naturally occurring error terms by their absolute values without showing genuine cancellations, then one can only hope to obtain 'Type I' and

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

'Type II' estimates in certain restricted ranges depending on the density of the sequence. Heath-Brown [13] obtains an asymptotic in a sparser sequence precisely because he is able to treat the error terms arising in a nontrivial manner. We discuss this further in Section 11.

With more care, one could give a quantitative bound to the $o(1)$ error term appearing in Theorem 1.1.

2. Outline of the proof

In the interest of clarity, we prove Theorems 1.1 and 1.2 together in the case of $K = \mathbb{Q}(\sqrt[n]{\theta})$ in Sections 5–11, and then in Section 12, we sketch the few modifications to the argument required to obtain Theorem 1.1 in the general case of $K = \mathbb{Q}(\omega)$.

We now give a broad outline of the key steps in the proof; what we say here should be thought of as a heuristic motivation and not interpreted precisely.

Given a small quantity $\eta_1 > 0$ and large quantities X_i of size about X, we let

$$\mathscr{A} = \{\mathbf{a} \in \mathbb{Z}^{n-k} : a_i \in [X_i, X_i + \eta_1 X_i]\}.$$

We establish a suitable estimate for the number of times $N_K(\mathbf{a})$ is prime for $\mathbf{a} \in \mathscr{A}$ for each of these smaller sets individually. For each $\mathbf{a} \in \mathscr{A}$, there is a principal ideal $(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})$ with the same norm. Provided all elements of $\mathscr{A}$ have norm of size $\gg X^n$ (as will be the case for typical choices of the X_i) and provided η_1 is sufficiently small, this ideal is unique (since units are a discrete group in K). Thus, we wish to count the number of degree 1 prime ideals in $\mathfrak{A} = \{(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}) : \mathbf{a} \in \mathscr{A}\}$, and so we can use the unique factorization of ideals in K.

In Section 6, we apply a combinatorial decomposition to $\mathfrak{A}$ based on Buchstab's identity and Harman's sieve [11]. In the case when $n > 4k$, this takes the simple form

$$\begin{aligned}\#\{\text{prime ideals in } \mathfrak{A}\} = {} & \#\{\mathfrak{a} \in \mathfrak{A} \text{ with no prime factor of norm} < X^{n-3k-4\epsilon}\} \\ & - \sum_{X^{n-3k-4\epsilon} < N(\mathfrak{p}) < X^{n/2+2\epsilon}} \#\{\mathfrak{a} \in \mathfrak{A} \text{ with } \mathfrak{p} \text{ the factor of smallest norm}\}.\end{aligned}$$

The point of this decomposition is that we will be able to appropriately estimate terms when every ideal counted has a factor whose norm is in the interval $[X^{k+\epsilon}, X^{n-2k-\epsilon}]$. This is clearly the case with the second term on the right-hand side above. The first term can be repeatedly decomposed by further Buchstab iterations so that all terms count ideals with a prime factor of norm in

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

the interval $[R, RX^{n-3k-4\epsilon}]$ (for any suitable choice of R) or simply count the number of ideals in $\mathscr{A}$ which are a multiple of some divisor $\mathfrak{d}$.

Thus, it suffices to obtain suitable asymptotic estimates (at least on average) for the number of ideals in $\mathfrak{A}$ which are a multiple of some ideal $\mathfrak{d}$ or the number of ideals in $\mathfrak{A}$ with a particular type of prime factorization whenever this prime factorization ensures the existence of a conveniently sized factor. These estimates are the so-called 'Type I' (linear) and 'Type II' (bilinear) estimates which provide the key arithmetic content.

Our Type I estimate of Section 7 states that

$$\sum_{N(\mathfrak{d})\in[D,2D]}\left|\#\{\mathfrak{a}\in\mathfrak{A}:\mathfrak{d}|\mathfrak{a}\}-\frac{\rho(\mathfrak{d})\#\mathfrak{A}}{N(\mathfrak{d})}\right|\ll X^{n-k-1}D^{1/(n-k)+\epsilon}+D,$$

where ρ is the function defined by

$$\rho(\mathfrak{d})=\frac{\#\{\mathbf{x}\in[1,N(\mathfrak{d})]^{n-k}:\mathfrak{d}|(\sum_{i=1}^{n-k}x_i\sqrt[n]{\theta^{i-1}})\}}{N(\mathfrak{d})^{n-k-1}}.$$

This allows us to accurately count the number of ideals in $\mathfrak{A}$ which are a multiple of an ideal of norm $O(X^{n-k-\epsilon})$ on average. Since $\#\mathfrak{A}\approx X^{n-k}$, we see that this range is essentially best possible.

If $\mathfrak{d}=(\sum_{i=1}^{n}d_i\sqrt[n]{\theta^{i-1}})$ is a principal ideal in $\mathbb{Z}[\sqrt[n]{\theta}]$ and if $\mathbb{Z}[\sqrt[n]{\theta}]=\mathcal{O}_K$, then we see that the number of ideals $\mathfrak{a}=\mathfrak{e}\mathfrak{d}$ in $\mathfrak{A}$ which are a multiple of $\mathfrak{d}$ is given by

$$\#\left\{\mathbf{e}\in\mathbb{Z}^n:\sum_{i=1}^{n-k}e_i\sqrt[n]{\theta^{i-1}}\times\sum_{i=1}^{n}d_i\sqrt[n]{\theta^{i-1}}\in\mathscr{A}\right\}$$
$$=\{\mathbf{e}\in\mathbb{Z}^n:\mathbf{d}^{(i)}\cdot\mathbf{e}\in[X_i,X_i+\eta_1X_i]\text{ for }i\leqslant n-k,\mathbf{d}^{(i)}\cdot\mathbf{e}=0\text{ for }i>n-k\},$$

where $\mathbf{d}^{(i)}$ is the ith row in the multiplication-by-$\sum_{i=1}^{n}d_i\sqrt[n]{\theta^{i-1}}$ matrix with respect to the basis $\{\sqrt[n]{\theta^{i-1}}\}_{1\leqslant i\leqslant n}$. But this is counting vectors in the lattice defined by $\mathbf{d}^{(i)}\cdot\mathbf{e}=0$ for $i>n-k$ in the bounded region defined by $\mathbf{d}^{(i)}\cdot\mathbf{e}\in[X_i,X_i+\eta_1X_i]$ for $i\leqslant n-k$. By estimates from the geometry of numbers, the number of such points is approximately the volume of the bounded region divided by the lattice discriminant, provided the lattice and the bounded region are not too skewed. Our Type I estimate then follows from showing that the number of skewed lattices is rare. (Small technical modifications are made to deal with $\mathfrak{d}$ in other ideal classes and if $\mathbb{Z}[\sqrt[n]{\theta}]\neq\mathcal{O}_K$.)

The argument then relies on establishing a suitable Type II estimate, which is the main part of the paper. Given integers $\ell'\leqslant\ell$ and a polytope $\mathcal{R}\subseteq\mathbb{R}^\ell$ such

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

that any $\mathbf{e} \in \mathcal{R}$ has $e_i \geqslant \epsilon^2$ for all i and $k + \epsilon \leqslant \sum_{i=1}^{\ell'} e_i \leqslant n - 2k - \epsilon$, our Type II estimate obtains an asymptotic for the sum

$$\sum_{\mathfrak{a} \in \mathfrak{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}),$$

where

$$\mathbf{1}_{\mathcal{R}}(\mathfrak{a}) = \begin{cases} 1, & \mathfrak{a} = \mathfrak{p}_1 \dots \mathfrak{p}_\ell,\ N(\mathfrak{p}_i) = X^{e_i},\ (e_1, \dots, e_\ell) \in \mathcal{R}, \\ 0, & \text{otherwise.} \end{cases}$$

This sum counts ideals in $\mathfrak{A}$ with a given number of prime factors each of a given size, and the condition that $k + \epsilon \leqslant \sum_{i=1}^{\ell'} e_i \leqslant n - 2k - \epsilon$ implies that $\mathfrak{a}$ has a 'conveniently sized' ideal factor. By performing a decomposition to $\mathcal{R}$, we may assume that $\mathcal{R} = \mathcal{R}_1 \times \mathcal{R}_2$ for two polytopes $\mathcal{R}_1, \mathcal{R}_2$ with $\mathcal{R}_2$ corresponding to the conveniently sized factor. We are left to estimate the bilinear sum

$$\sum_{\mathfrak{a}\mathfrak{b} \in \mathfrak{A}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}).$$

We estimate this sum using a combination of L^1 and L^2 bounds. We introduce an approximation $\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$ to $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})$, which is a sieve weight designed to have the same distributional properties as $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})$. The sums $\sum_{\mathfrak{a}\mathfrak{b} \in \mathfrak{A}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$ can then be estimated using our Type I estimates, and they give the expected asymptotic.

To show that the error in this approximation is small, we use Linnik's dispersion method to exploit the bilinear structure. By Cauchy–Schwarz and using $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \leqslant 1$, we are left to bound

$$\sum_{\mathfrak{a}} \Bigg| \sum_{\mathfrak{b} : \mathfrak{a}\mathfrak{b} \in \mathfrak{A}} (\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})) \Bigg|^2.$$

Writing $g_{\mathfrak{b}} = \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$, expanding the square and swapping the order of summation, we are left to estimate

$$\sum_{\mathfrak{b}_1, \mathfrak{b}_2} g_{\mathfrak{b}_1} g_{\mathfrak{b}_2} \sum_{\substack{\mathfrak{a} \\ \mathfrak{b}_1\mathfrak{a}, \mathfrak{b}_2\mathfrak{a} \in \mathfrak{A}}} 1.$$

If $\mathfrak{b}_1, \mathfrak{b}_2$ are both principal and $\mathfrak{a} = (\sum_{i=1}^{n} a_i \sqrt[n]{\theta^{i-1}})$ for some $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n$, then each condition $\mathfrak{a}\mathfrak{b}_1 \in \mathfrak{A}$ and $\mathfrak{a}\mathfrak{b}_2 \in \mathfrak{A}$ imposes k linear constraints on $\mathbf{a}$ (with coefficients of these linear constraints depending on the coefficients of the $\mathfrak{b}_i$). For generic $\mathfrak{b}_1, \mathfrak{b}_2$, these constraints are linearly independent, and so $\mathbf{a}$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

will be constrained to lie in a bounded region in a rank $n-2k$ lattice. Using the geometry of numbers again, the number of such $\mathbf{a}$ is roughly the volume of the region divided by the lattice discriminant, provided neither are too skewed. An iterative argument shows that the number of skewed lattices here is acceptably small.

To finish the estimate, we have to show suitable cancellation in the sum

$$\sum_{\mathfrak{b}_1,\mathfrak{b}_2}\frac{g_{\mathfrak{b}_1}g_{\mathfrak{b}_2}\operatorname{vol}(\mathscr{R}_{\mathfrak{b}_1,\mathfrak{b}_2})}{\det(\Lambda_{\mathfrak{b}_1,\mathfrak{b}_2})},$$

where $\mathscr{R}_{\mathfrak{b}_1,\mathfrak{b}_2}$ and $\Lambda_{\mathfrak{b}_1,\mathfrak{b}_2}$ are the bounded region and lattice which $\mathbf{a}$ was constrained to. The volume of the bounded region is continuous and so plays a minor role. More significant complications occur in showing that those $\mathfrak{b}_1$, $\mathfrak{b}_2$ for which $\det(\Lambda_{\mathfrak{b}_1,\mathfrak{b}_2})$ is small make a negligible contribution.

The determinant can be small if a certain vector of polynomials in the coefficients of $\mathfrak{b}_1$, $\mathfrak{b}_2$ is small in either the Euclidean metric or a p-adic metric. To show that this is only rarely the case, we obtain a (sharp) bound on the dimension of the corresponding variety given by these polynomials. We obtain this by elementary algebraic means by exploiting the simple explicit description of multiplication of elements in the order $\mathbb{Z}[\sqrt[n]{\theta}]$.

Having shown that only those $\mathfrak{b}_1$, $\mathfrak{b}_2$ for which $\Lambda_{\mathfrak{b}_1,\mathfrak{b}_2}$ has determinant almost as large as possible make a contribution, we can localize the coefficients of $\mathfrak{b}_1$, $\mathfrak{b}_2$ to a small region in the Euclidean metric and p-adic metrics for small p. Once localized in this way, the denominator no longer plays an important role. The remaining sum then factors, and so we are ultimately left to show cancellation in

$$\sum_{\mathfrak{b}}{}'g_{\mathfrak{b}},$$

where $\sum'$ indicates that the coefficients are localized to a small box and an arithmetic progression. Recalling that $g_{\mathfrak{b}}=\mathbf{1}_{\mathcal{R}_1}(\mathfrak{b})-\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$, we can show such an estimate using a Siegel–Walfisz-type bound for Hecke characters. In avoiding some algebraic considerations, we require uniformity in the conductor to be slightly larger than a fixed power of a logarithm in the norm of the ideals considered, and this requires us to take explicit account of possible fluctuations caused by a Siegel zero throughout the argument.

3. Notation

We view n, k, θ and $K=\mathbb{Q}(\sqrt[n]{\theta})$ (or $K=\mathbb{Q}(\omega)$ in Section 12) as fixed throughout the paper. All constants implied by $O(\cdot)$, $o(\cdot)$, $\ll$ and $\gg$ may depend

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

on θ (and, hence, may depend on n and k since $3k+1 \leqslant n$ and n is the degree of θ). All asymptotic notation should be interpreted in the limit as $X \to \infty$.

Throughout the paper, we let ϵ be a small but fixed (that is, independent of X) positive constant which is always assumed to be sufficiently small in terms of n and k. Our implied constants will not depend on ϵ unless explicitly stated, but we will assume $\epsilon \geqslant 1/\log\log X$ to avoid too many dependencies in our error terms. We let Δ_K be the discriminant of the field K, $\phi_K(\mathfrak{a}) = N(\mathfrak{a})\prod_{\mathfrak{p}|\mathfrak{a}}(1-N(\mathfrak{p})^{-1})$, and γ_K the residue of $\zeta_K(s)$ at $s=1$.

By abuse of notation, we write $N = N_{K/\mathbb{Q}}$ for the norm form on ideals of K and for algebraic integers of K. We let $N_K(\mathbf{x})$ be the polynomial in $n-k$ variables $x_1, \dots, x_{n-k}$ which coincides with $N_{K/\mathbb{Q}}(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})$ on integers.

We use lower Gothic script (for example, $\mathfrak{a}$, $\mathfrak{b}$, ...) to denote integral ideals of K and $\mathfrak{p}$ to denote a prime ideal of K. Algebraic integers in $\mathcal{O}_K$ will typically be written in Greek lower case (for example, $\alpha, \beta, \dots$) and (α) is used to denote the principal ideal generated by α. Vectors will be denoted by roman bold lower case (for example, $\mathbf{a}$, $\mathbf{b}$, ...), and we have endeavoured to use consistent notation across vectors, integers and ideals referring to related objects so that $\mathfrak{b} = (\beta)$ for the principal ideal generated by $\beta = \sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}}$ for some vector $\mathbf{b}$. We let $\|\mathbf{b}\| = \sqrt{\sum_i b_i^2}$ denote the usual Euclidean norm.

4. Basic estimates

We recall some results from the geometry of numbers and Minkowski's theory of successive minima. We recall that a *lattice* in $\mathbb{R}^k$ is a discrete subgroup of the additive group $\mathbb{R}^k$.

LEMMA 4.1 (Minkowski-reduced basis). *Let $\Lambda \subseteq \mathbb{R}^k$ be a lattice. Then there is a set $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ of linearly independent vectors in $\mathbb{R}^k$ such that*

(1) *$\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is a basis:*

$$\Lambda = \mathbf{v}_1\mathbb{Z} + \cdots + \mathbf{v}_r\mathbb{Z}.$$

(2) *The $\mathbf{v}_i$ are quasiorthogonal: For any $x_1, \dots, x_r \in \mathbb{R}$, we have*

$$\|x_1\mathbf{v}_1 + \cdots + x_r\mathbf{v}_r\| \asymp \sum_{i=1}^{r} \|x_i\mathbf{v}_i\|.$$

(3) *The sizes of the $\mathbf{v}_i$ are controlled by successive minima: If $\lambda_1 \leqslant \lambda_2 \cdots \leqslant \lambda_r$ are the successive minima of Λ, then $\|\mathbf{v}_i\| \asymp \lambda_i$ for all i. In particular,*

$$\|\mathbf{v}_1\| \cdots \|\mathbf{v}_r\| \asymp \det(\Lambda).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The implied constants above depend only on the ambient dimension k. Here, $\det(\Lambda)$ *is the r-dimensional volume of the fundamental parallelepiped, given by*

$$\left\{\sum_{i=1}^{r} x_i \mathbf{v}_i : x_1, \dots, x_r \in [0,1]\right\},$$

*and the j*th *successive minimum is the smallest quantity λ_j such that Λ contains j linearly independent vectors of norm at most λ_j.*

Proof. This follows from [6, Page 110] or [3, Ch. 1]. Explicitly, let $\mathbf{a}_1$, $\dots \mathbf{a}_r$ be chosen in turn such that $\mathbf{a}_j$ is the shortest vector of Λ which is linearly independent from $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}$. Then $\|\mathbf{a}_i\| = \lambda_i$ and the $\mathbf{a}_i$ are linearly independent by definition. By [3, Page 13, Corollary 2], there is then an integral basis $\mathbf{v}_1, \dots, \mathbf{v}_r$ of Λ with $\mathbf{v}_j = \sum_{i=1}^{j-1} \mu_{i,j}\mathbf{v}_i + \mu_{j,j}\mathbf{a}_j$ for some constants $|\mu_{i,j}| \leqslant 1$. In particular, $\|\mathbf{v}_i\| \ll \lambda_i$ by the triangle inequality. Since $\{\mathbf{v}_i\}$ is a basis, $\mathbf{v}_j$ is linearly independent of $\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}$. Thus, $\|\mathbf{v}_j\| \geqslant \lambda_j$ by minimality of $\|\mathbf{a}_j\|$, and so we have $\|\mathbf{v}_j\| \asymp \lambda_j$. By Minkowski's second Theorem (see, for example, [3, Page 205, Theorem 1]), we have that $\det(\Lambda) \asymp \lambda_1 \cdots \lambda_r$, so $\det(\Lambda) \asymp \|\mathbf{v}_1\| \cdots \|\mathbf{v}_r\|$. Trivially, we have that $\|\sum_{i=1}^{r} x_i \mathbf{v}_i\| \ll \sum_{i=1}^{r} \|x_i \mathbf{v}_i\| \ll \|x_j \mathbf{v}_j\|$ for some $1 \leqslant j \leqslant r$. Let $\mathbf{v}_j = \mathbf{v}_j' + \mathbf{v}_j''$ where $\mathbf{v}_j'' \in \mathbb{R}^k$ is linearly dependent on the other $\mathbf{v}_i$ and where $\mathbf{v}_j' \in \mathbb{R}^k$ is orthogonal to the other $\mathbf{v}_i$. We then have that

$$\lambda_1 \cdots \lambda_r \asymp \det(\Lambda) = \det(\mathbf{v}_1| \cdots |\mathbf{v}_r) = \det(\mathbf{v}_1| \cdots |\mathbf{v}_j'| \cdots |\mathbf{v}_r) \ll \|\mathbf{v}_j'\| \prod_{i \neq j} \lambda_i.$$

Thus, $\|\mathbf{v}_j'\| \gg \lambda_j \gg \|\mathbf{v}_j\|$. But since $\mathbf{v}_j'$ is orthogonal to the other $\mathbf{v}_i$, we have $\|\sum_{i=1}^{r} x_i \mathbf{v}_i\| \geqslant \|x_j \mathbf{v}_j'\| \gg \|x_j \mathbf{v}_j\|$, as required. Together, this gives the result. □

We see that the properties of the Minkowski-reduced basis above indicate that each generating vector $\mathbf{v}_i$ has a positive proportion of its length in a direction orthogonal to all the other basis vectors.

LEMMA 4.2 (Well-sized generators). *Let $\mathfrak{a}$ be a principal ideal. Then there is a generator α of $\mathfrak{a}$ such that*

$$|\alpha^\sigma| \ll N(\mathfrak{a})^{1/n}$$

for all embeddings $\sigma : k \hookrightarrow \mathbb{C}$. In particular, $\alpha = (\theta n)^{-n} \sum_{i=1}^{n} a_i \sqrt[n]{\theta^{i-1}}$ for some integers $a_i \ll N(\mathfrak{a})^{1/n}$.

Proof. Let $\alpha \in \mathcal{O}_K$ and let K have r_1 real embeddings and r_2 complex ones. The Minkowski embedding sends α to the vector $(\log|\alpha^\sigma|)_\sigma \in \mathbb{R}^{r_1+r_2}$ indexed by

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

embeddings σ of K, where one only considers one of the two complex conjugate embeddings. The set of units of $\mathcal{O}_K$ is sent to a rank $r_1 + r_2 - 1$ lattice of determinant $O(1)$ in the trace 0 hyperplane. $\alpha \in \mathcal{O}_K$ is sent to a point $\mathbf{x}$ with trace $\log N(\alpha)$. Note that $\mathbf{x} - \log N(\alpha)/n$ is a point of trace 0. By Minkowski's convex body theorem (or Lemma 7.1), there is a point $\mathbf{e}$ in the lattice such that

$$\left|\mathbf{e} + \mathbf{x} - \frac{\log N(\alpha)}{n}\right| \leqslant C,$$

for some suitably large constant C. We then see that $\mathbf{e}$ is the image of a unit ϵ, and $\alpha\epsilon$ satisfies the required properties. □

LEMMA 4.3 (Prime ideal theorem). *There is a constant $c > 0$ such that*

$$\sum_{N(\mathfrak{a}) \leqslant X} \Lambda(\mathfrak{a}) = X + O(X \exp(-c\sqrt{\log X})).$$

LEMMA 4.4 (Zero-free region apart from Siegel zeros). *There is at most one modulus $\mathfrak{d}^*$ with $N(\mathfrak{d}^*) \leqslant \exp(\sqrt{\log X})$ and at most one primitive Hecke character $\chi_{\mathfrak{d}^*} \pmod{\mathfrak{d}^*}$ such that the Hecke L-function $L(s, \chi_{\mathfrak{d}^*})$ has a zero in the region*

$$\left\{s = \sigma + it : \sigma \geqslant 1 - \frac{c}{\sqrt{\log(X(2+|t|))}}\right\}.$$

Here $c > 0$ is a fixed small constant. This character, if it exists, must be a real quadratic character and the corresponding L-function has a unique real simple zero $\beta_{\mathfrak{d}^*}$ in the above region. The modulus $\mathfrak{d}^*$ in this case must satisfy $N(\mathfrak{d}^*) > (\log X)^{\epsilon}$ and $\mathfrak{d}^*$ must be square-free apart from a factor of norm $O(1)$.

LEMMA 4.5 (Prime ideal theorem with Hecke characters). *Let $\chi \neq \chi_{\mathfrak{d}^*}$ be a nontrivial primitive Hecke character with $\chi = \chi_1\chi_2$, where χ_1 is the torsion part of χ and χ_2 is torsion-free. Letting $\lambda_1, \ldots, \lambda_{n-1}$ be a basis of the torsion-free characters, we have that $\chi_2 = \prod_{i=1}^{n-1} \lambda_i^{m_i}$ for some integers m_i. Then we have*

$$\sum_{N(\mathfrak{a}) \leqslant X} \Lambda(\mathfrak{a})\chi(\mathfrak{a}) \ll X \exp(-c\sqrt{\log X})$$

uniformly over all such primitive $\chi = \chi_1\chi_2 \neq \chi_{\mathfrak{d}^*}$ of conductor $\leqslant \exp(\sqrt{\log X})$ and with $m_i \ll \exp(\sqrt{\log X})$ for all $1 \leqslant i \leqslant n-1$. In the case $\chi = \chi_{\mathfrak{d}^*}$, we have

$$\sum_{N(\mathfrak{a}) \leqslant X} \Lambda(\mathfrak{a})\chi_{\mathfrak{d}^*}(\mathfrak{a}) = \frac{-X^{\beta_{\mathfrak{d}^*}}}{\beta_{\mathfrak{d}^*}} + O(X \exp(-c\sqrt{\log X})).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Proof of Lemmas 4.3, 4.4 and 4.5. See [21, Theorem 1.9], for example. □

LEMMA 4.6 (Growth of Hecke L-series). *Let r_1 and $2r_2$ denote the number of real and complex embeddings of K, and let $\lambda_1, \dots, \lambda_{r_1+r_2-1}$ be a basis for the torsion-free Hecke characters. Let χ be a Hecke character of conductor $\mathfrak{q}$, and let $q = N(\mathfrak{q})$. Then χ factors as $\chi = \chi_1\lambda_1^{m_1}\dots\lambda_{r_1+r_2-1}^{m_{r_1+r_2-1}}$, where χ_1 is a class character mod q and $m_1, \dots, m_{r_1+r_2-1} \in \mathbb{Z}$. Then we have*

$$L(1-\sigma+it,\chi) \ll_\epsilon \left(\left(1+|t|+\sum_{i=1}^{r_1+r_2-1}|m_i|\right)q\right)^{n\sigma/2+\epsilon}$$

for $|\sigma+it| \geqslant 1/10$. The implied constant depends on K and the choice of basis $\lambda_1, \dots, \lambda_{r_1+r_2-1}$.

Proof. This follows from the Phragmén–Lindelöf principle—see [7, Equation (1.2.8)], for example. □

LEMMA 4.7 (Lower bound in zero-free-type region). *There is a constant $c_K > 0$ such that for $\sigma > 1 - c_K/\log t$, we have*

$$\frac{1}{\zeta_K(\sigma+it)} \ll \log(2+|t|) + \frac{1}{|1-\sigma-it|}.$$

Proof. This follows from [20, Lemma b and γ] and [4], for example. □

5. Initial manipulations

We begin our first steps in the proof of Theorems 1.1 and 1.2 for $K = \mathbb{Q}(\sqrt[n]{\theta})$. Here we use a simple decomposition to reduce our problem to counting principal prime ideals whose generators are localized. We may assume, without loss of generality, that θ is a positive integer if n is odd. We note that $N_K(X, 1, 0, \dots, 0) = X^n - \theta$ has no fixed prime divisor, and so N_K does not have a fixed prime divisor (in particular, $\mathfrak{S} \neq 0$). We wish to reduce the proof to the following proposition, where we set

$$\eta_1 = (\log X)^{-100}. \tag{5.1}$$

PROPOSITION 5.1 (Localized prime ideal counts). *Let $\mathscr{R} = \{\mathbf{x} \in \mathbb{R}^{n-k} : x_i \in [X_i, (1+\eta_1)X_i]\}$ be a hyperrectangle fully contained in $\{\mathbf{x} \in \mathbb{R}^k : \epsilon X \leqslant x_i \leqslant X,$*

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$N_K(\mathbf{x}) \geqslant \epsilon X^n\}$. *Let*

$$\mathcal{A}(\mathbf{a}_0) = \left\{ \left(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}} \right) : \mathbf{a} \in \mathscr{R} \cap \mathbb{Z}^{n-k},\ \mathbf{a} \equiv \mathbf{a}_0 \pmod{q^*} \right\}.$$

Then if $n \geqslant 4k$, *we have*

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \#\{\mathfrak{a} \in \mathcal{A}(\mathbf{a}_0) : \mathfrak{p}|\mathfrak{a} \Rightarrow N(\mathfrak{p}) > X^{n/2+\epsilon}\} = (\mathfrak{S} + O(\epsilon^{1/n})) \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X},$$

and if $n \geqslant 22k/7$, *we have*

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \#\{\mathfrak{a} \in \mathcal{A}(\mathbf{a}_0) : \mathfrak{p}|\mathfrak{a} \Rightarrow N(\mathfrak{p}) > X^{n/2+\epsilon}\} \gg \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{\log X}.$$

We note that since we are summing over all relevant choices of $\mathbf{a}_0$, the restrictions mod q^* in Proposition 5.1 are somewhat artificial. We have included them since we will consider each $\mathbf{a}_0$ separately in our later analysis. As an intermediate step, we establish the following lemma from Proposition 5.1.

LEMMA 5.2 (Localized prime counts in $\mathcal{O}_K$). *Let* $\mathscr{R} = \{\mathbf{x} \in \mathbb{R}^{n-k} : x_i \in [X_i, (1+\eta_1)X_i]\}$ *be a hyperrectangle fully contained in* $\{\mathbf{x} \in \mathbb{R}^k : \epsilon X \leqslant x_i \leqslant X, N_K(\mathbf{x}) \geqslant \epsilon X^n\}$. *Let* $\mathscr{A}'(\mathbf{a}_0) \subseteq \mathcal{O}_K$ *be given by*

$$\mathscr{A}'(\mathbf{a}_0) = \left\{ \sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}} : \mathbf{a} \in \mathscr{R} \cap \mathbb{Z}^{n-k},\ \mathbf{a} \equiv \mathbf{a}_0 \pmod{q^*} \right\}.$$

Then if $n \geqslant 4k$, *we have*

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \#\{\alpha \in \mathscr{A}'(\mathbf{a}_0) : N(\alpha)\ \textit{prime}\} = (\mathfrak{S} + O(\epsilon^{1/n})) \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X},$$

and if $n \geqslant 22k/7$, *we have*

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \#\{\alpha \in \mathscr{A}'(\mathbf{a}_0) : N(\alpha)\ \textit{prime}\} \gg \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{\log X}.$$

Proof of Lemma 5.2 assuming Proposition 5.1. This is simply a question of converting a count of prime algebraic integers to counting principal prime ideals. We define

$$\mathcal{A}(\mathbf{a}_0) = \left\{ \left(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}} \right) : \mathbf{a} \in \mathscr{A}'(\mathbf{a}_0) \right\}$$

to be the set of principal ideals generated by elements of $\mathscr{A}'(\mathbf{a}_0)$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We claim that every ideal in $\mathcal{A}(\mathbf{a}_0)$ has a unique generator in $\mathscr{A}'(\mathbf{a}_0)$. If $\mathbf{x} \in \mathscr{R}$, then $N_K(\mathbf{x}) \geqslant \epsilon X^n$ and $x_i \leqslant X$ for all i. Thus, it follows that $|\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}^{\sigma}| \ll X$ for all embeddings σ. This gives (letting ι denote the identity embedding, and $\Sigma(K/\mathbb{Q})$ the set of embeddings of $K/\mathbb{Q}$)

$$\left|\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\right| = \frac{N_K(\mathbf{x})}{\prod_{\substack{\sigma \in \Sigma(K/\mathbb{Q}) \\ \sigma \neq \iota}} |\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}^{\sigma}|} \gg \epsilon X.$$

In particular, if $\mathbf{y}, \mathbf{x} \in \mathscr{R}$, then $\mathbf{y} = \mathbf{x} + O(\eta_1 X)$, and so $\sum_{i=1}^{n-k} y_i \sqrt[n]{\theta^{i-1}} / \sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}} = 1 + O_\epsilon(\eta_1)$, which cannot be a nontrivial unit when η_1 is sufficiently small (since the units of $\mathcal{O}_K$ distinct from 1 are bounded uniformly away from 1). Thus, there are no two associates $\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}$ and $\sum_{i=1}^{n-k} y_i \sqrt[n]{\theta^{i-1}}$ in $\mathscr{A}'(\mathbf{a}_0)$. We therefore see that $\mathcal{A}(\mathbf{a}_0)$ is indeed in bijection with $\mathscr{A}'(\mathbf{a}_0)$.

There are $O(X^{n/2})$ prime ideals $\mathfrak{p}$ with $N(\mathfrak{p}) < X$ not prime (i.e a prime ideal of degree greater than 1), so it suffices to simply count prime ideals in $\mathcal{A}$ at the cost of an error term of size $O(X^{n/2})$. Putting this together, we see that

$$\begin{aligned}&\#\{\alpha \in \mathscr{A}'(\mathbf{a}_0) : N(\alpha) \text{ prime}\}\\ &\quad = \#\{\mathfrak{a} \in \mathcal{A}(\mathbf{a}_0) : \mathfrak{p}|\mathfrak{a} \Rightarrow N(\mathfrak{p}) > X^{n/2+\epsilon}\} + O(X^{n/2}).\end{aligned}$$

We now see that the statements of Lemma 5.2 follow immediately from Proposition 5.1, giving the result. □

Proof of Theorems 1.2 and 1.1 for $K = \mathbb{Q}(\sqrt[n]{\theta})$ assuming Proposition 5.1. We aim to reduce the statement of Theorems 1.2 and 1.1 to that of Lemma 5.2 by considering the contribution from small regions separately.

The measure of $\mathbf{t} \in [1, X]^{n-k}$ such that $N_K(\mathbf{t}) \leqslant \kappa X^n$ is $O(\kappa^{1/n} X^{n-k})$ uniformly in κ, and so we see that

$$\int_{\substack{\mathbf{t} \in [1,X]^{n-k} \\ N_K(\mathbf{t}) \geqslant 2}} \cdots \int \frac{dt_1 \dots dt_{n-k}}{\log N_K(\mathbf{t})} = (1 + o(1)) \frac{X^{n-k}}{n \log X}.$$

Thus, it suffices to show that if $n \geqslant 4k$, we have

$$\begin{aligned}&\#\{\mathbf{a} \in \mathbb{Z}^{n-k} : 1 \leqslant a_i \leqslant X,\ N_K(\mathbf{a}) \text{ prime}\}\\ &\quad = (\mathfrak{S} + O(\epsilon^{1/2n})) \int_{\substack{\mathbf{t} \in [1,X]^{n-k} \\ N_K(\mathbf{t}) \geqslant 2}} \cdots \int \frac{dt_1 \dots dt_{n-k}}{\log N_K(\mathbf{t})},\end{aligned} \tag{5.2}$$

and if $n \geqslant 22k/7$, then the left-hand side of (5.2) is bounded below by a positive constant times the right-hand side.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We first consider the region

$$\mathscr{E} = \{\mathbf{x} \in \mathbb{R}^{n-k} : 0 \leqslant x_i \leqslant \epsilon X \text{ for some } i\}.$$

By a simple sieve upper bound (see [9, Theorem 5.1] or Lemma 7.9 of Section 7), the number of prime values of $N_K(a_1, \dots, a_{n-k})$ for $\mathbf{a} \in \mathscr{E} \cap \mathbb{Z}^{n-k}$ is $O(\epsilon X^{n-k}/\log X)$. The contribution of $\mathbf{t} \in \mathscr{E}$ to the integral on the right-hand side of (5.2) is also $O(\epsilon X^{n-k}/\log X)$. Thus, we may restrict $\mathbf{a}$ and $\mathbf{t}$ to lie outside of $\mathscr{E}$, and so in the region where $x_i > \epsilon X$ for all i.

We recall from (5.1) that $\eta_1 = (\log X)^{-100}$. We cover the region $\{\mathbf{x} \in \mathbb{R}^{n-k} : \epsilon X \leqslant x_i \leqslant X\}$ with $O(\epsilon^{-o(1)}\eta_1^{-(n-k)})$ disjoint hyperrectangles of the form $\{\mathbf{x} \in \mathbb{R}^{n-k} : x_i \in (X_i, X_i + \eta_1 X_i]\}$. Again a sieve upper bound shows that the number of prime values of $N_K(a_1, \dots, a_{n-k})$ for integer vectors $\mathbf{a}$ in such a hyperrectangle is $O(\eta_1^{n-k} X^{n-k}/\log X)$. Thus, the total number of prime values of N_K from the $O_\epsilon(\eta_1^{-(n-k-1)})$ hyperrectangles not entirely contained within our region $\{\mathbf{x} \in \mathbb{R}^{n-k} : \epsilon X < x_i \leqslant X\}$ is $O_\epsilon(\eta_1 X^{n-k}/\log X)$. Similarly, we see that the total contribution to the integral on the right-hand side over real vectors $\mathbf{t}$ in the union of such boundary hyperrectangles is $O_\epsilon(\eta_1 X^{n-k}/\log X)$. Thus, we may restrict our attention to hyperrectangles fully contained in the region $\{\mathbf{x} \in \mathbb{R}^{n-k} : \epsilon X \leqslant x_i \leqslant X\}$.

We can clearly discard any hyperrectangles for which the norm is always negative since they make no contribution to either side of (5.2). We note that $\frac{\partial}{\partial x_j} N_K(x_1, \dots, x_{n-k}) \ll X^{n-1}$ on $[1, X]^{n-k}$ for all $j \in \{1, \dots, n-k\}$. Thus, if $|N_K(\mathbf{x})| \leqslant \epsilon X^n$, then all points $\mathbf{y}$ in the same hyperrectangle as $\mathbf{x}$ satisfy $|N_K(\mathbf{y})| \leqslant 2\epsilon X^n$. But there are $O(\epsilon^{1/n} X^{n-k})$ integer points $\mathbf{a} \in [1, X]^{n-k}$ for which $|N_K(\mathbf{a})| \leqslant 2\epsilon X^n$ since, given any choice of $a_2, \dots, a_{n-k} \leqslant X$, $N_K(\mathbf{a})$ is a nonzero integer polynomial of degree n in a_1, and we see that a_1 must lie within $O(\epsilon^{1/n} X)$ of one of the (complex) roots of this polynomial. Thus, there are $O(\epsilon^{1/n-o(1)}\eta_1^{-(n-k)})$ hyperrectangles containing a point $\mathbf{x}$ with $|N_K(\mathbf{x})| \leqslant \epsilon X^n$, and the total contribution from these hyperrectangles is $O(\epsilon^{1/n-o(1)} X^{n-k}/\log X)$. Similarly, the contribution to the integral on the right-hand side from $\mathbf{t}$ in the union of such hyperrectangles is $O(\epsilon^{1/n-o(1)} X^{n-k}/\log X)$. Thus, we may further restrict our attention to hyperrectangles with $N_K(\mathbf{x}) \geqslant \epsilon X^n$ for all $\mathbf{x}$ in the hyperrectangle.

Thus, we only need to consider hyperrectangles fully contained in the region $\{\mathbf{x} \in \mathbb{R}^{n-k} : \epsilon X \leqslant x_i, N_K(\mathbf{x}) > \epsilon X^n\}$. But for such hyperrectangles, the result follows immediately from Lemma 5.2 since $N(\sum_{i=1}^k a_i \sqrt[n]{\theta^{i-1}}) = N_K(\mathbf{a})$. □

Thus, we are left to establish Proposition 5.1.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

6. Sieve decomposition

In this section, we give a combinatorial decomposition of the number of primes in $\mathcal{A}$ based on Harman's sieve [11] and reduce our result to establishing suitable Type I and Type II estimates.

6.1. Initial setup. It will be notationally convenient to fix an (slightly artificial) ordering of ideals in K for this section. We first fix an ordering of prime ideals of K such that $\mathfrak{p}_1 < \mathfrak{p}_2$ if $N(\mathfrak{p}_1) < N(\mathfrak{p}_2)$, and we choose an arbitrary ordering of prime ideals of the same norm. We extend this to a total ordering of all ideals so that $\mathfrak{a} < \mathfrak{b}$ if $N(\mathfrak{a}) < N(\mathfrak{b})$, whilst if $N(\mathfrak{a}) = N(\mathfrak{b})$, we have $\mathfrak{a} < \mathfrak{b}$ if the least prime ideal factor of $\mathfrak{a}/\gcd(\mathfrak{a}, \mathfrak{b})$ is less than the least prime ideal factor of $\mathfrak{b}/\gcd(\mathfrak{a}, \mathfrak{b})$. Given a set of ideals $\mathcal{C}$ and an ideal $\mathfrak{a}$, we let

$$\mathcal{C}_{\mathfrak{a}} = \{\mathfrak{b} : \mathfrak{a}\mathfrak{b} \in \mathcal{C}\},$$
$$S(\mathcal{C}, \mathfrak{a}) = \#\{\mathfrak{b} \in \mathcal{C} : \mathfrak{p}|\mathfrak{b} \Rightarrow \mathfrak{p} > \mathfrak{a}\}.$$

For convenience, we let $\varpi = 0.3182$, and we fix ideals $\mathfrak{r}_1, \mathfrak{r}_2$ chosen maximally with respect to this ordering such that

$$N(\mathfrak{r}_1) \leqslant \begin{cases} X^{n(1-3\varpi)}, & n < 4k, \\ X^{n-3k-4\epsilon}, & n \geqslant 4k, \end{cases} \tag{6.1}$$

$$N(\mathfrak{r}_2) \leqslant X^{n(1/2+\epsilon)}. \tag{6.2}$$

In particular, we see that

$$\#\{\mathfrak{a} \in \mathcal{A}(\mathbf{a}_0) : \mathfrak{p}|\mathfrak{a} \Rightarrow N(\mathfrak{p}) > X^{n(1/2+\epsilon)}\} = S(\mathcal{A}(\mathbf{a}_0), \mathfrak{r}_2).$$

We now wish to decompose $S(\mathcal{A}(\mathbf{a}), \mathfrak{r}_2)$ into various terms such that each term can either be estimated asymptotically or the term is positive and can be dropped for a lower bound. To ease notation, we suppress the dependence of $\mathcal{A}(\mathbf{a}_0)$ on $\mathbf{a}_0$, and so write $\mathcal{A} = \mathcal{A}(\mathbf{a}_0)$. Roughly speaking, we will be able to asymptotically estimate terms of the form $S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1)$ when $N(\mathfrak{d}) < X^{n-k-4\epsilon}$ and terms $S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r})$ for fairly arbitrary ideals $\mathfrak{r}$ if $X^{k+\epsilon} \leqslant N(\mathfrak{d}) \leqslant X^{n-2k-\epsilon}$ (this latter type we refer to as the 'Type II range'). To make this precise, we introduce some further notation.

To keep track of the decomposition for $\mathcal{A}$, we perform the identical decompositions to a simpler set $\mathcal{B}$, which we use to compare to $\mathcal{A}$. To account for the impact of a possible exceptional character χ^*, we consider ideals with a fixed value of a real Hecke character so that the number of prime ideals in $\mathcal{B}$ fluctuates in the same manner as those in $\mathcal{A}$. Let $\mathfrak{a}_0 = (\sum_{i=1}^{n-k} (\mathbf{a}_0)_i \sqrt[n]{\theta^{i-1}})$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

be the ideal generated by the algebraic integer corresponding to $\mathbf{a}_0$, let χ^* be a real Hecke character on ideals with modulus $\mathfrak{q}^*$ and let $q^* = N(\mathfrak{q}^*)$. χ^* will be taken to be an exceptional character, if one exists, and an arbitrary such character otherwise, and q^* will satisfy $(\log x)^\epsilon \ll q^* \ll \exp(\sqrt[4]{\log X})$. $\mathfrak{q}^*$ will be square-free as an ideal, apart from a possible factor of norm $O(1)$. We see that χ^* takes values in $\{0, 1, -1\}$ and factors on principal ideals as $\chi^*((\alpha)) = \chi_f^*(\alpha)\chi_\infty^*(\alpha)$ as its finite and infinite components. Since all elements of $\mathcal{A}$ are principal and their coordinates are localized such that no norms are small, χ_∞^* takes a constant value on $\mathcal{A}$; let us call this $\chi_\infty^*(\mathcal{A})$. Since all elements of $\mathcal{A}$ come from a vector $\mathbf{a} \equiv \mathbf{a}_0 \pmod{q^*}$, we also have that $\chi_f^*(\alpha)$ is constant and equal to $\chi_f^*(\sum_{i=1}^{n-k}(\mathbf{a}_0)_i \sqrt[n]{\theta^{i-1}})$ for all ideals $(\alpha) \in \mathcal{A}$. Let $N_0 \asymp_\epsilon X$ be such that the smallest norm of an ideal in $\mathcal{A}$ is N_0^n. We then define the set $\mathcal{B}$ of ideals of $\mathcal{O}_K$ by

$$\begin{aligned}\mathcal{B} &= \mathcal{B}(\mathbf{a}_0)\\ &= \{\text{ideals } \mathfrak{b} \text{ of } \mathcal{O}_K : N(\mathfrak{b}) \in [N_0^n, (1+\eta_1)N_0^n],\ \chi^*(\mathfrak{b}) = \chi_\infty^*(\mathcal{A})\chi_f^*(\alpha_0)\}.\end{aligned}$$

Here $\alpha_0 = \sum_{i=1}^{n-k}(\mathbf{a}_0)_i \sqrt[n]{\theta^{i-1}}$.

By a *polytope* in $\mathbb{R}^\ell$, we mean a bounded region defined by a set of linear inequalities, where the inequalities can be strict, weak or a combination of strict and weak inequalities. Given a polytope $\mathcal{R} \subseteq \mathbb{R}^\ell$ for some ℓ, we define

$$\mathbf{1}_{\mathcal{R}}(\mathfrak{a}) = \begin{cases} 1, & \mathfrak{a} = \mathfrak{p}_1 \dots \mathfrak{p}_\ell \text{ with } N(\mathfrak{p}_i) = X^{e_i},\ (e_1, \dots, e_\ell) \in \mathcal{R}, \\ 0, & \text{otherwise.} \end{cases} \tag{6.3}$$

We see that $\mathbf{1}_{\mathcal{R}}$ is the indicator function of ideals with a particular type of prime ideal factorization, given by the polytope $\mathcal{R}$. Since we are only concerned with $\mathbf{1}_{\mathcal{R}}(\mathfrak{a})$ for $\mathfrak{a} \in \mathcal{A}$ or $\mathfrak{a} \in \mathcal{B}$, we will only consider points with $N(\mathfrak{a}) \in [N_0^n, (1 + O(\eta_1))N_0^n]$, and so we could restrict our attention to polytopes $\mathcal{R}$ with $\sum_{i=1}^{\ell} e_i = n \log N_0 / \log X + O(\eta_1)$. For technical reasons, we find it useful to actually consider larger $\mathcal{R}$ without this restriction which are independent of X, although it is useful to keep in mind the fact that only these points will actually contribute to our final estimates. With this setup, we are now able to state our two key propositions and the main lemmas.

6.2. Key propositions and lemmas.

PROPOSITION 6.1 (Type II sums). *Let $\mathcal{R} \subseteq [\epsilon^2, 2n]^\ell$ be a polytope in $\mathbb{R}^\ell$ such that $(e_1, \dots, e_\ell) \in \mathcal{R} \Rightarrow k + \epsilon \leqslant \sum_{j=1}^{\ell'} e_j \leqslant n - 2k - \epsilon$ for some $\ell' \leqslant \ell$ and*

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

such that $\mathcal{R}$ contains points $\mathbf{x}, \mathbf{y}$ with $\sum_{i=1}^{\ell} x_i > n+\epsilon$, $\sum_{i=1}^{\ell} y_i \leqslant n-\epsilon$. Then we have

$$\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) \ll_{\mathcal{R}} \eta_1^{1/2}\#\mathcal{A}.$$

Here

$$\tilde{\mathfrak{S}} = \prod_{p\nmid q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)\left(1-\frac{\nu_2(p)}{p^n}\right)^{-1},$$

$$\nu(p) = \#\{1\leqslant a_1,\dots,a_{n-k}\leqslant p : N_K(\mathbf{a})\equiv 0 \pmod p\},$$

$$\nu_2(p) = \#\left\{1\leqslant a_1,\dots,a_n\leqslant p : N\left(\sum_{i=1}^{n} a_i\sqrt[n]{\theta^{i-1}}\right)\equiv 0 \pmod p\right\}.$$

PROPOSITION 6.2 (Sieve asymptotic terms). *Let $\mathcal{R}\subseteq[\epsilon^2,2n]^{\ell}$ be a polytope in $\mathbb{R}^{\ell}$ such that $(e_1,\dots e_{\ell})\in\mathcal{R}\Rightarrow\sum_{i=1}^{\ell} e_i < n-k-4\epsilon$, and $\mathcal{R}$ contains points $\mathbf{x}, \mathbf{y}$ with $\sum_{i=1}^{\ell} x_i > n+\epsilon$, $\sum_{i=1}^{\ell} y_i\leqslant n-\epsilon$. Let $X^{\epsilon^2} < N(\mathfrak{a}_1)\leqslant X^{n-3k-4\epsilon}$. Then we have*

$$\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d})\left(S(\mathcal{A}_{\mathfrak{d}},\mathfrak{a}_1) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}S(\mathcal{B}_{\mathfrak{d}},\mathfrak{a}_1)\right)$$
$$\ll_{\mathcal{R}} \frac{\exp(-\epsilon^{-1/2})\#\mathcal{A}}{\log X}\prod_{p\mid q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)^{-1}.$$

Here $\tilde{\mathfrak{S}}$ and $\nu(p)$ are as in Proposition 6.1.

Assuming these two propositions, it is fairly straightforward to establish Proposition 5.1 when $n\geqslant 4k$. We first record a couple of estimates for the set $\mathcal{B}=\mathcal{B}(\mathbf{a}_0)$.

LEMMA 6.3.

$$\#\mathcal{B} = \frac{\gamma_K}{2}\frac{\phi_K((q^*))}{q^{*n}}\eta_1 N_0^n + O(N_0^{n-1+o(1)}).$$

In the lemma above, $\phi_K(\mathfrak{a}) = \#\{\mathfrak{b} \pmod{\mathfrak{a}} : \gcd(\mathfrak{b},\mathfrak{a}) = 1\}$ is Euler's ϕ function for ideals of K.

Proof. This is a simple exercise in counting via Perron's formula, using the bound $L(1-\sigma+it,\chi^{*2})$, $L(1-\sigma+it,\chi^*)\ll((1+|t|)q^*)^{n\sigma/2+\epsilon}$ for $|\sigma+it|\geqslant 1/10$ from Lemma 4.6. Let $c=1+1/\log N_0$ and $T=N_0$. Moving the line of

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

integration to $\Re(s) = 1/2$ gives

$$\begin{aligned}\#\mathcal{B} &= \sum_{\substack{\mathfrak{b} \\ N(\mathfrak{b})\in[N_0^n,(1+\eta_1)N_0^n]}} \frac{\chi^*(\mathfrak{b})^2 + \chi^*(\mathfrak{b})\chi_f^*(\mathbf{a}_0)\chi_\infty^*(\mathcal{A})}{2} \\ &= \frac{1}{2\pi i}\int_{c-iT}^{c+iT} (L(s,(\chi^*)^2) + \chi_f^*(\mathbf{a}_0)\chi_\infty^*(\mathcal{A})L(s,\chi^*))\frac{N_0^{ns}((1+\eta)^s-1)ds}{2s} \\ &\quad + O\left(\frac{N_0^n(\log N_0)^2}{T}\right) \\ &= \frac{\eta_1 N_0^n}{2}\operatorname*{Res}_{s=1}(L(s,(\chi^*)^2)) + O(N_0^{n-1+o(1)}) \\ &= \frac{\gamma_K}{2}\frac{\phi_K((q^*))}{q^{*n}}\eta_1 N_0^n + O(N_0^{n-1+o(1)}).\end{aligned}$$

Here γ_K is the residue of $\zeta_K(s)$ at $s = 1$, and the first summation is over all ideals of $\mathcal{O}_K$ with the norm restriction. □

Trivially, we have that the size of $\mathcal{A}$ is given by $\#\mathcal{A} = (1 + o(1))\,\eta_1^{n-k}q^{*-(n-k)}\prod_{i=1}^{n-k} X_i$ for any choice of $\mathbf{a}_0$.

Finally, we have the following lemmas which show that if we sum over all $\mathbf{a}_0 \in [1, q^*]^{n-k}$, then we remove any distortions caused by a possible exceptional character from primes in $\mathcal{B}$. We delay the proof of Lemma 6.4 to Section 7.

LEMMA 6.4. *Let $\mathcal{R} \subseteq [\epsilon^2, 2n]^\ell$ be a closed polytope which contains points $\mathbf{x}, \mathbf{y}$ with $\sum_{i=1}^\ell x_i > n + \epsilon, \sum_{i=1}^\ell y_i \leqslant n - \epsilon$. Then*

$$\sum_{\substack{\mathbf{a}_0\in[1,q^*]^{n-k} \\ \gcd(N_K(\mathbf{a}_0),q^*)=1}} \sum_{\mathfrak{b}\in\mathcal{B}(\mathbf{a}_0)} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) = \frac{q^{*(n-k)}\eta_1 N_0^n}{2\log X}\prod_{p\mid q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)(I_{\mathcal{R}} + o_{\mathcal{R}}(1)),$$

where

$$I_{\mathcal{R}} = n\int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R} \\ \sum_{i=1}^\ell e_i = n}} \cdots \int \frac{de_1\dots de_{\ell-1}}{e_1\dots e_\ell}.$$

If $\ell = 1$, then $I_{\mathcal{R}}$ is interpreted as 1 if $n \in \mathcal{R}$ and 0 otherwise.

LEMMA 6.5. *Let $\mathcal{R} \subseteq [\epsilon^2, 2n]^\ell$ be a closed polytope. Then*

$$\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}} \tilde{\mathfrak{S}}\frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} \sum_{\mathfrak{b}\in\mathcal{B}(\mathbf{a}_0)} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) = \mathfrak{S}\frac{\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{n\log X}(I_{\mathcal{R}} + o_{\mathcal{R}}(1)),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

where $I_{\mathcal{R}}$ is as in Lemma 6.4. In particular, choosing $\mathcal{R} = [n(1/2+\epsilon), 2n]$, we have

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \tilde{\mathfrak{S}} \frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} S(\mathcal{B}(\mathbf{a}_0), \mathfrak{r}_2) = (1+o(1))\mathfrak{S} \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X}.$$

Proof of Lemma 6.5 assuming Lemma 6.4. We recall that

$$\#\mathcal{A}(\mathbf{a}_0) = (1+o(1)) \frac{\eta_1^{n-k} \prod_{i=1}^{n-k} X_i}{q^{*(n-k)}} = (1+o(1)) \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{q^{*(n-k)}}$$

for all choices of $\mathbf{a}_0$ and that $\#\mathcal{B}(\mathbf{a}_0) = (\gamma_K/2 + o(1))\eta_1 N_0^n \phi_K((q^*))/q^{*n}$ by Lemma 6.3. Since $\mathbf{1}_{\mathcal{R}}(\mathfrak{b})$ is supported on ideals with no factors of small norm, we see that there is no contribution from $\mathbf{a}_0$ with $\gcd(N_K(\mathbf{a}_0), q^*) \neq 1$. We see that the number of choices of $\mathbf{a}_0 \in [1, q^*]^{n-k}$ such that $\mathfrak{a}_0 = (\sum_{i=1}^{n-k} (\mathbf{a}_0)_i \sqrt[n]{\theta^{i-1}})$ has no common ideal factor with (q^*) is given by $q^{*(n-k)} \prod_{p|q^*}(1 - \nu(p)/p^{n-k})$. Thus, by Lemma 6.4 and our estimates for $\#\mathcal{A}(\mathbf{a}_0)$ and $\#\mathcal{B}(\mathbf{a}_0)$, we have that

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \tilde{\mathfrak{S}} \frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} \sum_{\mathfrak{b} \in \mathcal{B}(\mathbf{a}_0)} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})$$
$$= \frac{(I_{\mathcal{R}} + o_{\mathcal{R}}(1)) q^{*n} \tilde{\mathfrak{S}}}{\gamma_K \phi_K((q^*))} \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X} \prod_{p|q^*} \left(1 - \frac{\nu(p)}{p^{n-k}}\right).$$

We recall that $\gamma_K = \prod_p (1 - \nu_2(p) p^{-n})^{-1}(1-p^{-1})$ is the residue at $s=1$ of $\zeta_K(s)$ and so find that

$$\frac{q^{*n}}{\gamma_K \phi_K((q^*))} \tilde{\mathfrak{S}} \prod_{p|q^*} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) = \prod_p \left(1 - \frac{\nu(p)}{p^{n-k}}\right)\left(1 - \frac{1}{p}\right)^{-1} = \mathfrak{S}.$$

Thus, we find that

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \tilde{\mathfrak{S}} \frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} \sum_{\mathfrak{b} \in \mathcal{B}(\mathbf{a}_0)} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) = (I_{\mathcal{R}} + o_{\mathcal{R}}(1)) \mathfrak{S} \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X}. \qquad \square$$

Proof of Proposition 5.1 assuming Propositions 6.1, 6.2, Lemma 6.4 and $n \geqslant 4k$. We first consider $n > 6k$. In this case, $n - 3k - 4\epsilon > n/2 + \epsilon$, so it follows from Proposition 6.2 that (explicitly putting in our dependence on $\mathbf{a}_0$)

$$S(\mathcal{A}(\mathbf{a}_0), \mathfrak{r}_2) = \tilde{\mathfrak{S}} \frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} S(\mathcal{B}(\mathbf{a}_0), \mathfrak{r}_2)$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$+ O\Bigg(\frac{\exp(-\epsilon^{-1/2})\#\mathcal{A}(\mathbf{a}_0)}{\log X}\prod_{p\mid q^*}\Big(1-\frac{\nu(p)}{p^{n-k}}\Big)^{-1}\Bigg).$$

We now sum over the choices of $\mathbf{a}_0$, noting that there is only a contribution from those such that $\mathfrak{a}_0 = (\sum_{i=1}^{n-k}(\mathbf{a}_0)_i\sqrt[n]{\theta^{i-1}})$ has no common ideal factor with (q^*). The number of such $\mathbf{a}_0$ is $(q^*)^{n-k}\prod_{p\mid q^*}(1-\nu(p)/p^{n-k})$, and we recall that $\#\mathcal{A}(\mathbf{a}_0) = (1+o(1))\#(\mathscr{R}\cap\mathbb{Z}^{n-k})q^{*-(n-k)}$. Thus, we obtain

$$\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}} S(\mathcal{A}(\mathbf{a}_0),\mathfrak{r}_2) = \sum_{\mathbf{a}_0\in[1,q^*]^{n-k}} \tilde{\mathfrak{S}}\frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)}S(\mathcal{B}(\mathbf{a}_0),\mathfrak{r}_2)) + O\Bigg(\frac{\exp(-\epsilon^{-1/2})\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{\log X}\Bigg).$$

Lemma 6.5 gives an asymptotic estimate for the main term, giving

$$\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}} S(\mathcal{A}(\mathbf{a}_0),\mathfrak{r}_2) = \mathfrak{S}\frac{\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{n\log X}(1+O(\exp(-\epsilon^{-1/2}))).$$

This gives the result in the case $n > 6k$.

We now consider $6k \geqslant n > 4k$. We see that by Buchstab's identity (this simply applies inclusion–exclusion according to the smallest prime factor), we have

$$S(\mathcal{A},\mathfrak{r}_2) = S(\mathcal{A},\mathfrak{r}_1) - \sum_{\mathfrak{r}_1<\mathfrak{p}\leqslant\mathfrak{r}_2} S(\mathcal{A}_\mathfrak{p},\mathfrak{p}).$$

Applying the same decomposition to $\mathcal{B}$ and subtracting the difference weighted by $\tilde{\mathfrak{S}}\#\mathcal{A}/\#\mathcal{B}$, we see that

$$S(\mathcal{A},\mathfrak{r}_2) = \frac{\tilde{\mathfrak{S}}\#\mathcal{A}}{\#\mathcal{B}}S(\mathcal{B},\mathfrak{r}_2) + \Big(S(\mathcal{A},\mathfrak{r}_1) - \frac{\tilde{\mathfrak{S}}\#\mathcal{A}}{\#\mathcal{B}}S(\mathcal{B},\mathfrak{r}_1)\Big) - \Bigg(\sum_{\mathfrak{r}_1<\mathfrak{p}\leqslant\mathfrak{r}_2} S(\mathcal{A}_\mathfrak{p},\mathfrak{p}) - \frac{\tilde{\mathfrak{S}}\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{r}_1<\mathfrak{p}\leqslant\mathfrak{r}_2} S(\mathcal{B}_\mathfrak{p},\mathfrak{p})\Bigg).$$

By Proposition 6.2, the first term in parentheses is negligible. The second term in parentheses counts ideals with $O(1)$ prime ideal factors, one of which lies between $\mathfrak{r}_1$ and $\mathfrak{r}_2$ and all of which are larger than $\mathfrak{r}_1$. Therefore, splitting the sum according to the number of prime factors, it can be written as a sum of $O(1)$ terms of the form

$$\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_\mathcal{R}(\mathfrak{a}) - \frac{\tilde{\mathfrak{S}}\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}}\mathbf{1}_\mathcal{R}(\mathfrak{b})$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

for some polytope $\mathcal{R}$ satisfying the conditions of Proposition 6.1. Explicitly, we can choose

$$\mathcal{R}_1 = \{\mathbf{e} \in \mathbb{R}^2 : n - 3k - 4\epsilon \leqslant e_1 \leqslant n(1/2+\epsilon),\ e_1 \leqslant e_2,\ e_2 \leqslant n\},$$
$$\mathcal{R}_2 = \{\mathbf{e} \in \mathbb{R}^3 : n - 3k - 4\epsilon \leqslant e_1 \leqslant n(1/2+\epsilon),\ e_1 \leqslant e_2 \leqslant e_3,\ e_3 \leqslant n\},$$
$$\mathcal{R}_3 = \{\mathbf{e} \in \mathbb{R}^4 : n - 3k - 4\epsilon \leqslant e_1 \leqslant n(1/2+\epsilon),\ e_1 \leqslant e_2 \leqslant e_3 \leqslant e_4,\ e_4 \leqslant n\}.$$

(We note that since $n > 4k$ elements with all prime ideal factors bigger than $\mathfrak{r}_1$ can have at most four prime ideal factors.) In particular, it follows from Proposition 6.1 that these terms are negligible. Using Lemma 6.5, we see that this gives (making explicit the dependence of $\mathcal{A}$ on $\mathbf{a}_0$)

$$\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} S(\mathcal{A}(\mathbf{a}_0), \mathfrak{r}_2) = \mathfrak{S}\frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{n \log X}(1 + O(\exp(-\epsilon^{-1/2})))$$

for $n > 4k$.

Finally, we consider the case when $n = 4k$. In this case, we cannot estimate terms $S(\mathcal{A}_\mathfrak{p}, \mathfrak{p})$ with $N(\mathfrak{p}) \in \mathcal{E}$, where $\mathcal{E} = [X^{n-3k-4\epsilon}, X^{k+\epsilon}] \cup [X^{2k-\epsilon}, X^{n/2+\epsilon}]$, since this lies outside the range of our Type II estimates. However, bounding these terms by $0 \leqslant S(\mathcal{A}_\mathfrak{p}, \mathfrak{p}) \leqslant S(\mathcal{A}_\mathfrak{p}, \mathfrak{r}_1)$ introduces a negligible error term to the final estimates since this range of $\mathfrak{p}$ is short. Specifically, letting $\lambda = \lambda(\mathbf{a}_0) = \tilde{\mathfrak{S}}\#\mathcal{A}/\#\mathcal{B}$, we have

$$\begin{aligned}\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} S(\mathcal{A}, \mathfrak{r}_2) = &\sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \lambda S(\mathcal{B}, \mathfrak{r}_2) + \sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} (S(\mathcal{A}, \mathfrak{r}_1) - \lambda S(\mathcal{B}, \mathfrak{r}_1))\\ &- \sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \sum_{X^{k+\epsilon} < N(\mathfrak{p}) \leqslant X^{2k-\epsilon}} (S(\mathcal{A}_\mathfrak{p}, \mathfrak{p}) - \lambda S(\mathcal{B}_\mathfrak{p}, \mathfrak{p}))\\ &+ \sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \sum_{N(\mathfrak{p}) \in \mathcal{E}} O(S(\mathcal{A}_\mathfrak{p}, \mathfrak{r}_1) + \lambda S(\mathcal{B}_\mathfrak{p}, \mathfrak{r}_1)).\end{aligned}$$

As before, the first term in parentheses on the right-hand side is negligible by Proposition 6.2, and the second term in parentheses is negligible by Proposition 6.1. Finally, by Proposition 6.2 and Lemma 6.5, the last term is

$$\ll \sum_{\mathbf{a}_0 \in [1,q^*]^{n-k}} \sum_{N(\mathfrak{p}) \in \mathcal{E}} \frac{\tilde{\mathfrak{S}}\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)} S(\mathcal{B}_\mathfrak{p}(\mathbf{a}_0), \mathfrak{r}_1) \ll \epsilon \frac{\#(\mathscr{R} \cap \mathbb{Z}^{n-k})}{\log X}.$$

Thus, this is negligible, and so using Lemma 6.5 for the main term, we obtain the result. □

When $n < 4k$, we require a more complicated decomposition based on the use of Harman's sieve. Here we discard some terms through positivity, which

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

restricts us to obtaining a lower bound of the correct order of magnitude. Because the ranges of our 'Type I' and 'Type II' estimates are the same as those used in Harman's work on the problem of Diophantine approximation by primes, we could use precisely the same decomposition as Harman uses in [10]. The only minor difference is that in our case, the summations are over prime ideals rather than rational primes, but this does not affect the final estimates since they both have the same density. Instead, since Harman's decomposition is not fully explicit, we have included an explicit description of an adequate decomposition in the appendix to this article along with a Mathematica file performing the relevant numerical computations for this decomposition. The result of this is the following proposition.

PROPOSITION 6.6 (Sieve decomposition for $n < 4k$). *Let $\varpi n > k$. There exist sets $\mathcal{S}_1, \dots, \mathcal{S}_5$ of polytopes which are independent of X such that for any set $\mathcal{C}$ of ideals $\mathfrak{a}$ with $\epsilon X^n < N(\mathfrak{a}) \ll X^n$, we have*

$$S(\mathcal{C}, \mathfrak{r}_2) = \sum_{\mathcal{R}\in\mathcal{S}_1} \sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{r}_1) - \sum_{\mathcal{R}\in\mathcal{S}_2} \sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{r}_1) + \sum_{\mathcal{R}\in\mathcal{S}_3} \sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \sum_{\mathcal{R}\in\mathcal{S}_4} \sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) + \sum_{\mathcal{R}\in\mathcal{S}_5} \sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}).$$

Moreover, the sets $\mathcal{S}_1, \dots, \mathcal{S}_5$ satisfy

(1) $\#\mathcal{S}_i \ll 1$ *for each i.*

(2) *(All terms involve a bounded number of primes factors) Each polytope $\mathcal{R} \in \bigcup_{i=1}^5 \mathcal{S}_i$ lies in $\mathbb{R}^\ell$ for some $\ell \leqslant 1/\epsilon^2$ (but different polytopes may be of different dimensions).*

(3) *(No term involves small prime factors) If $\mathcal{R} \in \bigcup_{i=1}^5 \mathcal{S}_i$ and $(e_1, \dots, e_\ell) \in \mathcal{R}$, then $e_j \geqslant \epsilon^2$ for all $j \in \{1, \dots, \ell\}$.*

(4) *($\mathcal{R}$ does not depend too much on the norms) Each polytope $\mathcal{R} \in \bigcup_{i=1}^5 \mathcal{S}_i$ contains a point $\mathbf{x}$ and a point $\mathbf{y}$ with $\sum_{i=1}^\ell x_i > n + \epsilon$ and $\sum_{i=1}^\ell y_i < n - \epsilon$.*

(5) *($\mathcal{S}_1$ and $\mathcal{S}_2$ correspond to simpler sieve terms) If $\mathcal{R} \in \mathcal{S}_1 \cup \mathcal{S}_2$ and $(e_1, \dots, e_\ell) \in \mathcal{R}$, then $\sum_{i=1}^\ell e_i < n - k - 4\epsilon$.*

(6) *($\mathcal{S}_3$ and $\mathcal{S}_4$ correspond to Type II terms) If $\mathcal{R} \in \mathcal{S}_3 \cup \mathcal{S}_4$ and $(e_1, \dots, e_\ell) \in \mathcal{R}$, then there is some ℓ' such that*

$$k + \epsilon \leqslant \sum_{i=1}^{\ell'} e_i \leqslant n - 2k - \epsilon.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

(7) *(The terms from $\mathcal{S}_5$ do not contribute too much) We have that all $\mathcal{R} \in \mathcal{S}_5$ are closed and*

$$\sum_{\mathcal{R}\in\mathcal{S}_5} I_{\mathcal{R}} < 0.99,$$

where

$$I_{\mathcal{R}} = n \int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R}\\ \sum_{i=1}^{\ell} e_i = n}} \cdots \int \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell}.$$

As mentioned above, the proof of Proposition 6.6 essentially follows from the work of Harman [10], but in the interests of explicitness and verifiability, we have included an alternative proof in the appendix. Since the full decomposition is complicated to write down (and requires nontrivial numerical computation), we just highlight some key details here.

In general, we use two means of transforming terms $S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{a})$ in our decomposition:

(1) Buchstab iterations: Given ideals $\mathfrak{a}_1 < \mathfrak{a}_2$ and $\mathfrak{d}$ with $N(\mathfrak{d}) < X^{n-k-4\epsilon}/N(\mathfrak{a}_2)$, we can apply two Buchstab iterations, which gives

$$S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{a}_2) = S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{a}_1) - \sum_{\mathfrak{a}_1 < \mathfrak{p}_1 \leqslant \mathfrak{a}_2} S(\mathcal{C}_{\mathfrak{d}\mathfrak{p}_1}, \mathfrak{a}_1) + \sum_{\mathfrak{a}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{a}_2} S(\mathcal{C}_{\mathfrak{d}\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{p}_2).$$

If $\mathfrak{a}_1 = \mathfrak{r}_1$, then the first two sums correspond to polytopes in $\mathcal{S}_1$ and $\mathcal{S}_2$. Some of the terms in the final sum will involve factors which fall into our Type II range and so correspond to polytopes in $\mathcal{S}_3$ and $\mathcal{S}_4$; we are left to obtain a suitable estimate for the remaining terms.

(2) Reversal of roles: If $\mathcal{T}$ is a set of ideals $\mathfrak{t}$ satisfying $\mathfrak{b} < \mathfrak{t} < \mathfrak{b}^2$, we can write

$$\sum_{\mathfrak{p}\in\mathcal{T} \text{ prime}} S(\mathcal{A}_{\mathfrak{d}\mathfrak{p}}, \mathfrak{a}) = \sum_{\mathfrak{u}\in\mathcal{U}} S(\mathcal{A}^*_{\mathfrak{u}\mathfrak{d}}, \mathfrak{b}),$$

where $\mathcal{A}^*_{\mathfrak{u}\mathfrak{d}} = \{\mathfrak{t} \in \mathcal{A}_{\mathfrak{u}\mathfrak{d}} : \mathfrak{t} \in \mathcal{T}\}$ and $\mathcal{U} = \{\mathfrak{u} : \mathfrak{p}|\mathfrak{u} \Rightarrow \mathfrak{p} > \mathfrak{a}\}$.

Having applied these transformations in some combination a finite number of times, we produce a decomposition of the required shape

$$S(\mathcal{C}, \mathfrak{r}_2) = \sum_{\mathcal{R}\in\mathcal{S}_1}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{r}_1) - \sum_{\mathcal{R}\in\mathcal{S}_2}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{C}_{\mathfrak{d}}, \mathfrak{r}_1) + \sum_{\mathcal{R}\in\mathcal{S}_3}\sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \sum_{\mathcal{R}\in\mathcal{S}_4}\sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) + \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{a}\in\mathcal{C}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

for some explicit sets $\mathcal{S}_1, \dots, \mathcal{S}_5$ of polytopes $\mathcal{R}$ independent of X and with $\#\mathcal{S}_j \ll 1$. Here we recall that $\mathbf{1}_{\mathcal{R}}(\mathfrak{a})$ is the indicator function of ideals which have a particular shape of prime factorization determined by $\mathcal{R}$. It then requires a numerical verification that for this particular choice of decomposition, we have $\sum_{\mathcal{R}\in\mathcal{S}_5} I_{\mathcal{R}} < 0.99$.

Proof of Proposition 5.1 assuming Propositions 6.1, 6.2, Lemma 6.4 and $n < 4k$. Applying Proposition 6.6 to $\mathcal{A}$, we obtain

$$S(\mathcal{A}, \mathfrak{r}_2) = \sum_{\mathcal{R}\in\mathcal{S}_1}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1) - \sum_{\mathcal{R}\in\mathcal{S}_2}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1) + \sum_{\mathcal{R}\in\mathcal{S}_3}\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \sum_{\mathcal{R}\in\mathcal{S}_4}\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) + \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}). \tag{6.4}$$

The point of this decomposition is that we can obtain asymptotic estimates for the terms coming from polytopes in $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$ and $\mathcal{S}_4$ by a combination of our 'Type I' and 'Type II' estimates, and so we obtain a lower bound for $S(\mathcal{A}, \mathfrak{r}_2)$ by dropping the terms coming from $\mathcal{S}_5$ through positivity. Specifically, the terms from $\mathcal{S}_1$ and $\mathcal{S}_2$ can be estimated using Proposition 6.2 and the terms from $\mathcal{S}_3$ and $\mathcal{S}_4$ can be estimated using Proposition 6.1. It will turn out that since $\sum_{\mathcal{R}\in\mathcal{S}_5} I_{\mathcal{R}} < 1$, we still obtain a positive lower bound for $S(\mathcal{A}, \mathfrak{r}_2)$, giving the result.

Applying the same decomposition of Proposition 6.6 to $\mathcal{B}$ and subtracting these terms multiplied by a constant $\lambda = \tilde{\mathfrak{S}}\#\mathcal{A}/\#\mathcal{B}$ from (6.4) gives

$$\begin{aligned}
S(\mathcal{A}, \mathfrak{r}_2) &= \lambda S(\mathcal{B}, \mathfrak{r}_2) + \sum_{\mathcal{R}\in\mathcal{S}_1}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d})(S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1) - \lambda S(\mathcal{B}_{\mathfrak{d}}, \mathfrak{r}_1)) \\
&\quad - \sum_{\mathcal{R}\in\mathcal{S}_2}\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d})(S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1) - \lambda S(\mathcal{B}_{\mathfrak{d}}, \mathfrak{r}_1)) \\
&\quad + \sum_{\mathcal{R}\in\mathcal{S}_3}\left(\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \lambda \sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\right) \\
&\quad - \sum_{\mathcal{R}\in\mathcal{S}_4}\left(\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \lambda \sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\right) + \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \lambda \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) \\
&\geqslant \lambda\left(S(\mathcal{B}, \mathfrak{r}_2) - \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\right) \\
&\quad - \sum_{\mathcal{R}\in\mathcal{S}_1\cup\mathcal{S}_2}\left|\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d})(S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1) - \lambda S(\mathcal{B}_{\mathfrak{d}}, \mathfrak{r}_1))\right| \\
&\quad - \sum_{\mathcal{R}\in\mathcal{S}_3\cup\mathcal{S}_4}\left|\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \lambda \sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\right|.
\end{aligned} \tag{6.5}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Here we have dropped the nonnegative terms $\sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{R}}(\mathfrak{a})$ for a lower bound.

By Proposition 6.2, the second term on the right-hand side of (6.5) involving a sum over $\mathcal{R}\in\mathcal{S}_1\cup\mathcal{S}_2$ is negligible since $\mathcal{S}_1$ and $\mathcal{S}_2$ only contain polytopes with sum of coordinates at most $n-k-4\epsilon$. Similarly, the last term on the right-hand side of (6.5) involving $\mathcal{R}\in\mathcal{S}_3\cup\mathcal{S}_4$ is negligible by Proposition 6.1 since they only involve polytopes where a subset of the coordinates lies in the Type II range. This gives us the lower bound for $k/\varpi+8\epsilon<n<4k$

$$S(\mathcal{A},\mathfrak{r}_2)\geqslant\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\Big(S(\mathcal{B},\mathfrak{r}_2)-\sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{b}\in\mathcal{B}}\mathbf{1}_{\mathcal{R}}(\mathfrak{b})\Big)+O\Big(\frac{\epsilon\#\mathcal{A}}{\log X}\prod_{p|q^*}\Big(1-\frac{\nu(p)}{p^{n-k}}\Big)^{-1}\Big).$$

We now sum over $\mathbf{a}_0\in[1,q^*]^{n-k}$ such that $(\sum_{i=1}^{n-k}(\mathbf{a}_0)_i\sqrt[n]{\theta^{i-1}})$ has no ideal factor in common with q^*. By Lemma 6.5, we have

$$\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}}\tilde{\mathfrak{S}}\frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)}S(\mathcal{B}(\mathbf{a}_0),\mathfrak{r}_2)=(1+o(1))\mathfrak{S}\frac{\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{n\log X},$$

and for each $\mathcal{R}\in\mathcal{S}_5$,

$$\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}}\tilde{\mathfrak{S}}\frac{\#\mathcal{A}(\mathbf{a}_0)}{\#\mathcal{B}(\mathbf{a}_0)}\sum_{\mathfrak{b}\in\mathcal{B}(\mathbf{a}_0)}\mathbf{1}_{\mathcal{R}}(\mathfrak{b})=(1+o(1))\mathfrak{S}\frac{\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{n\log X}I_{\mathcal{R}}.$$

Putting these estimates together, we obtain

$$\begin{aligned}\sum_{\mathbf{a}_0\in[1,q^*]^{n-k}}S(\mathcal{A}(\mathbf{a}_0),\mathfrak{r}_2)&\geqslant\mathfrak{S}\frac{\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{n\log X}\Big(1-\sum_{\mathcal{R}\in\mathcal{S}_5}I_{\mathcal{R}}+O(\epsilon)\Big)\\&\gg\frac{\mathfrak{S}\#(\mathscr{R}\cap\mathbb{Z}^{n-k})}{\log X}\end{aligned}$$

since, by Proposition 6.6, we have that $\sum_{\mathcal{R}\in\mathcal{S}_5}I_{\mathcal{R}}<0.99$. This gives the result whenever $n>k/\varpi+8\epsilon>22k/7$, as required. □

Thus, to establish Proposition 5.1 and, hence, Theorems 1.1 and 1.2 for $K=\mathbb{Q}(\sqrt[n]{\theta})$, it suffices to prove Lemma 6.4 and Propositions 6.1 and 6.2.

7. Type I sums

In this section, we establish Lemma 6.4 and Proposition 6.2 under the assumption of Proposition 6.1 by using estimates from the geometry of numbers.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

LEMMA 7.1 (Geometry of numbers). *Let $\mathcal{R} \subseteq \mathbb{R}^\ell$ be a region such that any line parallel to the coordinate axes intersects $\mathcal{R}$ in $O(1)$ intervals. Then we have*

$$\#\{\mathbf{a} \in \mathbb{Z}^\ell \cap \mathcal{R}\} = \operatorname{vol} \mathcal{R} + O\Bigg(1 + \sum_{j=1}^{\ell-1} V_j\Bigg),$$

where V_j is the sum of all the $(\ell - j)$-dimensional volumes of the projections of $\mathcal{R}$ formed by equating j coordinates to zero. In particular, if $\mathcal{R}$ is contained in an ℓ-dimensional hypercube of side length B and $\Lambda \subseteq \mathbb{Z}^\ell$ is a rank ℓ lattice with successive minima $Z_1 \leqslant \cdots \leqslant Z_\ell$, then we have

$$\#\{\mathbf{a} \in \Lambda \cap \mathcal{R}\} = \frac{\operatorname{vol} \mathcal{R}}{\det(\Lambda)} + O\Bigg(1 + \sum_{j=1}^{\ell-1} \frac{B^j}{\prod_{i=1}^{j} Z_i}\Bigg).$$

Proof. The first statement is Davenport's theorem [5]. For the second statement, there is a basis $\mathbf{z}_1, \dots, \mathbf{z}_\ell$ of Λ with $\|\mathbf{z}_i\| \asymp Z_i$ and $\|\sum_{i=1}^{\ell} a_i \mathbf{z}_i\| \gg \sum_{i=1}^{\ell} \|a_i \mathbf{z}_i\|$ for any $\mathbf{a} \in \mathbb{R}^\ell$ by Lemma 4.1. Letting M be the $\ell \times \ell$ matrix with columns $\mathbf{z}_1, \dots, \mathbf{z}_\ell$, we see that counting $\mathbf{x} \in \Lambda \cap \mathcal{R}$ is the same as counting $\mathbf{x}' \in \mathbb{Z}^\ell \cap M^{-1}\mathcal{R}$. This region has volume $\operatorname{vol} \mathcal{R}/\det(M) = \operatorname{vol} \mathcal{R}/\det(\Lambda)$. Any point $\mathbf{a} = \sum_{i=1}^{\ell} a_i \mathbf{z}_i$ must be a distance $O(B)$ from the centre $\mathbf{c} = \sum_{i=1}^{\ell} c_i \mathbf{z}_i$ of the hypercube containing $\mathcal{R}$, and so $\sum_{i=1}^{\ell} \|(a_i - c_i)\mathbf{z}_i\| \ll \|\sum_{i=1}^{\ell} (a_i - c_i)\mathbf{z}_i\| \ll B$. This means that a_i is constrained to lie in an interval of length $O(B/Z_i)$, and, hence, in this case, $V_j = O(B^{\ell-j}/\prod_{i=1}^{\ell-j} Z_i)$. □

LEMMA 7.2. *Given $\mathbf{d}, \mathbf{e} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$, let $\mathbf{e} \diamond \mathbf{d}$ be the vector $\mathbf{b}$ such that*

$$\sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}} = \sum_{i=1}^{n} e_i \sqrt[n]{\theta^{i-1}} \times \sum_{i=1}^{n} d_i \sqrt[n]{\theta^{i-1}}$$

and let $\Lambda_{\mathbf{d}}$ be the lattice

$$\Lambda_{\mathbf{d}} = \{\mathbf{e} \in \mathbb{Z}^n : (\mathbf{d} \diamond \mathbf{e})_j = 0 \text{ for } n - k < j \leqslant n\}.$$

Then for any $\mathbf{d} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$,

(1) *$\Lambda_{\mathbf{d}}$ is a rank $n - k$ lattice.*

(2) *$\det(\Lambda_{\mathbf{d}}) \ll \|\mathbf{d}\|^k$.*

Proof. We see that the jth component of $\mathbf{e} \diamond \mathbf{d}$ is $\mathbf{v}_{j,\mathbf{d}} \cdot \mathbf{e}$, where $\mathbf{v}_{j,\mathbf{d}}$ is the jth row in the multiplication-by-$\sum_{i=1}^{n} d_i \sqrt[n]{\theta^{i-1}}$ matrix with respect to the basis $\{\sqrt[n]{\theta^{i-1}}\}_{i=1}^{n}$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The multiplication-by-$\sum_{i=1}^n d_i \sqrt[n]{\theta^{i-1}}$ matrix has determinant $N(\sum_{i=1}^n d_i \sqrt[n]{\theta^{i-1}})$, and so is nonzero for any $\mathbf{d} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$. Thus, the vectors $v_{n-k+1,\mathbf{d}}, \dots, v_{n,\mathbf{d}}$ are k linearly independent vectors, so $\Lambda_{\mathbf{d}}$ is a lattice of rank $n-k$.

Since the components of $v_{j,\mathbf{d}}$ have size $O(\|\mathbf{d}\|)$, the lattice has determinant $\det \Lambda_{\mathbf{d}} \ll \prod_{j=n-k+1}^n \|\mathbf{v}_{j,\mathbf{d}}\| \ll \|\mathbf{d}\|^k$. (This bound follows from considering the dual lattice or is an immediate consequence of Lemma 10.1.) □

LEMMA 7.3. *Let $\mathbf{d} \in (\mathbb{Z}^n \backslash \{\mathbf{0}\}) \cap [-D, D]^n$ and $\Lambda_{\mathbf{d}}$ be as in Lemma 7.2. Let $\mathbf{z}_1(\mathbf{d})$ denote the shortest nonzero vector in $\Lambda_{\mathbf{d}}$. Then we have $\|\mathbf{z}_1(\mathbf{d})\| \ll D^{k/(n-k)}$ and*

$$\#\{\mathbf{d} \in [1, D]^n : \|\mathbf{z}_1(\mathbf{d})\| \leqslant Z\} \ll D^{n-k+o(1)} Z^{n-k}.$$

In particular,

$$\sum_{\|\mathbf{d}\| \leqslant D} \frac{1}{\|\mathbf{z}_1(\mathbf{d})\|^{n-k-1}} \ll D^{n-k+k/(n-k)+o(1)}.$$

Proof. By Lemma 7.2, $\Lambda_{\mathbf{d}}$ has rank $n-k$ and determinant $O(D^k)$ when $\mathbf{d} \in [-D, D]^n$. By Lemma 4.1, if $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{n-k}$ are the successive minima of $\Lambda_{\mathbf{d}}$, then

$$\|\mathbf{z}_1(\mathbf{d})\|^{n-k} = \lambda_1^{n-k} \leqslant \lambda_1 \cdots \lambda_{n-k} \ll \det(\Lambda_{\mathbf{d}}) \ll D^k,$$

so $\|\mathbf{z}_1(\mathbf{d})\| \ll D^{k/(n-k)}$. This gives the first claim.

Since $\mathbf{z}_1(\mathbf{d}) \in \Lambda_{\mathbf{d}}$, we have $(\mathbf{d} \diamond \mathbf{z}_1(\mathbf{d}))_j = 0$ for $n-k < j \leqslant n$. By Lemma 7.8, given $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$, there are at most $\tau(\sum_{i=1}^n x_i \sqrt[n]{\theta^{i-1}}) \ll \|\mathbf{x}\|^{o(1)}$ choices of $\mathbf{d}$ and $\mathbf{z}$ such that $\mathbf{z} \diamond \mathbf{d} = \mathbf{x}$ since $\sum_{i=1}^n z_i \sqrt[n]{\theta^{i-1}}$ and $\sum_{i=1}^n d_i \sqrt[n]{\theta^{i-1}}$ must be divisors of $\sum_{i=1}^n x_i \sqrt[n]{\theta^{i-1}}$. Moreover, such a $\mathbf{x}$ must have $x_j = 0$ for $j > n-k$. Putting this together, for any choice of $Z > 0$, we find that

$$\begin{aligned}\sum_{\substack{\mathbf{d} \in [1,D]^n \\ \|\mathbf{z}_1(\mathbf{d})\| \leqslant Z}} 1 &\leqslant \sum_{\substack{\mathbf{z} \in \mathbb{Z}^n \\ \|\mathbf{z}\| \leqslant Z}} \sum_{\substack{\mathbf{d} \in [1,D]^n \\ (\mathbf{d} \diamond \mathbf{z})_j = 0 \text{ if } j > n-k}} 1 \\ &\leqslant \sum_{\substack{\mathbf{x} \in \mathbb{Z}^{n-k} \\ \|\mathbf{x}\| \ll DZ}} \tau\left(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\right) \\ &\leqslant D^{n-k+o(1)} Z^{n-k+o(1)}.\end{aligned}$$

This gives the second claim.

By considering $\|\mathbf{z}_1(\mathbf{d})\|$ in dyadic intervals $[Z, 2Z]$, we find

$$\sum_{\mathbf{d} \in [1,D]^n} \frac{1}{\|\mathbf{z}_1(\mathbf{d})\|^{n-k-1}} \ll \log D \sup_{Z \ll D^{k/(n-k)}} \frac{1}{Z^{n-k-1}} \sum_{\substack{\mathbf{d} \in [1,D]^n \\ Z \leqslant \|\mathbf{z}_1(\mathbf{d})\| \leqslant 2Z}} 1$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll \sup_{Z \ll D^{k/(n-k)}} D^{n-k+o(1)} Z$$
$$\ll D^{n-k+k/(n-k)+o(1)}.$$

This gives the final claim. □

LEMMA 7.4 (Weak Type I estimate). *Let $\mathfrak{d}$ be an ideal of $\mathcal{O}_K$ with $N(\mathfrak{d})$ coprime to Q. Let $\mathcal{R} \subset [-X, X]^{n-k}$ satisfy the conditions of Lemma 7.1. Then we have*

$$\#\left\{\mathbf{a} \in \mathbb{Z}^{n-k} \cap \mathcal{R} : \mathfrak{d} \mid \left(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}\right),\ \mathbf{a} \equiv \mathbf{a}_0 \pmod{Q}\right\}$$
$$= \frac{\rho(\mathfrak{d}) \operatorname{vol} \mathcal{R}}{N(\mathfrak{d}) Q^{n-k}} + O(N(\mathfrak{d})^n X^{n-k-1}).$$

Here ρ is the function defined by

$$\rho(\mathfrak{d}) = \frac{\#\{\mathbf{a} \in [1, N(\mathfrak{d})]^{n-k} : \mathfrak{d} \mid (\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})\}}{N(\mathfrak{d})^{n-k-1}}.$$

Proof. We split the count into residue classes modulo $QN(\mathfrak{d})$. We note that if $\mathbf{a} \equiv \mathbf{b} \pmod{N(\mathfrak{d})}$, then $\mathfrak{d} \mid (\sum_{i=1}^{n-k} b_i \sqrt[n]{\theta^{i-1}})$ if and only if $\mathfrak{d} \mid (\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})$. Therefore,

$$\sum_{\substack{\mathbf{a} \in \mathbb{Z}^{n-k} \cap \mathcal{R} \\ \mathbf{a} \equiv \mathbf{a}_0 \pmod{Q} \\ \mathfrak{d} \mid (\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})}} 1 = \sum_{\substack{\mathbf{b} \in [1, QN(\mathfrak{d})]^{n-k} \\ \mathbf{b} \equiv \mathbf{a}_0 \pmod{Q} \\ \mathfrak{d} \mid (\sum_{i=1}^{n-k} b_i \sqrt[n]{\theta^{i-1}})}} \sum_{\mathbf{a} \equiv \mathbf{b} \pmod{QN(\mathfrak{d})}} 1.$$

By letting $\mathbf{a} = \mathbf{b} + \mathbf{a}_2 QN(\mathfrak{d})$, we see that the inner sum is over $\mathbf{a}_2 \in \mathbb{Z}^{n-k} \cap \mathcal{R}'$ where $\mathcal{R}' = (\mathcal{R} - \mathbf{b})/QN(\mathfrak{d})$ is a translated and scaled copy of $\mathcal{R}$. Since $\mathcal{R}'$ is contained in a hypercube of side length $X/QN(\mathfrak{d})$, by Lemma 7.1, the inner sum is given by

$$\operatorname{vol} \mathcal{R}' + O\left(1 + \frac{X^{n-k-1}}{Q^{n-k-1} N(\mathfrak{d})^{n-k-1}}\right)$$
$$= \frac{\operatorname{vol} \mathcal{R}}{Q^{n-k} N(\mathfrak{d})^{n-k}} + O\left(1 + \frac{X^{n-k-1}}{Q^{n-k-1} N(\mathfrak{d})^{n-k-1}}\right).$$

Since Q and $N(\mathfrak{d})$ are coprime, there are precisely $N(\mathfrak{d})^{n-k-1} \rho(\mathfrak{d}) \ll N(\mathfrak{d})^n$ terms in the sum over $\mathbf{b}$ by the Chinese remainder theorem. This gives the result. □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

PROPOSITION 7.5 (Type I estimate). *Let $\mathcal{R} = \mathcal{R}(X) \subseteq [-X, X]^{n-k}$ be a region such that any line parallel to the coordinate axes intersects $\mathcal{R}$ in $O(1)$ intervals. Given a vector $\mathbf{a}_0 \in \mathbb{Z}^{n-k}$, and a quantity $Q \leqslant X^{1/2}$, we define the set*

$$\mathscr{C} = \left\{ \sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}} : \mathbf{a} \in \mathbb{Z}^{n-k} \cap \mathcal{R},\ \mathbf{a} \equiv \mathbf{a}_0 \pmod{Q} \right\}.$$

We let $\mathscr{C}_{\mathfrak{d}} = \{\kappa \in \mathscr{C} : \mathfrak{d} | (\kappa)\}$ be the elements of $\mathscr{C}$ which generate an ideal which is a multiple of $\mathfrak{d}$. Then we have

$$\sum_{\substack{N(\mathfrak{d}) \in [D, 2D] \\ \gcd(N(\mathfrak{d}), Q) = 1}} \left| \#\mathscr{C}_{\mathfrak{d}} - \frac{\rho(\mathfrak{d}) \operatorname{vol} \mathcal{R}}{Q^{n-k} N(\mathfrak{d})} \right| \ll X^{n-k-1} Q^{n+o(1)} D^{1/(n-k)+o(1)} + D Q^{n+o(1)}.$$

In particular, taking $\mathcal{R} = [X_1, X_1 + \eta_1 X_1] \times \cdots \times [X_{n-k}, X_{n-k} + \eta_1 X_{n-k}]$, $\mathbf{x}_0 = \mathbf{a}_0$ and $Q = q^*$, we have

$$\sum_{\substack{N(\mathfrak{d}) \in [D, 2D] \\ \gcd(N(\mathfrak{d}), q^*) = 1}} \left| \#\mathcal{A}_{\mathfrak{d}} - \frac{\rho(\mathfrak{d}) \#\mathcal{A}}{N(\mathfrak{d})} \right| \ll X^{n-k-1+o(1)} D^{1/(n-k)+o(1)} + D X^{o(1)}.$$

Here ρ is the function defined by

$$\rho(\mathfrak{d}) = \frac{\#\{\mathbf{a} \in [1, N(\mathfrak{d})]^{n-k} : \mathfrak{d} | (\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})\}}{N(\mathfrak{d})^{n-k-1}}.$$

Proof. We consider separately the contribution from ideals $\mathfrak{d}$ occurring in each class $\mathcal{C} \in Cl_K$. Given a class $\mathcal{C}$, we fix a representative integral ideal $\mathfrak{c} \in \mathcal{C}$ with $\gcd(N(\mathfrak{c}), Q) = 1$. We can choose such an ideal with $N(\mathfrak{c}) = Q^{o(1)}$. (Since Q has $O(\log Q)$ prime factors, there must be a prime ideal in $\mathcal{C}$ with norm coprime to Q amongst the first $O(\log Q)$ prime ideals in $\mathcal{C}$.) We let $(\delta_{\mathfrak{c}})$ be the principal fractional ideal $\mathfrak{d}\mathfrak{c}^{-1}$, where the generator $\delta_{\mathfrak{c}} = \sum_{i=1}^{n} d_i \sqrt[n]{\theta^{i-1}}/(\theta n N(\mathfrak{c}))^n$ is chosen such that $d_i \in \mathbb{Z}$ with $d_i \ll D^{1/n} Q^{o(1)}$. (The d_i can be taken as integers since $\mathfrak{d}\mathfrak{c}^{-1}(N(\mathfrak{c}))$ is integral and $\mathbb{Z}[\sqrt[n]{\theta}]$ is an order in $\mathcal{O}_K$ of index dividing $(\theta n)^n$. The d_i can be chosen to be of size $O(D^{1/n} Q^{o(1)})$ by Lemma 4.2.) We note that $|\delta_{\mathfrak{c}}^{\sigma_0}| = N(\delta_{\mathfrak{c}}) / \prod_{\sigma \neq \sigma_0} |\delta_{\mathfrak{c}}^{\sigma}| \gg D^{1/n} Q^{o(1)}$ for any embedding σ_0.

We see that

$$\begin{aligned} \#\{\alpha \in \mathscr{C} : \mathfrak{d} | (\alpha)\} &= \#\{\alpha \in \mathscr{C} : (\alpha) = \mathfrak{a}' \mathfrak{d} = \mathfrak{a}' \mathfrak{c} \mathfrak{d} \mathfrak{c}^{-1} \text{ for some integral } \mathfrak{a}'\} \\ &= \#\{\beta \in \mathcal{O}_K : \delta_{\mathfrak{c}} \beta \in \mathscr{C},\ \mathfrak{c} | (\beta)\}. \end{aligned}$$

Here we have put β as a generator of the principal ideal $\mathfrak{a}'\mathfrak{c}$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We let $\beta = (\theta n)^{-n}\sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}}$ with $\mathbf{b} \in \mathbb{Z}^n$. All such β have such a representation since $\mathbb{Z}[\sqrt[n]{\theta}]$ is an order in $\mathcal{O}_K$ of index dividing $(\theta n)^n$. Moreover, since $\mathbb{Z}[\sqrt[n]{\theta}] \subseteq \mathcal{O}_K$, provided $\mathbf{b}$ lies in a suitable residue class $\pmod{(\theta n)^n}$, we have that $(\theta n)^{-n}\sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}} \in \mathcal{O}_K$, and so with this restriction on residue classes, the representation is then bijective. We may introduce a further restriction $\pmod{N(\mathfrak{c})^n(\theta n)^{2n}}$ to ensure $\beta\delta_{\mathfrak{c}} \in \mathbb{Z}[\sqrt[n]{\theta}]$ and $\mathfrak{c}|(\beta)$. We now split the count into residue classes mod $q = QN(\mathfrak{c})^n(\theta n)^{2n}$ so that we are left to estimate

$$\sum_{\mathbf{b}_0}{}' \sum_{\substack{\mathbf{b}\in\mathbb{Z}^n \\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q \\ \delta_{\mathfrak{c}}\beta\in\mathscr{C}}} 1. \tag{7.1}$$

Here $\sum'_{\mathbf{b}_0}$ indicates that we sum over $\mathbf{b}_0 \in [1, q]^n$ restricted to the residue classes $\pmod{N(\mathfrak{c})^n(\theta n)^{2n}}$ described above and also such that the coefficient of $\sqrt[n]{\theta^{i-1}}$ in $\beta_0\delta_{\mathfrak{c}} \in \mathbb{Z}[\sqrt[n]{\theta}]$ is congruent to 0 $\pmod{Q}$ for $i > n-k$ and congruent to $(\mathbf{a}_0)_i \pmod{Q}$ for $i \leqslant n-k$.

We concentrate on the inner sum. Recall that $\mathbf{d} \diamond \mathbf{b}$ denotes the vector $\mathbf{e}$ such that $\sum_{i=1}^{n} e_i \sqrt[n]{\theta^{i-1}} = \sum_{i=1}^{n} d_i \sqrt[n]{\theta^{i-1}} \times \sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}}$. Since $\delta_{\mathfrak{c}}\beta \in \mathscr{C}$, we must have that $(\mathbf{d} \diamond \mathbf{b})_j = 0$ for $n-k < j \leqslant n$. Thus, $\mathbf{b}$ is restricted to lie in the lattice $\Lambda_{\mathbf{d}}$ described by Lemma 7.2. If there is no vector $\mathbf{b}^{(1)} \in \Lambda_{\mathfrak{d}}$ such that $\mathbf{b}^{(1)} \equiv \mathbf{b}_0 \pmod q$, then the inner sum of (7.1) is clearly empty. If there is such a vector, we write $\mathbf{b} = \mathbf{b}^{(1)} + q\mathbf{b}^{(2)}$, giving

$$\sum_{\mathbf{b}_0}{}' \sum_{\substack{\mathbf{b}\in\mathbb{Z}^n \\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q \\ \delta_{\mathfrak{c}}\beta\in\mathscr{C}}} 1 = \sum_{\mathbf{b}_0}{}'' \sum_{\substack{\mathbf{b}^{(2)}\in\Lambda_{\mathfrak{d}} \\ \delta_{\mathfrak{c}}\beta_1+q\delta_{\mathfrak{c}}\beta_2\in\mathscr{C}}} 1.$$

Here $\sum''$ indicates that we have the additional condition that such a vector $\mathbf{b}^{(1)}$ exists, and we have put $\beta_1 = (\theta n)^{-n}\sum_{i=1}^{n} b_i^{(1)} \sqrt[n]{\theta^{i-1}}$. The conditions $\mathbf{b}^{(2)} \in \Lambda_{\mathfrak{d}}$ and $\delta_{\mathfrak{c}}\beta_1 + q\delta_{\mathfrak{c}}\beta_2 \in \mathscr{C}$ are equivalent to $\mathbf{b}^{(2)} \in \Lambda_{\mathfrak{d}} \cap \mathcal{R}'$ for some region $\mathcal{R}'$. Since, for any embedding σ, we have $|\delta_{\mathfrak{c}}^{\sigma}| \gg D^{1/n}q^{o(1)}$ and any $\alpha \in \mathscr{C}$ has $|\alpha^{\sigma}| \ll X$, we see that $|(\beta_2 + \beta_1/q)^{\sigma}| \ll XD^{-1/n}q^{-1+o(1)}$. In particular, $\mathcal{R}'$ is contained in a hypercube of side length $XD^{-1/n}q^{-1+o(1)}$. Thus, by Lemma 7.1, we have

$$\sum_{\mathbf{b}^{(2)}\in\Lambda_{\mathfrak{d}}\cap\mathcal{R}'} 1 = \frac{\operatorname{vol}\mathcal{R}'}{\det(\Lambda_{\mathfrak{d}})} + O\left(1 + \frac{X^{n-k-1}}{\|\mathbf{z}_1(\mathfrak{d})\|^{n-k-1}D^{(n-k-1)/n}}\right), \tag{7.2}$$

where $\mathbf{z}_1(\mathfrak{d})$ is the shortest nonzero vector in $\Lambda_{\mathfrak{d}}$. We recall that $\mathcal{R}'$ is the region for $\mathbf{b}^{(2)}$ from the condition that $\delta_{\mathfrak{c}}\beta_1 + q\delta_{\mathfrak{c}}\beta_2 \in \mathscr{C}$. From this, we see that $\operatorname{vol}\mathcal{R}' = \operatorname{vol}\mathcal{R}/f_{\mathfrak{d},q}$ for some quantity $f_{\mathfrak{d},q}$ independent of X. Thus, after summing over

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$\mathbf{b}_0$, we find

$$\#\{\alpha \in \mathscr{C} : \mathfrak{d}|(\alpha)\} = \operatorname{vol}\mathcal{R} \sideset{}{''}\sum_{\mathbf{b}_0} \frac{1}{f_{\mathfrak{d},q}} + O\left(q^n + \frac{q^n X^{n-k-1}}{\|\mathbf{z}_1(\mathfrak{d})\|^{n-k-1} D^{(n-k-1)/n}}\right).$$

We note that the sum over $\mathbf{b}_0$ above depends on q and $\mathfrak{d}$ but not on X or $\mathcal{R}$. Since this holds for all X and $\mathcal{R}$, if X is large compared with q, $\mathfrak{d}$ and $\mathcal{R}$ is the hypercube $[1, X]^{n-k}$, we see that the main term must match that of Lemma 7.4, and so we must have that

$$\sideset{}{''}\sum_{\mathbf{b}_0} \frac{1}{f_{\mathfrak{d},q}} = \frac{\rho(\mathfrak{d})}{N(\mathfrak{d})Q^{n-k}}.$$

Since the above equation is independent of X and $\mathcal{R}$, it must, in fact, hold regardless of the choice of X and $\mathcal{R}$. Thus,

$$\#\{\alpha \in \mathscr{C} : \mathfrak{d}|(\alpha)\} = \operatorname{vol}\mathcal{R}\frac{\rho(\mathfrak{d})}{N(\mathfrak{d})Q^{n-k}} + O\left(q^n + \frac{q^n X^{n-k-1}}{\|\mathbf{z}_1(\mathfrak{d})\|^{n-k-1} D^{(n-k-1)/n}}\right). \tag{7.3}$$

By Lemma 7.3, when summing over $N(\mathfrak{d}) \in [D, 2D]$, the error term in (7.3) contributes a total

$$\begin{aligned}
&\ll \sum_{\mathbf{d} \ll D^{1/n}Q^{o(1)}} \left(q^n + \frac{q^n X^{n-k-1}}{\|\mathbf{z}_1(\mathfrak{d})\|^{n-k-1} D^{(n-k-1)/n}}\right)\\
&\ll Dq^n Q^{o(1)} + q^n \frac{X^{n-k-1}}{D^{(n-k-1)/n}} \sum_{\mathbf{d} \ll D^{1/n}Q^{o(1)}} \frac{1}{\|\mathbf{z}_1(\mathfrak{d})\|^{n-k-1}}\\
&\ll Dq^n Q^{o(1)} + q^n X^{n-k-1} D^{1/(n-k)+o(1)} Q^{o(1)}.
\end{aligned}$$

Recalling that $q \ll Q^{1+o(1)}$, we see that this is

$$\ll X^{n-k-1} Q^{n+o(1)} D^{1/(n-k)+o(1)} + DQ^{n+o(1)}.$$

This gives the result. □

LEMMA 7.6.

$$\sum_{N(\mathfrak{d}) \leqslant D} \left|\#\mathcal{B}_{\mathfrak{d}} - \frac{\#\mathcal{B}}{N(\mathfrak{d})}\right| \ll X^{n-1+o(1)} D^{1/n}.$$

Proof. The proof of Lemma 6.3 shows that the number of ideals $\mathfrak{a}$ of norm at most $Y > X^{\epsilon}$ with $\chi^*(\mathfrak{a}) = \chi^*(\mathfrak{a}_0)$ is (recalling $q^* = X^{o(1)}$)

$$\frac{\gamma_K \phi_K((q^*))}{2q^{*n}} Y + O(Y^{1-1/n+o(1)}),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

where $\gamma_K = \mathrm{Res}_{s=1}\,\zeta_K(s)$. Letting $\mathfrak{b} = \mathfrak{a}\mathfrak{d} \in \mathcal{B}_{\mathfrak{d}}$, this gives

$$\#\mathcal{B}_{\mathfrak{d}} = \#\left\{\mathfrak{a} : \frac{N_0^n}{N(\mathfrak{d})} \leqslant N(\mathfrak{a}) \leqslant \frac{N_0^n(1+\eta_1)}{N(\mathfrak{d})},\ \chi^*(\mathfrak{a}\mathfrak{d}) = \chi^*(\mathfrak{a}_0)\right\}$$
$$= \frac{\eta_1\gamma_K\phi_K((q^*))N_0^n}{2q^{*n}N(\mathfrak{d})} + O\left(\frac{X^{n-1+o(1)}}{N(\mathfrak{d})^{1-1/n}}\right).$$

Applying this also with $\mathfrak{d} = (1)$, we see that the main term above is $(\#\mathcal{B} + O(X^{n-1+o(1)}))/N(\mathfrak{d})$. Summing over $\mathfrak{d}$ then gives the result. □

Recall from Proposition 7.5 that ρ is defined by

$$\rho(\mathfrak{d}) = \frac{\#\{\mathbf{a} \in [1, N(\mathfrak{d})]^{n-k} : \mathfrak{d} | (\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}})\}}{N(\mathfrak{d})^{n-k-1}}.$$

We wish to establish some basic properties of this function.

LEMMA 7.7. (i) $\rho(\mathfrak{p}) = 1$ *for any degree one prime ideal* $\mathfrak{p} \nmid (\theta n)$.

(ii) *We have*

$$\#\left\{\mathbf{x} \in [1, p^2]^{n-k} : p^2 | N\left(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\right)\right\} \ll p^{2n-2k-2}.$$

In particular, for any ideal with $N(\mathfrak{e})$ *a power of* p, *we have*

$$\frac{\rho(\mathfrak{e})}{N(\mathfrak{p})} \ll \frac{1}{p^2}$$

unless $\mathfrak{e}$ *is a degree 1 prime ideal above* p.

(iii) $\rho(\mathfrak{a}\mathfrak{b}) = \rho(\mathfrak{a})\rho(\mathfrak{b})$ *if* $\gcd(N(\mathfrak{a}), N(\mathfrak{b})) = 1$.

Proof. (i) Let $N(\mathfrak{p}) = p \nmid \theta n$, so $\mathbb{Z}[\sqrt[n]{\theta}]/p\mathbb{Z}[\sqrt[n]{\theta}] \cong \mathcal{O}_K/p\mathcal{O}_K$. There exists $\mathbf{a} \in [1, p]^n$ such that $\mathfrak{p} | (\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}})$ but $p^2 \nmid N(\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}})$ since there are asymptotically more ideals which are a multiple of $\mathfrak{p}$ than there are ideals having norm a multiple of p^2, by Lemma 7.6. But then the multiplication-by-$\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}$ matrix (with respect to the basis $\{1, \sqrt[n]{\theta}, \ldots, \sqrt[n]{\theta^{n-1}}\}$) has determinant a multiple of p but not of p^2, and so has rank $n-1$ over $\mathbb{F}_p$. This means the p^{n-1} distinct multiples of $\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}$ in $\mathbb{Z}[\sqrt[n]{\theta}]/p\mathbb{Z}[\sqrt[n]{\theta}]$ are all the elements of $\mathcal{O}_K/p\mathcal{O}_K$ which generate an ideal which is a multiple of $\mathfrak{p}$. In addition, the condition that $\sum_{i=1}^n b_i \sqrt[n]{\theta^{i-1}}$ is congruent modulo p to a multiple

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

of $\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}$ is equivalent to $\mathbf{c}_{\mathfrak{p}} \cdot \mathbf{b} \equiv 0 \pmod p$ for some integer vector $\mathbf{c}_{\mathfrak{p}}$ since the multiplication-by-$\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}$ matrix has kernel of rank 1. Therefore, $\rho(\mathfrak{p})$ counts the number of $\mathbf{x} \in [1, p]^{n-k} \times \{0\}^k$ such that $\mathbf{c}_{\mathfrak{p}} \cdot \mathbf{x} = 0 \pmod p$. But $N(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})$ has no fixed prime divisor, so $\mathbf{c}_{\mathfrak{p}} \cdot \mathbf{x}$ cannot vanish for all $\mathbf{x} \in [1, p]^{n-k} \times \{0\}^k$. Thus, there are exactly p^{n-k-1} such $\mathbf{x}$, giving $\rho(\mathfrak{p}) = 1$.

(ii) $N_K(\mathbf{x})$ is a nonzero polynomial in x_1 since the leading term is x_1^n. Moreover, the resultant of $N_K(\mathbf{x})$ and $\frac{\partial}{\partial x_1} N_K(\mathbf{x})$ (viewed as polynomials in x_1) is a nonzero polynomial in $x_2, \dots, x_{n-k}$ since N_K is separable. Both of these are therefore nonzero polynomials over $\mathbb{F}_p$ for p sufficiently large. Thus, there are $O(p^{n-k-1})$ choices of $x_2, \dots, x_{n-k} \pmod p$ such that the resultant is $0 \pmod p$, and for any such choice, there are $O(1)$ values of $x_1 \pmod p$ with $N_K(\mathbf{x}) \equiv 0 \pmod p$. These constraints give rise to $O(p^{2n-2k-2})$ choices of $\mathbf{x} \pmod{p^2}$. Alternatively, if the resultant is nonzero, then for any such choice of $x_2, \dots, x_{n-k}$, there are $O(1)$ choices of $x_1 \pmod p$ such that $N(\mathbf{x}) \equiv 0 \pmod p$, and all of these choices of x_1 lift (by Hensel's lemma) to a unique $x_1 \pmod{p^2}$ such that $N_K(\mathbf{x}) \equiv 0 \pmod{p^2}$. Thus, in either case, there are $O(p^{2n-2k-2})$ choices. The result follows.

(iii) This follows immediately from the Chinese remainder theorem. □

LEMMA 7.8 (Divisor bound for $\mathcal{O}_K$). *For any positive integer m, we have*

$$\sum_{\substack{\|\mathbf{x}\| \ll X \\ x_j = 0 \text{ if } j > n-k}} \tau\left(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\right)^m \ll X^{n-k} (\log X)^{O_m(1)}.$$

Proof. By [13, Lemma 4.4], given any integer $r > 0$, an ideal $\mathfrak{a}$ has an ideal factor $\mathfrak{b} | \mathfrak{a}$ with $N(\mathfrak{b}) \leqslant N(\mathfrak{a})^{1/r}$ and $\tau(\mathfrak{a}) \leqslant 2^{r-1} \tau(\mathfrak{b})^{2r-1}$. Thus, taking $r = n^2$, we have

$$\begin{aligned}
\sum_{\substack{\|\mathbf{x}\| \ll X \\ x_j = 0 \text{ if } j > n-k}} \tau\left(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\right)^m &\ll \sum_{N(\mathfrak{d}) \ll X^{1/n}} \tau(\mathfrak{d})^{2mn^2} \sum_{\substack{\|\mathbf{x}\| \ll X \\ x_j = 0 \text{ if } j > n-k \\ \mathfrak{d} | (\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})}} 1 \\
&\ll \sum_{N(\mathfrak{d}) \ll X^{1/n}} \tau(\mathfrak{d})^{2mn^2} \rho(\mathfrak{d}) N(\mathfrak{d})^{n-k-1} \left(\frac{X^{n-k}}{N(\mathfrak{d})^{n-k}} + O(X^{n-k-1}) \right) \\
&\ll X^{n-k} \sum_{N(\mathfrak{d}) < X^{1/n}} \frac{\tau(\mathfrak{d})^{2mn^2} \rho(\mathfrak{d})}{N(\mathfrak{d})}.
\end{aligned}$$

Here we bounded the number of $\mathbf{x}$ with $\mathfrak{d} | (\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})$ trivially by splitting

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

the x_i into arithmetic progressions $\pmod{N(\mathfrak{d})}$. The sum over $\mathfrak{d}$ is then bounded by

$$\prod_{N(\mathfrak{p})\leqslant X^{1/n}}\left(1+\frac{2^{2mn^2}\rho(\mathfrak{p})}{N(\mathfrak{p})}+O\left(\frac{1}{N(\mathfrak{p})^2}\right)\right)\ll\prod_{p<X^{1/n}}\left(1+\frac{2^{2mn^2}\nu_p}{p}+O\left(\frac{1}{p^2}\right)\right)\ll(\log X)^{O_m(1)},$$

by Lemma 7.7. □

LEMMA 7.9 (Fundamental lemma). *Let $\mathfrak{z}_0$ be chosen maximally with $N(\mathfrak{z}_0)\leqslant X^{\epsilon^2}$. Then we have*

$$\sum_{\substack{N(\mathfrak{d})<X^{n-k-\epsilon}\\ \mathfrak{p}|\mathfrak{d}\Rightarrow\mathfrak{p}>\mathfrak{z}_0}}\tau(\mathfrak{d})\left|S(\mathcal{A}_\mathfrak{d},\mathfrak{z}_0)-\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}S(\mathcal{B}_\mathfrak{d},\mathfrak{z}_0)\right|\ll\frac{\exp(-\epsilon^{-2/3})}{\log X}\#\mathcal{A}\prod_{p|q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)^{-1}.$$

Here $\tilde{\mathfrak{S}}=\prod_{p\nmid q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)\left(1-\frac{\nu_2(p)}{p^n}\right)^{-1}$ is as in Proposition 6.1.

Proof. We first relate the estimate to a sieving problem over $\mathbb{Q}$, where the result then follows from the classical 'fundamental lemma' of sieve methods. We have

$$\begin{aligned}S(\mathcal{A}_\mathfrak{d},\mathfrak{z}_0)&=\#\{\mathfrak{a}\in\mathcal{A}_\mathfrak{d}:\mathfrak{p}|\mathfrak{a}\Rightarrow\mathfrak{p}>\mathfrak{z}_0\}\\&=\#\{\mathfrak{a}\in\mathcal{A}_\mathfrak{d}:p|N(\mathfrak{a})\Rightarrow p>X^{\epsilon^2}\}\\&\quad+O\Bigg(\sum_{p\in[X^{\epsilon^2/n},X^{\epsilon^2}]}\#\{\mathfrak{a}\in\mathcal{A}_\mathfrak{d}:p^2|N(\mathfrak{a})\}\Bigg).\end{aligned}$$

By Proposition 7.5 and Lemma 7.7, the final term is $O(X^{n-k-\epsilon^2/n})$. The first term is a classical sieve quantity.

Define a function ρ_2 on primes by

$$\rho_2(p)=\frac{\#\{\mathbf{a}\in[1,p^n]^{n-k}:\ p|N(\sum_{i=1}^{n-k}a_i\sqrt[n]{\theta^{i-1}})\}}{p^{n(n-k)}}=\frac{\nu(p)}{p^{n-k}}$$

and extend ρ_2 to a function on square-free integers by multiplicativity. By inclusion–exclusion, we have that

$$\rho_2(p)=\sum_{\substack{\mathfrak{p}\\ p|N(\mathfrak{p})}}\frac{\rho(\mathfrak{p})}{N(\mathfrak{p})}-\sum_{\substack{\mathfrak{p}_1<\mathfrak{p}_2\\ p|N(\mathfrak{p}_1),N(\mathfrak{p}_2)}}\frac{\rho(\mathfrak{p}_1\mathfrak{p}_2)}{N(\mathfrak{p}_1\mathfrak{p}_2)}+\cdots.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

For a square-free integer e satisfying $\gcd(e, N(\mathfrak{d})) = 1$ and $\gcd(q^*, eN(\mathfrak{d})) = 1$, we define $R_{\mathfrak{d}}(e)$ by

$$R_{\mathfrak{d}}(e) = \#\{\mathfrak{a} \in \mathcal{A}_{\mathfrak{d}} : e | N(\mathfrak{a})\} - \frac{\rho_2(e)\rho(\mathfrak{d})\#\mathcal{A}}{N(\mathfrak{d})}.$$

We see from the inclusion–exclusion formula above that

$$R_{\mathfrak{d}}(e) \ll \sum_{\substack{\mathfrak{e} \\ e|N(\mathfrak{e}) \\ N(\mathfrak{e})|e^n}} \mu^2(\mathfrak{e}) \left| \mathcal{A}_{\mathfrak{d}\mathfrak{e}} - \frac{\rho(\mathfrak{d}\mathfrak{e})}{N(\mathfrak{d}\mathfrak{e})}\#\mathcal{A} \right|.$$

Thus, by Proposition 7.5, the error terms $R_{\mathfrak{d}}(e)$ satisfy

$$\begin{aligned} &\sum_{\substack{N(\mathfrak{d}) < X^{n-k-\epsilon} \\ \mathfrak{p}|\mathfrak{d} \Rightarrow \mathfrak{p} > \mathfrak{z}_0}} \sum_{\substack{e < X^{\epsilon/(2n^2)} \\ \gcd(e, q^* N(\mathfrak{d}))=1}} \tau(\mathfrak{d})\mu^2(e)|R_{\mathfrak{d}}(e)| \\ &\qquad \leqslant \sum_{N(\mathfrak{d}) < X^{n-k-\epsilon}} \sum_{\substack{N(\mathfrak{e}) < X^{\epsilon/2n} \\ \gcd(N(\mathfrak{e}), q^* N(\mathfrak{d}))=1}} X^{o(1)} \left| \#\mathcal{A}_{\mathfrak{d}\mathfrak{e}} - \frac{\rho(\mathfrak{d}\mathfrak{e})}{N(\mathfrak{d}\mathfrak{e})}\#\mathcal{A} \right| \\ &\qquad \ll X^{n-k-\epsilon/2n+o(1)}. \end{aligned} \tag{7.4}$$

Here we used the divisor bound $\tau(\mathfrak{d}) < X^{o(1)}$ in the second line and Proposition 7.5 in the final line. We note that $\rho_2(p) = \nu(p)/p^{n-k} = \nu_p/p + O(p^{-2})$ by Lemma 7.7, where ν_p is the number of degree one prime ideals of $\mathcal{O}_K$ above p. By the fundamental lemma of sieve methods (see, for example, [**8**, Theorem 6.9]) and the bound (7.4), we have

$$\begin{aligned} &\sum_{\substack{N(\mathfrak{d}) < X^{n-k-\epsilon} \\ \mathfrak{p}|\mathfrak{d} \Rightarrow \mathfrak{p} > \mathfrak{z}_0}} \tau(\mathfrak{d}) \left| S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{z}_0) - \frac{\rho(\mathfrak{d})\#\mathcal{A}}{N(\mathfrak{d})} \prod_{\substack{p < X^{\epsilon^2} \\ p \nmid q^*}} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) \right| \\ &\qquad \ll \exp(-\epsilon^{-1}) \prod_{\substack{p < X^{\epsilon^2} \\ p \nmid q^*}} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) \#\mathcal{A} \sum_{\substack{N(\mathfrak{d}) < X^{n-k-\epsilon} \\ \mathfrak{p}|\mathfrak{d} \Rightarrow \mathfrak{p} > \mathfrak{z}_0}} \frac{\tau(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})} \\ &\qquad + O(X^{n-k-\epsilon/2n+o(1)}). \end{aligned} \tag{7.5}$$

Here we used the fact that $\mathfrak{d}$ has no prime factors with norm $\leqslant N(\mathfrak{z}_0)$ and so must satisfy $\gcd(q^*e, N(\mathfrak{d})) = 1$ since $q^* \leqslant X^{o(1)}$, and we only consider e with prime factors $p \leqslant N(\mathfrak{z}_0)$.

The sum over $\mathfrak{d}$ in the final bound is then easily seen to be $O(\epsilon^{-4})$ by an Euler product upper bound and Lemma 7.7.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We now replace $\rho(\mathfrak{d})$ with the constant 1 in the main term of (7.5). Since $\rho(\mathfrak{p}) = 1$ on degree 1 prime ideals and $\mathfrak{d}$ is restricted to prime factors $\mathfrak{p} > \mathfrak{z}_0$, by Lemma 7.7, we have that

$$\sum_{\substack{N(\mathfrak{d})<X^{n-k-\epsilon}\\ \mathfrak{p}\mid\mathfrak{d}\Rightarrow\mathfrak{p}>\mathfrak{z}_0}} \frac{\tau(\mathfrak{d})|\rho(\mathfrak{d})-1|}{N(\mathfrak{d})} \ll X^{o(1)}\Bigg(\prod_{p>X^{\epsilon^2/n}}\Big(1+\frac{O(1)}{p^2}\Big)-1\Bigg) \ll X^{-\epsilon^2/2n},$$

so this change introduces a negligible error term. Thus, since $\mathfrak{S} = \prod_p(1-\nu(p)p^{-(n-k)})(1-p^{-1})^{-1} = O(1)$, we have

$$\sum_{\substack{N(\mathfrak{d})<X^{n-k-\epsilon}\\ \mathfrak{p}\mid\mathfrak{d}\Rightarrow\mathfrak{p}>\mathfrak{z}_0}} \tau(\mathfrak{d})\Bigg|S(\mathcal{A}_{\mathfrak{d}},\mathfrak{z}_0)-\frac{\#\mathcal{A}}{N(\mathfrak{d})}\prod_{\substack{p<X^{\epsilon^2}\\ p\nmid q^*}}\Big(1-\frac{\nu(p)}{p^{n-k}}\Big)\Bigg| \ll \frac{\exp(-\epsilon^{-2/3})\#\mathcal{A}}{\log X}\prod_{p\mid q^*}\Big(1-\frac{\nu(p)}{p^{n-k}}\Big)^{-1}.$$

An identical argument works for the sets $\mathcal{B}_{\mathfrak{d}}$, with $\nu_2(p)/p^n$ instead of $\nu(p)/p^{n-k}$. Subtracting these expressions and noting the main terms cancel, we have

$$\sum_{\substack{N(\mathfrak{d})<X^{n-k-\epsilon}\\ \mathfrak{p}\mid\mathfrak{d}\Rightarrow\mathfrak{p}>\mathfrak{z}_0}} \tau(\mathfrak{d})\Big|S(\mathcal{A}_{\mathfrak{d}},\mathfrak{z}_0)-\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}S(\mathcal{B}_{\mathfrak{d}},\mathfrak{z}_0)\Big| \ll \frac{\exp(-\epsilon^{-2/3})}{\log X}\#\mathcal{A}\prod_{p\mid q^*}\Big(1-\frac{\nu(p)}{p^{n-k}}\Big)^{-1}.$$

□

Using Lemma 7.9, we can now prove Proposition 6.2, assuming Proposition 6.1.

Proof of Proposition 6.2 assuming Proposition 6.1. To ease notation, let $\mathfrak{a}_0$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ be chosen maximally with respect to our ordering of ideals subject to $N(\mathfrak{a}_0) \leqslant X^{\epsilon^2}$, $N(\mathfrak{a}_2) \leqslant X^{k+2\epsilon}$ and $N(\mathfrak{a}_3) \leqslant X^{n-2k-2\epsilon}$, and let $\mathfrak{a}_1$ be as in the statement of the proposition. We see from this choice that $\mathfrak{a}_0 = \mathfrak{z}_0$ defined previously and that $N(\mathfrak{a}_1) \leqslant X^{n-3k-4\epsilon}$ so that $\mathfrak{a}_1\mathfrak{a}_2 \leqslant \mathfrak{a}_3$. We first consider the contribution from $\mathfrak{d} < \mathfrak{a}_2$. Given a set of ideals $\mathcal{C}$, we let

$$T_m(\mathcal{C};\mathfrak{d}) = \sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m\leqslant\mathfrak{a}_2}} S(\mathcal{C}_{\mathfrak{p}_1\ldots\mathfrak{p}_m},\mathfrak{a}_0),$$

$$U_m(\mathcal{C};\mathfrak{d}) = \sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m\leqslant\mathfrak{a}_2}} S(\mathcal{C}_{\mathfrak{p}_1\ldots\mathfrak{p}_m},\mathfrak{p}_m),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$V_m(\mathcal{C};\mathfrak{d}) = \sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{a}_2<\mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m\leqslant\mathfrak{a}_2\mathfrak{p}_m}} S(\mathcal{C}_{\mathfrak{p}_1\ldots\mathfrak{p}_m},\mathfrak{p}_m).$$

Since $\mathfrak{a}_2\mathfrak{a}_1\leqslant\mathfrak{a}_3$, all products $\mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m$ occurring in $V_m(\mathcal{C};\mathfrak{d})$ lie in our Type II range between $\mathfrak{a}_2$ and $\mathfrak{a}_3$.

By Buchstab's identity, we have that

$$U_m(\mathcal{C};\mathfrak{d}) = T_m(\mathcal{C};\mathfrak{d}) - U_{m+1}(\mathcal{C};\mathfrak{d}) - V_{m+1}(\mathcal{C};\mathfrak{d}).$$

We define $T_0(\mathcal{C};\mathfrak{d}) = S(\mathcal{C};\mathfrak{a}_0)$ and $V_0(\mathcal{C};\mathfrak{d}) = 0$. This gives

$$S(\mathcal{C},\mathfrak{a}_1) = T_0(\mathcal{C};\mathfrak{d}) - V_1(\mathcal{C};\mathfrak{d}) - U_1(\mathcal{C};\mathfrak{d}) = \sum_{m\geqslant 0}(-1)^m(T_m(\mathcal{C};\mathfrak{d}) + V_m(\mathcal{C};\mathfrak{d})).$$

We apply the above decomposition to $\mathcal{A}_\mathfrak{d}$. This gives an expression with $O(\epsilon^{-2})$ terms since trivially $T_m(\mathcal{A}_\mathfrak{d}) = U_m(\mathcal{A}_\mathfrak{d}) = V_m(\mathcal{A}_\mathfrak{d}) = 0$ if $m > n/\epsilon^2$. Applying the same decomposition to $S(\mathcal{B}_\mathfrak{d},\mathfrak{a}_1)$, subtracting the difference weighted by $\lambda = \tilde{\mathfrak{S}}\#\mathcal{A}/\#\mathcal{B}$ and summing over $\mathfrak{d}<\mathfrak{a}_2$ with $\mathbf{1}_\mathcal{R}(\mathfrak{d})\neq 0$, we obtain

$$\begin{aligned}\sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_\mathcal{R}(\mathfrak{d})&(S(\mathcal{A}_\mathfrak{d},\mathfrak{a}_1) - \lambda S(\mathcal{B}_\mathfrak{d},\mathfrak{a}_1))\\ &\ll \sum_{0\leqslant m\leqslant n/\epsilon^2}\sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_\mathcal{R}(\mathfrak{d})|T_m(\mathcal{A}_\mathfrak{d};\mathfrak{d}) - \lambda T_m(\mathcal{B}_\mathfrak{d};\mathfrak{d})|\\ &\quad + \sum_{0\leqslant m\leqslant n/\epsilon^2}\Big|\sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_\mathcal{R}(\mathfrak{d})(V_m(\mathcal{A}_\mathfrak{d};\mathfrak{d}) - \lambda V_m(\mathcal{B}_\mathfrak{d};\mathfrak{d}))\Big|. \qquad (7.6)\end{aligned}$$

For the first term on the right-hand side of (7.6), we expand T_m as a sum, giving

$$\begin{aligned}\sum_{\mathfrak{d}<\mathfrak{a}_2}&\mathbf{1}_\mathcal{R}(\mathfrak{d})|T_m(\mathcal{A}_\mathfrak{d};\mathfrak{d}) - \lambda T_m(\mathcal{B}_\mathfrak{d};\mathfrak{d})|\\ &\leqslant \sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_\mathcal{R}(\mathfrak{d})\sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m\leqslant\mathfrak{a}_2}}|S(\mathcal{A}_{\mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m},\mathfrak{a}_0) - \lambda S(\mathcal{B}_{\mathfrak{d}\mathfrak{p}_1\ldots\mathfrak{p}_m},\mathfrak{a}_0)|.\end{aligned}$$

We put $\mathfrak{d}' = \mathfrak{p}_1\ldots\mathfrak{p}_m\mathfrak{d}$ and note that any given $\mathfrak{d}'$ occurs at most $\epsilon^{-2}\tau(\mathfrak{d}')$ times in the sum above and satisfies $\mathfrak{d}'\leqslant\mathfrak{a}_2$. Thus, using Lemma 7.9, we have

$$\begin{aligned}\sum_{\mathfrak{d}<\mathfrak{a}_2}&\mathbf{1}_\mathcal{R}(\mathfrak{d})|T_m(\mathcal{A}_\mathfrak{d};\mathfrak{d}) - \lambda T_m(\mathcal{B}_\mathfrak{d};\mathfrak{d})|\\ &\ll \sum_{\substack{\mathfrak{d}'\leqslant\mathfrak{a}_2\\ \mathfrak{p}|\mathfrak{d}'\Rightarrow\mathfrak{p}>\mathfrak{a}_0}}\epsilon^{-2}\tau(\mathfrak{d}')|S(\mathcal{A}_{\mathfrak{d}'},\mathfrak{a}_0) - \lambda S(\mathcal{B}_{\mathfrak{d}'},\mathfrak{a}_0)|\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll \epsilon^{-2}\frac{\exp(-\epsilon^{-2/3})\#\mathcal{A}}{\log X}\prod_{p|q^*}\left(1-\frac{\nu(p)}{p^{n-k}}\right)^{-1}.$$

For the second term on the right-hand side of (7.6), we expand V_m and $S(\mathcal{A}_{\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m}, \mathfrak{p}_m)$. For the part of the inner sum involving $\mathcal{A}_{\mathfrak{d}}$, this gives

$$\begin{aligned}\sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})V_m(\mathcal{A}_{\mathfrak{d}};\mathfrak{d}) &= \sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})\sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{a}_2<\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m\leqslant\mathfrak{a}_2\mathfrak{p}_m}} S(\mathcal{A}_{\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m},\mathfrak{p}_m)\\ &= \sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})\sum_{\substack{\mathfrak{a}_0<\mathfrak{p}_m\leqslant\cdots\leqslant\mathfrak{p}_1\leqslant\mathfrak{a}_1\\ \mathfrak{a}_2<\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m\leqslant\mathfrak{a}_2\mathfrak{p}_m}}\sum_{\substack{\mathfrak{a}\\ \mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m\mathfrak{a}\in\mathcal{A}\\ \mathfrak{p}|\mathfrak{a}\Rightarrow\mathfrak{p}>\mathfrak{p}_m}} 1.\end{aligned}$$

Since $\mathfrak{a}$ occurring in the sum above has all prime ideal factors bigger than $\mathfrak{a}_0$, it has $O(\epsilon^{-2})$ prime factors constrained only to be larger than $\mathfrak{p}_m$. Thus, we may rewrite the above expression as

$$\sum_{\mathcal{R}'}\sum_{\mathfrak{a}'\in\mathcal{A}}\mathbf{1}_{\mathcal{R}'}(\mathfrak{a}),$$

where $\mathcal{R}'$ ranges over $O(\epsilon^{-2})$ polytopes describing the possible prime factorizations of $\mathfrak{a}$, all independent of X. Each polytope is in $[\epsilon^2, 2n]^\ell$ for some $\ell \ll \epsilon^{-2}$. Moreover, by ordering the coordinates such that the first ℓ' coordinates correspond to the factor $\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m$, we see that $(e_1,\dots,e_\ell)\in\mathcal{R}\Rightarrow k+\epsilon\leqslant\sum_{i=1}^{\ell'}e_i\leqslant n-2k-\epsilon$ since $\mathfrak{a}_2<\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m\leqslant\mathfrak{a}_3$. Applying the same manipulations to $\lambda V_m(\mathcal{B}_{\mathfrak{d}};\mathfrak{d})$, we find

$$\sum_{\mathfrak{d}<\mathfrak{a}_2}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})(V_m(\mathcal{A}_{\mathfrak{d}};\mathfrak{d})-\lambda V_m(\mathcal{B}_{\mathfrak{d}};\mathfrak{d}))\ll\sum_{\mathcal{R}'}\left|\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{R}'}(\mathfrak{a})-\lambda\sum_{\mathfrak{b}\in\mathcal{B}}\mathbf{1}_{\mathcal{R}'}(\mathfrak{b})\right|.$$

By Proposition 6.1, this is $O(\#\mathcal{A}(\log X)^{-10})$. This completes the proof for $\mathfrak{d}\leqslant\mathfrak{a}_2$.

The contribution from $\mathfrak{d}$ with $\mathfrak{d}\geqslant\mathfrak{a}_3$ and $N(\mathfrak{d})\leqslant X^{2k+2\epsilon}$ can be handled by an essentially identical argument. Let $\mathfrak{b}_2,\mathfrak{b}_3$ be chosen maximally such that $N(\mathfrak{b}_2)\leqslant X^{2k+2\epsilon}$ and $N(\mathfrak{b}_3)\leqslant X^{n-k-2\epsilon}$ and let T_m', U_m', V_m' be T_m, U_m, V_m with $\mathfrak{a}_2$ replaced by $\mathfrak{b}_2$ in the conditions on the summation. Applying an analogous decomposition to the argument above, it suffices to handle only the terms corresponding to T_m' and V_m'. Since $\mathfrak{b}_2\mathfrak{a}_1\leqslant\mathfrak{b}_3$, all products $\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m$ occurring in V_m' lie in the range $[\mathfrak{b}_2,\mathfrak{b}_3]$. In particular, if $\mathfrak{a}\in\mathcal{A}_{\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m}$ for such a product $\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m$, then $\mathfrak{a}=\mathfrak{a}'\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m$ for some ideal $\mathfrak{a}'$ with $N(\mathfrak{a}')\in[X^{k+\epsilon},X^{n-2k-\epsilon}]$. Such sums can be handled by our Type II estimate given by Proposition 6.1. Similarly, any product

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m$ occurring in T_m' satisfies $\mathfrak{d}\mathfrak{p}_1\dots\mathfrak{p}_m \leqslant \mathfrak{b}_2$, and so the terms T_m' can be handled by our Type I estimate given by Proposition 7.5.

Finally, the contribution from $\mathfrak{d}$ with $\mathfrak{a}_2 \leqslant \mathfrak{d} \leqslant \mathfrak{a}_3$ or $X^{2k+2\epsilon} \leqslant N(\mathfrak{d}) \leqslant X^{n-k-2\epsilon}$ is negligible without any Buchstab decompositions since it can be written as a sum over $O(\epsilon^{-2})$ polytopes to which Proposition 6.1 applies. This gives the result. □

LEMMA 7.10 (Pólya–Vinogradov-type inequality). *Let $\mathfrak{q}$ be an ideal with a prime ideal factor of norm at least $\log\log\log X$ and $q = N(\mathfrak{q})$. Let χ_f be a character of $\mathcal{O}_K$ with modulus $\mathfrak{q}$ and no infinite component (that is, χ_f factors through $(\mathcal{O}_K/\mathfrak{q}\mathcal{O}_K)^\times$). Then we have*

$$\sum_{\substack{\mathbf{a}\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}),q)=1}} \chi_f\Big(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}\Big) = o\Big(\sum_{\substack{\mathbf{a}\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}),q)=1}} 1\Big). \tag{7.7}$$

Proof. This follows from a Pólya–Vinogradov-type inequality for $\mathbb{Z}[\sqrt[n]{\theta}]/q\mathbb{Z}[\sqrt[n]{\theta}]$, but there are some technical complications relating the restrictions on the algebraic integers α_0 appearing to ideals and the modulus $\mathfrak{q}$ of χ_f. We let $\mathfrak{q}_2'$ be a prime ideal factor of $\mathfrak{q}$ of largest norm, and factor $\mathfrak{q} = \mathfrak{q}_1\mathfrak{q}_2$ with $\mathfrak{q}_1$ the largest factor of $\mathfrak{q}$ with norm coprime to $N(\mathfrak{q}_2')$. By assumption, we have that $N(\mathfrak{q}_2') \gg \log\log\log X$ and is a prime power of exponent at most n, so $\mathfrak{q}_2$ is coprime to the ideal generated by $n\theta$. Correspondingly, we factor $\chi_f = \chi_1\chi_2$ into characters modulo $\mathfrak{q}_1$ and $\mathfrak{q}_2$. Letting $q_2 = N(\mathfrak{q}_2)$, we see that $\mathfrak{q}_2|(q_2)$ and so we can view χ_2 as a character on $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}] \cong \mathcal{O}_K/q_2\mathcal{O}_K$. (We have $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}] \cong \mathcal{O}_K/q_2\mathcal{O}_K$ since $\mathfrak{q}_2$ is coprime to the ideal generated by $n\theta$.) Finally, we note that we have $q_2 \gg N(\mathfrak{q}_2') \gg \log\log\log X$. By writing $\mathbf{a}_0 = q_2\mathbf{a}_1 + q_1\mathbf{a}_2$ (where $q_1 = N(\mathfrak{q}_1)$) and using the Chinese remainder theorem, we see that it is sufficient to show that

$$\sum_{\mathbf{a}\in[1,q_2]^{n-k}} \chi_2\Big(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}\Big) = o(q_2^{n-k}).$$

Finally, we let ψ be the additive character of $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$ given by $\psi(\sum_{j=1}^n a_j\sqrt[n]{\theta^{j-1}}) = \exp(2\pi i a_n/q_2)$ and $\hat{\chi}_2$ be the Fourier transform of χ_2 given by

$$\hat{\chi}_2(\beta) = \frac{1}{q_2^n} \sum_{\gamma\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} \chi_2(\gamma)\psi(\beta\gamma).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We have

$$\sum_{\mathbf{a}\in[1,q_2]^{n-k}} \chi_2\Big(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}\Big) = \sum_{\beta\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} \hat{\chi}_2(\beta) \sum_{\mathbf{a}\in[1,q_2]^{n-k}} \psi(-\alpha\beta),$$

where $\alpha = \sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}}$, viewed as an element of $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$. The inner sum is 0 unless the final $n-k$ components of β are equal to 0, in which case it is q_2^{n-k}. Thus, we are left to show

$$\sum_{b_1,\dots,b_k\in\mathbb{Z}/q_2\mathbb{Z}} \hat{\chi}_2\Big(\sum_{i=1}^{k} b_i \sqrt[n]{\theta^{i-1}}\Big) = o(1).$$

We note that

$$\begin{aligned} q_2^n \hat{\chi}_2(\beta) &= \sum_{\gamma\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} \chi_2(\gamma)\psi(\beta\gamma) \\ &= \sum_{\alpha\in\beta\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} \psi(\alpha) \sum_{\substack{\lambda\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}] \\ \beta\lambda=\alpha}} \chi_2(\lambda). \end{aligned}$$

Denote the inner sum by $f_\beta(\alpha)$. We then see that $\chi_2(\mu) f_\beta(\alpha) = f_\beta(\mu\alpha)$ for any invertible $\mu \in \mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$. But if $\mu \equiv 1 \pmod{(q_2)/\gcd((q_2),(\alpha))}$, then $\mu\alpha = \alpha$ in $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$, and so $f_\beta(\alpha) = 0$ unless $\chi_2(\mu) = 1$ for all invertible $\mu \equiv 1 \pmod{(q_2)/\gcd((q_2),(\alpha))}$. Here the ideals are viewed as ideals in $\mathcal{O}_K$, noting that the choice of lift of $\alpha \in \mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$ does not affect the ideal $\gcd((\alpha),(q_2))$ (recall that $\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}] \cong \mathcal{O}_K/q_2\mathcal{O}_K$). But χ_2 is induced by a primitive character $\pmod{\mathfrak{q}_2}$, and so this only occurs if $\mathfrak{q}_2 | (q_2)/\gcd((q_2),(\alpha))$, that is, if $\mathfrak{q}_2 \nmid (\alpha)$ (since $\mathfrak{q}_2'$ is a prime ideal of large norm, and so lies above an unramified rational prime). Thus, $\hat{\chi}_2(\beta) = 0$ if $\mathfrak{q}_2 | (\beta)$.

We also note that $\hat{\chi}_2(\mu\beta) = \overline{\chi_2(\mu)}\hat{\chi}_2(\beta)$ for any invertible μ, so $\hat{\chi}_2$ is of constant magnitude $c_\mathfrak{d}$ on all β such that $\gcd((\beta),(q_2)) = \mathfrak{d}$. By Parseval's identity, we have

$$\sum_{\alpha\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} |\hat{\chi}_2(\alpha)|^2 = \frac{1}{q_2^n} \sum_{\beta\in\mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]} |\chi_2(\beta)|^2 = \phi_K((q_2))/q_2^n \ll 1.$$

Thus, since there are $O(q_2^n/N(\mathfrak{d}))$ elements $\beta \in \mathbb{Z}[\sqrt[n]{\theta}]/q_2\mathbb{Z}[\sqrt[n]{\theta}]$ with $\gcd((\beta),(q_2)) = \mathfrak{d}$, we see that $c_\mathfrak{d}^2 q_2^n/N(\mathfrak{d}) \ll 1$, which gives

$$|\hat{\chi}_2(\beta)| \leqslant \frac{N(\gcd((\beta),(q_2)))^{1/2}}{q_2^{n/2}}.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Thus, recalling that $\hat{\chi}_2(\beta) = 0$ if $\mathfrak{q}_2|(\beta)$, we have

$$\Big|\sum_{b_1,\dots,b_k\in\mathbb{Z}/q_2\mathbb{Z}} \hat{\chi}_2\Big(\sum_{i=1}^k b_i\sqrt[n]{\theta^{i-1}}\Big)\Big| \leqslant \frac{1}{q_2^{n/2}}\sum_{\mathfrak{d}|(q_2)/\mathfrak{q}_2} N(\mathfrak{d})^{1/2} \sum_{\substack{1\leqslant b_1,\dots,b_k\leqslant q_2\\ \mathfrak{d}|(\sum_{i=1}^k b_i\sqrt[n]{\theta^{i-1}})}} 1.$$

We see that the final sum is counting points in a bounded region in a lattice of rank k. Any point $(b_1,\dots,b_k)$ with $\mathfrak{d}|(\sum_{i=}^k b_i\sqrt[n]{\theta^{i-1}})$ must have $\sum_{i=1}^k |b_i| \gg N(\sum_{i=1}^k b_i\sqrt[n]{\theta^{i-1}})^{1/n} \geqslant N(\mathfrak{d})^{1/n}$. Thus, all nonzero vectors in the lattice, and in particular the basis vectors, must have length $\gg N(\mathfrak{d})^{1/n}$. Therefore, the number of points is $O(1+q_2^k/N(\mathfrak{d})^{k/n})$. This gives

$$\begin{aligned}\Big|\sum_{b_1,\dots,b_k\in\mathbb{Z}/q_2\mathbb{Z}} \hat{\chi}_2\Big(\sum_{i=1}^k b_i\sqrt[n]{\theta^{i-1}}\Big)\Big| &\leqslant \frac{1}{q_2^{n/2}}\sum_{\mathfrak{d}|(q_2)/\mathfrak{q}_2}\Big(N(\mathfrak{d})^{1/2}+q_2^k N(\mathfrak{d})^{1/2-k/n}\Big)\\ &\ll_\epsilon \frac{q_2^\epsilon}{N(\mathfrak{q}_2)^{1/2}}+\frac{q_2^\epsilon}{N(\mathfrak{q}_2)^{1/2-k/n}}.\end{aligned}$$

Here we used the divisor bound in the final line. Since $2k < n$, this is $o(1)$, as required. □

We finish this section with a proof of Lemma 6.4.

Proof of Lemma 6.4. We recall the definition of $\mathcal{B}(\mathbf{a}_0)$:

$$\mathcal{B}(\mathbf{a}_0) = \{\text{ideals } \mathfrak{b} \text{ of } \mathcal{O}_K : N(\mathfrak{b}) \in [N_0^n, (1+\eta_1)N_0^n],\ \chi^*(\mathfrak{b}) = \chi_\infty^*(\mathcal{A})\chi_f^*(\alpha_0)\},$$

where here, and throughout the lemma, $\alpha_0 = \sum_{i=1}^{n-k}(\mathbf{a}_0)_i\sqrt[n]{\theta^{i-1}}$. To ease notation, let $q = q^*$. We have

$$\begin{aligned}\sum_{\substack{\mathbf{a}_0\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}_0),q)=1}}\ \sum_{\mathfrak{b}\in\mathcal{B}(\mathbf{a}_0)} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) &= \sum_{\substack{\mathbf{a}_0\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}_0),q)=1}}\ \sum_{\substack{N(\mathfrak{b})\in[N_0^n,(1+\eta_1)N_0^n]\\ \chi^*(\mathfrak{b})=\chi_\infty^*(\mathcal{A})\chi_f^*(\alpha_0)}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\\ &= \sum_{N(\mathfrak{b})\in[N_0^n,(1+\eta_1)N_0^n]} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) \sum_{\substack{\mathbf{a}_0\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}_0),q)=1}} \frac{1+\chi^*(\mathfrak{b})\chi_\infty^*(\mathcal{A})\chi_f^*(\alpha_0)}{2}\\ &= \Big(\sum_{\substack{\mathbf{a}_0\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}_0),q)=1}}\frac{1}{2}\Big)\Big(\sum_{N(\mathfrak{b})\in[N_0^n,(1+\eta_1)N_0^n]}\mathbf{1}_{\mathcal{R}}(\mathfrak{b})\Big)\\ &\quad + O\Big(\frac{\eta_1 N_0^n}{\log X}\Big|\sum_{\substack{\mathbf{a}_0\in[1,q]^{n-k}\\ \gcd(N_K(\mathbf{a}_0),q)=1}}\chi_f^*(\alpha_0)\Big|\Big).\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

In the second line, we have used the fact that $\mathbf{1}_{\mathcal{R}}$ is supported on ideals with norm coprime to q and so on ideals with $(\chi^*)^2 = 1$. In the last line, we have separated the summations and used a simple sieve bound for the sum over $\mathfrak{b}$ in the error term. Recalling the definition of $\nu(p)$, the first term in parentheses is

$$\sum_{\substack{\mathbf{a}_0 \in [1,q]^{n-k} \\ \gcd(N_K(\mathbf{a}_0),q)=1}} \frac{1}{2} = \frac{q^{n-k}}{2} \prod_{p|q} \left(1 - \frac{\nu(p)}{p^{n-k}}\right).$$

By the prime ideal theorem (Lemma 4.3), the second term in parentheses is

$$\begin{aligned}
&\sum_{N(\mathfrak{b}) \in [N_0^n, (1+\eta_1)N_0^n]} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) \\
&= \frac{q^{n-k}}{\log X} \prod_{p|q} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) \left(\int_{\substack{(e_1,\dots,e_\ell) \in \mathcal{R} \\ X^{\sum_{i=1}^\ell e_i} \in [N_0^n, (1+\eta_1)N_0^n]}} \cdots \int \frac{X^{\sum_{i=1}^\ell e_i} de_1 \dots de_\ell}{e_1 \dots e_\ell} + o(1) \right) \\
&= \frac{q^{n-k} N_0^n}{\log X} \prod_{p|q} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) \left(\int_{\substack{(e_1,\dots,e_\ell) \in \mathcal{R} \\ X^{\sum_{i=1}^\ell e_i} \in [N_0^n, (1+\eta_1)N_0^n]}} \cdots \int \frac{de_1 \dots de_\ell}{e_1 \dots e_\ell} + o_{\mathcal{R}}(1) \right).
\end{aligned}$$

Since we have an error term which depends on $\mathcal{R}$, we may think of $\mathcal{R}$ as fixed and η_1 as small. Since $\mathcal{R}$ is closed and $\log N_0^n / \log X = n + o(1)$, we see that the integral is equal to

$$\eta_1 \int_{\substack{(e_1,\dots,e_\ell) \in \mathcal{R} \\ \sum_{i=1}^\ell e_i = n}} \cdots \int \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell} + o_{\mathcal{R}}(1) = I_{\mathcal{R}} + o(1).$$

Finally, we recall that $\mathfrak{q}^*$ is square-free apart from an ideal factor of norm $O(1)$ and $N(\mathfrak{q}^*) \gg (\log x)^\epsilon$, and so χ_f^* satisfies the conditions of Lemma 7.10. Thus, we have that

$$\frac{\eta_1 N_0^n}{\log X} \left| \sum_{\substack{\mathbf{a}_0 \in [1,q]^{n-k} \\ \gcd(N_K(\mathbf{a}_0),q)=1}} \chi_f^*(\alpha_0) \right| = o\left(\frac{q^{n-k} \eta_1 N_0^n}{\log X} \prod_{p|q} \left(1 - \frac{\nu(p)}{p^{n-k}}\right) \right).$$

This gives the result. □

Thus, we are left to establish Proposition 6.1.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

8. Type II Estimate: The L^1 Bounds

In this section, we introduce an approximation $\tilde{\mathbf{1}}_{\mathcal{R}} \approx \mathbf{1}_{\mathcal{R}}$ in our Type II sums and establish various L^1 estimates based on this. Much of this section is a generalization of the corresponding estimates of Heath-Brown [13]. The aim of this section is to reduce the proof of Proposition 6.1 to Proposition 8.7.

We wish to establish Proposition 6.1, namely that

$$\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b}) \ll_{\mathcal{R}} \eta_1^{1/2}\#\mathcal{A}, \tag{8.1}$$

where $\mathcal{R} \subseteq [\epsilon^2, 2n]^\ell$ is a polytope such that there is an $\ell' \leqslant \ell$ so that any $\mathbf{e} \in \mathcal{R}$ satisfies $k + \epsilon \leqslant \sum_{i=1}^{\ell'} e_i \leqslant n - 2k - \epsilon$. We recall that $\eta_1 = (\log X)^{-100}$ and that

$$\mathbf{1}_{\mathcal{R}}(\mathfrak{a}) = \begin{cases} 1, & \mathfrak{a} = \mathfrak{p}_1 \dots \mathfrak{p}_\ell \text{ with } N(\mathfrak{p}_i) = X^{e_i},\ (e_1, \dots, e_\ell) \in \mathcal{R}, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mathcal{A} = \left\{ \left(\sum_{i=1}^{n-k} a_i \sqrt[n]{\theta^{i-1}} \right) : X_i \leqslant a_i \leqslant X_i + \eta_1 X_i,\ a_i \equiv (\mathbf{a}_0)_i \pmod{q^*} \right\},$$

$$\mathcal{B} = \{\mathfrak{b} : N(\mathfrak{b}) \in [N_0^n, (1+\eta_1)N_0^n],\ \chi^*(\mathfrak{b}) = \chi_\infty^*(\mathcal{A})\chi_f^*(\alpha_0)\},$$

$$\tilde{\mathfrak{S}} = \prod_{p\nmid q^*} \left(1 - \frac{\nu(p)}{p^{n-k}}\right)\left(1 - \frac{\nu_2(p)}{p^n}\right)^{-1},$$

with $N_0^n \geqslant \epsilon X^n$ the smallest norm of an ideal in $\mathcal{A}$. Since the implied constant is allowed to depend on $\mathcal{R}$, we may assume that $\mathcal{R}$ is defined by a bounded number of linear inequalities, none of which depends on our underlying parameter X. We will therefore suppress the dependence on $\mathcal{R}$ for the rest of this section.

We now wish to reduce Proposition 6.1 to the following statement.

LEMMA 8.1. *Let $\mathcal{R}$ satisfy the assumptions of Proposition 6.1. Given a hypercube $\mathcal{C}$, write $\mathcal{C} = \mathcal{R}_1 \times \mathcal{R}_2$ with $\mathcal{R}_2$ representing the first ℓ' coordinates and $\mathcal{R}_1$ the final $\ell - \ell'$ coordinates. Then for any set of nonoverlapping hypercubes of side length η_1^2 which covers $\mathcal{R}$, we have*

$$\sum_{\substack{\mathcal{C}=\mathcal{R}_1\times\mathcal{R}_2 \\ \mathcal{C}\cap\mathcal{R}\neq\emptyset}} \left(\sum_{\substack{\mathfrak{a}_1,\mathfrak{a}_2 \\ \mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{\substack{\mathfrak{b}_1,\mathfrak{b}_2 \\ \mathfrak{b}_1\mathfrak{b}_2\in\mathcal{B}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{b}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}_2) \right) \ll_{\mathcal{R}} \eta_1^{1/2}\#\mathcal{A},$$

$$\sum_{\substack{\mathcal{C}=\mathcal{R}_1\times\mathcal{R}_2 \\ \mathcal{C}\subseteq\mathcal{R}}} \left(\sum_{\substack{\mathfrak{a}_1,\mathfrak{a}_2 \\ \mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{\substack{\mathfrak{b}_1,\mathfrak{b}_2 \\ \mathfrak{b}_1\mathfrak{b}_2\in\mathcal{B}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{b}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}_2) \right) \ll_{\mathcal{R}} \eta_1^{1/2}\#\mathcal{A}.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We note that $\mathbf{1}_{\mathcal{R}_2}$ is supported on ideals $\mathfrak{b}$ with $N(\mathfrak{b}) \in [X^{k+\epsilon/2}, X^{n-2k-\epsilon/2}]$ from our bounds on $\sum_{i=1}^{\ell'} e_i$.

Proof of Proposition 6.1 assuming Lemma 8.1. We cover $\mathcal{R}$ by $O(\eta_1^{-2\ell})$ nonoverlapping hypercubes $\mathcal{C}$ so that each of $e_1, \dots, e_\ell$ lies in intervals of side length η_1^2. We see that

$$\sum_{\mathcal{C}\subseteq\mathcal{R}} \mathbf{1}_{\mathcal{C}}(\mathfrak{a}) \leqslant \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) \leqslant \sum_{\mathcal{C}\cap\mathcal{R}\neq\emptyset} 1_{\mathcal{C}}(\mathfrak{a}).$$

Thus, first upper bounding the sum over $\mathfrak{a} \in \mathcal{A}$ and lower bounding the sum over $\mathfrak{b} \in \mathcal{B}$, and then lower bounding the sum over $\mathfrak{a}$ and upper bounding the sum over $\mathfrak{b}$, we obtain

$$\begin{aligned}\left|\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{R}}(\mathfrak{b})\right| &\leqslant \left|\sum_{\mathcal{C}\cap\mathcal{R}\neq\emptyset}\left(\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{C}}(\mathfrak{a}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{C}}(\mathfrak{b})\right)\right| \\ &\quad + \left|\sum_{\mathcal{C}\subseteq\mathcal{R}}\left(\sum_{\mathfrak{a}\in\mathcal{A}} \mathbf{1}_{\mathcal{C}}(\mathfrak{a}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{C}}(\mathfrak{b})\right)\right| \\ &\quad + \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\substack{\mathcal{C}\cap\mathcal{R}\neq\emptyset\\ \mathcal{C}\not\subseteq\mathcal{R}}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{C}}(\mathfrak{b}). \qquad (8.2)\end{aligned}$$

By the prime ideal theorem (Lemma 4.3), we have for $e_1, \dots, e_\ell \geqslant \epsilon^2$,

$$\sum_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_\ell\\ N(\mathfrak{p}_i)\in[X^{e_i}, X^{e_i+\eta_1^2}]}} 1 \ll_\epsilon \eta_1^{2\ell} X^{\sum_{i=1}^{\ell} e_i}.$$

Thus, since $\mathcal{B}$ is supported on ideals $\mathfrak{b}$ with $N(\mathfrak{b}) \in [N_1, (1+\eta_1)N_1]$, we see that for any hypercube $\mathcal{C}$ under consideration,

$$\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{C}}(\mathfrak{b}) \ll_\epsilon \eta_1^{2\ell} N_1 \ll \eta_1^{2\ell-1}\#\mathcal{B}. \qquad (8.3)$$

There are $O_{\mathcal{R}}(\eta_1^{-2(\ell-1)})$ hypercubes $\mathcal{C}$ intersecting the boundary of $\mathcal{R}$ since $\mathcal{R}$ is a polytope defined by $O_{\mathcal{R}}(1)$ inequalities. Therefore, by (8.3), the final term on the right-hand side of (8.2) contributes

$$\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\substack{\mathcal{C}\cap\mathcal{R}\neq\emptyset\\ \mathcal{C}\not\subseteq\mathcal{R}}}\sum_{\mathfrak{b}\in\mathcal{B}} \mathbf{1}_{\mathcal{C}}(\mathfrak{b}) \ll_\epsilon \#\mathcal{A}\sum_{\substack{\mathcal{C}\cap\mathcal{R}\neq\emptyset\\ \mathcal{C}\not\subseteq\mathcal{R}}} \eta_1^{2\ell-1} \ll_{\epsilon,\mathcal{R}} \eta_1\#\mathcal{A},$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

which is negligible. Thus, it suffices to show

$$\Big|\sum_{\mathcal{C}\cap\mathcal{R}\neq\emptyset}\Big(\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{C}}(\mathfrak{a})-\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}}\mathbf{1}_{\mathcal{C}}(\mathfrak{b})\Big)\Big|\ll_{\mathcal{R}}\eta_1^{1/2}\#\mathcal{A},$$

and similarly when summing over all $\mathcal{C}$ with $\mathcal{C}\subseteq\mathcal{R}$.

Any hypercube $\mathcal{C}$ can be identified with $\mathcal{R}_1\times\mathcal{R}_2$ with $\mathcal{R}_2$ representing the first ℓ' coordinates of $\mathcal{C}$. Call $\mathcal{C}$ *good* if $\mathcal{C}\cap\mathcal{R}\neq\emptyset$ and $\mathcal{C}$ does not contain any point $\mathbf{e}$ such that $|e_i-e_j|\ll\eta_1^2$ for some $1\leqslant i<j\leqslant\ell$. If $\mathcal{C}$ is good, then any $\mathfrak{a}$ with $\mathbf{1}_{\mathcal{C}}(\mathfrak{a})\neq 0$ has a unique representation as $\mathfrak{a}=\mathfrak{a}_1\mathfrak{a}_2$ with $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)=\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2)=1$. If $\mathcal{C}$ does contain a point $\mathbf{e}$ such that $|e_i-e_j|\ll\eta_1^2$, then there can be between 1 and n different representations $\mathfrak{a}=\mathfrak{a}_1\mathfrak{a}_2$. Thus,

$$\begin{aligned}&\sum_{\mathcal{C}\cap\mathcal{R}\neq\emptyset}\Big(\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{C}}(\mathfrak{a})-\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}\in\mathcal{B}}\mathbf{1}_{\mathcal{C}}(\mathfrak{b})\Big)\\&\ll\Big|\sum_{\mathcal{C}\cap\mathcal{R}\neq\emptyset}\Big(\sum_{\mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}}\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2)-\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\mathfrak{b}_1\mathfrak{b}_2\in\mathcal{B}}\mathbf{1}_{\mathcal{R}_1}(\mathfrak{b}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}_2)\Big)\Big|\\&\quad+O\Big(\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}}\sum_{\substack{\mathcal{C}\cap\mathcal{R}\neq\emptyset\\ \mathcal{C}\text{ not good}}}\Big|\sum_{\mathfrak{b}\in\mathcal{C}}\mathbf{1}_{\mathcal{C}}(\mathfrak{b})\Big|\Big),\end{aligned}\tag{8.4}$$

and similarly when considering all $\mathcal{C}\subseteq\mathcal{R}$.

There are $O(\eta_1^{-2(\ell-1)})$ hypercubes which contain a point $\mathbf{e}$ with $|e_i-e_j|\ll\eta_1^2$ for some $1\leqslant i<j\leqslant\ell$. By (8.3), each such hypercube contributes $O_\epsilon(\eta_1^{2\ell-1}\#\mathcal{B})$ to the inner sum above. Thus, the contribution from hypercubes which are not good is $O_\epsilon(\eta_1\#\mathcal{A})$.

Finally, Lemma 8.1 shows that the first term on the right-hand side of (8.4) is $O_{\mathcal{R}}(\eta_1^{1/2}\#\mathcal{A})$, giving Proposition 6.1. □

It will be convenient to split the sum to localize the size of the norm of $\mathfrak{a}_1\mathfrak{a}_2$ and $\mathfrak{b}_1\mathfrak{b}_2$. We let

$$\begin{aligned}&\eta_2=\eta_1^{10\ell},\\&\mathcal{A}'=\Big\{\Big(\sum_{i=1}^{n-k}a_i\sqrt[n]{\theta^{i-1}}\Big):X_i\leqslant a_i\leqslant X_i+\eta_1X_i,\ a_i\equiv(\mathbf{a}_0')_i\ (\mathrm{mod}\ J!q^*),\\&\qquad N\Big(\sum_{i=1}^{n-k}a_i\sqrt[n]{\theta^{i-1}}\Big)\in[X_0^n,X_0^n+\eta_2X_0^n]\Big\},\\&\mathcal{B}'=\{\mathfrak{b}\in\mathcal{B}:N(\mathfrak{b})\in[X_0^n,X_0^n+\eta_2X_0^n]\}.\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Here we have extended the congruence conditions in $\mathcal{A}$ from $\mathbf{a} \equiv \mathbf{a}_0 \pmod{q^*}$ to $\mathbf{a} \equiv \mathbf{a}_0' \pmod{J!q^*}$, for a suitable constant $J \ll 1$ which will be chosen later do be large enough in terms of n and k. We consider separately all $\mathbf{a}_0'$ such that $\mathbf{a}_0' \equiv \mathbf{a}_0 \pmod{q^*}$ and $\mathbf{a} \equiv \mathbf{a}_0' \pmod{J!} \implies p \nmid N(\sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}) \forall p \leqslant J$. (It is sufficient to only consider such $\mathbf{a}_0'$ since $\mathbf{1}_{\mathcal{R}}$ is supported on ideals with no small factors.) The key estimate we wish to establish is the following.

PROPOSITION 8.2. *Let $\mathcal{R}$ satisfy the assumptions of Proposition 6.1. Uniformly for $X_0 \in [N_0, (1 + O(\eta_1))N_0]$ and over all hypercubes $\mathcal{C} = \mathcal{R}_1 \times \mathcal{R}_2 \cap \mathcal{R} \neq \emptyset$ occurring in Lemma 8.1, we have*

$$\sum_{\substack{\mathfrak{a}_1, \mathfrak{a}_2 \\ \mathfrak{a}_1\mathfrak{a}_2 \in \mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) = \frac{q_0^n \tilde{\mathfrak{S}} c_{\mathcal{R}_1 \times \mathcal{R}_2}(X_0^n) \# \mathcal{A}'}{\phi_K((q_0))\gamma_K} \left(1 + \frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell X_0^{n-n\beta^*}}\right) + O(\eta_2^{1/3} \# \mathcal{A}'),$$

where $q_0 = J!q^*$, and

$$\sum_{\substack{\mathfrak{b}_1, \mathfrak{b}_2 \\ \mathfrak{b}_1\mathfrak{b}_2 \in \mathcal{B}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{b}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}_2) = \frac{q^{*n} c_{\mathcal{R}_1 \times \mathcal{R}_2}(X_0^n) \# \mathcal{B}'}{\phi_K((q^*))\gamma_K} \left(1 + \frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell X_0^{n-n\beta^*}}\right) + O(\eta_2^{1/3} \# \mathcal{B}'),$$

where $\beta^* \in [0, 1]$ is a quantity depending only on X and where for a set $\mathcal{S} \subset \mathbb{R}^\ell$

$$c_{\mathcal{S}}(t) = \int_{\substack{(e_1, \dots, e_\ell) \in \mathcal{S} \\ \sum_{i=1}^\ell e_i \in \mathcal{I}_t}} \cdots \int \frac{de_1 \dots de_\ell}{\eta_2^{1/2} \prod_{i=1}^\ell e_i},$$

$$\mathcal{I}_t = \left[\frac{\log t}{\log X}, \frac{\log(t + \eta_2^{1/2} t)}{\log X}\right].$$

Here β^* will be a possible exceptional zero if one exists, and 0 otherwise.

We note that for any set $\mathcal{S} \subseteq [\epsilon^2, 2n]^\ell$, we have the following Lipschitz bounds.

LEMMA 8.3. *Let $\mathcal{S} \subseteq [\epsilon^2, 2n]^\ell$, and let $s^+ = \sup\{\sum_{i=1}^\ell e_i : \mathbf{e} \in \mathcal{S}\}$ and $s^- = \inf\{\sum_{i=1}^\ell e_i : e_i \in \mathcal{S}\}$.*

(i) *We have*

$$c_{\mathcal{S}}(t + \delta) - c_{\mathcal{S}}(t) \ll_\epsilon \frac{\delta}{\eta_2^{1/2} t}.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

(ii) *If $\mathcal{S}$ is a polytope and $\log t/\log X \in [s^- + \epsilon, s^+ - \epsilon]$, then we have*

$$c_{\mathcal{S}}(t+\delta) - c_{\mathcal{S}}(t) \ll_{\epsilon,\mathcal{S}} \frac{\delta}{t}.$$

(iii) *If $\mathcal{S}$ is a hypercube (with edges parallel to the coordinate axes) and $\ell > 1$, then*

$$c_{\mathcal{S}}(t+\delta) - c_{\mathcal{S}}(t) \ll_{\epsilon} \frac{\delta}{t}.$$

All implied constants may depend on n and ℓ.

We note that the implied constant in the first bound is independent of $\mathcal{S}$, whereas the implied constant in the second bound depends on $\mathcal{S}$.

Proof. The first bound is straightforward. For any choice of $e_1, \dots, e_{\ell-1}$, we have

$$\Bigg| \int\limits_{\substack{e_\ell \\ (e_1,\dots,e_\ell)\in\mathcal{S} \\ \sum_{i=1}^{\ell} e_i \in \mathcal{I}_t}} \frac{de_\ell}{e_\ell} - \int\limits_{\substack{e_\ell \\ (e_1,\dots,e_\ell)\in\mathcal{S} \\ \sum_{i=1}^{\ell} e_i \in \mathcal{I}_{t+\delta}}} \frac{de_\ell}{e_\ell} \Bigg| \ll_\epsilon \#(\mathcal{I}_{t+\delta} \setminus \mathcal{I}_t) + \#(\mathcal{I}_t \setminus \mathcal{I}_{t+\delta}) \ll \frac{\delta}{t}.$$

Expanding $c_{\mathcal{S}}(t+\delta) - c_{\mathcal{S}}(t)$ by the integral definition and substituting this bound then gives the first claim.

We now consider the second claim of the lemma. The result is trivial if $\delta > \epsilon^3$, so we may assume $\delta \leqslant \epsilon^3$. Since $\mathcal{S}$ is a polytope, the $(\ell-1)$-dimensional region $\mathcal{S}_u$ of $\mathbf{e} \in \mathcal{S}$ with $\sum_{i=1}^{\ell} e_i = u$ is a polytope depending on u. After translating $\mathcal{S}_u$ by $O(v)$, we see that it differs from $\mathcal{S}_{u+v}$ by a region of ($(\ell-1)$-dimensional) volume $O_{\mathcal{S}}(v)$, unless $\mathcal{S}$ has a face contained in $\sum_{i=1}^{\ell-1} e_i = u_0$ for some $u_0 \in [u, u+v]$. But $\mathcal{S}$ cannot contain such a face for $u \in [s^-, s^+ - v]$ since it is convex. Therefore, for $u \in [s^-, s^+ - \epsilon]$ and $v \leqslant \epsilon^3$, we have

$$\int \cdots \int_{(e_1,\dots,e_\ell)\in\mathcal{S}_u} \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell} = \int \cdots \int_{(e_1,\dots,e_\ell)\in\mathcal{S}_{u+v}} \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell} + O_{\epsilon,\mathcal{S}}(v).$$

Here we used the fact that if $\mathbf{e} \in \mathcal{S}$, then $e_i \geqslant \epsilon^2$. Thus, we find that $|c_{\mathcal{S}}(t+\delta) - c_{\mathcal{S}}(t)|$ is

$$\ll \int_{u\in\mathcal{I}_t} \frac{1}{\eta_2^{1/2}} \Bigg(\int\limits_{\substack{\mathbf{e}\in\mathcal{S} \\ \sum_{i=1}^{\ell} e_i = u}} \cdots \int \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell} - \int\limits_{\substack{\mathbf{e}\in\mathcal{S} \\ \sum_{i=1}^{\ell} e_i = u + \frac{\log(1+\delta/t)}{\log X}}} \cdots \int \frac{de_1 \dots de_{\ell-1}}{e_1 \dots e_\ell} \Bigg) du$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll_{\epsilon,\mathcal{S}} \int_{u\in\mathcal{I}_t} \frac{1}{\eta_2^{1/2}} \frac{\delta}{t} \ll_{\mathcal{S}} \frac{\delta}{t}.$$

This gives the second claim.

Finally, if $\mathcal{S} \subseteq [\epsilon^2, 2n]^\ell$ is a hypercube with edges parallel to the coordinate axes and $\ell > 1$, then the $(\ell-1)$-dimensional volume of $\mathbf{e} \in \mathcal{S}$ with $\sum_{i=1}^{\ell} e_i = u$ is a region which varies in a Lipschitz manner as described above, with Lipschitz constant $O(1)$ independent of $\mathcal{S}$, since all faces of $\mathcal{S}$ are at an angle $\gg 1$ from the hyperplanes $\sum_{i=1}^{\ell} e_i = u$. Using this in the bound above gives the final claim. □

We first show that Proposition 8.2 gives Lemma 8.1, and so Proposition 6.1. We will then go on to establish Proposition 8.2.

Proof of Lemma 8.1 assuming Proposition 8.2. Summing the first estimate of Proposition 8.2 over all hypercubes $\mathcal{C} \subseteq \mathcal{R}$ under consideration, we obtain

$$\begin{aligned}\sum_{\mathcal{R}_1\times\mathcal{R}_2=\mathcal{C}\subseteq\mathcal{R}} \sum_{\substack{\mathfrak{a}_1,\mathfrak{a}_2\\ \mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2)&\\ = \frac{q_0^n\tilde{\mathfrak{S}}\#\mathcal{A}'}{\phi_K((q_0))\gamma_K}\left(1+\frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell X_0^{n-n\beta^*}}\right)\sum_{\mathcal{C}\subseteq\mathcal{R}} c_{\mathcal{C}}(X_0^n)&\\ + O(\eta_2^{1/3}\eta_1^{-2(\ell-1)}\#\mathcal{A}').&\end{aligned}$$

Since $\mathcal{R}$ is convex and contains points with sum of coordinates bigger than $n+\epsilon$ and smaller than $n-\epsilon$, there are $O_{\mathcal{R}}(\eta_1^{-2(\ell-2)})$ hypercubes $\mathcal{C} = [a_1, a_1+\eta_1^2) \times \cdots \times [a_\ell, a_\ell+\eta_1^2)$ intersecting the boundary of $\mathcal{R}$ with $\sum_{i=1}^{\ell} a_i = n\log X_0/\log X + O(\eta_1^2)$. For each such hypercube $\mathcal{C}$, we see $c_{\mathcal{C}}(X_0^n) \ll \eta_1^{2\ell-2}$. Therefore, we see that

$$\sum_{\mathcal{C}\subseteq\mathcal{R}} c_{\mathcal{C}}(X_0^n) = c_{\mathcal{R}}(X_0^n) + O_{\mathcal{R}}\left(\eta_1^{-2\ell+4}\sup_{\mathcal{C}\cap\mathcal{R}\neq\emptyset} c_{\mathcal{C}}(X_0^n)\right) = c_{\mathcal{R}}(X_0^n) + O_{\mathcal{R}}(\eta_1^2).$$

Thus, we have

$$\begin{aligned}\sum_{\mathcal{R}_1\times\mathcal{R}_2=\mathcal{C}\subseteq\mathcal{R}} \sum_{\substack{\mathfrak{a}_1,\mathfrak{a}_2\\ \mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) = \frac{q_0^n\tilde{\mathfrak{S}}\#\mathcal{A}'}{\phi_K((q_0))\gamma_K}\left(1+\frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell X_0^{n-n\beta^*}}\right)c_{\mathcal{R}}(X_0^n)&\\ + O_{\mathcal{R}}(\eta_1^2\#\mathcal{A}').&\end{aligned}$$

We note that for all $N_0^n \leqslant X_0^n \leqslant (1+O(\eta_1))N_0^n$, we have $c_{\mathcal{R}}(X_0^n) = c_{\mathcal{R}}(N_0^n) + O_{\mathcal{R}}(\eta_1)$ by Lemma 8.3, and we have $1/X_0^{n-n\beta^*} = (1+O(\eta_1))/N_0^{n-n\beta^*}$. We recall

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

that $q_0 \leqslant N_0$ and $\eta_1 = (\log X)^{-100}$, so $q_0^n/\phi_K((q)) < \eta_1^{-1/100}$. Thus, inserting these bounds and summing over a suitable set of disjoint choices of $\mathcal{A}'$ covering $\mathcal{A}$, noting that there are $\phi_K((q_0))/\phi_K((q^*))$ choices of $\mathbf{a}_0'$, we obtain

$$\sum_{\mathcal{R}_1\times\mathcal{R}_2=\mathcal{C}\subseteq\mathcal{R}}\sum_{\substack{\mathfrak{a}_1,\mathfrak{a}_2\\ \mathfrak{a}_1\mathfrak{a}_2\in\mathcal{A}}}\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}_1)\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) = \frac{q^{*n}\tilde{\mathfrak{S}}c_{\mathcal{R}}(N_0^n)\#\mathcal{A}}{\phi_K((q^*))\gamma_K}\left(1+\frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell N_0^{n-n\beta^*}}\right) + O_{\mathcal{R}}(\eta_1^{9/10}\#\mathcal{A}).$$

We obtain an entirely analogous result for $\mathcal{B}$ which is larger by a factor $\#\mathcal{B}/(\tilde{\mathfrak{S}}\#\mathcal{A})$. This gives the second claim of Lemma 8.1. The first claim is entirely analogous, but we sum over $\mathcal{C}\cap\mathcal{R}\neq\emptyset$ instead of $\mathcal{C}\subseteq\mathcal{R}$. □

Thus, we are left to establish Proposition 8.2, which we will do over the next two sections.

We first wish to replace $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a})$ with a more easily controlled approximation $\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{a})$. To do this, we will take into account the possible effect of an exceptional character distorting the distribution of prime ideals in residue classes $\pmod{\mathfrak{q}}$, and so we recall the results on zero-free regions for Hecke L-functions given by Lemmas 4.4 and 4.5. This also makes precise the choice of q^*, χ^* in the definitions of $\mathcal{A}$, $\mathcal{B}$ which so far have been treated as arbitrary quantities and the quantity β^* appearing in Proposition 8.2.

We now describe how we define χ^*, $\mathfrak{q}^*$ and β^* and our approximation $\tilde{\mathbf{1}}_{\mathcal{R}_2}$. If an exceptional character $\chi_{\mathfrak{d}^*}$ does exist (in the sense of Lemma 4.4) and $N(\mathfrak{d}^*)\leqslant \exp(\sqrt[4]{\log X})$, then we let $\chi^*=\chi_{\mathfrak{d}^*}$ with corresponding modulus $\mathfrak{q}^*=\mathfrak{d}^*$ and real zero $\beta^*=\beta_{\mathfrak{d}^*}$. If $\chi_{\mathfrak{d}^*}$ does not exist or if $N(\mathfrak{d}^*)>\exp(\sqrt[4]{\log X})$, then we make an arbitrary choice of $\mathfrak{q}^*$ and χ^* such that χ^* is a nontrivial primitive real character to a square-free modulus $\mathfrak{q}^*$ with $N(\mathfrak{q}^*)\asymp\exp(\sqrt[5]{\log X})$, and we take $\beta^*=1/2$.

With this choice of $\mathfrak{q}^*$, χ^*, β^*, regardless of which situation we are in, we recall the consequences of Lemma 4.5: we have that

$$\sum_{N(\mathfrak{a})\leqslant X}\Lambda(\mathfrak{a})\chi(\mathfrak{a}) \ll X\exp(-c\sqrt{\log X}) \tag{8.5}$$

uniformly over all nontrivial primitive Hecke characters $\chi=\chi_1\prod_{i=1}^{n-1}\lambda_i^{m_i}\neq\chi^*$ with torsion part χ_1 of conductor $\leqslant q^{*(\log\log X)^2}\exp(\sqrt[5]{\log X})$ and with $m_i\leqslant q^{*(\log\log X)^2}\exp(\sqrt[5]{\log X})$ for all $1\leqslant i\leqslant n-1$. If instead $\chi=\chi^*$, we have

$$\sum_{N(\mathfrak{a})\leqslant X}\Lambda(\mathfrak{a})\chi^*(\mathfrak{a}) = \frac{-X^{\beta^*}}{\beta^*}+O(X\exp(-c\sqrt{\log X})). \tag{8.6}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

If $\beta^* = 1/2$, then all the terms involving χ^* or β^* will be negligible and can be ignored.

We then define

$$\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}) = c_{\mathcal{R}_2}(N(\mathfrak{b}))\left(1 + \frac{\chi^*(\mathfrak{b})}{(-\beta^*)^{\ell'} N(\mathfrak{b})^{1-\beta^*}}\right) \sum_{\mathfrak{d}|\mathfrak{b}} \lambda_{\mathfrak{d}}, \tag{8.7}$$

where

$$R = X^{\epsilon^2},$$
$$\lambda_{\mathfrak{d}} = \begin{cases} \mu(\mathfrak{d}) \log \dfrac{R}{N(\mathfrak{d})}, & N(\mathfrak{d}) < R, \\ 0, & \text{otherwise}, \end{cases}$$

and we recall the definition of $c_{\mathcal{R}_2}(t)$ from Proposition 8.2.

The sum $\sum_{\mathfrak{d}|\mathfrak{b}} \lambda_{\mathfrak{d}}$ should be thought of as a sieve weight which approximates the indicator function of ideals with no prime ideal factors of norm less than R, whilst the $c_{\mathcal{R}_2}(N(\mathfrak{b}))$ factor represents the density of $\mathbf{1}_{\mathcal{R}_2}$ on ideals of norm approximately $N(\mathfrak{b})$.

We will now proceed to show that the first estimate of Proposition 8.2 holds with $\tilde{\mathbf{1}}_{\mathcal{R}_2}$ in place of $\mathbf{1}_{\mathcal{R}_2}$ and establish the second estimate directly. This then reduces the problem to showing $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{a}_2) \approx \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{a}_2)$ for $\mathfrak{a}_1\mathfrak{a}_2 \in \mathcal{A}'$, which we do by our L^2 estimate in the next section.

To ease notation, we fix $\mathfrak{a}_0$ such that $\chi^*(\mathfrak{a}_0) = \chi^*_\infty(\mathcal{A})\chi^*_f(\alpha_0)$.

LEMMA 8.4. *Let $\mathcal{C} = \mathcal{R}_1 \times \mathcal{R}_2$ be as in Proposition 8.2. Then*

$$\sum_{\substack{\mathfrak{a},\mathfrak{b} \\ \mathfrak{a}\mathfrak{b}\in\mathcal{B}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) = \frac{q^{*n} c_{\mathcal{R}_1\times\mathcal{R}_2}(X_0^n)\#\mathcal{B}'}{\phi_K((q^*))\gamma_K}\left(1 + \frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^{\ell} X_0^{n-n\beta^*}}\right) + O(\eta_2 \#\mathcal{B}').$$

Proof. This essentially follows from the prime ideal theorem. We recall that $\mathcal{B}' = \{\mathfrak{a} : N(\mathfrak{a}) \in \mathcal{I}, \chi^*(\mathfrak{a}) = \chi^*(\mathfrak{a}_0)\}$, where $\mathcal{I}$ is the interval $[X_0^n, X_0^n + \eta_2 X_0^n]$. Since $\chi^*(\mathfrak{a}\mathfrak{b})^2 = 1$ if $\gcd(\mathfrak{a}\mathfrak{b}, \mathfrak{q}^*) = 1$, which occurs on the support of $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})$, we have

$$\sum_{\substack{\mathfrak{a},\mathfrak{b} \\ \mathfrak{a}\mathfrak{b}\in\mathcal{B}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) = \frac{1}{2} \sum_{\substack{\mathfrak{a},\mathfrak{b} \\ N(\mathfrak{a}\mathfrak{b})\in\mathcal{I}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})(1 + \chi^*(\mathfrak{a}\mathfrak{b})\chi^*(\mathfrak{a}_0)).$$

By the prime ideal theorem (Lemma 4.3), partial summation and Lemma 8.3, we

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

have

$$\begin{aligned}
&\sum_{\substack{\mathfrak{a},\mathfrak{b}\\ N(\mathfrak{a}\mathfrak{b})\in\mathcal{I}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})\\
&\quad= \int\cdots\int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R}_1\times\mathcal{R}_2\\ X^{\sum_{i=1}^\ell e_i}\in\mathcal{I}}} \frac{X^{\sum_{i=1}^\ell e_i}de_1\dots de_\ell}{\prod_{i=1}^\ell e_i} + O\Big(X_0^n\exp(-\frac{c}{2}\sqrt{\log X_0})\Big)\\
&\quad= X_0^n\log X\int_{X^t\in\mathcal{I}}\Big(\int\cdots\int_{\substack{\mathbf{e}\in\mathcal{R}_1\times\mathcal{R}_2\\ \sum_{i=1}^\ell e_i=t}}\frac{de_1\dots de_{\ell-1}}{e_1\dots e_\ell}\Big)dt + O(\eta_2^2X_0^n),\\
&\quad= \eta_2X_0^nc_{\mathcal{R}_1\times\mathcal{R}_2}(X_0^n) + O(\eta_2^2X_0^n).
\end{aligned}$$

Similarly, using (8.6), we have

$$\begin{aligned}
&\sum_{\substack{\mathfrak{a},\mathfrak{b}\\ N(\mathfrak{a}\mathfrak{b})\in\mathcal{I}}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b})\chi^*(\mathfrak{a}\mathfrak{b})\\
&\quad= \frac{1}{(-\beta^*)^\ell}\int\cdots\int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R}_1\times\mathcal{R}_2\\ X^{\sum_{i=1}^\ell e_i}\in\mathcal{I}}}\frac{X^{\beta^*\sum_{i=1}^\ell e_i}de_1\dots de_\ell}{\prod_{i=1}^\ell e_i} + O(\eta_2^2X_0^n)\\
&\quad= \frac{\eta_2}{(-\beta^*)^\ell}X_0^{n\beta^*}c_{\mathcal{R}_1\times\mathcal{R}_2}(X_0^n) + O(\eta_2^2X_0^n).
\end{aligned}$$

Since $\#\mathcal{B}' = \phi_K((q^*))\gamma_K\eta_2X_0^n/2q^{*n} + O(X_0^{n-1})$, this gives the result. □

LEMMA 8.5. *There is a constant $c>0$ such that for any integer q, we have*

$$\sum_{\substack{N(\mathfrak{d})<R\\ \gcd(N(\mathfrak{d}),q)=1}}\frac{\mu(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})}\log\frac{R}{N(\mathfrak{d})} = \frac{q^n\tilde{\mathfrak{S}}}{\phi_K((q))\gamma_K} + O\Big(q^{o(1)}\exp(-c\sqrt{\log R})\Big).$$

Proof. This is an application of counting via complex analysis and the zero-free region of $\zeta_K(s)$. By Perron's formula, we have (noting that the integrals converge absolutely)

$$\sum_{\substack{N(\mathfrak{d})<R\\ \gcd(N(\mathfrak{d}),q)=1}}\frac{\mu(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})}\log\frac{R}{N(\mathfrak{d})} = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty}\frac{R^s}{s^2}\Big(\sum_{\gcd(N(\mathfrak{d}),q)=1}\frac{\mu(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})^{1+s}}\Big)ds$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$= \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} \frac{R^s}{s^2\zeta_K(1+s)} f(1+s)ds, \quad (8.8)$$

where $f(s)$ is given by the Euler product

$$\begin{aligned} f(s) &= \prod_{\mathfrak{p}\nmid(q)}\left(1-\frac{\rho(\mathfrak{p})}{N(\mathfrak{p})^s}\right)\left(1-\frac{1}{N(\mathfrak{p})^s}\right)^{-1}\prod_{\mathfrak{p}|(q)}\left(1-\frac{1}{N(\mathfrak{p})^s}\right)^{-1} \\ &= \prod_{p\nmid q}\left(1-\frac{\nu_p}{p^s}+O(p^{-2\Re(s)})\right)\left(1-\frac{\nu_p}{p^s}+O(p^{-2\Re(s)})\right)^{-1} \\ &\quad \times \prod_{\mathfrak{p}|(q)}\left(1-\frac{1}{N(\mathfrak{p})^s}\right)^{-1} \\ &= \prod_{p\nmid q}\left(1+O(p^{-2\Re(s)})\right)\prod_{\mathfrak{p}|(q)}\left(1-\frac{1}{N(\mathfrak{p})^s}\right)^{-1}. \end{aligned}$$

Here we have made use of Lemma 7.7 to bound the error terms in the Euler product and assumed that $\Re(s) \geqslant 3/4$. In particular, $f(1+s)$ converges absolutely for $\Re(s) \geqslant -1/4$ and is of size $O(q^{o(1)})$ in this region.

We first move the line of integration in (8.8) to $\Re(s) = 1/\log R$ (covering a region where the integrand is analytic), giving

$$\sum_{\substack{N(\mathfrak{d})<R \\ \gcd(N(\mathfrak{d}),q)=1}} \frac{\mu(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})}\log\frac{R}{N(\mathfrak{d})} = \frac{1}{2\pi i}\int_{1/\log R-i\infty}^{1/\log R+i\infty} \frac{R^s}{s^2\zeta_K(1+s)} f(1+s)ds.$$

Using the bound $\zeta_K(1+1/\log R+it)^{-1} \ll \log(2+|t|)$ for $|t| \geqslant 1$ from Lemma 4.7, we see that the contribution from $|\Im(s)| > T := \exp(\sqrt{\log R})$ is

$$\ll \int_{t>T} \frac{q^{o(1)}\log t}{t^2}dt \ll \frac{q^{o(1)}\log T}{T}.$$

Thus, we may discard this part of the integral at the cost of a negligible error. We now move the truncated contour of integration to the left again, to $\Re(s) = -2c/\log T$, where $c = c_K/2 > 0$ is defined in terms of the constant of Lemma 4.7, so we have the bound $\zeta_K(s)^{-1} \ll \log(2+|s|)$ within this region. This introduces a term from the pole at $s=0$, an integral over the line $\Re(s) = -2c/\log T$, and small contour integrals along the lines $\Im(s) = \pm T$. The contours integrals with $|\Im(s)| = T$ contribute $O(q^{o(1)}\log T/T^2)$, and so are negligible. The contour integral with $\Re(s) = -2c/\log T$ contributes

$$\frac{1}{2\pi i}\int_{-2c/\log T-iT}^{-2c/\log T+iT} \frac{R^s f(1+s)}{s^2\zeta_K(1+s)}ds \ll q^{o(1)}(\log T)^2 R^{-2c/\log T}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll q^{o(1)} \exp(-c\sqrt{\log R}).$$

Thus, only the residue at $s = 0$ makes a nonnegligible contribution, and we have

$$\sum_{\substack{N(\mathfrak{d})<R \\ \gcd(N(\mathfrak{d}),q)=1}} \frac{\mu(\mathfrak{d})\rho(\mathfrak{d})}{N(\mathfrak{d})} \log \frac{R}{N(\mathfrak{d})} = \operatorname*{Res}_{s=0} \frac{R^s f(1+s)}{s^2 \zeta_K(1+s)} + O\left(q^{o(1)} \exp(-c\sqrt{\log R})\right)$$

$$= \gamma_K^{-1} f(1) + O\left(q^{o(1)} \exp(-c\sqrt{\log R})\right).$$

The result follows on noting that $f(1) = q^n \tilde{\mathfrak{S}}/\phi_K((q))$. □

LEMMA 8.6. *We have*

$$\sum_{\substack{\mathfrak{a},\mathfrak{b} \\ \mathfrak{a}\mathfrak{b}\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}) = \frac{q_0^n \#\mathcal{A}' \tilde{\mathfrak{S}} c_{\mathcal{R}_1\times\mathcal{R}_2}(X_0^n)}{\phi_K((q_0))\gamma_K}\left(1+\frac{\chi^*(\mathfrak{a}_0)}{(-\beta^*)^\ell X_0^{n-n\beta^*}}\right)+O(\eta_2^{1/3}\#\mathcal{A}'),$$

where $q_0 = J!q^*$.

Proof. This is a sieve calculation, relying on Proposition 7.5 and the prime ideal theorem in the form (8.5) and (8.6). We substitute the definition (8.7) of $\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$ and swap the order of summation to give

$$\sum_{\substack{\mathfrak{a},\mathfrak{b} \\ \mathfrak{a}\mathfrak{b}\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$$

$$= \sum_{\mathfrak{a}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \sum_{N(\mathfrak{d})<R} \lambda_{\mathfrak{d}} \sum_{\substack{\mathfrak{u}\in\mathcal{A}' \\ \mathfrak{a}\mathfrak{d}|\mathfrak{u}}} c_{\mathcal{R}_2}(N(\mathfrak{u}/\mathfrak{a}))\left(1+\frac{\chi^*(\mathfrak{u}/\mathfrak{a})}{(-\beta^*)^{\ell'} N(\mathfrak{u}/\mathfrak{a})^{1-\beta^*}}\right). \quad (8.9)$$

We wish to replace $c_{\mathcal{R}_2}(N(\mathfrak{u}/\mathfrak{a}))$ with $c_{\mathcal{R}_2}(X_0^n/N(\mathfrak{a}))$. Since all ideals in $\mathcal{A}'$ have norm $X_0^n + O(\eta_2 X_0^n)$ with $X_0^n \gg X^n$, we have that $c_{\mathcal{R}_2}(N(\mathfrak{u}/\mathfrak{a})) = c_{\mathcal{R}_2}(X_0^n/N(\mathfrak{a})) + O(\eta_2^{1/2})$ by Lemma 8.3. This error term contributes

$$\ll \eta_2^{1/2} \log X \sum_{\mathfrak{a}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \sum_{N(\mathfrak{d})<R} \frac{|\lambda_{\mathfrak{d}}|}{\log X} \#\mathcal{A}'_{\mathfrak{a}\mathfrak{d}}.$$

We recall $\mathbf{1}_{\mathcal{R}_1}$ is supported on ideals $\mathfrak{a}$ with $N(\mathfrak{a}) \ll X^{n-k-\epsilon}$ and with all prime factors $\mathfrak{p}$ of $\mathfrak{a}$ satisfying $N(\mathfrak{p}) \geqslant X^{\epsilon^2} = R$. Thus, $\mathfrak{d}, \mathfrak{a}$ must be coprime if they make a contribution to the sum. We let $\mathfrak{e} = \mathfrak{a}\mathfrak{d}$ and recall that $|\lambda_{\mathfrak{d}}| \ll \log X$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Putting this together, the error in replacing $c_{\mathcal{R}_2}(N(\mathfrak{u}/\mathfrak{a}))$ with $c_{\mathcal{R}_2}(X_0^n/N(\mathfrak{a}))$ contributes a total

$$\ll \eta_2^{1/2} \log X \sum_{N(\mathfrak{e}) \ll RX^{n-k-\epsilon}} \#\mathcal{A}'_{\mathfrak{e}} \ll \eta_2^{1/2} \log X \#\mathcal{A}' \sum_{N(\mathfrak{e}) < X^{n-2k-\epsilon/2}} \frac{\rho(\mathfrak{e})}{N(\mathfrak{e})} + X^{n-k-\epsilon/2n}$$

by Proposition 7.5, noting that if $\gcd(N(\mathfrak{e}), q_0) \neq 1$, then $\#\mathcal{A}'_{\mathfrak{e}} = 0$. The sum here is $O(\log X)$ by an Euler product upper bound and Lemma 7.7, so the total error is $O(\eta_2^{1/3}\#\mathcal{A}')$.

An essentially identical argument shows that we can replace $N(\mathfrak{u}/\mathfrak{a})^{1-\beta^*}$ in (8.9) with $X_0^{n-n\beta^*}/N(\mathfrak{a})^{1-\beta^*}$ at the cost of an error term of size $O(\eta_2^{1/2}\#\mathcal{A}')$.

Since all elements $\mathfrak{u}$ of $\mathcal{A}'$ have $\chi^*(\mathfrak{u}) = \chi^*(\mathfrak{a}_0)$, we are left to evaluate

$$\sum_{\mathfrak{a}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) c_{\mathcal{R}_2}\left(\frac{X_0^n}{N(\mathfrak{a})}\right)\left(1 + \frac{\chi^*(\mathfrak{a})\chi^*(\mathfrak{a}_0)N(\mathfrak{a})^{1-\beta^*}}{(-\beta^*)^{\ell'} X_0^{n-n\beta^*}}\right) \sum_{N(\mathfrak{d})<R} \lambda_{\mathfrak{d}} \#\mathcal{A}'_{\mathfrak{a}\mathfrak{d}}.$$

Since all elements of $\mathcal{A}'$ have norm coprime to q_0, we can restrict to $\gcd(N(\mathfrak{d}), q_0) = 1$. Using Proposition 7.5, again, we may then replace $\#\mathcal{A}'_{\mathfrak{a}\mathfrak{d}}$ with $\rho(\mathfrak{a}\mathfrak{d})\#\mathcal{A}'/N(\mathfrak{a}\mathfrak{d})$ at the cost of an error $O(X^{n-k-\epsilon/2n})$, which is negligible. Thus, we are left to evaluate

$$\#\mathcal{A}' \sum_{\mathfrak{a}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) c_{\mathcal{R}_2}\left(\frac{X_0^n}{N(\mathfrak{a})}\right)\left(1 + \frac{\chi^*(\mathfrak{a})\chi^*(\mathfrak{a}_0)N(\mathfrak{a})^{1-\beta^*}}{(-\beta^*)^{\ell'} X_0^{n-n\beta^*}}\right) \sum_{\substack{N(\mathfrak{d})<R \\ \gcd(N(\mathfrak{d}),q_0)=1}} \frac{\lambda_{\mathfrak{d}}\rho(\mathfrak{a}\mathfrak{d})}{N(\mathfrak{a}\mathfrak{d})}.$$

Any pairs $\mathfrak{a}$, $\mathfrak{d}$ making a contribution must be coprime since $\mathbf{1}_{\mathcal{R}_1}$ is supported on ideals with all factors having norm at least R. Thus, we may replace $\rho(\mathfrak{a}\mathfrak{d})$ with $\rho(\mathfrak{d})\rho(\mathfrak{a})$, and so the double sum factorizes as

$$\left(\sum_{\mathfrak{a}} \frac{\rho(\mathfrak{a})\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})}{N(\mathfrak{a})} c_{\mathcal{R}_2}\left(\frac{X_0^n}{N(\mathfrak{a})}\right)\left(1 + \frac{\chi^*(\mathfrak{a})\chi^*(\mathfrak{a}_0)N(\mathfrak{a})^{1-\beta^*}}{(-\beta^*)^{\ell'} X_0^{n-n\beta^*}}\right)\right) \times \left(\sum_{\substack{N(\mathfrak{d})<R \\ \gcd(N(\mathfrak{d}),q_0)=1}} \frac{\lambda_{\mathfrak{d}}\rho(\mathfrak{d})}{N(\mathfrak{d})}\right).$$

By Lemma 8.5, we have that the second factor is $q_0^n \widetilde{\mathfrak{S}}/\gamma_K \phi_K((q_0)) + O(q_0^{o(1)} \exp(-c\sqrt{\log R}))$.

Since all degree 1 prime ideals have $\rho(\mathfrak{p}) = 1$, we see that $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\rho(\mathfrak{a}) = \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})$ unless $p^2 | N(\mathfrak{a})$ for some $p > X^{\epsilon^2}$. Thus, we can replace $\rho(\mathfrak{a})$ with the constant 1 in the first factor at the cost of an error

$$\ll \sum_{p > X^{\epsilon^2}} \sum_{\substack{N(\mathfrak{a}) < X^{n-k-\epsilon/2} \\ p^2 | N(\mathfrak{a})}} \frac{\rho(\mathfrak{a})}{N(\mathfrak{a})} \ll \sum_{p > X^{\epsilon^2}} \frac{1}{p^2} \prod_{N(\mathfrak{p})<X} \left(1 + \frac{\rho(\mathfrak{p})}{N(\mathfrak{p})}\right) \ll X^{-\epsilon^2} \log X,$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

by Lemma 7.7. This is negligible, and we can evaluate the resulting expressions by partial summation, the prime ideal theorem (Lemma 4.3) and (8.6). We have

$$\sum_{\mathfrak{a}} \frac{\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})}{N(\mathfrak{a})} c_{\mathcal{R}_2}\left(\frac{X_0^n}{N(\mathfrak{a})}\right) = \int \cdots \int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R}_1\times\mathcal{R}_2 \\ \sum_{i=1}^{\ell} e_i \in \mathcal{I}_{X_0^n}}} \frac{de_1 \dots de_\ell}{\eta_2^{1/2} \prod_{i=1}^{\ell} e_i} + O(\eta_2),$$

$$\sum_{\mathfrak{a}} \frac{\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})\chi^*(\mathfrak{a})}{N(\mathfrak{a})^{\beta^*}} c_{\mathcal{R}_2}\left(\frac{X_0^n}{N(\mathfrak{a})}\right) = \int \cdots \int_{\substack{(e_1,\dots,e_\ell)\in\mathcal{R}_1\times\mathcal{R}_2 \\ \sum_{i=1}^{\ell} e_i \in \mathcal{I}_{X_0^n}}} \frac{de_1 \dots de_\ell}{(-\beta^*)^{\ell-\ell'} \eta_2^{1/2} \prod_{i=1}^{\ell} e_i} + O(\eta_2).$$

Combining these estimates gives the result. □

With these lemmas in place, we can reduce the proof of Proposition 8.2 to the following proposition.

PROPOSITION 8.7. *Let $\mathfrak{c}$ be a fixed ideal. Uniformly over all hypercubes $\mathcal{R}_1 \times \mathcal{R}_2$ intersecting $\mathcal{R}$ and uniformly over all $\mathcal{A}' \subseteq \mathcal{A}$, we have*

$$\sum_{\substack{\mathfrak{a},\ \mathfrak{b} \text{ principal} \\ \mathfrak{c}|\mathfrak{b},\ \mathfrak{c}'|\mathfrak{a} \\ \mathfrak{a}\mathfrak{b}/N(\mathfrak{c})\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}')(\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})) \ll \eta_2^{1/2} \#\mathcal{A}'. \tag{8.10}$$

Proof of Proposition 8.2 assuming Proposition 8.7. Lemma 8.4 gives the second statement of Proposition 8.2, and Lemma 8.6 gives the first statement with $\tilde{\mathbf{1}}_{\mathcal{R}_2}$ in place of $\mathbf{1}_{\mathcal{R}_2}$. We are therefore left to show that the error introduced by replacing $\mathbf{1}_{\mathcal{R}_2}$ with $\tilde{\mathbf{1}}_{\mathcal{R}_2}$ in the first statement is suitably small. In particular, it is sufficient to show that uniformly over all hypercubes $\mathcal{C} = \mathcal{R}_1 \times \mathcal{R}_2$ with $\mathcal{C} \cap \mathcal{R} \neq \emptyset$ and all sets $\mathcal{A}'$

$$\sum_{\mathfrak{a}\mathfrak{b}\in\mathcal{A}'} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a})(\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})) \ll \eta_2^{1/2} \#\mathcal{A}'.$$

We split the sum over $\mathfrak{b}$ into ideal classes $\mathcal{C} \in Cl_K$. We let $\mathfrak{c}$ be an ideal in $\mathcal{C}$, and $\mathfrak{c}' = (N(\mathfrak{c}))/\mathfrak{c}$. Then $\mathfrak{a}\mathfrak{c}'$ and $\mathfrak{b}\mathfrak{c}$ are both principal integral ideals, and so can be written as (α), (β) say with $\mathfrak{c}'|(\alpha)$ and $\mathfrak{c}|(\beta)$. The above estimate now follows immediately from Proposition 8.7. □

Thus, we are left to establish Proposition 8.7.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

9. Localized ideal counts

The aim of this section is to show that $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) \approx \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})$ when $\mathfrak{b}$ is localized to a particular ideal class, residue mod q and angle of Hecke character. The main result of this section is Proposition 9.7, which will be important in establishing Proposition 8.7 (and, hence, Theorems 1.1 and 1.2) in the later sections.

LEMMA 9.1. *Let $\lambda_1, \dots, \lambda_{n-1}$ be a fixed basis of the torsion-free Hecke characters of K. Let $\Delta > 0$. Let $\alpha \in \mathcal{O}_K$ and $\mathfrak{a} = (\alpha)$. Let $\mathfrak{b}$ be a principal ideal such that for each $j \in \{1, \dots, n-1\}$, we have*

$$|\lambda_j(\mathfrak{b}) - \lambda_j(\mathfrak{a})| \leqslant \Delta$$

and such that $|N(\mathfrak{b}) - N(\mathfrak{a})| \leqslant \Delta N(\mathfrak{a})$.

Then there is a generator β of $\mathfrak{b}$ such that

$$\beta = \alpha(1 + O(\Delta)).$$

We caution that the implied constant above may depend on the choice of basis, but for the purposes of this paper, we just consider a single fixed basis.

Proof. This fact is given, for example, in [7, Section 3.2]. Alternatively, it follows from the characterizations of torsion-free characters from [19, Ch. 7, §6]. A torsion-free character (that is, of pure infinity type) takes the form

$$\chi((\gamma)) = \exp\left(\sum_\sigma \left(p_\sigma \log\left(\frac{\gamma^\sigma}{|\gamma^\sigma|}\right) + i q_\sigma \log|\gamma^\sigma|\right)\right),$$

where the sum is over embeddings σ, $p_\sigma \in \mathbb{Z}$ satisfy $p_\sigma p_{\overline{\sigma}} = 0$ and $q_\sigma \in \mathbb{R}$ satisfy $q_\sigma = q_{\overline{\sigma}}$ and $\sum_\sigma q_\sigma = 0$. Provided the right-hand side is trivial on units, this is a well-defined character on principal ideals.

The result follows from Lemma 4.2 if $\Delta \gg 1$, so we may assume that Δ is sufficiently small. Let $\epsilon_1, \dots, \epsilon_{r_1+r_2-1}$ be a basis for the torsion-free units in $\mathcal{O}_K$. Given $\mathbf{m} = (m_1, \dots, m_{r_1+r_2-1})$, there is a choice of coefficients q_σ such that

$$\sum_\sigma q_\sigma \log|\epsilon_j^\sigma| = 2\pi m_j.$$

(This is a system of linearly independent linear equations—the linear independence follows from the nonvanishing of the regulator.) By considering

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$\mathbf{m}$ as the standard basis vectors of $\mathbb{Z}^{r_1+r_2-1}$, we see that there are choices $q_{j,\sigma}$ for $1 \leqslant j \leqslant r_1 + r_2 - 1$ such that

$$i \sum_{\sigma} q_{j,\sigma} \log |\epsilon_r^{\sigma}| = \begin{cases} 2\pi i & \text{if } r = j, \\ 0 & \text{otherwise,} \end{cases}$$

and these give rise to Hecke characters $\chi_1, \dots, \chi_{r_1+r_2-1}$ such that

$$\chi_j((\gamma)) = \exp\left(i \sum_{\sigma} q_{j,\sigma} \log |\gamma^{\sigma}| \right)$$

since the right-hand side is invariant under multiplication of γ by units. These χ_j are torsion-free, and so of the form $\lambda_1^{e_1} \dots \lambda_{n-1}^{e_{n-1}}$ for some $e_1, \dots, e_{n-1} \ll_{\mathbf{m}} 1$. Thus, since $\lambda_j((\alpha)) = \lambda_j((\beta)) + O(\Delta)$, we have $\chi_j((\alpha)) = \chi_j((\beta)) + O(\Delta)$, and so

$$i \sum_{\sigma} q_{j,\sigma} \log |\alpha^{\sigma}| = i \sum_{\sigma} q_{j,\sigma} \log |\beta^{\sigma}| + O(\Delta) \pmod{2\pi i}.$$

But, by construction, we see that we can find $\beta_2 = \epsilon_1^{m_1} \dots \epsilon_{r_1+r_2-1}^{m_{r_1+r_2-1}} \beta$ for suitable $\mathbf{m} \in \mathbb{Z}$ such that

$$i \sum_{\sigma} q_{j,\sigma} \log |\alpha^{\sigma}| = i \sum_{\sigma} q_{j,\sigma} \log |\beta_2^{\sigma}| + O(\Delta).$$

Moreover, since $N(\beta_2) = N(\alpha)(1 + O(\Delta))$, we also have $\sum_{\sigma} \log |\beta_2^{\sigma}| = \sum_{\sigma} \log |\alpha^{\sigma}|(1 + O(\Delta))$. Thus, we see that since the $(q_{j,\sigma})_{\sigma}$ are linearly independent, $|\beta_2^{\sigma}| = (1 + O(\Delta))|\alpha^{\sigma}|$ for all σ.

Similarly, we can choose $p_{\sigma_0} = 1$ for a complex embedding σ_0, and $p_{\sigma} = 0$ for all other embeddings, and then find constants q_{σ} such that

$$\chi_{\sigma_0}((\gamma)) = \exp\left(\log\left(\frac{\gamma^{\sigma_0}}{|\gamma^{\sigma_0}|} \right) + \sum_{\sigma} q_{\sigma} \log |\gamma^{\sigma}| \right)$$

is a Hecke character. Again, we must have $\chi_{\sigma_0}(\alpha) = \chi_{\sigma_0}(\beta_2)(1 + O(\Delta))$. But since $|\alpha^{\sigma}| = |\beta_2^{\sigma}|(1 + O(\Delta))$ for all σ, we see that this implies $\alpha^{\sigma_0} = \beta_2^{\sigma_0}(1 + O(\Delta))$. Thus, we have that $\alpha^{\sigma} = \beta_2^{\sigma}(1 + O(\Delta))$ for all complex embeddings σ and that $|\alpha^{\sigma}| = |\beta_2^{\sigma}|(1 + O(\Delta))$ for all real embeddings. From this, we see that $\alpha = \beta_2(1 + O(\Delta))$. □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

LEMMA 9.2. *Let $\lambda_1, \dots, \lambda_{n-1}$ be a basis of the torsion-free Hecke characters, and define*

$$W(\mathfrak{a}; \mathfrak{b}; \Delta) = \begin{cases} \displaystyle\prod_{j=1}^{n-1}\left(1 - \frac{1}{2\pi\Delta}\left|\arg\left(\frac{\lambda_j(\mathfrak{a})}{\lambda_j(\mathfrak{b})}\right)\right|\right), & \text{if } \left|\arg\left(\frac{\lambda_j(\mathfrak{a})}{\lambda_j(\mathfrak{b})}\right)\right| \leqslant 2\pi\Delta \ \forall j, \\ 0, & \text{otherwise.} \end{cases}$$

Let $A \asymp B^n$ and $\Delta > A^{-\epsilon^2/2n}$. Then we have

$$\sum_{\substack{A \leqslant N(\mathfrak{a}) \leqslant A + \Delta A \\ \mathfrak{a} \text{ principal}}} W(\mathfrak{a}; \mathfrak{b}; \Delta) = \frac{\gamma_K \Delta^n A}{h_K}(1 + O(\Delta)).$$

(Here we use the branch of $\arg(x)$ *such that* $\arg(x) \in [-\pi, \pi)$.*)*

Proof. The result is trivial if $\Delta \gg 1$, so we assume that Δ is sufficiently small. By Fourier expansion, if $|z| = 1$, then

$$2\pi\Delta \sum_{m \in \mathbb{Z}} z^m \left(\frac{\sin \pi m \Delta}{\pi m \Delta}\right)^2 = \begin{cases} \left(1 - \dfrac{1}{2\pi\Delta}|\arg(z)|\right), & \text{if } |\arg(z)| \leqslant 2\pi\Delta, \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$\begin{aligned} W(\mathfrak{a}; \mathfrak{b}; \Delta) &= \Delta^{n-1} \sum_{\mathbf{m} \in \mathbb{Z}^{n-1}} \prod_{j=1}^{n-1} \frac{\lambda_j(\mathfrak{a})^{m_j}}{\lambda_j(\mathfrak{b})^{m_j}} \left(\frac{\sin \pi m_j \Delta}{\pi m_j \Delta}\right)^2 \\ &= \Delta^{n-1} \sum_{\mathbf{m} \in \mathbb{Z}^{n-1}} \frac{\chi^{\mathbf{m}}(\mathfrak{a})}{\chi^{\mathbf{m}}(\mathfrak{b})} \hat{w}(\mathbf{m}). \end{aligned} \tag{9.1}$$

Here $\chi^{\mathbf{m}}(\mathfrak{a}) = \prod_{j=1}^{n-1} \lambda_j^{m_j}(\mathfrak{a})$, $\hat{w}(\mathbf{m}) = \prod_{j=1}^{n-1} (\sin \pi m_j \Delta / \pi m_j \Delta)^2$, and we take $\sin \pi m_j \Delta / \pi m_j \Delta$ to be 1 when $m_j = 0$.

We note that

$$\sum_{\substack{A \leqslant N(\mathfrak{a}) \leqslant A + \Delta A \\ \mathfrak{a} \text{ principal}}} W(\mathfrak{a}; \mathfrak{b}; \Delta) = \frac{\Delta^{n-1}}{h_K} \sum_{\xi} \sum_{\mathbf{m} \in \mathbb{Z}^{n-1}} \chi^{\mathbf{m}}(\mathfrak{b})^{-1} \hat{w}(\mathbf{m}) \sum_{A \leqslant N(\mathfrak{a}) \leqslant A + \Delta A} \chi^{\mathbf{m}}(\mathfrak{a}) \xi(\mathfrak{a}), \tag{9.2}$$

where ξ runs over all characters of the class group Cl_K.

Since $\hat{w}(\mathbf{m}) \ll \prod_{j=1}^{n-1} \min(1, (m_j \Delta)^{-2})$, those terms with $m_j > M_0$ for some j contribute $O(\Delta^{-n+2} A / M_0)$ in total to (9.2). Choosing $M_0 = \Delta^{-2n}$ shows that these contribute $O(\Delta^{n+2} A)$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

If $\|\mathbf{m}\| \ll M_0 < A^{\epsilon^2}$ and $\chi^{\mathbf{m}}\xi$ is nontrivial, then the inner sum over $\mathfrak{a}$ in (9.2) is $O(A^{1-\epsilon})$ by Perron's formula and the bound $L_K(s, \chi^{\mathbf{m}}\xi) \ll O(|s|+\|\mathbf{m}\|)^{n(1-\sigma)/2}$ from Lemma 4.6. Thus, these terms contribute $O(\Delta^{n-1}M_0^{n-1}A^{1-\epsilon}) = O(\Delta^{2n}A)$ in total to (9.2).

Finally, the term with $\chi^{\mathbf{m}}\xi = 1$ contributes $\gamma_K \Delta^n A(1 + O(\Delta))/h_K$. Putting these estimates together gives the result. □

LEMMA 9.3. *Let $\mathfrak{c}$ be a fixed ideal and $q \ll q^{*\log\log B}\exp(\sqrt[6]{\log B})$ with $(\theta n)^n N(\mathfrak{c})|q$. Let $\beta_0, \alpha \in \mathcal{O}_K$ be such that $\gcd((q), (\beta_0)) = \mathfrak{c}$ and $\beta_0 = \alpha(1 + O(\delta_0))$. Let $\Delta = \delta_0^n$ and $N_K(\alpha) \asymp B^k$. Define*

$$V(\alpha) = \sum_{\substack{\beta\in\mathcal{O}_K\\ |\beta-\alpha|\leqslant\Delta|\alpha|/\delta_0^{1/2n}\\ \beta\equiv\beta_0 \pmod q\\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a})}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})W(\mathfrak{a};\mathfrak{b};\Delta).$$

Then if $\mathfrak{q}^*|(q)/\mathfrak{c}$, we have

$$V(\alpha) = \frac{\Delta^{n-1}}{h_K\phi_K((q)/\mathfrak{c})} \times \sum_{\substack{\mathfrak{b}\\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a})\\ \mathfrak{c}|\mathfrak{b}}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})\big(1 + \chi^*(\mathfrak{b}/\mathfrak{c})\overline{\chi^*(\mathfrak{b}_0/\mathfrak{c})}\big) + O(\delta_0^{1/2}\Delta^n B^n).$$

If instead $\mathfrak{q}^* \nmid (q)/\mathfrak{c}$, we have

$$V(\alpha) = \frac{\Delta^{n-1}}{h_K\phi_K((q)/\mathfrak{c})} \sum_{\substack{\mathfrak{b}\\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a})\\ \mathfrak{c}|\mathfrak{b}}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) + O(\delta_0^{1/2}\Delta^n B^n).$$

Here $\mathfrak{b}$ denotes the ideal (β) generated by $\beta \in \mathcal{O}_K$. Similarly, $\mathfrak{a} = (\alpha)$ denotes the ideal generated by α and $\mathfrak{b}_0 = (\beta_0)$ the ideal generated by β_0.

Proof. We first detect $\beta \equiv \beta_0' \pmod q$ by characters $\chi_{\mathfrak{f}}$ of the multiplicative group $(\mathcal{O}_K/\mathfrak{f})^\times$ where $\mathfrak{f} = (q)/\mathfrak{c}$. Since $\gcd((q), (\beta_0)) = \mathfrak{c}$, we see that β/β_0' can be viewed as an element of $\mathcal{O}_K/\mathfrak{f}$ if $\mathfrak{c}|\mathfrak{b}$. We see that $\#(\mathcal{O}_K/\mathfrak{f})^\times = \phi_K((q)/\mathfrak{c})$, and so

$$V(\alpha) = \frac{1}{\phi_K((q)/\mathfrak{c})}\sum_{\chi_{\mathfrak{f}}} \sum_{\substack{\beta\in\mathcal{O}_K\\ |\beta-\alpha|\leqslant\Delta|\alpha|/\delta_0^{1/2n}\\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta\\ \mathfrak{c}|\mathfrak{b}}} \chi_{\mathfrak{f}}(\beta/\beta_0')\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})W(\mathfrak{a},\mathfrak{b};\Delta), \tag{9.3}$$

where $\sum_{\chi_{\mathfrak{f}}}$ is a sum over all characters of $(\mathcal{O}_K/\mathfrak{f})^\times$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The characters $\chi_{\mathfrak{f}}$ are not characters of ideals, and so we first translate them to this setting. Given a character $\chi_{\mathfrak{f}}$ on $(\mathcal{O}_K/\mathfrak{f})^\times$, as in the proof of Lemma 9.1, there is a choice of constants $p_{\sigma,\chi_{\mathfrak{f}}}, q_{\sigma,\chi_{\mathfrak{f}}} \ll 1$ for each embedding $\sigma : K \hookrightarrow \mathbb{C}$ such that

$$\tilde{\chi}_{\mathfrak{f}}(\gamma) = \chi_{\mathfrak{f}}(\gamma) \exp\Big(\sum_\sigma q_{\sigma,\chi_{\mathfrak{f}}} \log|\gamma^\sigma|\Big) \prod_\sigma (\gamma^\sigma/|\gamma^\sigma|)^{p_{\sigma,\chi_{\mathfrak{f}}}}$$

is trivial on units of $\mathcal{O}_K$. This then defines a character on principal ideals coprime to $\mathfrak{f}$, which we can lift to a character on all ideals coprime to $\mathfrak{f}$. The resulting character is not unique since there are $O(1)$ possible choices of the constants $p_{\sigma,\chi_{\mathfrak{f}}}, q_{\sigma,\chi_{\mathfrak{f}}}$ and the lift is only unique up to multiplication by Hilbert characters. This lack of uniqueness is irrelevant to us, so we arbitrarily fix a lift for each $\chi_{\mathfrak{f}}$, which we also denote by $\tilde{\chi}_{\mathfrak{f}}$.

We would like to replace $\chi_{\mathfrak{f}}(\beta/\beta_0')$ by $\tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{b}_0')$ in (9.3) so that we have characters of ideals. Since $\mathbf{b}, \mathbf{b}_0' \in \mathcal{C}$, we have $\mathbf{b} = \mathbf{b}_0'(1 + O(\delta_0))$, and so, since log is continuous, $\chi_{\mathfrak{f}}(\beta/\beta_0') = \tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{b}_0')(1 + O(\delta_0))$. This error term $O(\delta_0)$ contributes

$$\ll \frac{1}{\phi_K((q)/\mathfrak{c})} \sum_{\chi_{\mathfrak{f}}} \sum_{\beta=\alpha(1+O(\Delta/\delta_0^{1/2n}))} \delta_0 \ll \delta_0^{1/2} \Delta^n B^n$$

to (9.3), which is negligible. Thus,

$$V(\alpha) = \frac{1}{\phi_K((q)/\mathfrak{c})} \times \sum_{\chi_{\mathfrak{f}}} \sum_{\substack{\beta\in\mathcal{O}_K \\ |\beta-\alpha|\leqslant\Delta|\alpha|/\delta_0^{1/2n} \\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta \\ \mathfrak{c}|\mathfrak{b}}} \tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{b}_0')\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) W(\mathfrak{a},\mathfrak{b};\Delta) + O(\delta_0^{1/2}\Delta^n B^n).$$

Since all β in the above sum satisfy $\beta = \alpha(1 + O(\Delta/\delta_0^{1/2n}))$ and that $\Delta/\delta_0^{1/2n}$ is sufficiently small, we see that no two terms appearing are associates. Therefore, (β) ranges over a set of principal prime ideals $\mathfrak{b}$ with $|N(\mathfrak{b}) - N(\mathfrak{a})| \leqslant \Delta N(\mathfrak{a})$. Since $W(\mathfrak{a},\mathfrak{b};\Delta) = 0$ unless $|\lambda_j(\mathfrak{a}) - \lambda_j(\mathfrak{b})| \ll \Delta$, we may restrict the summation over β such that this holds. But then by Lemma 9.1, every such ideal $\mathfrak{b}$ occurs exactly once in the above sum. Therefore,

$$V(\alpha) = \frac{1}{\phi_K((q)/\mathfrak{c})} \sum_{\chi_{\mathfrak{f}}} \sum_{\substack{\mathfrak{b} \text{ principal} \\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta \\ \mathfrak{c}|\mathfrak{b}}} \tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{b}_0')\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) W(\mathfrak{a},\mathfrak{b};\Delta) + O(\delta_0^{1/2}\Delta^n B^n).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We use characters ξ of the class group Cl_K to detect the condition that $\mathfrak{b}$ is principal and insert the Fourier expansion (9.1) of W. This gives

$$V(\alpha) = \frac{\Delta^{n-1}}{h_K\phi_K((q)/\mathfrak{c})} \sum_{\mathbf{m}\in\mathbb{Z}^{n-1}} \hat{w}(\mathbf{m}) \sum_{\chi_{\mathfrak{f}}} \sum_{\xi} \chi^{\mathbf{m}}(\mathfrak{a})\chi^{-\mathbf{m}}(\mathfrak{c})\xi(\mathfrak{c})\overline{\tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}_0/\mathfrak{c})}$$
$$\times \sum_{\substack{\mathfrak{c}|\mathfrak{b}\\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta}} \chi^{-\mathbf{m}}(\mathfrak{b}/\mathfrak{c})\xi(\mathfrak{b}/\mathfrak{c})\tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{c})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) + O(\delta_0^{1/2}\Delta^n B^n). \tag{9.4}$$

By partial summation and (8.5), we have that if $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$ is nontrivial (that is, takes values not in $\{0, 1\}$) and not induced by an exceptional character χ^*, then there is a constant $c_0 > 0$ such that

$$\sum_{N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a})} \chi^{-\mathbf{m}}(\mathfrak{b}/\mathfrak{c})\xi(\mathfrak{b}/\mathfrak{c})\tilde{\chi}_{\mathfrak{f}}(\mathfrak{b}/\mathfrak{c})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) \ll B^n \exp(-c_0\sqrt{\log B})$$

uniformly for q, $\|\mathbf{m}\| \leqslant q^{*(\log\log B)^2}\exp(\sqrt[5]{\log B})$. This implies that the total contribution to (9.4) from all such characters $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$ with $\|\mathbf{m}\| \ll M_0 = \Delta^{-2n} \ll q^{*2n\log\log B}\exp(2n\sqrt[6]{\log B})$ is

$$\ll \Delta^{-1}B^n M_0^{n-1}\exp(-c_0\sqrt{\log B}),$$

which is negligible. Thus, we only need to consider characters with $\|\mathbf{m}\| > M_0$ or when $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$ is a finite order character induced by 1 or χ^*.

As before, using the trivial bound $\hat{w}(\mathbf{m}) \ll \prod_j \min(1, (m_j\Delta)^{-2})$, those characters with $\|\mathbf{m}\| \geqslant M_0$ contribute $\ll \Delta^{-2n+2}B^n M_0^{-1} \ll \delta_0\Delta^n B^n$, which is negligible. We are therefore left only with the contribution from when $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$ is induced by the trivial character 1 or is induced by χ^*.

By considering the finite part of $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$, we see that this character can only be induced by χ^* if $\mathfrak{q}^*|(q)/\mathfrak{c}$, and in this case, there is a unique choice of $\tilde{\chi}_{\mathfrak{f}}$, ξ and $\mathbf{m} \ll 1$ such that $\xi\chi^{-\mathbf{m}}\tilde{\chi}_{\mathfrak{f}}$ is induced by χ^*. Similarly, there is a unique choice of $\tilde{\chi}_{\mathfrak{f}}$, ξ and $\mathbf{m} \ll 1$ such that $\xi\chi^{-\mathbf{m}}\tilde{\chi}_{\mathfrak{f}}$ is induced by 1.

Since $\mathbf{1}_{\mathcal{R}_2}$ is supported only on ideals coprime to q (because $q < X^{\epsilon^2}$), if $\chi^{-\mathbf{m}}\xi\tilde{\chi}_{\mathfrak{f}}$ is induced by χ^*, then we can replace it with χ^*, and if it is induced by 1, we can replace it by 1. We note that $\hat{w}(\mathbf{m}) = 1+O(\Delta)$ and $\chi^{\mathbf{m}}(\mathfrak{a}/\mathfrak{b}_0') = 1+O(\delta_0)$ if $\mathbf{m} \ll 1$ and recall that $\mathfrak{b}_0$ is principal so $\xi(\mathfrak{b}_0) = 1$.

Thus, putting the above estimates together, we find that if $\mathfrak{q}^*|(q)/\mathfrak{c}$, then

$$V(\alpha) = \frac{\Delta^{n-1}}{h_K\phi_K((q)/\mathfrak{c})} \sum_{\substack{\mathfrak{b}\\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta\\ \mathfrak{c}|\mathfrak{b}}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$+\frac{\Delta^{n-1}\overline{\chi^*(\mathfrak{b}_0/\mathfrak{c})}}{h_K\phi_K((q)/\mathfrak{c})}\sum_{\substack{\mathfrak{b}\\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta\\ \mathfrak{c}|\mathfrak{b}}}\chi^*(\mathfrak{b}/\mathfrak{c})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})+O(\delta_0^{1/2}\Delta^n B^n).$$

If instead $\mathfrak{q}^*\nmid(q)/\mathfrak{c}$, then we obtain the same expression but without the second summation. This gives the result. □

LEMMA 9.4. *Let $\mathfrak{c}$ be an integral ideal of norm $O(1)$. Let δ_0 and B be quantities satisfying $\exp(-\sqrt[6]{\log B})q^{*-\log\log B}\leqslant\delta_0\leqslant\eta_2$ and $X^{1/10}\leqslant B\leqslant X$. Let $\mathcal{C}\subseteq\mathbb{R}^n$ be a hypercube of side length $\delta_0 B$ which contains a point $\mathbf{b}_0\in\mathbb{Z}^n$ such that $\|\mathbf{b}_0\|\ll B$ and $\mathfrak{b}_0=((\theta n)^{-n}\sum_{i=1}^n(\mathbf{b}_0)_i\sqrt[n]{\theta^{i-1}})$ is an integral ideal which satisfies $N(\mathfrak{b}_0)=B_0^n\gg B^n$ and $\mathfrak{c}|\mathfrak{b}_0$.*

Then uniformly over all $q\ll q^{*\log\log B}\exp(\sqrt[6]{\log B})$ with $(\theta n)^n N(\mathfrak{c})|q$ and over all such $\mathcal{C}$, $\mathbf{b}_0$, we have the following:

- *If $\gcd((q),\mathfrak{b}_0)\neq\mathfrak{c}$, then*

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0\pmod q}}\mathbf{1}_{\mathcal{R}_2}\left(\frac{\mathfrak{b}}{\mathfrak{c}}\right)=0.$$

- *If $\gcd((q),\mathfrak{b}_0)=\mathfrak{c}$ and $\chi^*(\mathfrak{b}/\mathfrak{c})=\chi^*(\mathfrak{b}_0/\mathfrak{c})$ for all $\mathbf{b}\equiv\mathbf{b}_0\pmod q$, then*

$$\begin{aligned}&\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0\pmod q}}\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})\\ &=\frac{1}{\gamma_K\phi_K((q)/\mathfrak{c})N(\mathfrak{c})}\int\cdots\int_{\substack{\mathbf{a}\in\mathcal{C},\mathbf{e}\in\mathcal{R}_2\\ \sum_{i=1}^{\ell'}e_i=\log N(\mathfrak{a}/\mathfrak{c})/\log X}}\frac{de_1\dots de_{\ell'-1}d\mathbf{a}}{\log X\prod_{i=1}^{\ell'}e_i}+O(\delta_0^{n+1/2}B^n).\\ &\quad+\frac{\chi^*(\mathfrak{b}_0/\mathfrak{c})B_0^{n(\beta^*-1)}}{\gamma_K(-\beta^*)^{\ell'}\phi_K((q)/\mathfrak{c})N(\mathfrak{c})^{\beta^*}}\int\cdots\int_{\substack{\mathbf{a}\in\mathcal{C},\mathbf{e}\in\mathcal{R}_2\\ \sum_{i=1}^{\ell'}e_i=\log N(\mathfrak{a}/\mathfrak{c})/\log X}}\frac{de_1\dots de_{\ell'-1}d\mathbf{a}}{\log X\prod_{i=1}^{\ell'}e_i}.\end{aligned}$$

- *If $\gcd((q),\mathfrak{b}_0)=\mathfrak{c}$ but $\chi^*(\mathfrak{b}/\mathfrak{c})\neq\chi^*(\mathfrak{b}_0/\mathfrak{c})$ for some $\mathbf{b}\equiv\mathbf{b}_0\pmod q$, then*

$$\begin{aligned}&\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0\pmod q}}\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})\\ &=\frac{1}{\gamma_K\phi_K((q)/\mathfrak{c})N(\mathfrak{c})}\int\cdots\int_{\substack{\mathbf{a}\in\mathcal{C},\mathbf{e}\in\mathcal{R}_2\\ \sum_{i=1}^{\ell'}e_i=\log N(\mathfrak{a}/\mathfrak{c})/\log X}}\frac{de_1\dots de_{\ell'-1}d\mathbf{a}}{\log X\prod_{i=1}^{\ell'}e_i}+O(\delta_0^{n+1/2}B^n).\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Here $\mathfrak{b}$ *denotes the ideal* $((\theta n)^{-n}\sum_{i=1}^{n} b_i\sqrt[n]{\theta^{i-1}})$ *depending on the vector* $\mathbf{b}$. *All the implied constants are effectively computable.*

Proof. Fundamentally, this is an exercise in counting localized ideals via Hecke characters, although there are some technical complications passing conditions between the vectors $\mathbf{b}$, elements of the order $\mathbb{Z}[\sqrt[n]{\theta}]$, algebraic integers β and ideals $\mathfrak{b}$.

We note that the sum is 0 if the ideal $\mathfrak{b}_0/\mathfrak{c}$ is not coprime to (q) since $\mathbf{1}_{\mathcal{R}_2}$ is nonzero only when all prime ideal factors have norm at least $X^{\epsilon^2} > N(q)$, and this gives the first statement. Thus, we may assume $\gcd((q), \mathfrak{b}_0) = \mathfrak{c}$.

We first detect the condition $\mathbf{b} \in \mathcal{C}$ by Hecke characters. Since $\mathcal{C}$ has side length $\delta_0 B$ and contains a point $\mathbf{b}_0$ with $N(\mathfrak{b}_0) = B_0^n \asymp B^n$ (from the assumptions of the lemma), we have that $N(\mathfrak{b}) = B_0^n + O(\delta_0 B_0^n)$ for all $\mathbf{b} \in \mathcal{C}$. Here, and throughout, given $\mathbf{b} \in \mathbb{Z}^n$, we let $\beta = (\theta n)^{-n}\sum_{i=1}^{n} b_i\sqrt[n]{\theta^{i-1}}$ and $\mathfrak{b} = (\beta)$. By Lemma 9.2, choosing $A = N(\mathfrak{b})$ and $\Delta = \delta_0^n$, we have

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) = \frac{h_K}{\gamma_K \Delta^n B_0^n} \sum_{\mathfrak{a}\text{ principal}} \sum_{\substack{\mathbf{b}\in\mathcal{C}\\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) W(\mathfrak{a};\mathfrak{b};\Delta) + O(\delta_0^{n+1}B^n).$$

Here we used the fact that $\Delta = \delta_0^n \leqslant \delta_0$.

Let $\mathfrak{a} = (\alpha)$ with $\alpha = (\theta n)^{-n}\sum_{i=1}^{n} a_i\sqrt[n]{\theta^{i-1}}$ for some vector $\mathbf{a}$. We see that if $W(\mathfrak{a},\mathfrak{b};\Delta) \neq 0$, then $\lambda_j(\mathfrak{a}) = \lambda_j(\mathfrak{b})(1+O(\Delta))$ for all $j \in \{1,\dots,n-1\}$. Since we also have the condition $N(\mathfrak{a}) = N(\alpha) = N(\mathfrak{b})(1+O(\Delta))$, by Lemma 9.1, there is a generator α of $\mathfrak{a}$ such that $\alpha^\sigma = \beta^\sigma(1+O(\Delta))$ for all embeddings σ, and so $\mathbf{a} = \mathbf{b}(1+O(\Delta))$. Moreover, since $\mathbf{b} \in \mathcal{C}$, a hypercube of elements of norm $\gg B^n$ of side length $\delta_0 B$, all such α lie within a fundamental domain for the action by the unit group of $\mathcal{O}_K$. In particular, the α such that $\mathbf{a}$ is within $O(\Delta B)$ of $\mathcal{C}$ are in one-to-one correspondence with a set containing all the ideals $\mathfrak{a}$ making a nonzero contribution.

If the distance from $\mathbf{a}$ to the boundary of $\mathcal{C}$ is a sufficiently large multiple of ΔB, then the vectors $\mathbf{b}$ with $\mathbf{a} = \mathbf{b}(1+O(\Delta))$ are either all outside of $\mathcal{C}$ or all inside $\mathcal{C}$ depending on whether $\mathbf{a} \notin \mathcal{C}$ or $\mathbf{a} \in \mathcal{C}$. Since there are $O(\Delta B^n)$ vectors $\mathbf{a}$ within $O(\Delta B)$ of the boundary of $\mathcal{C}$, these $\mathbf{a}$ contribute a total

$$\ll \frac{h_K \Delta B^n}{\gamma_K \Delta^n B_0^n} \sup_{\|\mathbf{a}\|\ll B} \sum_{\mathbf{b}=\mathbf{a}+O(\Delta B)} 1 \ll \Delta B^n \ll \delta_0^{n+1} B^n.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Thus, we can restrict to $\mathbf{a} \in \mathcal{C}'$, a hypercube inside $\mathcal{C}$ with all points at least a certain multiple of ΔB from the boundary of $\mathcal{C}$. This leaves us with

$$\frac{h_K}{\gamma_K \Delta^n B_0^n} \sum_{\substack{\mathbf{a}\in\mathcal{C}' \\ \alpha\in\mathcal{O}_K}} \sum_{\substack{\mathbf{b}=\mathbf{a}(1+O(\Delta)) \\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q \\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) W(\mathfrak{a};\mathfrak{b};\Delta). \tag{9.5}$$

We can relax the condition $\mathbf{b} = \mathbf{a}(1 + O(\Delta))$ to $|\mathbf{b} - \mathbf{a}| \leqslant \Delta\|\mathbf{a}\|/\delta_0^{1/2n}$ since by our above discussion, the additional terms make no contribution.

We now consider the condition $\mathbf{b} \equiv \mathbf{b}_0 \pmod q$. We see that $\mathbf{b} \equiv \mathbf{b}_0 \pmod q$ is equivalent to $\beta = (\theta n)^{-n} \sum_{i=1}^n b_i \sqrt[n]{\theta^{i-1}} \in \mathcal{O}_K$ and $\beta \equiv \beta_0' \pmod q$ over $\mathcal{O}_K$ for one of $[(\theta n)^{-n}\mathbb{Z}[\sqrt[n]{\theta}] : \mathcal{O}_K] \ll 1$ different algebraic integers β_0'. (Here we are using the fact that $(\theta n)^n | q$ and $\mathfrak{b}_0$ is integral.) We may choose β_0' such that $\beta_0' = (\theta n)^{-n} \sum_{i=1}^n (b_0')_i \sqrt[n]{\theta^{i-1}}$ for some vector $\mathbf{b}_0' \in \mathcal{C}$. We consider each such β_0' separately. By Lemma 9.3, the inner sum depends on whether $\mathfrak{q}^* | (q)/\mathfrak{c}$ or not. We argue now in the case when this happens; if $\mathfrak{q}^* \nmid (q)/\mathfrak{c}$, the argument is identical with all terms involving χ^* simply omitted. By Lemma 9.3, we find that

$$\begin{aligned}\sum_{\substack{\beta\in\mathcal{O}_K \\ \beta\equiv\beta_0' \pmod q \\ 1\leqslant N(\mathfrak{a}/\mathfrak{b})\leqslant 1+\Delta}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) W(\mathfrak{a},\mathfrak{b};\Delta) &= O(\delta_0^{1/2}\Delta^n B^n) \\ &+ \frac{\Delta^{n-1}}{h_k \phi_K((q)/\mathfrak{c})} \sum_{\substack{\mathfrak{b} \\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a}) \\ \mathfrak{c}|\mathfrak{b}}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})\big(1 + \chi^*(\mathfrak{b}/\mathfrak{c})\overline{\chi^*(\mathfrak{b}_0/\mathfrak{c})}\big).\end{aligned}$$

We can estimate the inner sum of (9.4) by partial summation and Lemma 4.5, giving

$$\begin{aligned}\sum_{\substack{\mathfrak{b} \\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a}) \\ \mathfrak{c}|\mathfrak{b}}} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) &= \int_{\mathbf{e}\in\mathcal{R}_2\cap I(\mathfrak{a})} \cdots \int \frac{X^{\sum_{i=1}^{\ell'} e_i} de_1 \dots de_{\ell'}}{\prod_{i=1}^{\ell'} e_i} \\ &\quad + O(B^n \exp(-c_0\sqrt{\log B})), \\ \sum_{\substack{\mathfrak{b} \\ N(\mathfrak{a})/(1+\Delta)\leqslant N(\mathfrak{b})\leqslant N(\mathfrak{a}) \\ \mathfrak{c}|\mathfrak{b}}} \chi^*(\mathfrak{b}/\mathfrak{c})\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) &= \int_{\mathbf{e}\in\mathcal{R}_2\cap I(\mathfrak{a})} \cdots \int \frac{X^{\beta^* \sum_{i=1}^{\ell'} e_i} de_1 \dots de_{\ell'}}{(-\beta^*)^{\ell'} \prod_{i=1}^{\ell'} e_i} \\ &\quad + O(B^n \exp(-c_0\sqrt{\log B})),\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

where

$$I(\mathfrak{a}) = \left\{ \mathbf{e} \in \mathbb{R}^{\ell'} : \frac{\log N(\mathfrak{a}/\mathfrak{c})}{(1+\Delta)\log X} \leqslant \sum_{i=1}^{\ell'} e_i \leqslant \frac{\log N(\mathfrak{a}/\mathfrak{c})}{\log X} \right\}.$$

We note that $\hat{w}(\mathbf{m}) = 1 + O(\Delta)$ and $\chi^{\mathbf{m}}(\mathfrak{a}/\mathfrak{b}_0') = 1 + O(\delta_0)$ if $\mathbf{m} \ll 1$, and recall that $\mathfrak{b}_0$ is principal so $\xi(\mathfrak{b}_0) = 1$. Thus, (9.5) simplifies to give

$$\sum_{\substack{\mathbf{b}\in\mathcal{C} \\ \beta\equiv\beta_0' \pmod q}} \mathbf{1}_{\mathcal{R}_2}\left(\frac{\mathfrak{b}}{\mathfrak{c}}\right) = \frac{1}{\gamma_K \Delta B_0^n \phi_K((q)/\mathfrak{c})} \sum_{\substack{\mathbf{a}\in\mathcal{C}' \\ \alpha\in\mathcal{O}_K}} \int \cdots \int_{\mathbf{e}\in\mathcal{R}_2\cap I(\mathfrak{a})} \frac{X^{\sum_{i=1}^{\ell'} e_i} de_1 \dots de_{\ell'}}{\prod_{i=1}^{\ell'} e_i}$$

$$+ \frac{1}{\gamma_K \Delta B_0^n \phi_K((q)/\mathfrak{c})} \sum_{\substack{\mathbf{a}\in\mathcal{C}' \\ \alpha\in\mathcal{O}_K}} \frac{\chi^*(\mathfrak{b}_0'/\mathfrak{c})}{(-\beta^*)^{\ell'}} \int \cdots \int_{\mathbf{e}\in\mathcal{R}_2\cap I(\mathfrak{a})} \frac{X^{\beta^*\sum_{i=1}^{\ell'} e_i} de_1 \dots de_{\ell'}}{\prod_{i=1}^{\ell'} e_i}$$

$$+ O(\delta_0^{n+1/2} B^n).$$

The condition $\alpha \in \mathcal{O}_K$ is equivalent to a congruence condition on $\mathbf{a} \pmod{(\theta n)^n}$ which holds for a proportion $r_K^{-1} = [(\theta n)^{-n}\mathbb{Z}[\sqrt[n]{\theta}] : \mathcal{O}_K]^{-1}$ of the vectors $\mathbf{a}$ in a cube of side length $(\theta n)^n$. Using the fact that $X^{\sum_{i=1}^{\ell'} e_i} = (1 + O(\delta_0 \log X))B_0^n/N(\mathfrak{c})$, we see that partial summation shows that the right-hand side above is

$$\frac{1}{\gamma_K r_K N(\mathfrak{c})\phi_K((q)/\mathfrak{c})}\left(1 + \frac{\chi^*(\mathfrak{b}_0'/\mathfrak{c})}{(-\beta^*)^{\ell'} N(\mathfrak{c})^{\beta^*} B_0^{n-n\beta^*}}\right)$$
$$\times \int_{\mathbf{a}\in\mathcal{C}'} c_{\mathcal{R}_2}(N(\mathfrak{a})/N(\mathfrak{c}))d\mathbf{a} + O(\delta_0^{n+1/2}B^n).$$

We now sum over the r_K values of β_0'. (We recall that these are the elements of $\mathcal{O}_K/q\mathcal{O}_K$ of the form $\beta_0' = (\theta n)^{-n}\sum_{i=1}^n (\mathbf{b}_0')_i \sqrt[n]{\theta^{i-1}}$ with $\mathbf{b}_0' \equiv \mathbf{b}_0 \pmod q$.) We see that the terms involving $\chi^*(\mathfrak{b}_0'/\mathfrak{c})$ cancel unless all $\mathbf{b} \equiv \mathbf{b}_0 \pmod q$ have $\chi^*(\mathfrak{b}/\mathfrak{c}) = \chi^*(\mathfrak{b}_0/\mathfrak{c})$ since χ^* is primitive. The rest of the expression is independent of the β_0'. Thus, if $\chi^*(\mathfrak{b}/\mathfrak{c}) = \chi^*(\mathfrak{b}_0/\mathfrak{c})$ for all $\mathbf{b} \equiv \mathbf{b}_0 \pmod q$, we have

$$\sum_{\substack{\mathbf{b}\in\mathcal{C} \\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q}} \mathbf{1}_{\mathcal{R}_2}\left(\frac{\mathfrak{b}}{\mathfrak{c}}\right)$$
$$= \frac{1}{\gamma_K \phi_K((q)/\mathfrak{c})N(\mathfrak{c})} \int_{\mathbf{a}\in\mathcal{C}'} c_{\mathcal{R}_2}(N(\mathfrak{a})/N(\mathfrak{c}))d\mathbf{a} + O(\delta_0^{n+1/2}B^n)$$
$$+ \frac{\chi^*(\mathfrak{b}_0/\mathfrak{c})}{\gamma_K(-\beta^*)^{\ell'}\phi_K((q)/\mathfrak{c})N(\mathfrak{c})^{\beta^*}B_0^{n-n\beta^*}} \int_{\mathbf{a}\in\mathcal{C}'} c_{\mathcal{R}_2}(N(\mathfrak{a})/N(\mathfrak{c}))d\mathbf{a},$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

and if χ^* is not constant over these $\mathfrak{b}$, then we have the same expression with the final term removed.

Finally, extending the integration over $\mathbf{a}$ from $\mathcal{C}'$ to $\mathcal{C}$ introduces an error of size $O(\Delta B^n)$ since the integrand is of size $O(1)$, and this increases the volume of the region of integration by $O(\Delta B^n)$. This then gives the result. □

LEMMA 9.5. *Let $\mathfrak{d}$ be a square-free ideal with $\gcd(\mathfrak{d}, (q))|\mathfrak{b}_0$ and let $(\theta n)^{2n}|q$. Then we have*

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q\\ \mathfrak{d}|\mathfrak{b}}} 1 = \frac{\operatorname{vol}\mathcal{C}}{N(\operatorname{lcm}((q),\mathfrak{d}))} + O(B^{n-1}N(\mathfrak{d})^{n-1}q^{n(n-1)}).$$

Proof. Let $Q_1 = N(\operatorname{lcm}(\mathfrak{d}, (q))) \leqslant q^n N(\mathfrak{d})$. Splitting into residue classes $(\mathrm{mod}\ Q_1)$, we have that

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q\\ \mathfrak{d}|\mathfrak{b}}} 1 = \sum_{\substack{\mathbf{a} \pmod{Q_1}\\ \mathbf{a}\equiv\mathbf{b}_0 \pmod q\\ \mathfrak{d}|\mathfrak{a}}} \sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{a} \pmod{Q_1}}} 1.$$

Here we remind the reader again that $\mathfrak{a}$ is the ideal generated by $\alpha = (\theta n)^{-n}\sum_{i=1}^n a_i\sqrt[n]{\theta^{i-1}}$. Since $(\theta n)^{2n}|q$, the condition $\mathbf{a}\equiv\mathbf{b}\ (\mathrm{mod}\ q)$ is equivalent to $\alpha\equiv\beta'\ (\mathrm{mod}\ q)$ over $\mathcal{O}_K$ for one of $[(\theta n)^{-n}\mathbb{Z}[\sqrt[n]{\theta}] : \mathcal{O}_K]$ different β', all of which satisfy $\beta'\equiv\beta\ (\mathrm{mod}\ q')$ over $\mathcal{O}_K$ where $q' = q/(\theta n)^n$. Since q' has the same square-free part as q and $\mathfrak{d}$ is square-free, we then see that the outer sum has no terms unless $\gcd(\mathfrak{d}, (q))|\mathfrak{b}_0$, in which case there are Q_1^{n-1} terms in the outer sum. The inner sum is $(\operatorname{vol}\mathcal{C})/Q_1^n + O(\delta_0^{n-1}B^{n-1})$. □

LEMMA 9.6. *If $\gcd(\mathfrak{b}_0, (q)) = \mathfrak{c}$, then*

$$\sum_{\substack{\mathfrak{d}<R\\ \gcd(\mathfrak{d}\mathfrak{c},(q))|\mathfrak{b}_0}} \frac{\lambda_{\mathfrak{d}}}{\operatorname{lcm}(N(\mathfrak{d}\mathfrak{c}), N((q)))} = \frac{1}{\gamma_K\phi_K((q)/\mathfrak{c})N(\mathfrak{c})} + O(\delta_0),$$

and if $\gcd(\mathfrak{b}_0, (q)) \neq \mathfrak{c}$, then the left-hand side is $O(\delta_0)$.

Proof. We estimate that this in an analogous way to Lemma 8.5. We let $(q) = \mathfrak{c}\mathfrak{q}_1\mathfrak{q}_2$, with $\gcd(\mathfrak{q}_2, \mathfrak{b}_0/\mathfrak{c}) = 1$ and $\mathfrak{q}_1$ composed only of primes which divide $\mathfrak{b}_0/\mathfrak{c}$. Since $\lambda_{\mathfrak{d}} = 0$, if $\mathfrak{d}$ is not square-free and $\mathfrak{c}|\mathfrak{b}_0$, we may replace the condition

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$\gcd(\mathfrak{d}\mathfrak{c}, (q))|\mathfrak{b}_0$ with $\gcd(\mathfrak{d}, \mathfrak{q}_2) = 1$. The argument used to prove Lemma 8.5 then gives

$$\sum_{\substack{N(\mathfrak{d})<R\\ \gcd(\mathfrak{q}_2,\mathfrak{d})=1}} \frac{\mu(\mathfrak{d})\log\frac{R}{N(\mathfrak{d})}}{N(\mathrm{lcm}(\mathfrak{d}\mathfrak{c}, \mathfrak{c}\mathfrak{q}_1\mathfrak{q}_2))} = \frac{1}{2\pi i N(\mathfrak{q}_2\mathfrak{c})}\int_{1-i\infty}^{1+i\infty}\frac{R^s g(1+s)}{s^2\zeta_K(1+s)}ds$$
$$= \frac{1}{N(\mathfrak{q}_2\mathfrak{c})}\operatorname*{Res}_{s=0}\frac{R^s g(1+s)}{s^2\zeta_K(1+s)} + O\left(\exp(-c\sqrt{\log R})\right),$$

where

$$g(1+s) = \prod_{\mathfrak{p}|\mathfrak{q}_1}\frac{N(\mathfrak{p})^{-1} - N(\mathfrak{p})^{-1-s}}{1-N(\mathfrak{p})^{-1-s}}\prod_{\mathfrak{p}|\mathfrak{q}_2}\frac{1}{1-N(\mathfrak{p})^{-1-s}}.$$

We see that the residue is 0 if $\mathfrak{q}_1 \neq (1)$, whereas if $\mathfrak{q}_1 = (1)$ (so $\gcd(\mathfrak{b}_0, (q)) = \mathfrak{c}$), the residue is $\gamma_K^{-1}N(\mathfrak{q}_2)/\phi_K(\mathfrak{q}_2)$. Thus, if $\gcd(\mathfrak{b}_0, (q)) = \mathfrak{c}$, then

$$\sum_{\substack{\mathfrak{d}<R\\ \gcd(\mathfrak{d}\mathfrak{c},(q))|\mathfrak{b}_0}}\frac{\lambda_{\mathfrak{d}}}{\mathrm{lcm}(N(\mathfrak{d}\mathfrak{c}), N((q)))} = \frac{1}{\gamma_K\phi_K((q)/\mathfrak{c})N(\mathfrak{c})} + O(\delta_0),$$

and if $\gcd(\mathfrak{b}_0, (q)) \neq \mathfrak{c}$, then the left-hand side is $O(\delta_0)$. □

PROPOSITION 9.7. *Let $\mathfrak{c}, \delta_0, B, \mathcal{C}, \mathbf{b}_0$ be as in Lemma 9.4. Then uniformly over all $q \ll q^{*\log\log B}\exp(\sqrt[6]{\log B})$ with $N(\mathfrak{c})(\theta n)^n|q$ and over all such $\mathcal{C}, \mathbf{b}_0$, we have*

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q}}\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) = \sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod q}}\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) + O(\delta_0^{n+1/2}B^n).$$

Here $\mathfrak{b}$ denotes the ideal generated by $(\theta n)^{-n}\sum_{i=1}^n b_i\sqrt[n]{\theta^{i-1}}$. The implied constant is effectively computable.

Proof. If the result holds for any residue class $\mathbf{b}_0 \pmod{N(\mathfrak{c})(\theta n)^n q}$ instead of any residue class $\pmod q$, then (after perhaps adjusting the implied constants) by summing over all $\mathbf{b}_0$ in a given residue class $\pmod q$, we see that the result also holds for any residue class $\pmod q$. Thus, we may assume that $N(\mathfrak{c})^2(\theta n)^{2n}|q$.

We will evaluate the sum on the right-hand side, which is a standard sieve quantity, and show that it gives the same result as Lemma 9.4 gives for the left-hand side.

Substituting the definition (8.7) of $\tilde{\mathbf{1}}_{\mathcal{R}_2}$ and swapping the order of summation,

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

we have

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0' \pmod q}} \tilde{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) = \sum_{N(\mathfrak{d})<R} \lambda_{\mathfrak{d}} \sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0' \pmod q\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} c_{\mathcal{R}_2}(N(\mathfrak{b}/\mathfrak{c}))\left(1+\frac{\chi^*(\mathfrak{b}/\mathfrak{c})}{(-\beta^*)^{\ell'}N(\mathfrak{b}/\mathfrak{c})^{1-\beta^*}}\right).$$

We split $\mathcal{C}$ into $O(\delta_0^{-n})$ disjoint smaller hypercubes $\mathcal{C}'$ of side length $\delta_0^2 B$. Since $c_{\mathcal{R}_2}(N(\mathfrak{b}/\mathfrak{c}))$ satisfies the Lipschitz bound of Lemma 8.3, we can replace $c_{\mathcal{R}_2}(N(\mathfrak{b}/\mathfrak{c}))$ with

$$c_{\mathcal{R}_2}(\mathcal{C}') := \frac{1}{\operatorname{vol}\mathcal{C}'} \int\cdots\int_{\substack{\mathbf{a}\in\mathcal{C}',\mathbf{e}\in\mathcal{R}_2\\ \sum_{i=1}^{\ell'} e_i\in\mathcal{I}_{N(\mathfrak{a}/\mathfrak{c})}}} \frac{de_1\dots de_{\ell'}d\mathbf{a}}{\log X\prod_{i=1}^{\ell'} e_i}$$

on the hypercube $\mathcal{C}'$, at the cost of an error of total size

$$\ll \sum_{N(\mathfrak{d})<R} \log X \sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0' \pmod q\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} \frac{\delta_0}{\eta_2^{1/2}} \ll \sum_{N(\mathfrak{d})<R} \frac{\delta_0 \operatorname{vol}\mathcal{C}\log X}{\eta_2^{1/2}N(\mathfrak{d})} \ll \delta_0^{1/2}\operatorname{vol}\mathcal{C}.$$

Similarly, we can replace $N(\mathfrak{b}/\mathfrak{c})$ with $N(\mathfrak{b}_0/\mathfrak{c})$ at the cost of an error $O(\delta_0^{1/2}\operatorname{vol}\mathcal{C})$.

Thus, we are left to evaluate

$$\sum_{\mathcal{C}'} c_{\mathcal{R}_2}(\mathcal{C}') \sum_{N(\mathfrak{d})<R} \lambda_{\mathfrak{d}} \sum_{\substack{\mathbf{b}\in\mathcal{C}'\\ \mathbf{b}\equiv\mathbf{b}_0' \pmod q\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} \left(1+\frac{\chi^*(\mathfrak{b}/\mathfrak{c})}{(-\beta^*)^{\ell'}N(\mathfrak{b}_0/\mathfrak{c})^{1-\beta^*}}\right). \tag{9.6}$$

Recall that $\lambda_{\mathfrak{d}}$ is supported on square-free $\mathfrak{d}$. For such $\mathfrak{d}$, by Lemma 9.5, we see that provided $\gcd(\mathfrak{d}\mathfrak{c}, (q))\mid\mathfrak{b}_0$, we have

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}'\\ \mathbf{b}\equiv\mathbf{b}_0' \pmod q\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} 1 = \frac{\operatorname{vol}\mathcal{C}'}{N(\operatorname{lcm}((q),\mathfrak{d}\mathfrak{c}))} + O(B^{n-1}R^{n-1}), \tag{9.7}$$

and otherwise the sum is 0. The $O(B^{n-1}R^{n-1})$ error term makes a total contribution $O(B^{n-1}R^n\log X)$, which is negligible.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We now consider the terms involving χ^*. We have that $\chi^*(\mathfrak{b}/\mathfrak{c}) = 0$ if $\gcd(\mathfrak{q}^*, \mathfrak{b}/\mathfrak{c}) \neq 1$, and so there are no contributions from terms with $\gcd(\mathfrak{d}, \mathfrak{q}^*) \neq 1$. By splitting the sum into residue classes $(\mathrm{mod}\ Q_2)$ where $Q_2 = N(\mathrm{lcm}(\mathfrak{d}\mathfrak{c}\mathfrak{q}^*, (q))$, we see that

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}'\\ \mathbf{b}\equiv\mathbf{b}_0'\ (\mathrm{mod}\ q)\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} \chi^*(\mathfrak{b}/\mathfrak{c}) = \sum_{\substack{\mathbf{a}\ (\mathrm{mod}\ Q_2)\\ \mathbf{a}\equiv\mathbf{b}_0'\ (\mathrm{mod}\ q)\\ \mathfrak{d}\mid\mathfrak{a}/\mathfrak{c}}} \chi^*(\mathfrak{a}/\mathfrak{c}) \sum_{\substack{\mathbf{b}\in\mathcal{C}'\\ \mathbf{b}\equiv\mathbf{a}\ (\mathrm{mod}\ Q_2)}} 1.$$

By Lemma 9.5, the inner sum is $\mathrm{vol}\,\mathcal{C}'/Q_2^n + O(\delta_0^{2n-2}B^{n-1})$, and this error term makes a negligible total contribution. The remaining sum of $\chi^*(\mathfrak{a}/\mathfrak{c})$ is then seen to cancel completely unless $\chi^*(\mathfrak{b}) = \chi^*(\mathfrak{b}_0')$ for all $\mathbf{b} \equiv \mathbf{b}_0'\ (\mathrm{mod}\ q)$. If this is the case, then by Lemma 9.5, we have

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}'\\ \mathbf{b}\equiv\mathbf{b}_0'\ (\mathrm{mod}\ q)\\ \mathfrak{d}\mid\mathfrak{b}/\mathfrak{c}}} \chi^*(\mathfrak{b}/\mathfrak{c}) = \frac{\chi^*(\mathfrak{b}_0'/\mathfrak{c})\,\mathrm{vol}\,\mathcal{C}'}{N(\mathrm{lcm}((q), \mathfrak{d}\mathfrak{c}))} + O(B^{n-1}R^{n-1}), \tag{9.8}$$

and otherwise the sum is simply $O(B^{n-1}R^{n-1})$. Again, these $O(B^{n-1}R^{n-1})$ error terms make a total contribution $O(B^{n-1}R^n \log X)$, which is negligible.

Thus, to estimate (9.6), we see from (9.7) and (9.8) that it suffices to estimate

$$\sum_{\mathcal{C}'} c_{\mathcal{R}_2}(\mathcal{C}')\,\mathrm{vol}\,\mathcal{C}' \sum_{\substack{\mathfrak{d}<R\\ \gcd(\mathfrak{d}\mathfrak{c},(q))\mid\mathfrak{b}_0}} \frac{\lambda_{\mathfrak{d}}}{\mathrm{lcm}(N(\mathfrak{d}\mathfrak{c}), N((q)))}.$$

By Lemma 9.6, we have that the inner sum is

$$\frac{1}{\gamma_K\phi_K((q)/\mathfrak{c})N(\mathfrak{c})} + O(\delta_0),$$

provided $\gcd(\mathfrak{b}_0, (q)) = \mathfrak{c}$. Finally, we note that

$$\sum_{\mathcal{C}'} c_{\mathcal{R}_2}(\mathcal{C}')\,\mathrm{vol}\,\mathcal{C}' = c_{\mathcal{R}_2}(\mathcal{C})\,\mathrm{vol}\,\mathcal{C}.$$

Putting all these estimates together, we obtain an expression for

$$\sum_{\substack{\mathbf{b}\in\mathcal{C}\\ \mathbf{b}\equiv\mathbf{b}_0\ (\mathrm{mod}\ q)}} \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}),$$

which is identical to the estimates of Lemma 9.4. This gives the result. □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

10. Some lattice estimates

In this section, we collect some information about the structure of ideals $\mathfrak{b} \in \mathcal{A}'_{\mathfrak{a}}$ before we finish our Type II estimate in the next section. Here we exploit some of the simple structure from the fact $K = \mathbb{Q}(\sqrt[n]{\theta})$.

If $\mathfrak{a} = (\alpha)$ is principal, then $\mathfrak{b} \in \mathcal{A}'_{\mathfrak{a}}$ if $\mathfrak{b} = (\beta)$ with $(\beta\alpha) = (\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}})$ for some $\mathbf{x} \in \mathbb{Z}^n$ with $x_i \in [X_i, X_i + \eta_1 X_i]$ for $1 \leqslant i \leqslant n-k$ and $x_i = 0$ for $n-k < i \leqslant n$ and $\mathbf{x} \equiv \mathbf{x}_0 \pmod{q^*}$ and $N(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}) \in [X_0^n, X_0^n + \eta_2 X_0^n]$.

Since $\mathbb{Z}[\sqrt[n]{\theta}]$ is an order in $\mathcal{O}_K$ of finite index dividing $(\theta n)^n$, any principal ideal $\mathfrak{b}$ has a unique representation as (β) with $\beta = (\theta n)^{-n} \sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}}$ and $\mathbf{b} \in \mathbb{Z}^n \cap \mathcal{F}$ for a fundamental domain $\mathcal{F}$ by the action of the group of units $\mathcal{U}_K$, with $\mathbf{b}$ satisfying some integral linear congruence conditions $\mathbf{L}(\mathbf{b}) \equiv \mathbf{0} \pmod{(\theta n)^n}$. We have

$$\left(\sum_{i=1}^{n} b_i \sqrt[n]{\theta^{i-1}}\right)\left(\sum_{i=1}^{n} a_i \sqrt[n]{\theta^{i-1}}\right) = \left(\sum_{i=1}^{n} c_i \sqrt[n]{\theta^{i-1}}\right)$$

with

$$c_j = \left(\sum_{i=1}^{j} b_{j-i+1} a_i + \theta \sum_{i=j+1}^{n} b_{n+j-i+1} a_i\right) = T^{n-j}(\overleftarrow{\mathbf{b}}) \cdot \mathbf{a},$$

where $\cdot$ is the usual Euclidean dot product on $\mathbb{R}^n$, $\overleftarrow{\mathbf{v}}$ indicates the reverse of the coordinates of $\mathbf{v}$ (that is, $\overleftarrow{v}_j = v_{n+1-j}$) and T^i indicates the ith iterate of the linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$T(\mathbf{v})_j = \begin{cases} v_{j+1}, & j < n, \\ \theta v_1, & j = n. \end{cases}$$

We let $\diamond$ denote the above operation so that $\mathbf{c} = \mathbf{b} \diamond \mathbf{a}$. We note that

$$N(\mathbf{v}) = \det(T^0(\mathbf{v}) \mid T(\mathbf{v}) \mid \ldots \mid T^{n-1}(\mathbf{v})).$$

In particular, if $\mathbf{v} \neq 0$, then $T^j(\mathbf{v})$ are linearly independent for $0 \leqslant j < n$.

Thus, there is a bijection between pairs of principal ideals $\mathfrak{a}$, $\mathfrak{b}$ with $\mathfrak{a}\mathfrak{b}/N(\mathfrak{c}) \in \mathcal{A}'$, and vectors $\mathbf{a} \in \mathbb{Z}^n \cap \mathcal{F}$, $\mathbf{b} \in \mathbb{Z}^n$ (for any choice of fundamental domain $\mathcal{F}$ for the action of the unit group $\mathcal{O}_K^*$) with $\mathbf{L}(\mathbf{a}) \equiv \mathbf{L}(\mathbf{b}) \equiv \mathbf{0} \pmod{(\theta n)^n}$ and with $\mathbf{a} \diamond \mathbf{b} \in \mathcal{R}_X$, where $\mathcal{R}_X$ is given by

$$\begin{aligned} \mathcal{R}_X = \big\{ &\mathbf{x} \in \mathbb{R}^n : x_i \in [X'_i, X'_i + \eta_1 X'_i] \text{ for } i \leqslant n-k,\ x_i = 0 \text{ for } i > n-k, \\ &N(\textstyle\sum_{i=1}^{n} x_i \sqrt[n]{\theta^{i-1}}) \in [X_0'^n, X_0'^n + \eta_2 X_0'^n] \big\}. \end{aligned} \tag{10.1}$$

Here $X'_i = (\theta n)^{2n} N(\mathfrak{c}) X_i$, which still satisfy $X'_i \asymp_{\mathfrak{c}} X$. We see that, given $\mathbf{a} \in \mathbb{Z}^n$, the conditions $(\mathbf{b} \diamond \mathbf{a})_j = 0$ force $\mathbf{b}$ to satisfy k integral linear equations and,

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

hence, lie in a sublattice of $\mathbb{Z}^n$. With this in mind, we define the lattices

$$\begin{aligned}\Lambda_{\mathbf{v}} &= \{\mathbf{x}\in\mathbb{Z}^n : (\mathbf{x}\diamond\mathbf{v})_i = 0,\ n-k<i\leqslant n\}\\ &= \{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}\cdot T^i(\overleftarrow{\mathbf{v}}) = 0,\ 0\leqslant i\leqslant k-1\},\\ \Lambda_{\mathbf{v}_1,\mathbf{v}_2} &= \{\mathbf{x}\in\mathbb{Z}^n : (\mathbf{x}\diamond\mathbf{v}_1)_i = (\mathbf{x}\diamond\mathbf{v}_2)_i = 0,\ n-k<i\leqslant n\}\\ &= \{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}\cdot T^i(\overleftarrow{\mathbf{v}_1}) = \mathbf{x}\cdot T^i(\overleftarrow{\mathbf{v}_2}) = 0,\ 0\leqslant i\leqslant k-1\}. \qquad (10.2)\end{aligned}$$

We first establish some basic properties of these lattices.

LEMMA 10.1. *Let* $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2\in\mathbb{Z}^n\backslash\{\mathbf{0}\}$. *Let* $\wedge(\mathbf{v})\in\mathbb{Z}^{\binom{n}{k}}$ *be the vector of determinants of* $k\times k$ *submatrices of the* $k\times n$ *matrix formed by the* k *vectors* $T^0(\mathbf{v}),\dots,T^{k-1}(\mathbf{v})$. *Similarly, let* $\wedge(\mathbf{v}_1,\mathbf{v}_2)\in\mathbb{Z}^{\binom{n}{2k}}$ *be the vector of determinants of the* $2k\times 2k$ *submatrices of the* $2k\times n$ *matrix formed of the* $2k$ *vectors* $T^0(\mathbf{v}_1),\dots,T^{k-1}(\mathbf{v}_1)$ *and* $T^0(\mathbf{v}_2),\dots,T^{k-1}(\mathbf{v}_2)$. *Finally, let* $D_{\mathbf{v}}$ *be the largest integer* D *such that* $\wedge(\mathbf{v})\equiv\mathbf{0}\pmod D$ *and* $D_{\mathbf{v}_1,\mathbf{v}_2}$ *be the largest integer* D' *such that* $\wedge(\mathbf{v}_1,\mathbf{v}_2)\equiv\mathbf{0}\pmod{D'}$. *Then we have*

$$\det(\Lambda_{\mathbf{v}}) = \frac{\|\wedge(\mathbf{v})\|}{D_{\mathbf{v}}},$$
$$\det(\Lambda_{\mathbf{v}_1,\mathbf{v}_2}) = \frac{\|\wedge(\mathbf{v}_1,\mathbf{v}_2)\|}{D_{\mathbf{v}_1,\mathbf{v}_2}} \quad \text{if } \wedge(\mathbf{b}_1,\mathbf{b}_2)\neq\mathbf{0}.$$

Proof. Let $\mathbf{v}_1,\dots,\mathbf{v}_r$ be linearly independent vectors in $\mathbb{Z}^n$, and let $\Lambda=\{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}\cdot\mathbf{v}_1=\cdots=\mathbf{x}\cdot\mathbf{v}_r=0\}$. By [12, Lemma 1], $\det\Lambda=\det\Lambda^*$ where $\Lambda^*=\{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}=\sum_{i=1}^r c_i\mathbf{v}_i,\ c_i\in\mathbb{Q}\}$.

Let $D(\mathbf{v}_1,\dots,\mathbf{v}_r)$ be the largest integer such that the determinant of all $r\times r$ submatrices of the $n\times r$ matrix formed with linearly independent columns $\mathbf{v}_1,\dots,\mathbf{v}_r\in\mathbb{Z}^n$ vanish mod $D(\mathbf{v}_1,\dots,\mathbf{v}_r)$. (That is, the largest integer D such that $\mathbf{v}_1,\dots,\mathbf{v}_r$ are linearly dependent $\pmod D$.) We define a reduction procedure as follows. Given $\{\mathbf{x}_1,\dots,\mathbf{x}_r\}\in(\mathbb{Z}^n)^r$ with $D(\mathbf{x}_1,\dots,\mathbf{x}_r)\neq 1$, we choose (arbitrarily) a prime $p|D(\mathbf{x}_1,\dots,\mathbf{x}_r)$. By definition of $D(\cdot)$, this means that there are constants $c_1,\dots,c_r$ at least one of which is 1, such that $\sum_{i=1}^r c_i\mathbf{x}_i\equiv\mathbf{0}\pmod p$. We choose (arbitrarily) an index j such that $c_j=1$ and replace $\mathbf{x}_j$ with $\sum_{i=1}^r c_i\mathbf{x}_i/p\in\mathbb{Z}^n$ to produce a new set of vectors $(\mathbf{x}_1',\dots,\mathbf{x}_r')$, and we see that we must have $D(\mathbf{x}_1',\dots,\mathbf{x}_r')=D(\mathbf{x}_1,\dots,\mathbf{x}_r)/p$. By starting with $\{\mathbf{v}_1,\dots,\mathbf{v}_r\}$ and repeatedly performing this reduction, we arrive at a $\mathbb{Z}$-basis $\mathbf{z}_1,\dots,\mathbf{z}_r\in\mathbb{Z}^n$ for Λ^*. (This process clearly terminates as $D(\mathbf{x}_1,\dots,\mathbf{x}_r)$ decreases at each stage, and the resulting set is a basis since $D(\mathbf{z}_1,\dots,\mathbf{z}_r)=1$, so integral vectors in the $\mathbb{Q}$-span of $\mathbf{z}_1,\dots,\mathbf{z}_r$ lie in the $\mathbb{Z}$-span of $\mathbf{z}_1,\dots,\mathbf{z}_r$, and the $\mathbb{Q}$-span is clearly the whole lattice.) Moreover, we see that the $\mathbb{Z}$-span of $\mathbf{v}_1,\dots,\mathbf{v}_r$ is a lattice $\tilde{\Lambda}$ which is an index $D(\mathbf{v}_1,\dots,\mathbf{v}_r)$ sublattice of Λ^*.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Thus, $\det \Lambda = \det \Lambda^* = \det \tilde{\Lambda}/D(\mathbf{v}_1, \ldots, \mathbf{v}_r)$. But $\det \tilde{\Lambda}$ is simply the volume of the r-dimensional fundamental volume of $\tilde{\Lambda}$. If $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ are orthonormal vectors orthogonal to $\mathbf{v}_1, \ldots, \mathbf{v}_r$, then $\det \tilde{\Lambda}$ is given by the determinant of the $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n$. This is then seen to be the Euclidean norm of the exterior product of $\mathbf{v}_1, \ldots, \mathbf{v}_r$ (that is, the vector of all determinants of the $r \times r$ submatrices of the $r \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_r$) since both quantities are independent of a choice of orthonormal basis of $\mathbb{R}^n$ and agree on the orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ extending $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n$.

Applying the above argument to $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{T^0(\mathbf{v}), \ldots, T^{k-1}(\mathbf{v})\}$ gives the result for $\Lambda_{\mathbf{v}}$, whilst using $\{T^0(\mathbf{v}_1), \ldots, T^{k-1}(\mathbf{v}_1), T^0(\mathbf{v}_2), \ldots, T^{k-1}(\mathbf{v}_2)\}$ gives the result for $\Lambda_{\mathbf{v}_1,\mathbf{v}_2}$. □

LEMMA 10.2 (Vandermonde determinant). *Let $m_1, \ldots m_r$ be nonnegative integers and $n = r + \sum_{i=1}^r m_i$. Let $\lambda_1, \ldots, \lambda_r \in \mathbb{C}\backslash\{0\}$, and let $M = M(\lambda_1, \ldots \lambda_r, m_1, \ldots, m_r)$ be the $n \times n$ matrix*

$$\begin{pmatrix} \lambda_1 & \lambda_1 & \ldots & \lambda_1 & \lambda_2 & \ldots & \lambda_2 & \ldots & \lambda_r \\ \lambda_1^2 & 2\lambda_1^2 & \ldots & 2^{m_1}\lambda_1^2 & \lambda_2^2 & \ldots & 2^{m_2}\lambda_2^2 & \ldots & 2^{m_r}\lambda_r^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \lambda_1^n & n\lambda_1^n & \ldots & n^{m_1}\lambda_1^n & \lambda_2^n & \ldots & n^{m_2}\lambda_2^n & \ldots & n^{m_r}\lambda_r^n \end{pmatrix}$$

formed with entries in the jth row given by $j^m\lambda_i^j$ for $0 \leqslant m \leqslant m_i$ and $1 \leqslant i \leqslant r$.

Then we have

$$\det(M) = \left(\prod_{i=1}^r \prod_{m=1}^{m_i-1} m!\right)\left(\prod_{i=1}^r \lambda_i^{m_i(m_i+1)/2}\right)\left(\prod_{1\leqslant i<j\leqslant r} (\lambda_j - \lambda_i)^{m_i m_j}\right).$$

In particular, $\det(M) = 0$ *if and only if $\lambda_i = \lambda_j$ for some $i \neq j$.*

Proof. Let M be the matrix of the lemma. By subtracting a suitable linear combination of the first $j-1$ columns from the jth column, we see that $\det(M)$ is equal to $\det(M')$, where M' is the matrix with jth row given by $(j-1)\ldots(j-m)\lambda_i^j$ for $0 \leqslant m \leqslant m_i$ and $1 \leqslant i \leqslant r$ instead of $j^m\lambda_i^j$ (we interpret the expression as λ_i^j if $m = 0$). We see that the jth column of M' is a multiple of λ_1^j for all $1 \leqslant j \leqslant m_1$. Therefore, the determinant is a multiple of $\lambda_1^{m_1(m_1+1)/2}$, and similarly for the other λ_i by symmetry. We now wish to show that $(\lambda_1-\lambda_2)^{m_1m_2}$ divides the determinant. For $\ell = 0, \ldots, m_1m_2-1$, we consider

$$\left.\frac{\partial^\ell}{\partial\lambda_1^\ell}\right|_{\lambda_1=\lambda_2} \det(M) = \sum_{\substack{j_1,\ldots,j_{m_1+1}\geqslant 0 \\ j_1+\cdots+j_{m_1+1}=\ell}} \binom{n}{j_1, \ldots, j_{m_1+1}} \det(M^{(j_1,\ldots,j_{m_1+1})}),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Here $M^{(j_1,\dots,j_{m_1+1})}$ is the matrix formed by replacing the ith column v_i of M' with

$$\frac{\partial^{j_i}}{\partial \lambda_1^{j_i}}\bigg|_{\lambda_1=\lambda_2} v_i$$

for each $i \in \{1,\dots,m_1+1\}$. We see that this expression has jth entry $(j-1)\dots(j-i)\times(j-j_i+1)\dots j\lambda_2^{j-j_i}$. In particular, for $i \leqslant m_1+1$, we see that the ith column of $M^{(j_1,\dots,j_{m_1+1})}$ is a vector with jth entry $P(j)\lambda_2^j$ for some polynomial P of degree $i+j_i$. However, the columns $v_{m_1+2},\dots,v_{m_1+m_2+2}$ also have jth entry of the form $P(j)\lambda_2^j$ for some polynomial P of degree at most m_2. Thus, we have m_1+m_2+2 columns, and for each column, there is a polynomial P such that the jth entry of the column is $P(j)\lambda_2^j$ for all $1 \leqslant j \leqslant n$. But any $k+2$ vectors whose jth entry is of the form $P(j)\lambda^j$ for a polynomial P of degree at most k must be linearly dependent (this is seen by cancelling the highest coefficients in turn). Thus, we see that these columns are linearly independent only if for every $k \in \mathbb{N}$, there are at most $k+1$ columns involving a polynomial of degree at most k. But this requires that the sum of degrees of the m_1+m_2+2 polynomials be at least $(m_1+m_2+2)(m_1+m_2+1)/2$, which requires

$$\sum_{i=1}^{m_1+1}(j_i+i)+\sum_{i=1}^{m_2+1} i \geqslant \frac{(m_1+m_2+2)(m_1+m_2+1)}{2}.$$

This simplifies to

$$\ell = \sum_{i=1}^{m_1+1} j_i \geqslant m_1 m_2.$$

Thus, for all $\ell \in \{0,\dots,m_1m_2-1\}$, we see that $\det(M^{(j_1,\dots,j_{m_1+1})})=0$, and so we must have that $(\lambda_1-\lambda_2)^{m_1m_2}$ divides $\det(M')$. By symmetry, we therefore find that $\det(M')$ is a multiple of

$$\left(\prod_{i=1}^{r}\lambda_i^{m_i(m_i+1)/2}\right)\left(\prod_{1\leqslant i<j\leqslant r}(\lambda_j-\lambda_i)^{m_im_j}\right).$$

By expanding the determinant via rows, we see that the determinant is a homogeneous polynomial of degree $n(n+1)/2$ in the λ_i, and so $\det(M)$ must be proportional to the above expression. Finally, by considering the coefficient of $\lambda_1^{e_1}\lambda_2^{e_2}\dots\lambda_r^{e_r}$ with first e_1 minimal, then e_2 minimal, we see that the coefficient is

$$\prod_{i=1}^{r}\prod_{j_i=1}^{m_i+1}(j_i-1)!$$

This gives the result. □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

LEMMA 10.3 (Difference equations). *Let $c_1, \dots, c_r \in \mathbb{Q}$ with $c_1 \neq 0$ and $c_r \neq 0$. Let $x_1, \dots, x_J$ satisfy*

$$x_j = \sum_{i=1}^{r} c_i x_{j-i}$$

for $j > r$. Then there are constants $\lambda_1, \dots, \lambda_\ell \in \mathbb{C}$ and polynomials $P_1, \dots, P_\ell$ such that

$$x_j = \sum_{i=1}^{\ell} P_i(j)\lambda_i^j.$$

Moreover, $\sum_{i=1}^{\ell}(1 + \deg(P_i)) \leqslant r$, the λ_i lie in a finite extension of $\mathbb{Q}$ and the λ_i only depend on $c_1, \dots, c_r$.

Proof. Let M be the $r \times r$ matrix

$$M = \begin{pmatrix} c_1 & c_2 & c_3 & \dots & c_r \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix},$$

so if $\mathbf{x}_j = (x_j, x_{j-1}, \dots, x_{j-r+1})$, then $\mathbf{x}_{j+1} = M\mathbf{x}_j$ for $j \geqslant r$. In particular, $\mathbf{x}_j = M^{j-r}\mathbf{x}_r$ for all $j \geqslant r$. Since $c_r \neq 0$, we see that M is nonsingular. But M can be put into Jordan normal form after a change of basis, which means that $M = A^{-1}DA$ for some upper triangular matrix D formed of Jordan blocks. But then $M^j = A^{-1}D^jA$, and the entries of D^j are all of the form $P_i(j)\lambda_i^j$, where the λ_i are the eigenvalues of M and P_i is a polynomial of degree at most one less than the multiplicity of λ_i. This gives the result for the shape of the x_j. Since the λ_i are the eigenvalues of M and $\deg(P_i) + 1$ is at most the multiplicity of λ_i, we get the other claims of the lemma. □

LEMMA 10.4. *Let $n > 3k$. Let $\mathbf{b} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and let $\mathcal{L}$ be a linear subspace of $\mathbb{R}^n$ such that $\wedge(\mathbf{x}, \mathbf{b}) = \mathbf{0}$ for all $\mathbf{x} \in \mathcal{L}$. Then $\mathcal{L}$ has dimension at most k.*

Proof. If $\wedge(\mathbf{x}, \mathbf{b}) = \mathbf{0}$, then there exist constants $c_0, \dots, c_{k-1}, d_0, \dots, d_{k-1} \in \mathbb{Z}$ not all zero such that

$$\sum_{i=0}^{k-1} c_i T^i(\mathbf{x}) = \sum_{i=0}^{k-1} d_i T^i(\mathbf{b}).$$

Since $\mathbf{b} \neq 0$, we have that $\{T^i(\mathbf{b})\}_{i=0}^{n-1}$ are linearly independent vectors in $\mathbb{R}^n$. Thus, we cannot have $c_0 = \cdots = c_{k-1} = 0$ and we can write $\mathbf{x} = \sum_{i=0}^{n-1} x_i T^i(\mathbf{b})$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

With respect to this basis, the above equation implies that $\sum_{i=0}^{k-1} c_i x_{j-i} = 0$ for each $k \leqslant j < n$. Since the $c_0, \dots, c_{k-1}$ are not all zero, we let c_ℓ be the first nonzero element, so we have $x_j = \sum_{i=1}^{k-1-\ell} c'_{\ell+i} x_{j-i}$ for each $k \leqslant j < n$ with $c'_i = c_i/c_\ell$. For notational simplicity, we now restrict our argument to the case when $c_0, c_{k-1} \neq 0$; the other cases follow by an entirely analogous argument.

The equation $x_j = \sum_{i=1}^{k-1} c'_i x_{j-i}$ for $k \leqslant j < n$ is a difference equation and so, by Lemma 10.3, has solution $x_j = \sum_i P_i(j)\lambda_i^j$ for $1 \leqslant j < n$ for some polynomials $P_1, \dots, P_\ell$ with $\sum_i (1 + \deg(P_i)) \leqslant k-1$ and some constants λ_i in a finite extension of $\mathbb{Q}$, all of which may depend only on the constants c_i.

We will show that in any linear space $\mathcal{L} \subseteq \mathbb{R}^n$ containing only points with $\wedge(\mathbf{x}, \mathbf{b}) = \mathbf{0}$, at most $k-1$ different monomials $j^d\lambda_i^j$ can appear in such an expression over all possible choices of the c_i.

Assume the contrary for a contradiction. By taking linear combinations of these monomials, we see that there exists $\mathbf{x}, \mathbf{y} \in \mathcal{L}$ with $(\mathbf{x})_j = \sum_i P_i(j)\lambda_i^j$ and $(\mathbf{y})_j = \sum_m Q_m(j)\mu_m^j$ for $1 < j \leqslant n$, for some polynomials $P_i, Q_m \in \mathbb{C}[X]$ and some algebraic integers $\lambda_i, \mu_m \in \mathbb{C}$ such that $\sum_i (1 + \deg(P_i))$, $\sum_i (1 + \deg(Q_i)) \leqslant k-1$; but, in total, at least k different monomials $j^{m_1}\lambda_{m_2}^j$, $j^{m_3}\mu_{m_4}^j$ appear with nonzero coefficients across these two expressions. In particular, there is a real linear combination $a_1\mathbf{x} + a_2\mathbf{y}$ such that at least k different monomials appear. But $a_1\mathbf{x} + a_2\mathbf{y} \in \mathcal{L}$, so $(a_1\mathbf{x} + a_2\mathbf{y})_j$ can also be written as $\sum_i R_i(j)\gamma_i^j$ with at most $k-1$ different monomials $j^{m_1}\gamma_{m_2}^j$ appearing and $\sum_i (1 + \deg(R_i)) \leqslant k-1$. But then we have $\sum_i R_i(j)\gamma_i^j = a_1 \sum_i P_i(j)\lambda_i^j + a_2 \sum_i Q_i(j)\mu_i^j$ for all $1 < j \leqslant n$; so the monomials $j^{m_1}\gamma_{m_2}^j$, $j^{m_3}\mu_{m_4}^j$, $j^{m_5}\lambda_{m_6}^j$ satisfy a nonzero linear equation $\sum_i e_i M_i(j) = 0$ for all $1 < j \leqslant n$, for some constants e_i not all zero and distinct monomials $M_i(j)$ of the form $j^{m_1}\gamma_{m_2}^j$, $j^{m_3}\mu_{m_4}^j$ or $j^{m_5}\lambda_{m_6}^j$ (for some integers $m_1, \dots, m_6$). Moreover, since $\sum_i (1 + \deg(P_i))$, $\sum_i (1 + \deg(Q_i))$, $\sum_i (1 + \deg(R_i)) \leqslant k-1$, there are at most $3k-3$ monomials appearing in this expression. In matrix form, this set of equations is

$$\begin{pmatrix} M_1(1) & \dots & M_{3k-3}(1) \\ \vdots & & \vdots \\ M_1(n) & \dots & M_{3k-3}(n) \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_{3k-3} \end{pmatrix} = \mathbf{0}.$$

Since $n > 3k$, this includes the first $3k-3$ rows which form a $(3k-3) \times (3k-3)$ generalized Vandermonde matrix. By Lemma 10.2, the determinant of this matrix is nonzero. Thus, the vector $(e_1, \dots, e_{3k-3})$ must be zero, a contradiction to our assumption that it is nonzero. Thus, only $k-1$ different monomials can appear, and so $\mathcal{L}$ has dimension at most k (since x_0 is a free variable). □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

REMARK. The bound in Lemma 10.4 is tight since the subspace generated by the vectors $T^0(\mathbf{b}), \dots, T^{k-1}(\mathbf{b})$ has dimension k.

LEMMA 10.5. *Let $n > 3k$. Let $\mathbf{a} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ and $\Lambda_{\mathbf{a}}$ have successive minima $Z_1 \leqslant \cdots \leqslant Z_{n-k}$. Then $\Lambda_{\mathbf{a}}$ has a $\mathbb{Z}$-basis $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ such that*

- *for each $i \in \{1, \dots, n-k\}$ we have $Z_i \ll \|\mathbf{z}_i\| \ll Z_i$;*
- *$\wedge(\mathbf{z}_1, \mathbf{z}_{k+1}) \neq \mathbf{0}$;*
- *for any $\lambda_1, \dots, \lambda_{n-k} \in \mathbb{R}^{n-k}$, $\| \sum_{i=1}^{n-k} \lambda_i \mathbf{z}_i \| \gg \sum_{i=1}^{n-k} \|\lambda_i \mathbf{z}_i\|$.*

Proof. Since $T^0(\mathbf{a}), \dots, T^{k-1}(\mathbf{a})$ are linearly independent, we see that $\Lambda_{\mathbf{a}}$ has rank $n-k$. By Lemma 4.1, $\Lambda_{\mathbf{a}}$ has a Minkowski-reduced basis $\{\mathbf{z}_1, \dots, \mathbf{z}_{n-k}\}$. The space generated by $\mathbf{z}_1, \dots, \mathbf{z}_{k+1}$ is a linear space of dimension $k+1$, so by Lemma 10.4, we have that $\wedge(\mathbf{x}, \mathbf{z}_1)$ does not vanish for all $\mathbf{x}$ in this space. But since $\wedge(\cdot, \mathbf{z}_1) = \mathbf{0}$ is given by the vanishing of a system of homogeneous polynomials of degree $O(1)$, this means that there is a nonzero homogeneous polynomial $f \in \mathbb{Z}[X_1, \dots, X_{k+1}]$ of degree $O(1)$ such that $\wedge(\sum_{i=1}^{k+1} \lambda_i \mathbf{z}_i, \mathbf{z}_1) = \mathbf{0}$ only if $f(\lambda_1, \dots, \lambda_{k+1}) = 0$. But there is then a choice of $\lambda_1, \dots, \lambda_{k+1} \in \mathbb{Z}$ with $\lambda_{k+1} = 1$ and $\lambda_i \ll 1$ for all $1 \leqslant i \leqslant k$ such that $f(\lambda_1, \dots, \lambda_{k+1}) \neq 0$. Let $\mathbf{z}'_{k+1} = \sum_{i=1}^{k+1} \lambda_i \mathbf{z}_i$. We claim that $\{\mathbf{z}_1, \dots, \mathbf{z}_k, \mathbf{z}'_{k+1}, \mathbf{z}_{k+2} \dots, \mathbf{z}_{n-k}\}$ gives a basis with the required properties. Since $\mathbf{z}'_{k+1}$ is a linear combination of $\mathbf{z}_1, \dots, \mathbf{z}_{k+1}$ with $\mathbf{z}_{k+1}$-coefficient equal to 1, we see that this is indeed a basis since $\{\mathbf{z}_1, \dots, \mathbf{z}_{n-k}\}$ is. Since $f(\lambda_1, \dots, \lambda_{k+1}) \neq 0$, we have that $\wedge(\mathbf{z}'_{k+1}, \mathbf{z}_1) \neq 0$. Since $\lambda_i \ll 1$, we see that $\|\mathbf{z}'_{k+1}\| \asymp \sum_{i=1}^{k+1} \|\lambda_i \mathbf{z}_i\| \asymp Z_{k+1}$. Finally, since $\lambda_i \ll 1$, we have

$$\begin{aligned}\left\| a_{k+1}\mathbf{z}_{k+1} + \sum_{i \neq k+1} a_i \mathbf{z}_i \right\| &= \left\| \sum_{i=1}^{k} (a_i + O(a_{k+1}))\mathbf{z}_i + a_{k+1}\mathbf{z}_{k+1} + \sum_{i=k+2}^{n-k} a_i \mathbf{z}_i \right\| \\ &\asymp \sum_{i=1}^{k} |a_i + O(a_{k+1})| Z_i + |a_{k+1}| Z_{k+1} + \sum_{i=k+2}^{n-k} |a_i| Z_i \\ &\asymp \sum_{i=1}^{n-k} |a_i| Z_i.\end{aligned}$$

In the last line, we used the fact that if $|a_i + O(a_{k+1})| \gg a_i$, then the contribution is $\asymp |a_i| Z_i$, whereas if $|a_i + O(a_{k+1})| \ll a_{k+1}$, then the contribution is $O(a_{k+1} Z_{k+1})$ since $Z_1 \leqslant \cdots \leqslant Z_{k+1}$, and the (nonnegative) contribution is suitably bounded by the contribution from $a_{k+1} Z_{k+1}$. This gives the result. □

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

11. Type II estimate: the L^2 bound

In this section, we use the Linnik dispersion method and estimates from the geometry of numbers and elementary algebraic geometry to prove Proposition 8.7 and so finish the proof of our Type II estimate. We will make use of Proposition 9.7 and the estimates of Section 10. It is this section which involves the key new ideas behind our proof.

We recall from Proposition 8.7 that we wish to show that

$$\sum_{\substack{\mathfrak{a},\mathfrak{b}\text{ principal}\\ \mathfrak{c}|\mathfrak{b},\,\mathfrak{c}'|\mathfrak{a}\\ \mathfrak{a}\mathfrak{b}/N(\mathfrak{c})\in\mathcal{A}'}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}')(\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})-\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})) \ll \eta_2^{1/2}\#\mathcal{A}'.$$

Here $\eta_2=(\log X)^{-100(4\ell+2)}$ and

$$\mathcal{A}'=\Bigg\{\Bigg(\sum_{i=1}^{n-k}a_i\sqrt[n]{\theta^{i-1}}\Bigg): X_i\leqslant a_i\leqslant X_i+\eta_1X_i,\ a_i\equiv(\mathbf{a}_0')_i\ (\mathrm{mod}\ J!q^*),$$
$$N\Bigg(\sum_{i=1}^{n-k}a_i\sqrt[n]{\theta^{i-1}}\Bigg)\in[X_0^n,X_0^n+\eta_2X_0^n]\Bigg\}.$$

We first want to reduce this to the following proposition.

PROPOSITION 11.1. *Let* $\mathrm{mod}\ \tilde{q}=(\theta n)^nq^*N(\mathfrak{c})(J!)^J$ *and* $\epsilon_0=\tilde{q}^{-4n}\exp(-\sqrt[7]{\log X})$. *Let* $X^{k+\epsilon/2}\leqslant B\leqslant X^{n-2k-\epsilon/2}$ *and* $AB\asymp X$. *Let*

$$g_{\mathfrak{b}}=\begin{cases}\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})-\tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}), & \tau(\mathfrak{b})\leqslant\epsilon_0^{-2},\\ 0, & \textit{otherwise.}\end{cases}$$
$$\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}=\{\mathbf{a}\in\mathbb{R}^n:\|\mathbf{a}\|\in[A,2A],\ \mathbf{a}\diamond\mathbf{b}_1\in\mathcal{R}_X,\ \mathbf{a}\diamond\mathbf{b}_2\in\mathcal{R}_X\}.$$

Then we have

$$\sum_{\substack{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|\in[B,2B]\\ \mathbf{b}_1,\mathbf{b}_2\equiv\mathbf{b}_0\ (\mathrm{mod}\ \tilde{q})}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}}\sum_{\mathbf{a}\in\Lambda_{\mathbf{b}_1,\mathbf{b}_2}\cap\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}}1\ll\epsilon_0A^{n-2k}B^{2n-2k}.$$

We recall that the lattice $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ is defined in (10.2) and the region $\mathcal{R}_X$ is defined in (10.1).

Proof of Proposition 8.7 assuming Proposition 11.1. From the discussion at the beginning of Section 10, $\mathfrak{a}\mathfrak{b}/N(\mathfrak{c})\in\mathcal{A}'$ for principal $\mathfrak{a}$, $\mathfrak{b}$ is equivalent to $\mathfrak{a}=((\theta n)^{-n}\sum_{i=1}^n a_i\sqrt[n]{\theta^{i-1}})$ and $\mathfrak{b}=((\theta n)^{-n}\sum_{i=1}^n b_i\sqrt[n]{\theta^{i-1}})$ for some $\mathbf{a}\in\mathbb{Z}^n\cap\mathcal{F}$,

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$\mathbf{b} \in \mathbb{Z}^n$, for any choice of fundamental domain $\mathcal{F}$ of the action of the group of units U_K and with $\mathbf{a} \diamond \mathbf{b} \in \mathcal{R}_X$ satisfying some congruence condition $\tilde{\mathbf{L}}(\mathbf{a}, \mathbf{b}) \equiv \mathbf{0} \pmod{(\theta n)^n}$. Here we recall from (10.1) that

$$\mathcal{R}_X = \{\mathbf{x} \in \mathbb{R}^n : x_i \in [X_i', X_i' + \eta_1 X_i'] \text{ for } i \leqslant n-k,\ x_i = 0 \text{ for } i > n-k,\\ N(\textstyle\sum_{i=1}^n x_i \sqrt[n]{\theta^{i-1}}) \in [X_0'^n, X_0'^n + \eta_2 X_0'^n]\}.$$

We recall that we have localized the norms of the ideals appearing so that if $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}) \neq 0$, then $N(\mathfrak{a}) \in [A, 2A]$ for some quantity A, and if we also have $\mathfrak{a}\mathfrak{b}/N(\mathfrak{c}) \in \mathcal{A}'$, then $N(\mathfrak{b}) \in [B, 2B]$ for some quantity B with $X^{k+\epsilon/2} \leqslant B \leqslant X^{n-2k-\epsilon/2}$ and $X \ll AB \ll X$.

Any element $\mathbf{x} \in \mathcal{R}_X \cap \mathbb{Z}^n$ has $\|\mathbf{x}\| \ll X$, and so $\gamma = \sum_{i=1}^n x_i \sqrt[n]{\theta^{i-1}}$ has $|\gamma^\sigma| \ll X$ for all embeddings σ. Since $N(\gamma) = \prod_\sigma \gamma^\sigma \gg X^n$, this implies $|\gamma^\sigma| \gg X$ for all σ as well. We may choose a suitable fundamental domain $\mathcal{F}$ such that the vector $\mathbf{a}$ satisfies $\|\mathbf{a}\| \ll A$ by Lemma 4.2. This implies that $\alpha = (\theta n)^{-n} \sum_{i=1}^n a_i \sqrt[n]{\theta^{i-1}}$ has $|\alpha^\sigma| \ll A$ for all embeddings σ, and so any $\beta = \gamma/\alpha$ will then satisfy $|\beta^\sigma| \ll B$ for all σ. Thus, this choice of $\mathcal{F}$ allows us to restrict to $a_i \ll A$ and $b_i \ll B$ for all $1 \leqslant i \leqslant n$.

Thus, splitting $\mathbf{a}, \mathbf{b}$ into residue classes mod $\tilde{q} = (\theta n)^n q^* N(\mathfrak{c})(J!)^J$ (where J is the constant in the definition of $\mathcal{A}'$ which is $O(1)$ and will be eventually chosen large enough in terms of n and k), recalling that $q^* \leqslant \exp(\sqrt[4]{\log X})$ and letting $\epsilon_0 = \tilde{q}^{-4n} \exp(-\sqrt[7]{\log X})$, we see that it is sufficient to show that

$$\sum_{\substack{\mathbf{a} \in \mathbb{Z}^n \cap \mathcal{F} \\ \|\mathbf{a}\| \ll A \\ \mathbf{a} \equiv \mathbf{a}_0'' \pmod{\tilde{q}}}} \sum_{\substack{\mathbf{b} \in \Lambda_{\mathbf{a}} \\ \mathbf{b} \equiv \mathbf{b}_0 \pmod{\tilde{q}} \\ \mathbf{a} \diamond \mathbf{b} \in \mathcal{R}_X}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}')(\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})) \ll \epsilon_0^{1/2} A^{n-k} B^{n-k} \quad (11.1)$$

for any $\mathbf{a}_0''$, $\mathbf{b}_0$ with $p \nmid N_K(\mathbf{b}_0)$ for all $p \leqslant J$.

To sidestep some minor issues associated with $\tilde{\mathbf{1}}_{\mathcal{R}_2}$ occasionally being large if $\tau(\mathfrak{b})$ is large, we introduce a quantity $g_{\mathfrak{b}}$, defined by

$$g_{\mathfrak{b}} = \begin{cases} \mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}), & \tau(\mathfrak{b}) \leqslant \epsilon_0^{-2}, \\ 0, & \text{otherwise.} \end{cases}$$

We now replace $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c})$ with $g_{\mathfrak{b}}$. Since $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b}/\mathfrak{c}) \ll \tau(\sum_{i=1}^{n-k} b_i \sqrt[n]{\theta^{i-1}}) \log X$, the error introduced by this change is

$$O\Bigg(\sum_{\|\mathbf{a}\| \ll A} \sum_{\substack{\mathbf{b} \in \Lambda_{\mathbf{a}} \\ \|\mathbf{b}\| \ll B \\ \tau(\mathfrak{b}) > \epsilon_0^{-2}}} \tau(\mathfrak{b}) \log X\Bigg) \ll \sum_{\|\mathbf{a}\| \ll A} \sum_{\substack{\mathbf{b} \in \Lambda_{\mathbf{a}} \\ \|\mathbf{b}\| \ll B}} \epsilon_0^2 \tau(\mathfrak{b})^2 \log X$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll \sum_{\substack{\|\mathbf{x}\|\ll X\\ x_j=0 \text{ if } j>n-k}} \epsilon_0^2 \tau\Big(\sum_{i=1}^{n-k} x_i \sqrt[n]{\theta^{i-1}}\Big)^2$$
$$\ll \epsilon_0^2 X^{n-k} (\log X)^{O(1)},$$

by Lemma 7.8. Since $\epsilon_0 \leqslant \exp(-\sqrt[7]{\log X})$, this is $O(\epsilon_0 X^{n-k})$ and so negligible. Thus, in order to show (11.1), it is sufficient to show

$$\sum_{\substack{\mathbf{a}\in\mathbb{Z}^n\cap\mathcal{F}\\ \|\mathbf{a}\|\in[A,2A]\\ \mathbf{a}\equiv\mathbf{a}_0'' \pmod{\tilde{q}}}} \sum_{\substack{\mathbf{b}\in\Lambda_{\mathbf{a}}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod{\tilde{q}}\\ \mathbf{a}\diamond\mathbf{b}\in\mathcal{R}_X}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}') g_{\mathbf{b}} \ll \epsilon_0^{1/2} A^{n-k} B^{n-k}. \tag{11.2}$$

By Cauchy–Schwarz (dropping the constraints $\mathbf{a} \equiv \mathbf{a}_0'' \pmod{\tilde{q}}$ and $\mathbf{a} \in \mathcal{F}$ and upper bounding $\mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}')$ by 1), we have

$$\sum_{\substack{\mathbf{a}\in\mathbb{Z}^n\cap\mathcal{F}\\ \|\mathbf{a}\|\in[A,2A]\\ \mathbf{a}\equiv\mathbf{a}_0'' \pmod{\tilde{q}}}} \sum_{\substack{\mathbf{b}\in\Lambda_{\mathbf{a}}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod{\tilde{q}}\\ \mathbf{a}\diamond\mathbf{b}\in\mathcal{R}_X}} \mathbf{1}_{\mathcal{R}_1}(\mathfrak{a}/\mathfrak{c}') g_{\mathbf{b}}$$
$$\ll \Big(\sum_{\|\mathbf{a}\|\ll A} 1\Big)^{1/2} \Big(\sum_{\|\mathbf{a}\|\in[A,2A]} \Big|\sum_{\substack{\mathbf{b}\in\Lambda_{\mathbf{a}}\\ \mathbf{b}\equiv\mathbf{b}_0 \pmod{\tilde{q}}\\ \mathbf{a}\diamond\mathbf{b}\in\mathcal{R}_X}} g_{\mathbf{b}}\Big|^2\Big)^{1/2}.$$

The first sum in parentheses is $O(A^n)$, so it suffices to show that

$$\sum_{\substack{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|\in[B,2B]\\ \mathbf{b}_1,\mathbf{b}_2\equiv\mathbf{b}_0 \pmod{\tilde{q}}}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}} \sum_{\mathbf{a}\in\Lambda_{\mathbf{b}_1,\mathbf{b}_2}\cap\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}} 1 \ll \epsilon_0 A^{n-2k} B^{2n-2k}, \tag{11.3}$$

where

$$\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2} = \{\mathbf{a}\in\mathbb{R}^n : \|\mathbf{a}\|\in[A,2A],\ \mathbf{a}\diamond\mathbf{b}_1\in\mathcal{R}_X,\ \mathbf{a}\diamond\mathbf{b}_2\in\mathcal{R}_X\}.$$

This is precisely given by Proposition 11.1. □

Thus, we are left to establish Proposition 11.1.

If $\wedge(\mathbf{b}_1,\mathbf{b}_2) \neq \mathbf{0}$, then $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ is a rank $n-2k$ lattice, and we expect the inner sum in (11.3) to typically be (using Lemma 10.1)

$$\approx \frac{\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}}{\det\Lambda_{\mathbf{b}_1,\mathbf{b}_2}} = \frac{D_{\mathbf{b}_1,\mathbf{b}_2}\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}}{\|\wedge(\mathbf{b}_1,\mathbf{b}_2)\|} \approx \frac{cA^{n-2k}}{B^{2k}}$$

for some suitable constant $c = c_{\mathbf{b}_1,\mathbf{b}_2}$ of size ≈ 1 which varies continuously and slowly with $\mathbf{b}_1,\mathbf{b}_2$. The first approximation can fail if $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ is highly

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

skewed, whilst the second approximation can fail if $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ has an unusually small determinant. $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ can have small determinant either for Archimedean reasons (if $\| \wedge (\mathbf{b}_1, \mathbf{b}_2)\|$ is small) or for non-Archimedean reasons (if $D_{\mathbf{b}_1,\mathbf{b}_2}$ is large). To deal with these issues, we show that for most $\mathbf{b}_1, \mathbf{b}_2$, these complications do not occur.

REMARK. Usually one would introduce a smooth weight on the sum over $\mathbf{a}$ to allow for simpler or more precise analysis of the resulting inner sum. We have deliberately chosen not to smooth here because we wish to emphasize the elementary nature of the estimates we use from the geometry of numbers. In principal, smoothing would allow one to use exponential sums to widen the Type II ranges, but the author has not been able to get suitable control over the resulting exponential sums. Nontrivially estimating these sums for general n requires one to show equidistribution results for skewed lattices.

REMARK. The diagonal terms $\mathbf{b}_1 = \mathbf{b}_2$ contribute $A^{n-k}B^{n-k+o(1)}$ to the overall sum, and so we require $A^k < B^{n-k+o(1)}$. If we do not show cancellations in the error terms for the inner sum over $\mathbf{a}$ above, then we can only hope to gain an asymptotic if $A^{n-2k} > B^{2k}$ (but see the remark below). Together, these conditions force $X^k < B < X^{n-2k}$, and our Type II estimate applies in essentially the full range. Similar restrictions apply to any other sequence of density $1-k/n$, which is why the initial work on Diophantine approximation by primes had equivalent restrictions on the Type II range.

REMARK. We can obtain slightly more flexibility in our Type II estimates by restricting $\mathbf{b}$ to lie in a residue class $\pmod{Q}$ for a suitably sized modulus Q before applying Cauchy–Schwarz. This has the effect of increasing the contribution from the diagonal terms but enabling us to estimate the off-diagonal terms in a wider range. This has the potential to give an asymptotic formula for primes represented by an incomplete norm form of $\mathbb{Q}(\sqrt[n]{\theta})$ in the wider range $n > (2+\sqrt{2})k$. In the interests of brevity and clarity, we will not consider this further here, but we intend to address this in a future paper.

11.1. Archimedean estimates. We first consider complications when the lattice $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$ is skewed or has small determinant because $\| \wedge (\mathbf{b}_1, \mathbf{b}_2)\|$ is small.

We begin with a simple lemma counting the number of times a polynomial can be small. The key point is that this estimate is very uniform in the coefficients of f.

LEMMA 11.2. *Let* $f(x) = f_d x_1^d + f_{d-1}x^{d-1} + \cdots + f_0 \in \mathbb{Z}[x]$ *with* $f_d \neq 0$. *Let*

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$D \in \mathbb{Z}$ *be such that* $D/\gcd(D, f_d) = \prod_{i=1}^{\ell} p_i^{e_i}$. *Then we have*

$$\#\{n \in [1, y] : f(n) \equiv 0 \pmod{D},\ |f(n)| \leqslant B\} \leqslant d\tau_d(D)\left(1 + \min\left(y, \frac{B^{1/d}}{|f_d|^{1/d}}\right)\frac{1}{D'}\right),$$

where $D' = \prod_{i=1}^{\ell} p_i^{\lceil e_i/d \rceil} > D^{1/d}/\gcd(D, f_d)^{1/d}$.

Proof. Let $\tilde{D} = D/\gcd(D, f_d)$. Let f have (not necessarily distinct) roots $\alpha_{p,1}, \dots, \alpha_{p,d}$ in a suitable finite extension of $\mathbb{Q}_p$, and let $\|\cdot\|_p$ be the extension of the norm on $\mathbb{Q}_p$. Similarly, let f have roots $\alpha_{\infty,1}, \dots, \alpha_{\infty,d}$ over $\mathbb{C}$. If $f(n) \equiv 0 \pmod{D}$, then $\prod_{i=1}^{d} \|n - \alpha_{p,i}\|_p \leqslant \|\tilde{D}\|_p$ for all primes $p|\tilde{D}$, so certainly there exists a root $\alpha^{(p)}$ for each $p|\tilde{D}$ such that $\|n - \alpha^{(p)}\|_p \leqslant \|D'\|_p$ on recalling the definition of D'. Similarly, if $|f(n)| \leqslant B$, then certainly there is a root $\alpha^{(\infty)}$ such that $|n - \alpha^{(\infty)}| \leqslant (B/|f_d|)^{1/d}$.

Let us be given a root $\alpha^{(\infty)}$ over $\mathbb{C}$ and a root $\alpha^{(p)}$ over $\overline{\mathbb{Q}}_p$ for each prime $p|\tilde{D}$. Then integers n which satisfy $\|n - \alpha^{(p)}\|_p \leqslant \|D'\|_p$ for each $p|D'$ are simply integers in a single residue class modulo D' (by the Chinese remainder theorem), and those with $|n - \alpha^{(\infty)}| < (B/|f_d|)^{1/d}$ and $n \in [1, y]$ are integers in an interval of length $\ll \min(y, (B/|f_d|)^{1/d})$. Thus, there are at most $1 + \min(y, (B/|f_d|)^{1/d})/D'$ integers $n \leqslant y$ which satisfy $\|a - \alpha^{(p)}\|_p \leqslant \|D'\|_p$ for each $p|D'$ and $|n - \alpha^{(\infty)}| \leqslant (B/|f_d|)^{1/d}$. But there are at most d choices of $\alpha^{(\infty)}$ and at most $\tau_d(\tilde{D})$ possible choices of roots $\alpha^{(p)}$, so there are at most $d\tau_d(\tilde{D})(1 + \min(y, B^{1/d}|f_d|^{-1/d})/D')$ integers $n \leqslant y$ such that $f(n) \equiv 0 \pmod{D}$ and $|f(n)| \leqslant B$. □

LEMMA 11.3. *Let* $n > 3k$. *Let* $\Lambda_{\mathbf{a}}$ *have successive minima* $Z_1 \leqslant \cdots \leqslant Z_{n-k}$ *and a Minkowski-reduced basis* $\{\mathbf{z}_1, \dots, \mathbf{z}_{n-k}\}$. *Let* $\ell \leqslant 2k$ *be such that* $\kappa_2 = \|\wedge(\mathbf{z}_1, \mathbf{z}_\ell)\| Z_1^{-k} Z_\ell^{-k} > 0$. *Assume that* $Z_1 Z_\ell \ll BC$. *Finally, let*

$$S_{\mathbf{a}}(B, C; \kappa) = \#\{\mathbf{b}, \mathbf{c} \in \Lambda_{\mathbf{a}} : \|\mathbf{b}\| \leqslant B, \|\mathbf{c}\| \leqslant C, \|\wedge(\mathbf{b}, \mathbf{c})\| \leqslant \kappa B^k C^k\}.$$

Then we have

$$S_{\mathbf{a}}(B, C; \kappa) \ll \left(\frac{Z_1}{B} + \frac{Z_1 Z_\ell}{BC} + \min\left(1, \left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\frac{Z_1}{Z_\ell}\right) \times \frac{Z_\ell^\ell}{\prod_{i=1}^{\ell} Z_i}\left(\frac{BC}{Z_1 Z_\ell}\right)^{n-k} \log BC.$$

Proof. By symmetry, we may assume, without loss of generality, that $B \leqslant C$. We

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

may further assume that $Z_1 \ll B$ since otherwise there are no vectors $\mathbf{b} \in \Lambda_{\mathbf{a}}$ with $\|\mathbf{b}\| \leqslant B$ and so $S_{\mathbf{a}}(B, C; \kappa) = 0$.

We recall from Lemma 4.1 that we can write $\mathbf{b} = \sum_{i=1}^{n-k} b_i \mathbf{z}_i$, $\mathbf{c} = \sum_{i=1}^{n-k} c_i \mathbf{z}_i$ for integers $b_i \ll B/Z_i$ and $c_i \ll C/Z_i$. We have that $\| \wedge (\mathbf{b}, \mathbf{c})\|^2$ is given by an integer polynomial of degree $4k$ in the coefficients b_i, c_i, which is a polynomial of degree $2k$ in the b_i and degree $2k$ in the c_i. Since the coefficient of $b_1^{2k} c_\ell^{2k}$ is $\| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^2 \neq 0$, we have that this polynomial takes the form

$$b_1^{2k}(c_\ell^{2k} \| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^2 + f_2) + f_3,$$

where f_2 is a polynomial independent of b_1 and degree at most $2k-1$ in c_ℓ, and f_3 is a polynomial of degree at most $2k-1$ in b_1.

Let us be given a choice of $b_2, \dots, b_{n-k}, c_1, \dots, c_{\ell-1}, c_{\ell+1}, \dots, c_{n-k}$ and a quantity $U = 2^j \ll C^{2k} B^{2k}$. By Lemma 11.2, there are

$$\ll 1 + U^{1/2k} \| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^{-1/k}$$

possible values of c_ℓ such that $c_\ell^{2k} \| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^2 + f_2 \in [U, 2U]$. Here the implied constant does not depend on our choice of the other b_i, c_i or on U. For each such choice of c_ℓ, there are $O(1 + \kappa^{1/k} BCU^{-1/2k})$ possible choices of b_1 such that $\| \wedge (\mathbf{b}, \mathbf{c})\|^2 \ll \kappa^2 B^{2k} C^{2k}$ by Lemma 11.2 again. Thus, combining these bounds with the trivial bounds B/Z_1 and $1 + C/Z_\ell$ for the number of choices of b_1 and c_ℓ, respectively, we find that there are

$$\begin{aligned}
&\ll \min\left(1 + \frac{C}{Z_\ell}, 1 + \frac{U^{1/2k}}{\| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^{1/k}}\right) \min\left(\frac{B}{Z_1}, 1 + \frac{\kappa^{1/k} BC}{U^{1/2k}}\right) \\
&\ll 1 + \frac{B}{Z_1} + \frac{C}{Z_\ell} + \frac{\kappa^{1/k} BC}{\| \wedge (\mathbf{z}_1, \mathbf{z}_\ell)\|^{1/k}} \\
&\ll \frac{B}{Z_1} + \frac{C}{Z_\ell} + \frac{\kappa^{1/k} BC}{\kappa_2^{1/k} Z_1 Z_\ell}
\end{aligned}$$

possible choices of b_1, c_ℓ for this value of U. Since this bound does not depend on U, we can sum over all possible values of $U = 2^j$ with $1 \leqslant U \ll B^k C^k$ at the cost of a factor $O(\log BC)$. We also have the trivial bound where κ/κ_2 is replaced by 1. Thus, for any choice of $b_2, \dots, b_{n-k}, c_1, \dots, c_{\ell-1}, c_{\ell+1}, \dots, c_{n-k}$, we have

$$\ll \left(\frac{Z_1}{B} + \frac{Z_\ell}{C} + \min\left(1, \left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\right) \frac{BC \log BC}{Z_1 Z_\ell} \tag{11.4}$$

choices of b_1, c_ℓ such that $\| \wedge (\mathbf{b}, \mathbf{c})\| \leqslant \kappa B^k C^k$.

Let $j_B, j_C \leqslant n-k$ be chosen maximally such that $Z_{j_B} \leqslant B$ and $Z_{j_C} \leqslant C$. Then, since the number of choices of b_i is $O(1 + B/Z_i)$ (and similarly for c_i),

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

the number of choices of $b_2, \dots, b_{n-k}, c_1, \dots, c_{\ell-1}, c_{\ell+1}, \dots, c_{n-k}$ is

$$\ll \prod_{\substack{1\leqslant i\leqslant j_B\\ i\neq 1}} \frac{B}{Z_i} \prod_{\substack{1\leqslant i\leqslant j_C\\ i\neq \ell}} \frac{C}{Z_i}. \tag{11.5}$$

We recall that we assume $B \leqslant C$ so $j_B \leqslant j_C$. Thus, splitting into the three cases $j_C \geqslant j_B \geqslant \ell$, $j_C \geqslant \ell > j_B$ and $\ell > j_C \geqslant j_B$ and pulling out a factor $Z_\ell^\ell / \prod_{i=1}^{\ell} Z_i$, we see that (11.5) is

$$\ll Y = \frac{Z_\ell^\ell}{\prod_{i=1}^{\ell} Z_i} \times \begin{cases} \dfrac{B^{j_B-1}C^{j_C-1}}{Z_1^{\ell-2}Z_\ell^{j_B+j_C-\ell}}, & j_C \geqslant j_B \geqslant \ell, \\ \dfrac{B^{j_B-1}C^{j_C-1}}{Z_1^{j_B-1}Z_\ell^{j_C-1}}, & j_C \geqslant \ell > j_B, \\ \dfrac{B^{j_B-1}C^{j_C}}{Z_1^{j_B-1}Z_\ell^{j_C}}, & \ell > j_C \geqslant j_B. \end{cases}$$

Define a quantity F by

$$F = \begin{cases} \left(\dfrac{BC}{Z_1Z_\ell}\right)^{n-k-j_C} \left(\dfrac{Z_\ell}{Z_1}\right)^{j_B-\ell} \left(\dfrac{B}{Z_1}\right)^{j_C-j_B}, & j_C \geqslant j_B \geqslant \ell, \\ \left(\dfrac{BC}{Z_1Z_\ell}\right)^{n-k-j_C} \left(\dfrac{B}{Z_1}\right)^{j_C-j_B-1}, & j_C \geqslant \ell > j_B, \\ \left(\dfrac{BC}{Z_1Z_\ell}\right)^{n-k-j_C-1} \left(\dfrac{B}{Z_1}\right)^{j_C-j_B}, & \ell > j_C \geqslant j_B. \end{cases}$$

Since $BC/Z_1Z_\ell, Z_\ell/Z_1, B/Z_1 \gg 1$ and we have the bounds $\ell \leqslant 2k < n-k$ and $j_C \leqslant n-k$, we see that $F \geqslant 1$. Thus, we find that for all cases, we have

$$Y \leqslant YF \leqslant \frac{Z_\ell^\ell}{\prod_{i=1}^{\ell} Z_i} \frac{B^{n-k-1}C^{n-k-1}}{Z_1^{n-k-1}Z_\ell^{n-k-1}} \left(\frac{Z_1}{B} + \frac{Z_1}{Z_\ell}\right).$$

Combining this with our bound (11.4) on the number of choices of b_1, c_ℓ, we obtain that the total number of $\mathbf{b}, \mathbf{c}$ is

$$\ll \left(\frac{Z_1}{B} + \frac{Z_\ell}{C} + \min\left(1, \left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\right)\left(\frac{Z_1}{B} + \frac{Z_1}{Z_\ell}\right) \frac{Z_\ell^\ell B^{n-k}C^{n-k} \log BC}{Z_1^{n-k}Z_\ell^{n-k} \prod_{i=1}^{\ell} Z_i}.$$

Finally, we note that since $Z_1 \leqslant B$,

$$\frac{Z_1}{B}\left(\frac{Z_1}{B} + \frac{Z_\ell}{C} + \min\left(1, \left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\right) \ll \frac{Z_1Z_\ell}{BC} + \frac{Z_1}{B},$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

and since $Z_1 \leqslant Z_\ell$, $B \leqslant C$,

$$\frac{Z_1}{Z_\ell}\left(\frac{Z_1}{B}+\frac{Z_\ell}{C}+\min\left(1,\left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\right) \ll \frac{Z_1}{B}+\min\left(1,\left(\frac{\kappa}{\kappa_2}\right)^{1/k}\right)\frac{Z_1}{Z_\ell}.$$

These bounds give the result. □

LEMMA 11.4 (Determinant rarely small for Archimedean reasons). *Let $n > 3k$. Let*

$$S(A;B,C) = \{(\mathbf{a},\mathbf{b},\mathbf{c}) \in (\mathbb{Z}^n)^3 : \|\mathbf{a}\| \in [A,2A], \|\mathbf{b}\| \in [B,2B], \|\mathbf{c}\| \in [C,2C], \wedge(\mathbf{b},\mathbf{c}) \neq \mathbf{0},\ \mathbf{a} \in \Lambda_{\mathbf{b},\mathbf{c}}\}$$
$$S(A;B,C;\kappa) = \{(\mathbf{a},\mathbf{b},\mathbf{c}) \in S(A;B,C) : \|\wedge(\mathbf{b},\mathbf{c})\| \leqslant \kappa B^k C^k\}.$$

Then there is a constant $\delta = \delta(n,k) > 0$ and $G = G(n,k)$ such that

$$\#S(A;B,C;\kappa) \ll \left(\kappa^{\delta/k} + \min\left(1,\frac{(BC)^{1/2-\delta}}{B}\right)\right)A^{n-2k}B^{n-k}C^{n-k}\exp(G(\log\log BC)^2).$$

In particular, taking $\kappa = \epsilon_0^{8k/\delta} = \tilde{q}^{32kn/\delta}\exp(-8k\sqrt{\log X}/\delta)$ and $B = C \gg X^\delta$, we have

$$\#\{(\mathbf{a},\mathbf{b}_1,\mathbf{b}_2) \in \mathcal{S}(A;B,B) : 0 < \|\wedge(\mathbf{b}_1,\mathbf{b}_2)\| \ll \epsilon_0^{8k/\delta}B^{2k}\} \ll \epsilon_0^7 A^{n-2k}B^{2n-2k}.$$

Proof. We prove the result by induction on the size of BC. The lemma trivially holds if $BC \ll 1$. Assume that there is a constant G such that whenever $UV < 2^J$ with $U \leqslant V$, we have

$$\#S(A;U,V;\kappa) \leqslant G\left(\kappa^{\delta/k} + \min\left(1,\frac{(UV)^{1/2-\delta}}{U}\right)\right)A^{n-2k}U^{n-k}V^{n-k}\exp(G(\log\log UV)^2).$$

We now wish to bound $\#S(A;B,C;\kappa)$ using the same constant G for $BC \leqslant 2^{J+\delta J}$.

Given $\mathbf{b},\mathbf{c}$ with $\wedge(\mathbf{b},\mathbf{c}) \neq \mathbf{0}$ and $\|\mathbf{b}\| \in [B,2B]$ and $\|\mathbf{c}\| \in [C,2C]$, we have that any $\mathbf{a}$ such that $(\mathbf{a},\mathbf{b},\mathbf{c})$ is in $S(A;B,C,\kappa)$ satisfies $\|\mathbf{a}\| \in [A,2A]$ and $\mathbf{a} \in \Lambda_{\mathbf{b},\mathbf{c}}$. Since $\wedge(\mathbf{b},\mathbf{c}) \neq \mathbf{0}$, $\Lambda_{\mathbf{b},\mathbf{c}}$ is a lattice of rank $n-2k$ and determinant $\ll B^kC^k$. If $\mathbf{v} = \mathbf{v}(\mathbf{b},\mathbf{c})$ is the shortest vector in $\Lambda_{\mathbf{b},\mathbf{c}}$, then $\|\mathbf{v}\|^{n-2k} \ll B^kC^k$ and the number of $\mathbf{a} \in \Lambda_{\mathbf{b},\mathbf{c}}$ with $\|\mathbf{a}\| \in [A,2A]$ is $O(A^{n-2k}/\|\mathbf{v}\|^{n-2k})$. We recall

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

that $\mathbf{v} \in \Lambda_{\mathbf{b},\mathbf{c}}$ implies $\mathbf{b}, \mathbf{c} \in \Lambda_{\mathbf{v}}$. Thus, putting $\|\mathbf{v}\|$ in one of $O(\log BC)$ dyadic ranges $[V, 2V]$, we have

$$S(A; B, C; \kappa) \ll A^{n-2k} \sum_{\substack{\|\mathbf{b}\| \in [B,2B] \\ \|\mathbf{c}\| \in [C,2C] \\ \|\wedge(\mathbf{b},\mathbf{c})\| \leqslant \kappa BC}} \frac{1}{\|\mathbf{v}(\mathbf{b},\mathbf{c})\|^{n-2k}}$$

$$\ll A^{n-2k} (\log BC) \sup_{V^{n-2k} \ll B^k C^k} \sum_{\|\mathbf{v}\| \in [V,2V]} \frac{1}{V^{n-2k}} \sum_{\substack{\mathbf{b},\mathbf{c} \in \Lambda_{\mathbf{v}} \\ \|\mathbf{b}\| \in [B,2B] \\ \|\mathbf{c}\| \in [C,2C] \\ \|\wedge(\mathbf{b},\mathbf{c})\| \leqslant \kappa BC}} 1. \tag{11.6}$$

Since $\mathbf{v} \neq \mathbf{0}$, $\Lambda_{\mathbf{v}}$ is a lattice of rank $n-k$. Let this have successive minima $Z_1 \leqslant \cdots \leqslant Z_{n-k}$. We note that since $n > 3k$ and $V^{n-2k} \ll B^k C^k$, we have

$$Z_1^k Z_{k+1}^k \ll Z_1^k Z_{k+1}^{n-2k} \ll \prod_{i=1}^{n-k} Z_i \ll \det(\Lambda_{\mathbf{v}}) \ll V^k \ll (BC)^{k^2/(n-2k)}. \tag{11.7}$$

Thus, $Z_1 Z_{k+1} \ll (BC)^{1-2\delta}$ where $\delta = (n-3k)/(2n-4k) > 0$. By Lemma 11.3 (taking $\ell = k+1$), the inner sum in (11.6) is

$$\ll \left(\frac{Z_1}{B} + \frac{Z_1 Z_{k+1}}{BC} + \min\left(1, \left(\frac{\kappa}{\kappa_2}\right)^{\delta/k} \right) \frac{Z_1}{Z_{k+1}} \right) \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i} \left(\frac{BC}{Z_1 Z_{k+1}} \right)^{n-k} \log BC, \tag{11.8}$$

where $\kappa_2 = \sup_{\mathbf{z}_1, \mathbf{z}_{k+1}} Z_1^{-k} Z_{k+1}^{-k} \|\wedge(\mathbf{z}_1, \mathbf{z}_{k+1})\|$ and the supremum is over all $\mathbf{z}_1$, $\mathbf{z}_{k+1} \in \Lambda_{\mathbf{a}}$ which can be extended to a basis $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ with $\|\sum_{i=1}^{n-k} \lambda_i \mathbf{z}_i\| \asymp \sum_{i=1}^{n-k} |\lambda_i| Z_i$. We see that $\kappa_2 > 0$ by Lemma 10.5.

We note that there are

$$\gg \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i}$$

different vectors $\mathbf{y} \in \Lambda_{\mathbf{v}}$ with $Z_{k+1} \ll \|\mathbf{y}\| \ll Z_{k+1}$ such that $0 < \|\wedge(\mathbf{z}_1, \mathbf{y})\| \ll \kappa_2 Z_1^k Z_{k+1}^k$ since, given a basis $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ satisfying the properties of Lemma 10.5, all choices $\mathbf{y} = \mathbf{z}_{k+1} + \sum_{i=1}^k \lambda_i \mathbf{z}_i$ with $\|\wedge(\mathbf{y}, \mathbf{z}_1)\| \neq 0$ and $\lambda_i \ll Z_{k+1}/Z_i$ satisfy this by the maximality of κ_2. Thus, we may replace the factor $Z_{k+1}^k / \prod_{i=1}^k Z_i$ of (11.8) with a sum over all such $\mathbf{y}$. Putting $\|\mathbf{z}_1\|$, $\|\mathbf{y}\|$, κ_2 each in one of $O(\log BC)$ dyadic ranges $[Z, 2Z]$, $[Y, 2Y]$ and $[K, 2K]$, respectively, we have

$$S(A; B, C; \kappa) \ll A^{n-2k} B^{n-k} C^{n-k} (\log BC)^5$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\times \sup_{\substack{V^{n-2k}\ll B^kC^k\\ Z\ll Y\\ Z^kY^{n-2k}\ll V^k\\ (ZY)^{-k}\ll K\ll 1}} \sum_{\|\mathbf{v}\|\in[V,2V]} \frac{T_{Z,Y;B,C;\kappa,K}}{Z^{n-k}Y^{n-k}V^{n-2k}} \sum_{\substack{\mathbf{z}_1,\mathbf{y}\in\Lambda_{\mathbf{v}}\\ \|\mathbf{z}_1\|\in[Z,2Z]\\ \|\mathbf{y}\|\in[Y,2Y]\\ 0<\|\wedge(\mathbf{z}_1,\mathbf{y})\|\ll K(ZY)^k}} 1$$

$$\ll A^{n-2k}B^{n-k}C^{n-k}(\log BC)^5 \sup_{V,Z,Y,K} \frac{T_{Z,Y;B,C;\kappa,K}S(V;Z,Y;K)}{Z^{n-k}Y^{n-k}V^{n-2k}}, \tag{11.9}$$

where

$$T_{Z,Y;B,C;\kappa,K} = \left(\frac{Z}{B}+\frac{ZY}{BC}+\frac{Z}{Y}\min\left(1,\left(\frac{\kappa}{K}\right)^{\delta/k}\right)\right),$$

and where the supremum in the final line is over all V, Z, Y, K satisfying the constraints $V^{n-2k}\ll B^kC^k$, $Z\ll Y$, $Z^kY^{n-2k}\ll V^k$ and $(ZY)^{-k}\ll K\ll 1$.

If $BC < 2^{J+\delta J}$, then, since $ZY \ll (BC)^{1-2\delta}$, we have $ZY < 2^J$ if J is sufficiently large in terms of n, k. We can then apply the assumption of the lemma, giving

$$\frac{S(A;B,C;\kappa)}{A^{n-2k}B^{n-k}C^{n-k}(\log BC)^5}$$
$$\ll G \sup_{V,Z,Y,K} T_{Z,Y;B,C;\kappa,K}\left(K^{\delta/k}+\min\left(1,\frac{(ZY)^{1/2-\delta}}{Z}\right)\right)\exp(G(\log\log ZY)^2). \tag{11.10}$$

Since $Z\ll Y$ and $K\gg (ZY)^{-k}$, we have

$$\frac{Z}{Y}\min\left(1,\left(\frac{\kappa}{K}\right)^{\delta/k}\right)\left(K^{\delta/k}+\frac{(ZY)^{1/2-\delta}}{Z}\right) \ll \kappa^{\delta/k}\left(1+\frac{(ZY)^{\delta}(ZY)^{1/2-\delta}}{Y}\right)$$
$$\ll \kappa^{\delta/k}.$$

Since $K\ll 1$, $Z\ll B$, $ZY\ll (BC)^{1-2\delta}$ and $Z\ll (ZY)^{1/2}\ll (BC)^{1/2-\delta}$, we have

$$\left(\frac{Z}{B}+\frac{ZY}{BC}\right)\left(K^{\delta/k}+1\right) \ll \frac{Z}{B}+\frac{ZY}{BC}$$
$$\ll \min\left(1,\frac{(BC)^{1/2-\delta}}{B}\right)+\frac{1}{(BC)^{2\delta}}$$
$$\ll \min\left(1,\frac{(BC)^{1/2-\delta}}{B}\right).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Thus,

$$T_{Z,Y;B,C;\kappa,K}\left(K^{\delta/k}+\min\left(1,\frac{(ZY)^{1/2-\delta}}{Z}\right)\right)\ll\kappa^{\delta/k}+\min\left(1,\frac{(BC)^{1/2-\delta}}{B}\right).$$

Since $ZY\ll(BC)^{1-2\delta}$, we have $\log\log ZY<\log\log BC-2\delta$. Substituting these bounds into (11.10) gives

$$\frac{S(A;B,C;\kappa)}{A^{n-2k}B^{n-k}C^{n-k}(\log BC)^5}\ll G\left(\kappa^{\delta/k}+\min\left(1,\frac{(BC)^{1/2-\delta}}{B}\right)\right)\exp(G(\log\log BC-2\delta)^2).$$

Finally, $\exp(G(\log\log BC-2\delta)^2)\ll(\log BC)^{-6}\exp(G(\log\log BC)^2)$ for $G>2\delta^{-1}$, and so we obtain the claimed bound for $S(A;B,C;\kappa)$ if BC is large enough. □

LEMMA 11.5. *Let $n>3k$ and G' be sufficiently large in terms of n and k. Let $Z_i(\mathbf{a})$ be the ith successive minimum of $\Lambda_{\mathbf{a}}$. Then we have*

$$\sum_{0<\|\mathbf{a}\|\ll A}\frac{Z_{k+1}(\mathbf{a})^k}{Z_1(\mathbf{a})^{n-k}Z_{k+1}(\mathbf{a})^{n-k}\prod_{i=1}^k Z_i(\mathbf{a})}\ll A^{n-2k}\exp(G'(\log\log A)^2).$$

Proof. We already established a similar estimate in the course of the proof of Lemma 11.4. Let $\Lambda_{\mathbf{a}}$ have a basis $\mathbf{z}_1,\dots,\mathbf{z}_{n-k}$ with $\|\wedge(\mathbf{z}_1,\mathbf{z}_{k+1})\|\neq 0$ and $\|\sum_{i=1}^{n-k}\lambda_i\mathbf{z}_i\|\asymp\sum_{i=1}^{n-k}|\lambda_i|Z_i(\mathbf{a})$ for all $\lambda\in\mathbb{R}$. This basis exists by Lemma 10.5. There are

$$\gg\frac{Z_{k+1}^k(\mathbf{a})}{\prod_{i=1}^k Z_i(\mathbf{a})}$$

choices of $\mathbf{y}=\mathbf{z}_{k+1}+\sum_{i=1}^k\lambda_i\mathbf{z}_i$ with $\|\mathbf{y}\|\asymp Z_{k+1}(\mathbf{a})$ and $\wedge(\mathbf{z}_1,\mathbf{y})\neq\mathbf{0}$. Thus, using Lemma 11.4, we find

$$\begin{aligned}\sum_{0<\|\mathbf{a}\|\ll A}\frac{Z_{k+1}(\mathbf{a})^k}{Z_1(\mathbf{a})^{n-k}Z_{k+1}(\mathbf{a})^{n-k}\prod_{i=1}^k Z_i(\mathbf{a})}&\ll\sum_{0<\|\mathbf{a}\|\ll A}\sum_{\substack{\mathbf{z},\mathbf{y}\in\Lambda_{\mathbf{a}}\\ \|\mathbf{z}\|\asymp Z_1(\mathbf{a})\\ \|\mathbf{y}\|\asymp Z_{k+1}(\mathbf{a})\\ \wedge(\mathbf{z},\mathbf{y})\neq\mathbf{0}}}\frac{1}{\|\mathbf{z}\|^{n-k}\|\mathbf{y}\|^{n-k}}\\ &\ll(\log A)^3\sup_{0<Z,Y,A'\ll A}\sum_{\|\mathbf{y}\|\asymp Y}\sum_{\substack{\|\mathbf{z}\|\asymp Z\\ \wedge(\mathbf{y},\mathbf{z})\neq\mathbf{0}}}\sum_{\substack{\mathbf{a}\in\Lambda_{\mathbf{z},\mathbf{y}}\\ \|\mathbf{a}\|\asymp A'}}\frac{1}{Y^{n-k}Z^{n-k}}\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\ll (\log A)^3 \sup_{0<Z,Y,A'\ll A} \frac{S(A', Y, Z)}{Y^{n-k}Z^{n-k}}$$
$$\ll (\log A)^3 A^{n-2k} \exp(G(\log\log A^2)^2).$$

The result follows on taking $G' = 2G$. □

LEMMA 11.6. *Let $n > 3k$, and let $\delta > 0$ be sufficiently small in terms of n and k. Given a vector $\mathbf{a} \in \mathbb{Z}^n\backslash\{\mathbf{0}\}$, let $Z_{n-k}(\mathbf{a})$ be the $n-k$th successive minima of $\Lambda_{\mathbf{a}}$. Then if $A^{k/(1-\delta)} < B^{n-k}$, we have*

$$\#\{(\mathbf{a}, \mathbf{b}_1, \mathbf{b}_2) \in \mathcal{S}(A; B, B) : Z_{n-k}(\mathbf{a}) > B^{1-\delta/2}\} \ll A^{n-2k}B^{2n-2k-\delta/2}.$$

Proof. Let $Z_1, \dots, Z_{n-k}$ be the successive minima of $\Lambda_{\mathbf{a}}$. Since $A^{k/(1-\delta)} < B^{n-k}$, $n > 3k$ and $Z_1^k Z_{k+1}^{n-2k} \ll \det(\Lambda_{\mathbf{a}}) \ll A^k$, we have $Z_1 Z_{k+1} \ll B^{2-2\delta}$. Let j be chosen maximally such that $Z_j \leqslant B$. Then the number of $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}}$ is

$$\ll \frac{B^{2j}}{\prod_{i=1}^j Z_i^2} \ll \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i} \times \begin{cases} \dfrac{B^{2j}}{Z_1^j Z_{k+1}^j}, & j \leqslant k, \\ \dfrac{B^{2j}}{Z_1^k Z_{k+1}^{2j-k}}, & k+1 \leqslant j < n-k, \\ \dfrac{B^{2n-2k}}{Z_1^k Z_{k+1}^{2n-3k-2} Z_{n-k}^2}, & j = n-k. \end{cases}$$

Since $Z_{n-k} > B^{1-\delta/2}$ and $Z_1 Z_{k+1} \ll B^{2-2\delta}$ and $n > 3k$, we have that in each case, this is

$$\ll \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i} \frac{B^{2n-2k-\delta}}{Z_1^{n-k} Z_{k+1}^{n-k}}.$$

Thus, the number of triples $(\mathbf{a}, \mathbf{b}_1, \mathbf{b}_2)$ counted in the lemma is

$$\ll B^{2n-2k-\delta} \sum_{\|\mathbf{a}\|\in[A,2A]} \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i} \frac{1}{Z_1^{n-k} Z_2^{n-k}}.$$

But by Lemma 11.5, this is $\ll A^{n-2k}B^{2n-2k-\delta/2}$, as required. □

LEMMA 11.7 (Diagonal terms). *Let $n > 3k$ and let $\delta > 0$ be sufficiently small in terms of n and k. Then if $A^{k/(1-\delta)} < B^{n-k}$, we have*

$$\#\{(\mathbf{a}, \mathbf{b}_1, \mathbf{b}_2) \in S(A; B, B) : \wedge(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{0}\} \ll A^{n-2k}B^{2n-2k-\delta/3}.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Proof. Let $\mathbf{a}$ be given; so we wish to count $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}}$ with $\wedge(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{0}$. Let $\Lambda_{\mathbf{a}}$ have a basis $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ satisfying the properties of Lemma 10.5, and let $\mathbf{b}_1 = \sum_{i=1}^{n-k} \lambda_i \mathbf{z}_i$, $\mathbf{b}_2 = \sum_{i=1}^{n-k} \gamma_i \mathbf{z}_i$, with $\gamma_i, \lambda_i \ll B/Z_i$, where $Z_i = Z_i(\mathbf{a})$ are the successive minima of $\Lambda_{\mathbf{a}}$. By Lemma 11.6, we only need to count the contribution from $\mathbf{a}$ with $Z_{n-k}(\mathbf{a}) \ll B^{1-\delta/2}$. Since $\wedge(\mathbf{z}_1, \mathbf{z}_{k+1}) \neq \mathbf{0}$, we have that $\wedge(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{0}$ only if a nonzero polynomial (of degree $O(1)$) in the λ_i, γ_i vanishes. Thus, the number of choices of $\lambda_i, \gamma_i \ll B/Z_i$ such that $\wedge(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{0}$ is

$$\ll \left(1 + \frac{B}{Z_{n-k}}\right) \prod_{i=1}^{n-k-1} \left(1 + \frac{B}{Z_i}\right)^2.$$

Since $Z_{n-k} \ll B^{1-\delta/2}$ and $n > 3k$, this is

$$\ll Z_{n-k} \frac{B^{2n-2k-1}}{\prod_{i=1}^{n-k} Z_i^2} \ll \frac{B^{2n-2k-\delta/2}}{Z_1^{n-k} Z_{k+1}^{n-k}}.$$

But then by Lemma 11.5, this means the size of the set in the lemma is of size

$$\ll \sum_{\|\mathbf{a}\| \in [A, 2A]} \frac{Z_{k+1}^k}{\prod_{i=1}^k Z_i} \frac{B^{2n-2k-\delta/2}}{Z_1^{n-k} Z_{k+1}^{n-k}} \ll A^{n-2k} B^{2n-2k-\delta/3}. \qquad \square$$

11.2. Non-Archimedean estimates. We now consider $\mathbf{b}_1, \mathbf{b}_2$ for which the determinant of $\Lambda_{\mathbf{b}_1, \mathbf{b}_2}$ is small because $D_{\mathbf{b}_1, \mathbf{b}_2}$ is large. We first establish a couple of lemmas bounding the number of times a given polynomial $f \in \mathbb{Z}[X]$ can vanish $(\bmod D)$. The key point of these lemmas is that there is only a very weak dependence on the size of the coefficients of f.

LEMMA 11.8. *Let $\epsilon > 0$. Let $\mathbf{f} = (f_1, \dots, f_\ell) \in \mathbb{Z}[x_1, \dots, x_n]^\ell$ be a vector of $\ell \geqslant 2$ homogeneous polynomials of degree d with coefficients of size at most $F \geqslant 2$ in absolute value and no nonconstant common factor in $\mathbb{Z}[X_1, \dots, X_n]$ amongst all of them. For each prime p, let $e_p \in \mathbb{N}$ be such that not all the f_i take only the value $0 \pmod{p^{e_p}}$ on $(\mathbb{Z}/p^{e_p}\mathbb{Z})^n$ but that all the f_i only take the value $0 \pmod{p^{e_p - 1}}$ on $(\mathbb{Z}/p^{e_p-1}\mathbb{Z})^n$. Let $E = \prod_{p: e_p > 1} p^{e_p}$.*

Then for any reals $D_0 \geqslant 1$ and $1 \leqslant X_{\min} \leqslant X_1, \dots, X_n \leqslant X_{\max}$, we have

$$\#\{(\mathbf{x}, D) \in \mathbb{Z}^n \times \mathbb{Z},\ |x_i| \leqslant X_i,\ D > D_0,\ \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \pmod{D},\ \mathbf{f}(\mathbf{x}) \neq \mathbf{0}\}$$
$$\ll \left(\frac{1}{D_0^{1/2d}} + \frac{1}{X_{\min}}\right) (D_0 F X_{\max})^\epsilon E^n \prod_{i=1}^n X_i.$$

Here the implied constant depends only on ℓ, n, d, ϵ.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Proof. For this proof, we let all implied constants depend on ℓ, n, d and ϵ. Without loss of generality, we assume that $X_{\max} = X_1 \geqslant \cdots \geqslant X_n = X_{\min}$. We first want to show the existence of short vectors $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^n$ such that $\mathbf{f}(\mathbf{u})$ and $\mathbf{f}(\mathbf{v})$ have a small common divisor.

We choose $\mathbf{u} \in \mathbb{Z}^n$ such that $\|\mathbf{u}\| \ll E$, $u_n \neq 0$ and $\mathbf{f}(\mathbf{u}) \not\equiv \mathbf{0} \pmod{p^{e_p}}$ for any $p \leqslant d$ or any $p|E$. This is possible since $\mathbf{f}$ does not vanish on $(\mathbb{Z}/p^{e_p}\mathbb{Z})^n$.

We now choose $\mathbf{v}$ such that any integer dividing all components of $\mathbf{f}(\mathbf{u})$ and $\mathbf{f}(\mathbf{v})$ must divide E. For any prime p with $e_p = 1$, the fact that $\mathbf{f}$ does not vanish on $\mathbb{F}_p^n$ means that there is a polynomial $f_{p,1} \in \{f_1, \ldots, f_\ell\}$ such that $f_{p,1}$ has a nonzero coefficient over $\mathbb{F}_p$. Viewing $f_{p,1}(\mathbf{x})$ as a polynomial in x_1 and selecting a nonzero coefficient, we find a nonzero polynomial $f_{p,2}$ in $x_2, \ldots, x_n$ such that $f_{p,1}$ is a nonzero polynomial in x_1 if $f_{p,2} \not\equiv 0 \pmod{p}$. Repeating this, we obtain (possible constant) polynomials $f_{p,2}, \ldots, f_{p,n}$ with $f_{p,j}$ a nonzero polynomial in $x_j, \ldots x_n$, and $f_{p,j}$ is a nonzero polynomial in x_j if $f_{p,j+1} \not\equiv 0 \pmod{p}$. We then choose nonzero integers $v_n, \ldots, v_1$ in turn as small as possible such that $f_{p,j}(v_j, \ldots, v_n) \not\equiv 0 \pmod{p}$ for all $j \in \{1, \ldots, n\}$ and for any prime $p > d$ which divides all components of $\mathbf{f}(\mathbf{u})$ and has $e_p = 1$. This is possible since any nonzero polynomial of degree at most d can vanish at most d points over $\mathbb{F}_p$, and we only consider $p > d$.

Since $\mathbf{f}$ has coefficients of size $O(F)$ and $\|\mathbf{u}\| \ll E$, we have $\|\mathbf{f}(\mathbf{u})\| \ll (FE)^{O(1)}$. Thus, there are $O(\log FE)$ primes p which divide all components of $\mathbf{f}(\mathbf{u})$, and these must all satisfy $p > d$ if $e_p = 1$. Each of the polynomials can have at most d roots modulo any prime p under consideration. Therefore, each v_j is the least integer which avoids one of $O(1)$ residue classes mod p for $O(\log FE)$ different primes p. By the fundamental lemma of sieve methods, we have that $v_j \ll (\log FE)^{O(1)}$.

Thus, we have found $\mathbf{u}, \mathbf{v} \ll E(\log F)^{O(1)}$ such that any integer dividing all components of $\mathbf{f}(\mathbf{u})$ and $\mathbf{f}(\mathbf{v})$ must divide E. In particular, for any integer D, we have either $D/\gcd(D, \mathbf{f}(\mathbf{u})) > (D/E)^{1/2}$ or $D/\gcd(D, \mathbf{f}(\mathbf{v})) > (D/E)^{1/2}$. Thus, without loss of generality, it is sufficient to count pairs $(\mathbf{x}, D)$ as in the lemma with the extra condition that $D/\gcd(D, \mathbf{f}(\mathbf{w})) \geqslant (D/E)^{1/2}$ where $\mathbf{w} \in \mathbb{Z}^n$ is a fixed vector with $\|\mathbf{w}\| \ll E(\log F)^{O(1)}$, $w_n \neq 0$ and $\mathbf{f}(\mathbf{w}) \neq \mathbf{0}$. By replacing f_j with a suitable integral linear combination of the f_i we may moreover assume that $f_j(\mathbf{w})$ is the same for all j.

We now change variables. Since $w_n \neq 0$, $|w_i| \ll E(\log F)^{O(1)}$ and $X_1 \geqslant \cdots \geqslant X_n$, we can write any vector $\mathbf{x} \in \mathbb{Z}^n$ with $|x_i| \leqslant X_i$ as

$$w_n \mathbf{x} = \sum_{i=1}^{n-1} y_i \mathbf{e}_i + y_n \mathbf{w},$$

with $|y_i| \ll Y_i = X_i E(\log F)^{O(1)}$ for $i < n$ and $y_n \ll Y_n = X_n$, where $\mathbf{e}_i$ are

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

the standard basis vectors of $\mathbb{Z}^n$. Since the polynomials f_i are homogeneous, we have $\mathbf{f}(w_n\mathbf{x}) = w_n^d\mathbf{f}(\mathbf{x})$. Thus, it is sufficient to count pairs $(\mathbf{y}, D)$ with $D > D_0$, $D/\gcd(D, \mathbf{f}(\mathbf{w})) > (D/E)^{1/2}$, $|y_i| \ll Y_i$ and $\tilde{\mathbf{f}}(\mathbf{y}) \equiv \mathbf{0} \pmod D$ but $\tilde{\mathbf{f}}(\mathbf{y}) \neq \mathbf{0}$, where $\tilde{\mathbf{f}}(\mathbf{y}) = \mathbf{f}(\sum_{i=1}^{n-1} y_i\mathbf{e}_i + y_n\mathbf{w})$.

By the Euclidean algorithm (or calculating a suitable resultant), $D|\tilde{\mathbf{f}}(\mathbf{y})$ only if $D|g(y_1, \dots, y_{n-1})$ for some nonzero polynomial g independent of y_n and of degree at most d^2 and with coefficients of size at most $F^{O(1)}$ since the components of $\tilde{\mathbf{f}}$ have no nonconstant polynomial common factor. We consider separately the cases when $g(y_1, \dots, y_{n-1}) = 0$ and when it is nonzero.

There are $\ll \prod_{i=1}^{n-2} Y_i \ll X_{\min}^{-1}E^{n-2}(\log F)^{O(1)}\prod_{i=1}^{n-1} X_i$ choices of $y_1, \dots, y_{n-1}$ such that $g(y_1, \dots, y_{n-1}) = 0$. For any such choice, there are $\ll Y_n = X_n$ choices of y_n such that $\tilde{\mathbf{f}}(\mathbf{y}) \neq \mathbf{0}$ and then $O((X_{\max}F)^{\epsilon})$ choices of $D|\tilde{\mathbf{f}}(\mathbf{y})$. This gives the result in the case $g(y_1, \dots, y_{n-1}) = 0$.

There are $\ll \prod_{i=1}^{n-1} Y_i = E^{n-1}(\log F)^{O(1)}\prod_{i=1}^{n-1} X_i$ choices of $y_1, \dots, y_{n-1}$ such that $g(y_1, \dots, y_{n-1}) \neq 0$. Given such a choice, there are then $O((FX_{\max})^{\epsilon})$ choices of $D|g(y_1, \dots, y_{n-1})$. We now wish to count the number of choices of y_n such that $\tilde{\mathbf{f}}(\mathbf{y}) \equiv 0 \pmod D$. We recall that $f_j(\mathbf{w})$ is the same nonzero integer for all j, so f_1 is a polynomial of degree d in y_n with lead coefficient $f_1(\mathbf{w})$. Moreover, we only consider D with $D/\gcd(D, \mathbf{f}(\mathbf{w})) = D/\gcd(D, f_1(\mathbf{w})) > (D/E)^{1/2}$. In this case, using Lemma 11.2, we find that the number of choices of y_n such that $f_1(y_1, \dots, y_n) \equiv 0 \pmod D$ is

$$\ll \left(1 + X_n\left(\frac{D}{\gcd(D, f_1(\mathbf{w}))}\right)^{-1/d}\right)D^{\epsilon} \ll D_0^{\epsilon}E^{1/2d}X_n\left(\frac{1}{X_{\min}} + \frac{1}{D_0^{1/2d}}\right).$$

This gives the result. □

LEMMA 11.9. *We have*

$$\#\{\mathbf{b} \in \mathbb{F}_p^n : \wedge(\mathbf{b}) = \mathbf{0}\} \ll p^{k-1}.$$

Proof. We may assume that p is sufficiently large, so $\theta \not\equiv 0 \pmod p$ and $p > n$. We recall that if $\wedge(\mathbf{b}) = \mathbf{0} \in \mathbb{F}_p$, then there exist constants $c_0, \dots, c_{k-1}$ not all 0 such that

$$\sum_{i=0}^{k-1} c_i T^i(\mathbf{b}) = \mathbf{0}.$$

We argue in the case when this is the shortest linear relation of this type (so, in particular, $c_0, c_{k-1} \neq 0$); the other cases are entirely analogous. By inverting c_0, we have $\mathbf{b} = T^0(\mathbf{b}) = \sum_{i=1}^{k-1} c_i' T^i(\mathbf{b})$ for constants c_i' with $c_{k-1}' \neq 0$. Thus, letting $b_{n+j} = b_j/\theta$, we have $b_j = \sum_{i=1}^{k-1} c_i' b_{j-i}$ for all $j \in \mathbb{Z}$. Moreover, we may assume

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

that $\mathbf{b}$ does not satisfy any other recurrence equation of this type because in that case, we may take a linear combination and have $c'_{k-1} = 0$. This is a difference equation, and so $b_j = \sum_{i=1}^{k-1} P_i(j)\lambda_i^j$ for some polynomials $P_1, \dots, P_\ell$ with total degree at most $k-1$ and constants λ_i in a finite extension of $\mathbb{F}_p$. Moreover, the monomials $j^{m_1}\lambda_{m_2}^j$ uniquely determine $c'_1, \dots, c'_{k-1}$ as the coefficients of the monic polynomial $X^{k-1} - \sum_{i=1}^{k-1} c'_i X^{k-i-1} \in \mathbb{F}_p[X]$ of least degree which has λ_i as a root with multiplicity at least $\deg P_i$.

But then $\sum_i P_i(n+j)\lambda_i^{n+j} = b_{n+j} = b_j/\theta = \theta^{-1}\sum_i P_i(j)\lambda_i^j$ for all j. This gives a fixed linear combination of the monomials $j^{m_1}\lambda_{m_2}^j$ which vanishes for all j, and so as in Lemma 10.4, the coefficients of all monomials must be zero. Thus, on comparing the coefficient of $j^\ell\lambda_i^j$ and letting $p_{\ell,i}$ be the coefficient of x^ℓ in $P_i(x)$, we have $p_{\ell,i} = \theta\lambda_i^n \sum_{m\geqslant \ell} p_{m,i} n^{m-\ell}\binom{m}{\ell}$. By considering the coefficients in turn from the highest degree coefficients to the lowest degree, we see that either $P_i(x) = 0$ or $\lambda_i^n = \theta^{-1}$ and $p_{m,i} = 0$ for all $m \geqslant 2$.

Thus, we have $b_j = \sum_i p_{i,1}\lambda_i^j$ where for each i, we have $\lambda_i^n = \theta^{-1}$. But then there are $O(1)$ possibilities for the monomials appearing in b_j, and so $O(1)$ possible choices for the coefficients $c'_1, \dots, c'_{k-1}$. Since $\mathbf{b}$ is uniquely determined by $c'_1, \dots, c'_{k-1}$ and $b_1, \dots, b_{k-1}$, there are $O(p^{k-1})$ different possible choices of $\mathbf{b}$. □

REMARK. We expect the bound of Lemma 11.9 to be sharp for infinitely many p since it involves n equations in $n+k-1$ variables.

LEMMA 11.10. *Let $n > 3k$ and $\mathbf{a} \in \mathbb{Z}^n\backslash\{\mathbf{0}\}$ and p a prime. Then there exists $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}}$ such that $\wedge(\mathbf{b}_1, \mathbf{b}_2) \neq \mathbf{0} \pmod p$.*

Proof. Let $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ be a basis of $\Lambda_{\mathbf{a}}$. From the definition of $\Lambda_{\mathbf{a}}$, any $\mathbf{x} \in \mathbb{Z}^n$ which is in the $\mathbb{Q}$-span of $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ is in $\Lambda_{\mathbf{a}}$ and so must actually be in the $\mathbb{Z}$-span. Therefore, for any prime p, $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ are linearly independent $\pmod p$. After rearranging the coordinates, this means that the $(n-k)\times(n-k)$ matrix formed by taking the first $n-k$ components of $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ has nonzero determinant $\pmod p$, and so is invertible. But this means that given integers $b_1, \dots b_{n-k}$, there exists $\mathbf{x} \in \Lambda_{\mathbf{a}}$ such that $x_j \equiv b_j \pmod p$. In particular, there exists $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \Lambda_{\mathbf{a}}$ such that

$$\mathbf{x}_j^{(1)} \equiv \begin{cases} 1 \pmod p, & j = k, \\ 0 \pmod p, & 1 \leqslant j < 2k \text{ or } 2k < j \leqslant n-k, \end{cases}$$

$$\mathbf{x}_j^{(2)} \equiv \begin{cases} 1 \pmod p, & j = k, \\ 0 \pmod p, & 1 \leqslant j < 2k \text{ or } 2k < j \leqslant n-k. \end{cases}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We now consider the component of $\wedge(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ which is the determinant of the $2k \times 2k$ matrix formed by taking first $2k$ components of $\mathbf{x}^{(1)}, \dots, T^{k-1}(\mathbf{x}^{(1)})$, and $\mathbf{x}^{(2)}, \dots, T^{k-1}(\mathbf{x}^{(2)})$. We see that this matrix is a lower triangular with 1's on the diagonal and so has determinant 1. Therefore, $\wedge(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \neq \mathbf{0} \pmod p$, as required. □

LEMMA 11.11. *Let* $\mathbf{f} = (f_1, \dots, f_\ell) \in \mathbb{Z}[x_1, \dots, x_n]^\ell$ *be such that*

$$\#\{(a_1, \dots, a_n) \in [1, p]^n : \mathbf{f}(\mathbf{a}) \equiv \mathbf{0} \pmod p\} \ll p^{n-2}$$

for all primes p. *Then* $\mathbf{f}$ *has no nonconstant common factor.*

Proof. Imagine for a contradiction that there is a nonconstant polynomial $g \in \mathbb{Z}[x_1, \dots x_n]$ dividing all the f_i. Then there is a nonconstant polynomial g_1 dividing g defined over a finite extension of $\mathbb{Q}$ which is absolutely (that is, geometrically) irreducible ($g_1 = g$ if g is absolutely irreducible). By the Chebotarev density theorem, there are infinitely many primes p such that $g \pmod p$ has a factor $\overline{g}_1$ corresponding to g_1 which is defined over $\mathbb{F}_p$. It follows from the Hilbert Nullstellensatz (see, for example, [**18**, Proposition 7, page 157]) that $\overline{g}_1$ is absolutely irreducible over $\mathbb{F}_p$ for all but finitely many primes p. But the Lang–Weil bound implies that there are $(1 + o(1))p^{n-1}$ values $\mathbf{a} \in \mathbb{F}_p^n$ such that $\overline{g}_1(\mathbf{a}) = 0$ for any prime p for which $\overline{g}_1$ is defined over $\mathbb{F}_p$ and is absolutely irreducible over $\mathbb{F}_p$. In particular, there are $\gg p^{n-1}$ zeros of g over $\mathbb{F}_p$ for infinitely many primes p. This contradicts the assumption of the lemma, and so no such nonconstant polynomial g can exist. □

LEMMA 11.12 (Determinant rarely small for non-Archimedean reasons). *Let* $n > 3k$, $\delta > 0$ *and* $A^{k/(1-\delta)} < B^{n-k}$. *Then we have for any constant* $C > 0$

$$\#\{(\mathbf{a}, \mathbf{b}_1, \mathbf{b}_2) \in \mathcal{S}(A; B, B) : D_{\mathbf{b}_1,\mathbf{b}_2} > \epsilon_0^{-C}, \mathbf{b}_1 \equiv \mathbf{b}_2 \equiv \mathbf{b}_0 \pmod{\tilde{q}}\} \ll_C \epsilon_0^{C/20k} A^{n-2k} B^{2n-2k}.$$

Proof. By Lemma 11.6, we can restrict our attention to $\mathbf{a}$ such that $\Lambda_{\mathbf{a}}$ has all successive minima $Z_1, \dots, Z_{n-k} \ll B^{1-\delta/2}$, and by Lemma 11.7, to $\mathbf{b}_1, \mathbf{b}_2$ with $\wedge(\mathbf{b}_1, \mathbf{b}_2) \neq \mathbf{0}$. By Lemma 11.5 (since $\epsilon_0^{-C/20k} \gg \exp(G'(\log\log A)^2)$), it suffices to show for each such $\mathbf{a}$ that

$$1 \ll \frac{\epsilon_0^{C/10k} B^{2n-2k}}{\prod_{i=1}^{n-k} Z_i^2}. \tag{11.11}$$

$$\sum_{D > \epsilon_0^{-C}} \sum_{\substack{\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}} \\ \|\mathbf{b}_1\|, \|\mathbf{b}_2\| \in [B, 2B] \\ D \mid \wedge(\mathbf{b}, \mathbf{c}) \neq \mathbf{0}}}$$

We split our argument into different cases, depending on whether $D \leqslant B^{\delta/2}$ or

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$D > B^{\delta/2}$. We first consider $D \leqslant B^{\delta/2}$. We recall that $\wedge(\mathbf{b}_1, \mathbf{b}_2)$ is a vector of homogeneous integer polynomials in the coefficients of $\mathbf{b}_1, \mathbf{b}_2$ with coefficients of size $O(1)$ and degree at most $2k$. If $\wedge(\mathbf{b}_1, \mathbf{b}_2) \equiv \mathbf{0} \pmod p$, then there exists constants $c_0, \dots, c_{k-1}, d_0, \dots, d_{k-1} \in \mathbb{Z}$ at least one of which is 1 such that

$$\sum_{i=0}^{k-1} c_i T^i(\mathbf{b}_1) \equiv \sum_{i=0}^{k-1} d_i T^i(\mathbf{b}_2) \pmod p.$$

By symmetry, we may assume that one of the c_i is equal to 1. Given a choice of $c_0, \dots, c_{k-1}, d_0, \dots, d_{k-1}$ and $\mathbf{b}_2$, we see that we are counting solutions $\mathbf{b}_1 \in \Lambda_{\mathbf{a}}$ to a linear equation $M\mathbf{b}_1 \equiv \mathbf{v} \pmod p$ for some given $\mathbf{v} \in \mathbb{F}_p^n$ depending on $\mathbf{b}_2$ and $d_0, \dots, d_{k-1}$ and some given matrix M depending on $c_0, \dots, c_{k-1}$. The number of such solutions in $\mathbb{F}_p^n$ is at most the number of solutions of $M\mathbf{b}_1 \equiv \mathbf{0} \pmod p$ by linearity (it is the same if $\mathbf{v}$ is in the image of M). But the number of choices of $\mathbf{b}_1, c_0, \dots, c_{k-1}$ with one of the c_i equal to 1 and $M\mathbf{b}_1 \equiv \mathbf{0} \pmod p$ is the number of $\mathbf{b}_1 \in \mathbb{F}_p^n$ such that $\wedge(\mathbf{b}_1) \equiv \mathbf{0} \pmod p$. Thus, by Lemma 11.9, there are $O(p^{k-1})$ choices of $\mathbf{b}_1 \pmod p$ and $c_0, \dots, c_{k-1}$, given a choice of $\mathbf{b}_2$ and $d_0, \dots, d_{k-1}$.

Let $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ be a basis of $\Lambda_{\mathbf{a}}$ and $\overline{\Lambda_{\mathbf{a}}}$ be the reduction of $\Lambda_{\mathbf{a}} \pmod p$. Since the integer vectors in the $\mathbb{Q}$-span of $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ are in $\Lambda_{\mathbf{a}}$ from the definition of $\Lambda_{\mathbf{a}}$, they must, in fact, lie in the $\mathbb{Z}$-span of $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$. Thus, any basis $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ is linearly independent $\pmod p$, and so $\overline{\Lambda_{\mathbf{a}}}$ contains p^{n-k} points. Thus, there are $O(p^k)$ choices of $d_0, \dots, d_{k-1} \pmod p$ and $O(p^{n-k})$ choices of $\mathbf{b}_2 \in \overline{\Lambda_{\mathbf{a}}}$. Hence, in total, there are $O(p^{n+k-1}) \ll p^{2n-2k-2}$ choices of $\mathbf{b}_1, \mathbf{b}_2 \in \overline{\Lambda_{\mathbf{a}}}$ such that $\wedge(\mathbf{b}_1, \mathbf{b}_2) \equiv \mathbf{0} \pmod p$. But there are p^{2n-2k} choices of $\mathbf{b}_1, \mathbf{b}_2 \in \overline{\Lambda_{\mathbf{a}}}$. Thus, by Lemma 11.11, $\wedge(\sum_{i=1}^{n-k} a_i \mathbf{z}_i, \sum_{i=1}^{k} b_i \mathbf{z}_i)$ is a vector of polynomials in $\mathbf{a}, \mathbf{b}$ with no nonconstant common factor, and $\wedge(\mathbf{b}_1, \mathbf{b}_2)$ does not vanish on $\overline{\Lambda_{\mathbf{a}}}$ for p sufficiently large. If p is bounded by a constant, then by Lemma 11.10, we also have that $\wedge(\mathbf{b}_1, \mathbf{b}_2)$ does not vanish on $\overline{\Lambda_{\mathbf{a}}}$.

Let $D = \prod_{i=1}^{\ell} p_i^{e_i} = D_1 D_2$ with $D_1 = \prod_{i=1}^{\ell} p_i$, $D_2 = \prod_{i=1}^{\ell} p^{e_i - 1}$ be factorized into square-free and remaining parts. By the above discussion, there are $O(p_i^{2n-2k-2})$ choices of $\mathbf{b}_1, \mathbf{b}_2 \pmod{p_i}$ with $\wedge(\mathbf{b}_1, \mathbf{b}_2) \equiv \mathbf{0} \pmod{p_i}$ and $\mathbf{b}_1, \mathbf{b}_2 \in \overline{\Lambda_{\mathbf{a}}}$, and so certainly $O(p_i^{e_i(2n-2k)-2})$ choices $\pmod{p_i^{e_i}}$. Alternatively, by Lemma 11.2, there are $O(p_i^{e_i(2n-2k)-\lceil e_i/2k \rceil + o(e_i)})$ choices of $\mathbf{b}_1, \mathbf{b}_2 \pmod{p_i^{e_i}}$. (After a change of variable, one can assume that a homogeneous polynomial of degree d has a monomial cx_1^d, and so Lemma 11.2 applies for each choice of $x_2, \dots, x_n$.) Thus, by the Chinese remainder theorem, the total number of choices of possible residue classes for $\mathbf{b}_1, \mathbf{b}_2 \pmod D$ is

$$\ll \frac{D^{2n-2k}}{\prod_i p_i^{\max(2, \lceil e_i/2k \rceil + o(e_i))}} \ll \frac{D^{2n-2k}}{\prod_i (p_i^2)^{3/4} (p_i^{e_i/2k - 1 + o(e_i)})^{1/4}} \ll \frac{D^{2n-2k-1/10k}}{D_1^{1+1/50k} D_2^{1/50k+o(1)}}.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Since we are considering $D \leqslant B^{\delta/2} < B/Z_{n-k}$, the number of choices of $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}}$ with $\|\mathbf{b}_1\|, \|\mathbf{b}_2\|$ in any given residue class $\pmod D$ is $O(B^{2n-2k} D^{-(2n-2k)} / \prod_{i=1}^{n-k} Z_i^2)$. Thus, the total contribution from $\epsilon_0^{-C} < D < B^{\delta/2}$ is

$$\begin{aligned}
&\ll \sum_{\substack{D > \epsilon_0^{-C} \\ p \mid D_2 \Rightarrow p \mid D_1}} \frac{D^{2n-2k-1/10k}}{D_1^{1+1/50k} D_2^{1/50k+o(1)}} \cdot \frac{B^{2n-2k}}{D^{2n-2k} \prod_{i=1}^{n-k} Z_i^2} \\
&\ll \frac{\epsilon_0^{C/10k} B^{2n-2k}}{\prod_{i=1}^{n-k} Z_i^2} \sum_{\substack{D_1, D_2 \geqslant 1 \\ p \mid D_2 \Rightarrow p \mid D_1}} \frac{1}{D_1^{1+1/50k} D_2^{1/50k+o(1)}} \\
&\ll \frac{\epsilon_0^{C/10k} B^{2n-2k}}{\prod_{i=1}^{n-k} Z_i^2} \sum_{D_1 \geqslant 1} \frac{\tau(D_1)}{D_1^{1+1/50k}} \\
&\ll \frac{\epsilon_0^{C/10k} B^{2n-2k}}{\prod_{i=1}^{n-k} Z_i^2}. \qquad (11.12)
\end{aligned}$$

This is sufficient to give (11.11) when $D \leqslant B^{\delta/2}$.

Thus, we are left to consider the contributions when $D > B^{\delta/2}$. Let $\mathbf{z}_1, \dots, \mathbf{z}_{n-k}$ be a basis for $\Lambda_{\mathbf{a}}$ so that $\mathbf{b}_1 = \sum_{i=1}^{n-k} \lambda_i \mathbf{z}_i$, $\mathbf{b}_2 = \sum_{i=1}^{n-k} \gamma_i \mathbf{z}_i$ for some integers $\lambda_i, \gamma_i \ll B/Z_i$. From our above discussion, $\wedge(\sum_{i=1}^{n-k} \lambda_i \mathbf{z}_i, \sum_{i=1}^{n-k} \gamma_i \mathbf{z}_i)$ is a vector of homogeneous polynomials of degree $2k$ in $\lambda_1, \dots, \lambda_{n-k}, \gamma_1, \dots, \gamma_{n-k}$, with coefficients of size $O(B)$ and which does not vanish identically $\pmod p$ for p sufficiently large or $\pmod{p^J}$ for some fixed J for all other primes. Therefore, by Lemma 11.8, the number of triples $(\mathbf{b}_1, \mathbf{b}_2, D)$ with $D > B^{\delta/2}$ such that $\wedge(\mathbf{b}_1, \mathbf{b}_2) \equiv \mathbf{0} \pmod D$ but $\wedge(\mathbf{b}_1, \mathbf{b}_2) \neq \mathbf{0}$ is

$$\ll B^{-\delta/8k} \prod_{i=1}^{n-k} \frac{B^2}{Z_i^2}. \qquad (11.13)$$

Recalling that $\epsilon_0 = \tilde{q}^{-4n} \exp(-\sqrt[7]{\log X}) \geqslant \exp(-\sqrt[3]{\log X})$ and $B \gg X^{\delta}$, we see that (11.13) gives (11.11) in the remaining range $D > B^{\delta/2}$. □

11.3. Separation of variables and proof of Proposition 11.1. Finally, we are in a position to prove Proposition 11.1. We assume that $n > 3k$.

Proof of Proposition 11.1. We recall that we wish to show

$$\sum_{\substack{\|\mathbf{b}_1\|, \|\mathbf{b}_2\| \in [B, 2B] \\ \mathbf{b}_1, \mathbf{b}_2 \equiv \mathbf{b}_0 \pmod{\tilde{q}}}} g_{\mathbf{b}_1} \overline{g_{\mathbf{b}_2}} \sum_{\mathbf{a} \in \Lambda_{\mathbf{b}_1, \mathbf{b}_2} \cap \mathcal{R}_{\mathbf{b}_1, \mathbf{b}_2}} 1 \ll \epsilon_0 A^{n-2k} B^{2n-2k}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

for any choice of $\mathbf{b}_0$, where A, B satisfy $X \ll AB \ll X$ and $X^{k+\epsilon/2} \ll B \ll X^{n-2k-\epsilon/2}$. For δ sufficiently small in terms of ϵ, we see that this implies that $B^{2(1+\delta)k/(n-2k)} < A < B^{(1-\delta)(n-k)/k}$.

Combining Lemmas 11.4, 11.7 and 11.12 and recalling that $g_{\mathbf{b}} \ll \epsilon_0^{-2}$, we have

$$\sum_{\substack{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|\in[B,2B]\\ \|\wedge(\mathbf{b}_1,\mathbf{b}_2)\|\leqslant\epsilon_0^{8k/\delta}B^{2k} \text{ or } D_{\mathbf{b}_1,\mathbf{b}_2}>\epsilon_0^{-24k}}} |g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}}| \sum_{\mathbf{a}\in\Lambda_{\mathbf{b}_1,\mathbf{b}_2}\cap\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}} 1 \ll \epsilon_0^2 A^{n-2k}B^{2n-2k}.$$

Thus, we may restrict our attention to $\mathbf{b}_1, \mathbf{b}_2$ such that $\|\wedge(\mathbf{b}_1,\mathbf{b}_2)\| \geqslant \epsilon_0^{8k/\delta}B^{2k}$ and $D_{\mathbf{b}_1,\mathbf{b}_2} \leqslant \epsilon_0^{-30k}$.

We first deal with the $D_{\mathbf{b}_1,\mathbf{b}_2}$ factor. We note that

$$\sum_{\substack{D<\epsilon_0^{-30k}\\ D\mid D_{\mathbf{b}_1,\mathbf{b}_2}}}\sum_{\substack{d<\epsilon_0^{-60k}\\ d\mid D_{\mathbf{b}_1,\mathbf{b}_2}/D}}\mu(d) = \begin{cases} 1, & D_{\mathbf{b}_1,\mathbf{b}_2} < \epsilon_0^{-30k},\\ 0, & \epsilon_0^{-30k} \leqslant D_{\mathbf{b}_1,\mathbf{b}_2} < \epsilon_0^{-60k},\\ O(\tau(D_{\mathbf{b}_1,\mathbf{b}_2})^2) = \epsilon_0^{-o(1)}, & \epsilon_0^{-60k} \leqslant D_{\mathbf{b}_1,\mathbf{b}_2} < \epsilon_0^{-(100k)^2},\\ O(\epsilon_0^{-90k}), & \epsilon_0^{-(100k)^2} \leqslant D_{\mathbf{b}_1,\mathbf{b}_2}. \end{cases}$$

Thus, using Lemma 11.12, we see that we may replace the condition $D_{\mathbf{b}_1,\mathbf{b}_2} < \epsilon_0^{-30k}$ by the double sum on the left-hand side at the cost of a negligible error term coming from when $D_{\mathbf{b}_1,\mathbf{b}_2} > \epsilon_0^{-60k}$.

We are left with

$$\sum_{d<\epsilon_0^{-60k},\, D<\epsilon_0^{-30k}}\mu(d)\sideset{}{^*}\sum_{\substack{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|\in[B,2B]\\ \mathbf{b}_1\equiv\mathbf{b}_2\equiv\mathbf{b}_0 \pmod{\tilde{q}}\\ dD\mid D_{\mathbf{b}_1,\mathbf{b}_2}}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}} \sum_{\mathbf{a}\in\Lambda_{\mathbf{b}_1,\mathbf{b}_2}\cap\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}} 1,$$

where $\sum^*$ indicates that we have the condition that $\|\wedge(\mathbf{b}_1,\mathbf{b}_2)\| \geqslant \epsilon_0^{8k/\delta}B^{2k}$.

Splitting the sum over $\mathbf{b}_1, \mathbf{b}_2$ into residue classes modulo $D_1 = \operatorname{lcm}(dD, \tilde{q})$ and recalling $\tilde{q} = (\theta n)^n q^* N(\mathfrak{c}) \leqslant \epsilon_0^{-1} = \tilde{q}^{4n}\exp(\sqrt[7]{\log X})$, it suffices to show that

$$\sup_{\substack{D_1\ll\epsilon_0^{-100k}\\ \mathbf{d}_1,\mathbf{d}_2\in\mathbb{Z}^n}}\sideset{}{^*}\sum_{\substack{\|\mathbf{b}_1\|,\|\mathbf{b}_2\|\in[B,2B]\\ (\mathbf{b}_1,\mathbf{b}_2)\equiv(\mathbf{d}_1,\mathbf{d}_2) \pmod{D_1}}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}} \sum_{\mathbf{a}\in\Lambda_{\mathbf{b}_1,\mathbf{b}_2}\cap\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}} 1 \ll \epsilon_0^{400k^2}A^{n-2k}B^{2n-2k}.$$

By Lemmas 7.1 and 10.1, we have that the inner sum is

$$\frac{D_{\mathbf{b}_1,\mathbf{b}_2}\operatorname{vol}\mathcal{R}_A}{\|\wedge(\mathbf{b}_1,\mathbf{b}_2)\|} + O\left(1+\frac{A^{n-2k-1}}{V^{n-2k-1}}\right),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

where V is the length of the shortest vector in $\Lambda_{\mathbf{b}_1,\mathbf{b}_2}$. By Lemma 11.4, since $V^{n-2k} \ll \det \Lambda_{\mathbf{b}_1,\mathbf{b}_2} \ll B^{2k}$, this error term contributes

$$\ll B^{2n}\epsilon_0^{-4} + A^{n-2k-1}(\log B)\epsilon_0^{-4} \sup_{V^{n-2k}\ll B^{2k}} \frac{S(V;B,B)}{V^{n-2k-1}}$$
$$\ll B^{2n+o(1)} + A^{n-2k-1}B^{2n-2k+2k/(n-2k)+o(1)}.$$

Here we used the fact that $g_{\mathfrak{b}} \ll \epsilon_0^{-2}$ and that there are $O(\log B)$ choices of dyadic interval for $V \ll B^{2k/(n-2k)}$. This is $\ll A^{n-2k-\delta/2}B^{2n-2k}$ since $A \gg B^{2(1+\delta)k/(n-2k)}$ by assumption. Thus, we may restrict our attention to the main term.

We split the sum over $\mathbf{b}_1, \mathbf{b}_2$ into $O(\delta_0^{-2n})$ nonoverlapping hypercubes $\mathcal{C}_1, \dots, \mathcal{C}_\ell$ of side length $\delta_0 B$, for some suitable δ_0. There are $O(\delta_0^{-2n+1})$ hypercubes which do not have all points with norm either in $[B_0, 2B_0]$ or outside of this interval. Thus, on choosing

$$\delta_0 = \epsilon_0^{2000k^2/\delta}, \tag{11.14}$$

we see that these contribute a negligible amount. Thus, we are left to show

$$\sum_{1\leqslant i,j\leqslant r} \sum^{*}_{\substack{(\mathbf{b}_1,\mathbf{b}_2)\in\mathcal{C}_i\times\mathcal{C}_j \\ (\mathbf{b}_1,\mathbf{b}_2)\equiv(\mathbf{d}_1,\mathbf{d}_2) \pmod D}} \frac{g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}}\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}}{\| \wedge (\mathbf{b}_1,\mathbf{b}_2)\|} \ll \delta_0^{1/3}A^{n-2k}B^{2n-2k},$$

where $(\mathcal{C}_i)_{1\leqslant i\leqslant r}$ are the $O(\delta_0^{-n})$ hypercubes with all points in $\mathcal{C}_i$ having norm in $[B, 2B]$ (since $g_{\mathfrak{b}} = 0$ if $N(\mathfrak{b}) \notin [B, 2B]$).

Since the hypercubes have side length $\delta_0 B$ and $\| \wedge (\mathbf{b}_1, \mathbf{b}_2)\|$ is a continuous function in the components of $\mathbf{b}_1, \mathbf{b}_2$, with the derivative with respect to any component $O(B^{2k-1})$, we have that $\| \wedge (\mathbf{b}_1, \mathbf{b}_2)\|$ is almost constant on $\mathcal{C}_i \times \mathcal{C}_j$. Specifically, if $\| \wedge (\mathbf{b}_1', \mathbf{b}_2')\| \geqslant \delta_0^{1/2}B^{2k}$ for some $\mathbf{b}_1' \in \mathcal{C}_i$ and $\mathbf{b}_2' \in \mathcal{C}_j$, then $\| \wedge (\mathbf{b}_1, \mathbf{b}_2)\| = \| \wedge (\mathbf{b}_1', \mathbf{b}_2')\|(1 + O(\delta_0^{1/2}))$ for any $\mathbf{b}_1 \in \mathcal{C}_i$, $\mathbf{b}_2 \in \mathcal{C}_j$. Let $\mathbf{c}_i$ be the vector in the centre of $\mathcal{C}_i$. We now extend the sum to all pairs $(\mathbf{b}_1, \mathbf{b}_2) \in \mathcal{C}_i \times \mathcal{C}_j$ for which $\| \wedge (\mathbf{c}_i, \mathbf{c}_j)\| \geqslant \epsilon_0^{-8k/\delta}B^{2k}/2$. These additional terms can be shown to be negligible in an identical way to how we removed them originally. We are left to bound

$$\sum_{1\leqslant i,j\leqslant r} \frac{1}{\epsilon_0^{8k/\delta}B^{2k}} \left| \sum_{\substack{(\mathbf{b}_1,\mathbf{b}_2)\in\mathcal{C}_i\times\mathcal{C}_j \\ (\mathbf{b}_1,\mathbf{b}_2)\equiv(\mathbf{d}_1,\mathbf{d}_2) \pmod D}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}}\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2} \right|.$$

Similarly, $\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2}$ is the volume of a region whose dependence on $\mathbf{b}_1, \mathbf{b}_2$ is through constraints which are linear in the coefficients, and so $\operatorname{vol}\mathcal{R}_{\mathbf{b}_1,\mathbf{b}_2} \ll A^{n-2k}$ can vary by at most $O(\delta_0 A^{n-2k})$ on $\mathcal{C}_i \times \mathcal{C}_j$. This error contributes $O(\delta_0\epsilon_0^{-8k/\delta}A^{n-2k}B^{2n-2k})$ in total and so is negligible. Thus, we may replace

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

vol $\mathcal{R}_A$ with the volume evaluated at $\mathbf{c}_i, \mathbf{c}_j$, which we bound by $O(A^{n-2k})$. Thus, it suffices to show for any choice of $D < \delta_0^{-1/2}$ and any $i, j, \mathbf{d}_1, \mathbf{d}_2$

$$\sum_{\substack{(\mathbf{b}_1,\mathbf{b}_2)\in\mathcal{C}_i\times\mathcal{C}_j \\ (\mathbf{b}_1,\mathbf{b}_2)\equiv(\mathbf{d}_1,\mathbf{d}_2) \pmod D}} g_{\mathbf{b}_1}\overline{g_{\mathbf{b}_2}} \ll \delta_0^{2n+1/2}B^{2n}.$$

This sum factorizes as

$$\Bigg(\sum_{\substack{\mathbf{b}_1\in\mathcal{C}_i \\ \mathbf{b}_1\equiv\mathbf{d}_1 \pmod D}} g_{\mathbf{b}_1}\Bigg)\Bigg(\sum_{\substack{\mathbf{b}_2\in\mathcal{C}_j \\ \mathbf{b}_1\equiv\mathbf{d}_2 \pmod D}} \overline{g_{\mathbf{b}_2}}\Bigg).$$

We now replace $g_{\mathbf{b}}$ with the original coefficients $\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})$. As in Lemma 7.8, the error introduced by making this change is

$$\begin{aligned}
&\ll \sum_{\substack{\mathbf{b}\in\mathcal{C} \\ \tau(\mathfrak{b})>\epsilon_0^{-2}}} \tau(\mathfrak{b})\log X \ll \epsilon_0^{2000k^2/\delta}\sum_{\mathbf{b}\in\mathcal{C}}\tau(\mathfrak{b})^{1000k^2/\delta+2} \\
&\ll \delta_0 \sum_{N(\mathfrak{d})<B^{1/2}} \tau(N(\mathfrak{d}))^{O(1)} \sum_{\substack{\mathbf{d}\in(\mathbb{Z}/N(\mathfrak{d})\mathbb{Z})^n \\ \mathfrak{d}\mid\sum_{i=1}^n d_i\sqrt[n]{\theta^{i-1}}}} \sum_{\substack{\mathbf{b}\in\mathcal{C} \\ \mathbf{b}\equiv\mathbf{d} \pmod{N(\mathfrak{d})}}} 1 \\
&\ll \delta_0^{n+1}B^n \sum_{N(\mathfrak{d})<B^{1/2}} \frac{\tau(\mathfrak{d})^{O(1)}}{N(\mathfrak{d})} \\
&\ll \delta_0^{n+1}(\log B)^{O(1)}B^n.
\end{aligned}$$

Since the trivial bound for either sum is $\delta_0^n B^n \epsilon_0^{-2}$, this makes a negligible contribution. Thus, we are left to show that

$$\sum_{\substack{\mathbf{b}_1\in\mathcal{C}_i \\ \mathbf{b}_1\equiv\mathbf{d}_1 \pmod D}} (\mathbf{1}_{\mathcal{R}_2}(\mathfrak{b}) - \tilde{\mathbf{1}}_{\mathcal{R}_2}(\mathfrak{b})) \ll \delta_0^{n+1/2}B^{2n}.$$

We recall that $\delta_0 = q^{*-O(1)}\exp(-O(\sqrt[7]{\log X})) \geqslant q^{*-\log\log B}\exp(-\sqrt[6]{\log B})$ and that $D < \delta_0^{1/2} \ll q^{*\log\log B}\exp(\sqrt[6]{\log B})$. Thus, we may apply Proposition 9.7, which gives the desired result. This completes our proof of Proposition 11.1. □

Thus, we have established Theorem 1.2, and Theorem 1.1 in the case $K = \mathbb{Q}(\sqrt[n]{\theta})$.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

12. General $K = \mathbb{Q}(\omega)$

In this section, we sketch the changes in the argument required to generalize the above result to $K = \mathbb{Q}(\omega)$ for ω a root of a monic irreducible polynomial in $\mathbb{Z}[X]$ instead of $K = \mathbb{Q}(\sqrt[n]{\theta})$. Most of the arguments work with any occurrence of $\sqrt[n]{\theta^{i-1}}$ simply replaced by ω^{i-1}, but in a few places, we require some small modifications to the argument. We have used throughout the paper the fact that an element of $\mathcal{O}_K$ can be written as an element of $(\theta n)^{-n}\mathbb{Z}[\sqrt[n]{\theta}]$. For $K = \mathbb{Q}(\omega)$, we note that $\mathbb{Z}[\omega]$ is a finite index lattice in $\mathcal{O}_K$, and so $D_K^{-1}\mathbb{Z}[\omega] \subseteq \mathcal{O}_K \subseteq \mathbb{Z}[\omega]$ for a suitable constant D_K. Thus, we can simply replace $(\theta n)^n$ by D_K throughout.

We now consider the argument of Sections 8–11 which establishes the Type II estimate Proposition 6.1, where a couple of other changes are required. The argument of Section 8 is essentially unchanged as in no place did we use the explicit structure of K being of the form $\mathbb{Q}(\sqrt[n]{\theta})$.

In Section 10, we make use of the explicit multiplication rules in $\mathbb{Z}[\sqrt[n]{\theta}]$, and so we need to modify this for $\mathbb{Z}[\omega]$. We see that

$$\Big(\sum_{i=1}^{n} b_i\omega^{i-1}\Big)\Big(\sum_{i=1}^{n} a_i\omega^{i-1}\Big) = \Big(\sum_{i=1}^{n} c_i\omega^{i-1}\Big)$$

with

$$c_\ell = \Big(\sum_{i=1}^{\ell} b_{\ell+1-i}a_i + \sum_{i+j\geqslant n+2} \varepsilon_{i,j,\ell} b_i a_j\Big) = T_{n-\ell}(\mathbf{b})\cdot\mathbf{a},$$

where $\varepsilon_{i,j,\ell} \in \mathbb{Z}$ are some constants depending on the coefficients of the minimal polynomial f of ω. Here $T_0, \dots, T_{n-1}$ are linear maps with the property that $T_j(\mathbf{b})_\ell$ is equal to $b_{n+1-j-\ell}$ (or 0 if $n \leqslant j+\ell$) plus some integral linear combination of $b_{n-\ell+2}, \dots, b_n$ (if $\ell \geqslant 2$). Again, we let $\diamond$ denote the above operation so that $\mathbf{c} = \mathbf{b}\diamond\mathbf{a}$. We then have the corresponding definition of the lattices $\Lambda_{\mathbf{v}}$ and $\Lambda_{\mathbf{v}_1,\mathbf{v}_2}$

$$\begin{aligned}\Lambda_{\mathbf{v}} &= \{\mathbf{x}\in\mathbb{Z}^n : (\mathbf{x}\diamond\mathbf{v})_i = 0,\ n-k<i\leqslant n\}\\ &= \{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}\cdot T_i(\mathbf{v}) = 0,\ 0\leqslant i\leqslant k-1\},\\ \Lambda_{\mathbf{v}_1,\mathbf{v}_2} &= \{\mathbf{x}\in\mathbb{Z}^n : (\mathbf{x}\diamond\mathbf{v}_1)_i = (\mathbf{x}\diamond\mathbf{v}_2)_i = 0,\ n-k<i\leqslant n\}\\ &= \{\mathbf{x}\in\mathbb{Z}^n : \mathbf{x}\cdot T_i(\mathbf{v}_1) = \mathbf{x}\cdot T_i(\mathbf{v}_2) = 0,\ 0\leqslant i\leqslant k-1\},\end{aligned}$$

and Lemmas 10.1 and 11.10 then hold in an identical way with T^i replaced by T_i. In place of Lemmas 10.4 and 11.9, we have the following two simple lemmas.

LEMMA 12.1. *Given $\mathbf{b}\in\mathbb{Z}^n\backslash\{\mathbf{0}\}$, let $\mathcal{L}$ be a linear subspace of $\mathbb{R}^n$ such that $\wedge(\mathbf{x},\mathbf{b}) = \mathbf{0}$ for all $\mathbf{x}\in\mathcal{L}$. Then $\mathcal{L}$ has dimension at most $2k-1$.*

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Proof. We note that for $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$, we have $N(\sum_{i=1}^n x_i \omega^{i-1}) \neq 0$, so the columns $T_0(\mathbf{x}), \dots, T_{n-1}(\mathbf{x})$ in the multiplication-by-$\sum_{i=1}^n x_i \omega^{i-1}$ matrix are linearly independent. Thus, there are no constants $c_0, \dots, c_{k-1}$ not all zero such that $\sum_{i=0}^{k-1} c_i T_i(\mathbf{x}) = \mathbf{0}$. Thus , by linearity of the T_i, we see that given $c_0, \dots, c_{k-1}$ not all zero and given $d_0, \dots, d_{k-1}$, there is at most one $\mathbf{x} \in \mathbb{Z}^n$ such that

$$\sum_{i=0}^{k-1} c_i T_i(\mathbf{x}) = \sum_{i=0}^{k-1} d_i T_i(\mathbf{b}).$$

Hence, if $\wedge(\mathbf{x}, \mathbf{b}) = \mathbf{0}$, then $\mathbf{x}$ is given by vector of rational polynomial expressions in $c_0, \dots, c_{k-1}, d_0, \dots, d_{k-1}$. Since one of $c_0, \dots, c_{k-1}, d_0, \dots, d_{k-1}$ may be assumed to be 1, we see that $\mathbf{x}$ lies in a variety of dimension at most $2k-1$, and so any linear subspace containing only $\mathbf{x}$ of this form must have dimension at most $2k-1$. □

LEMMA 12.2. *We have*

$$\#\{\mathbf{b} \in \mathbb{F}_p^n : \wedge(\mathbf{b}) = \mathbf{0}\} \ll p^{2k-2}.$$

Proof. If $\wedge(\mathbf{b}) = \mathbf{0} \in \mathbf{F}_p^n$, then there are constants $c_0, \dots, c_{k-1}$ one of which is 1 such that $\sum_{i=0}^{k-1} c_i T_i(\mathbf{b}) = \mathbf{0}$. We argue in the case $c_{k-1} = 1$; the other cases are analogous. Looking at the ℓth component for $\ell \leqslant n-k+1$, we see that this gives $b_{n-k+2-\ell}$ in terms of $b_{n+3-k-\ell}, \dots, b_n$. In particular, $\mathbf{b}$ is uniquely determined by $b_{n-k+2}, \dots, b_n$ and $c_1, \dots, c_{k-2}$. Hence, there are at most p^{2k-2} choices of $\mathbf{b}$. □

Since we have a bound $2k-1$ in Lemma 12.1 instead of k of Lemma 10.4, we can only ensure that $\Lambda_{\mathbf{v}}$ has a basis satisfying the first and third conditions of Lemma 10.5 with $\wedge(\mathbf{z}_1, \mathbf{z}_{2k}) \neq 0$ instead of $\wedge(\mathbf{z}_1, \mathbf{z}_{k+1}) \neq 0$. This requires a number of small modifications throughout Section 11 with each instance of $\mathbf{z}_{k+1}$ replaced by $\mathbf{z}_{2k}$ (and some corresponding minor adjustments replacing $k+1$ with $2k$). This affects the argument when we establish (11.7) since, instead, we have

$$Z_1^{2k} Z_{2k}^{n-3k} \ll \det(\Lambda_{\mathbf{v}}) \ll V^k \ll (BC)^{k^2/(n-2k)},$$

and so to deduce that $Z_1 Z_{2k} \ll (BC)^{1-2\delta}$ for some $\delta > 0$, we require that $n > (5+\sqrt{5})k/2$. Similarly, for Lemma 11.6, to ensure that $Z_1 Z_{2k} \ll B^{2-2\delta}$ using $Z_1^{2k} Z_{2k}^{n-3k} \ll \det(\Lambda_{\mathbf{a}}) \ll A^k$, we require that $A^{k/(1-\delta)} \ll B^{2n-6k}$ as well as $A^{k/(1-\delta)} \ll B^{n-k}$. The rest of the Archimedean estimates go through as before.

For the non-Archimedean estimates, we use the bound of Lemma 12.2 instead of Lemma 11.9 in Lemma 11.12. In order to conclude that for $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda_{\mathbf{a}}$, we have $\wedge(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{0} \pmod p$ only if at least two nonzero polynomials with no

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

common factor vanish $(\mathrm{mod}\ p)$, we require that $n-k \geqslant 2k-2+k+2$ instead of $n-k \geqslant k-1+k+2$; that is, we require $n \geqslant 4k$. We note that $4k \geqslant (5+\sqrt{5})k/2$. With this restriction, the rest of the proof of the Type II estimate goes through as before.

Combining the above restrictions, we see that we have the Type II estimate, provided $n \geqslant 4k$, and any polytope $\mathcal{R} \subseteq [\epsilon^2, 2n]^\ell$ has

$$(\xi_1, \dots, \xi_\ell) \in \mathcal{R} \Rightarrow \max\left(k+\epsilon, \frac{kn+\epsilon}{2n-5k}\right) < \sum_{j=1}^{\ell'} \xi_j < n-2k-\epsilon$$

for some $\ell' \leqslant \ell$ (in addition to the assumptions already contained in Proposition 6.1). For $n < 5k$, this has reduced the range of our Type II estimate, and so we require a slightly different decomposition of $S(\mathcal{A}, \mathfrak{r}_2)$.

When $n \geqslant 4k$, we see that we can handle Type II terms if there is a factor with norm in the interval $[X^{n/3+\epsilon}, X^{n/2-\epsilon}]$. An identical argument then shows that we have an equivalent of Proposition 6.2 for sums $\sum_{\mathfrak{d}} S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1')$ instead of $\sum_{\mathfrak{d}} S(\mathcal{A}_{\mathfrak{d}}, \mathfrak{r}_1)$, where $\mathfrak{r}_1'$ is any ideal with $N(\mathfrak{r}_1') < X^{n/6-2\epsilon}$ (since this is the length $X^{n/2-\epsilon}/X^{n/3+\epsilon}$ of our new Type II range). We let $\mathfrak{r}_1', \mathfrak{r}_2', \mathfrak{r}_3', \mathfrak{r}_4', \mathfrak{r}_5', \mathfrak{r}_6'$ be chosen maximally (with respect to the ordering of ideals from Section 6) subject to $N(\mathfrak{r}_1') < X^{n/6-2\epsilon}$, $N(\mathfrak{r}_2') < X^{n/3+\epsilon}$, $N(\mathfrak{r}_3') < X^{n/2-\epsilon}$, $N(\mathfrak{r}_4') < X^{n/2+\epsilon}$, $N(\mathfrak{r}_5') < X^{2n/3-\epsilon}$, $N(\mathfrak{r}_6') < X^{2n/3+2\epsilon}$. By applying Buchstab's identity twice and splitting up some of the summations which appear, we have

$$\begin{aligned}
S(\mathcal{A}, \mathfrak{r}_4') &= S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{r}_1') - \sum_{\mathfrak{r}_1' < \mathfrak{p} \leqslant \mathfrak{r}_4'} S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p}) \\
&= S(\mathcal{A}, \mathfrak{r}_1') - \sum_{\mathfrak{r}_1' < \mathfrak{p}_1 \leqslant \mathfrak{r}_2'} S(\mathcal{A}_{\mathfrak{p}_1}, \mathfrak{r}_1') - \sum_{\mathfrak{r}_2' < \mathfrak{p} \leqslant \mathfrak{r}_3'} S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p}) - \sum_{\mathfrak{r}_3' < \mathfrak{p} \leqslant \mathfrak{r}_4'} S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p}) \\
&\quad + \sum_{\mathfrak{r}_1' < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{r}_2'} S(\mathcal{A}_{\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{p}_2) \\
&= S(\mathcal{A}, \mathfrak{r}_1') - \sum_{\mathfrak{r}_1' < \mathfrak{p}_1 \leqslant \mathfrak{r}_2'} S(\mathcal{A}_{\mathfrak{p}_1}, \mathfrak{r}_1') - \sum_{\mathfrak{r}_2' < \mathfrak{p} \leqslant \mathfrak{r}_3'} S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p}) - \sum_{\mathfrak{r}_3' < \mathfrak{p} \leqslant \mathfrak{r}_4'} S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p}) \\
&\quad + \sum_{\substack{\mathfrak{r}_1' < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{r}_2' \\ \mathfrak{r}_2' < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{r}_3' \text{ or } \mathfrak{r}_4' < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{r}_5'}} S(\mathcal{A}_{\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{p}_2) + \sum_{\substack{\mathfrak{r}_1' < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{r}_2' \\ \mathfrak{r}_1'^2 < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{r}_2' \text{ or } \mathfrak{r}_3' < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{r}_4' \text{ or } \mathfrak{r}_5' < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{r}_6'}} S(\mathcal{A}_{\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{p}_2).
\end{aligned}$$

The first three and the fifth terms in the decomposition above can be evaluated asymptotically by the equivalents of Propositions 6.1 and 6.2. The fourth and the final terms can be bounded in magnitude by replacing $S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{p})$ and $S(\mathcal{A}_{\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{p}_2)$ with $S(\mathcal{A}_{\mathfrak{p}}, \mathfrak{r}_1)$ and $S(\mathcal{A}_{\mathfrak{p}_1\mathfrak{p}_2}, \mathfrak{r}_1)$, respectively, and the equivalent of

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Proposition 6.2 then shows that these terms contribute $O(\epsilon)$ to the final estimate since the range of norms in the sums is of length $O(\epsilon)$ in the logarithmic scale. Thus, we have a decomposition where all terms can be evaluated asymptotically or contribute a negligible amount.

The final minor change is in the proof of Lemma 6.4. In establishing (7.7), we used the multiplicative structure of $\mathbb{Z}[\sqrt[n]{\theta}]$. However, recalling that $(\mathbf{a} \diamond \mathbf{b})_n = \mathbf{a} \cdot T_0(\mathbf{b})$ and $T_0(\mathbf{b})_\ell$ is equal to $b_{n+1-\ell}$, we see that

$$\sum_{\substack{\mathbf{a}\in[1,q]^n \\ a_j=0 \text{ if } j>n-k}} e(\mathbf{a}\cdot T_0(\mathbf{b})/q) = \begin{cases} q^{n-k}, & \text{if } b_n = \cdots = b_{k+1} = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, the proof goes through exactly as before.

Acknowledgements

The author is supported by a Clay Research Fellowship and is a Fellow by Examination of Magdalen College, Oxford. We thank Stanley Xiao for some useful comments and the anonymous referees for many helpful suggestions.

Appendix A. Explicit sieve decomposition

In our appendix, we give a description of an adequate sieve decomposition used in Section 6 in the case $n < 4k$. As mentioned previously, the work of Harman [10], in principle, gives a decomposition which is adequate for us, but in the interests of clarity, we give a different explicit decomposition here.

We recall that we have an ordering on ideals which respects the partial ordering by norm and that $\mathfrak{r}_2$ is maximal with $N(\mathfrak{r}_2) < X^{n(1/2+\epsilon)}$. We recall the notation $\mathcal{C}_\mathfrak{d}$ to denote the set of ideals $\mathfrak{c}$ such that $\mathfrak{c}\mathfrak{d}$ lies in the set $\mathcal{C}$ and the notation $S(\mathcal{C}, \mathfrak{z})$ to denote ideals in the set $\mathcal{C}$ with all ideal factors larger than $\mathfrak{z}$.

We wish to obtain a decomposition of $S(\mathcal{C}, \mathfrak{r}_2)$ of the type given by Proposition 6.6, which then allows us to obtain a lower bound for the number of primes in $\mathcal{A}$ by performing the same decomposition to $\mathcal{B}$, giving a lower bound of the form (6.5).

Rather than directly produce a decomposition of the form of Proposition 6.6 for a general set $\mathcal{C}$, it is more convenient and more conceptual for us to deal with $\mathcal{A}$ and $\mathcal{B}$ at the same time so that we can pay attention only to those terms which cannot be shown to be negligible by Propositions 6.1 and 6.2 since then the motivation for our decomposition is clear. With this in mind, we define

$$T(\mathfrak{d}, \mathfrak{z}) = S(\mathcal{A}_\mathfrak{d}, \mathfrak{z}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}} S(\mathcal{B}_\mathfrak{d}, \mathfrak{z}),$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

and we wish to make a decomposition of $T((1), \mathfrak{r}_2)$ into terms that can be shown to be negligible by Propositions 6.2 (giving the sets $\mathcal{S}_1$ and $\mathcal{S}_2$) and 6.1 (giving the sets $\mathcal{S}_3$ and $\mathcal{S}_4$) and some remaining terms for which we can produce an adequate lower bound (giving the set $\mathcal{S}_5$). It will be obvious from our construction that once we have obtained suitable decomposition of $T(\mathfrak{d}, \mathfrak{z})$, this immediately gives a suitable decomposition of the type in Proposition 6.6.

We see that Propositions 6.1 and 6.2 show that various averages of $T(\mathfrak{d}, \mathfrak{z})$ are negligible. Similarly, although the decomposition of Proposition 6.6 is given in terms of polytopes, we will deal just with sums of terms of $T(\mathfrak{d}, \mathfrak{z})$. Since all these expressions will be involving ideals with at most $1/(3\varpi - 1)$ prime factors with constraints only on the number and size of the prime factors, we see that they can be re-written in terms of polytopes to give a decomposition of the originally desired form.

We assume throughout that $0.25 < k/n + 4\epsilon < \varpi := 0.3182$ and note that $7/22 < 0.3182$. We will only use our Type II estimate of Proposition 6.1 to evaluate terms involving an ideal factor with norm in the interval $[X^{n\varpi}, X^{n(1-2\varpi)}]$ or $[X^{2n\varpi}, X^{n(1-\varpi)}]$, and we will only use Proposition 6.2 for sums $\sum_{\mathfrak{d}} \mathbf{1}_{\mathcal{R}}(\mathfrak{d}) S(\mathcal{A}_d, \mathfrak{r}_1)$ where the $\mathfrak{d}$ in the summation satisfies $N(\mathfrak{d}) < X^{n(1-\varpi)}$. (This corresponds to restricting to the conditions $\varpi n \leqslant \sum_{i=1}^{\ell'} e_i \leqslant n(1-2\varpi)$ and $\sum_{i=1}^{\ell} e_i < n(1-\varpi)$ in Propositions 6.1 and 6.2, respectively.) We note that the restriction to $k/n + 4\epsilon < \varpi$ implies that $N(\mathfrak{r}_1) < X^{n-3k-4\epsilon}$, as required by Proposition 6.2. Thus, our decomposition will be valid for all k, n satisfying $k/n < \varpi - 4\epsilon$.

We fix ideals $\mathfrak{z}_1 \leqslant \cdots \leqslant \mathfrak{z}_6$ chosen maximally (with respect to our ordering) subject to

$$N(\mathfrak{z}_1) \leqslant X^{n(1-3\varpi)}, \qquad N(\mathfrak{z}_2) \leqslant X^{n\varpi}, \qquad N(\mathfrak{z}_3) \leqslant X^{n(1-2\varpi)},$$
$$N(\mathfrak{z}_4) \leqslant X^{n(1/2+\epsilon)}, \qquad N(\mathfrak{z}_5) \leqslant X^{2n\varpi}, \qquad N(\mathfrak{z}_6) \leqslant X^{n(1-\varpi)}.$$

The quantities $\mathfrak{r}_1, \mathfrak{r}_2$ from Section 6 are equal to $\mathfrak{z}_1$ and $\mathfrak{z}_4$, respectively.

Since we can estimate $S(\mathcal{B}, \mathfrak{z}_4)$ by the prime ideal theorem (Lemma 4.3), we see that it suffices to get a suitable lower bound for $T((1), \mathfrak{z}_4)$ to produce the desired lower bound for $S(\mathcal{A}, \mathfrak{z}_4)$.

By Buchstab's identity,

$$\begin{aligned} T((1), \mathfrak{z}_4) &= T((1), \mathfrak{z}_1) - \sum_{\mathfrak{z}_1 < \mathfrak{p} \leqslant \mathfrak{z}_4} T(\mathfrak{p}, \mathfrak{p}) \\ &= T((1), \mathfrak{z}_1) - \sum_{\mathfrak{z}_1 < \mathfrak{p} \leqslant \mathfrak{z}_2} T(\mathfrak{p}, \mathfrak{z}_1) + \sum_{\mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2) \end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$- \sum_{\mathfrak{z}_2 < \mathfrak{p} \leqslant \mathfrak{z}_3} T(\mathfrak{p}, \mathfrak{p}) - \sum_{\mathfrak{z}_3 < \mathfrak{p} \leqslant \mathfrak{z}_4} T(\mathfrak{p}, \mathfrak{z}_1) + \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2)$$

$$=: T_1 - T_2 + T_3 - T_4 - T_5 + T_6. \tag{A.1}$$

By Proposition 6.2, the T_1, T_2 and T_5 terms are $O(\epsilon \#\mathcal{A}/\log X)$, which is acceptable. By Proposition 6.1, $T_4 = o(\#\mathcal{A}/\log X)$ and so is also negligible. Thus, we are left to consider T_3 and T_6.

We first split T_3 and T_6 into subsums $T_{3,1}$, $T_{3,2}$ and $T_{6,1}$, $T_{6,2}$ depending on whether $\mathfrak{p}_1\mathfrak{p}_2^2 \leqslant \mathfrak{z}_6$ or not. First we consider $T_{6,1}$, the terms from T_6 with $\mathfrak{p}_1\mathfrak{p}_2^2 \leqslant \mathfrak{z}_6$, where we can apply further Buchstab iterations. This gives

$$T_{6,1} := \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2) = \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{z}_1) - \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_3 \leqslant \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6 \\ N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3^2) \ll X^n}} T(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{z}_1)$$

$$+ \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6 \\ N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3^2) \ll X^n}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4).$$

Here we have the additional restriction $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3^2) \ll X^n$ since $T(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{p}_3) = 0$ otherwise. Since $\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_6$, Proposition 6.2 shows that the first two terms are negligible. This leaves

$$S_1 = \sum_{\substack{\mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4 \\ \mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6 \\ N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3^2) \ll X^n}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4). \tag{A.2}$$

Similarly, for $T_{3,1}$, the terms from T_3 with $\mathfrak{p}_1\mathfrak{p}_2^2 \leqslant \mathfrak{z}_6$, we find

$$\sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2) = \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{z}_1) - \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_3 \leqslant \mathfrak{p}_2 \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{z}_1)$$

$$+ \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_6}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4).$$

Since $\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_6$, by Proposition 6.2, the first two terms are negligible. If $\mathfrak{z}_2 < \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_3$, then the contribution is also negligible. Thus, the final term splits

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

as $S_2 + S_3 + o(\#\mathcal{A}/\log X)$, where

$$S_2 = \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_5 \\ \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_2}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) = \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_{1,1} \\ \mathfrak{p}_2\mathfrak{p}_1 \leqslant \mathfrak{z}_2}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4), \quad (A.3)$$

$$S_3 = \sum_{\substack{\mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1 \leqslant \mathfrak{z}_5 \\ \mathfrak{p}_1\mathfrak{p}_2 > \mathfrak{z}_3}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) = \sum_{\substack{\mathfrak{z}_{1,2} < \mathfrak{p}_1 \leqslant \mathfrak{z}_2 \\ \mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \cdots \leqslant \mathfrak{p}_1 \\ \mathfrak{p}_2\mathfrak{p}_1 > \mathfrak{z}_3}} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4). \quad (A.4)$$

Here $\mathfrak{z}_{1,1}$ and $\mathfrak{z}_{1,2}$ are chosen maximally such that $N(\mathfrak{z}_{1,1}) \leqslant X^{n(4\varpi-1)}$ and $N(\mathfrak{z}_{1,2}) \leqslant X^{n(1/2-\varpi)}$. We are left to consider S_1, S_2, S_3, $T_{3,2}$ and $T_{6,2}$.

We now consider $T_{6,2}$ and split it into $T_{6,2,1}$ and $T_{6,2,2}$ depending on whether $\mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_6$ or not. We first consider $T_{6,2,1}$, where we are dealing with terms with $\mathfrak{p}_1\mathfrak{p}_2^2 > \mathfrak{z}_6$ and $\mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_6$ and $\mathfrak{p}_1 > \mathfrak{z}_3$. Here we apply a reversal of roles. Over the collection of such $\mathfrak{p}_1, \mathfrak{p}_2$, we note that $T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2)$ is counting products $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{q}$ with $\mathfrak{p}|\mathfrak{q} \Rightarrow \mathfrak{p} > \mathfrak{p}_2$, with the size constraints on $(\mathfrak{p}_1, \mathfrak{p}_2)$

$$\mathfrak{p}_1\mathfrak{p}_2^2 > \mathfrak{z}_6, \quad \mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_6, \quad \mathfrak{z}_3 < \mathfrak{p}_1 \leqslant \mathfrak{z}_4, \quad \mathfrak{z}_1 < \mathfrak{p}_2 \leqslant \mathfrak{p}_1, \quad N(\mathfrak{p}_2^2\mathfrak{p}_1) \ll X^n. \quad (A.5)$$

Since the contribution with a factor $\mathfrak{a}$ satisfying $N(\mathfrak{a}) \in [Y, Y^{1+o(1)}]$ is negligible and $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{q}) \asymp X^n$, we see that we can translate these size constraints into constraints on the size of $\mathfrak{q}$ and $\mathfrak{p}_2$ at the cost of a negligible error. Therefore, letting $\mathfrak{z}(\mathfrak{q}, \mathfrak{p}_2)$ be maximal with norm at most $(X^{n+\epsilon}/N(\mathfrak{q}\mathfrak{p}_2))^{1/2+\epsilon}$, we find

$$\begin{aligned}\sum_{\mathfrak{p}_1,\mathfrak{p}_2} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2) &= \sum_{\mathfrak{q},\mathfrak{p}_2} T(\mathfrak{q}\mathfrak{p}_2, \mathfrak{z}(\mathfrak{q}, \mathfrak{p}_2)) \\ &= \sum_{\mathfrak{q},\mathfrak{p}_2} T(\mathfrak{q}\mathfrak{p}_2, \mathfrak{z}_1) - \sum_{\mathfrak{q},\mathfrak{p}_2} \sum_{\mathfrak{z}_1 < \mathfrak{p}_3 \leqslant \mathfrak{z}(\mathfrak{q},\mathfrak{p}_2)} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{p}_3).\end{aligned}$$

Here the summation over $\mathfrak{p}_1, \mathfrak{p}_2$ is constrained by (A.5), and the summation over $\mathfrak{q}, \mathfrak{p}_2$ is constrained by $((X)/\mathfrak{q}\mathfrak{p}_2, \mathfrak{p}_2)$ satisfying (A.5) in place of $(\mathfrak{p}_1, \mathfrak{p}_2)$ as well as $\mathfrak{p}|\mathfrak{q} \Rightarrow \mathfrak{p} > \mathfrak{p}_2$. The first term is negligible since $\mathfrak{p}_1 > \mathfrak{z}_2$ and so $\mathfrak{q}\mathfrak{p}_2 \leqslant \mathfrak{z}_6$. The second term is counting products $\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{q}_2$ with $\mathfrak{p}|\mathfrak{q} \Rightarrow \mathfrak{p} > \mathfrak{p}_2$ and $\mathfrak{p}|\mathfrak{q}_2 \Rightarrow \mathfrak{p} > \mathfrak{p}_3$. Thus, we may rewrite this term as

$$-\sum_{\mathfrak{q}_2,\mathfrak{p}_2,\mathfrak{p}_3} T(\mathfrak{q}_2\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{p}_2),$$

which is constrained by the conditions that $(\mathfrak{q}_2\mathfrak{p}_3, \mathfrak{p}_2)$ satisfies (A.5), $\mathfrak{z}_1 < \mathfrak{p}_3$ and $\mathfrak{p}|\mathfrak{q}_2 \Rightarrow \mathfrak{p} > \mathfrak{p}_3$. Applying another Buchstab iteration gives

$$-\sum_{\mathfrak{q}_2,\mathfrak{p}_2,\mathfrak{p}_3} T(\mathfrak{q}_2\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{p}_2) = -\sum_{\mathfrak{q}_2,\mathfrak{p}_2,\mathfrak{p}_3} T(\mathfrak{q}_2\mathfrak{p}_2\mathfrak{p}_3, \mathfrak{z}_1) + \sum_{\mathfrak{q}_2,\mathfrak{p}_2,\mathfrak{p}_3} \sum_{\mathfrak{z}_1 < \mathfrak{p}_4 \leqslant \mathfrak{p}_2} T(\mathfrak{q}_2\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4, \mathfrak{p}_4).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The first term is negligible since $\mathfrak{p}_1\mathfrak{p}_2 \leqslant \mathfrak{z}_6$ implies $\mathfrak{q}_2\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_6$. Thus, apart from an error term $O(\epsilon \#\mathcal{A}/\log X)$, we are left with

$$S_4 = \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{z}_1<\mathfrak{p}_4\leqslant\mathfrak{p}_2 \\ \mathfrak{p}|\mathfrak{q}\Rightarrow\mathfrak{p}>\mathfrak{p}_3 \\ (\mathfrak{q}\mathfrak{p}_3,\mathfrak{p}_2) \text{ satisfy (A.5)}}} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4, \mathfrak{p}_4). \tag{A.6}$$

Finally, we wish to consider $T_{3,2}$ and $T_{6,2,2}$, which is the terms with $\mathfrak{p}_1\mathfrak{p}_2^2 > \mathfrak{z}_6$ and either $\mathfrak{p}_1\mathfrak{p}_2 > \mathfrak{z}_6$ or $\mathfrak{p}_1 \leqslant \mathfrak{z}_2$ (note that these cannot simultaneously occur as $3\varpi < 1$). These contribute S_5 and S_6, respectively, where

$$S_5 = \sum_{\substack{\mathfrak{z}_3<\mathfrak{p}_1\leqslant\mathfrak{z}_4 \\ \mathfrak{z}_6<\mathfrak{p}_1\mathfrak{p}_2 \\ N(\mathfrak{p}_2^2\mathfrak{p}_1)\ll X^n}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2), \tag{A.7}$$

$$S_6 = \sum_{\substack{\mathfrak{p}_2\leqslant\mathfrak{p}_1\leqslant\mathfrak{z}_2 \\ \mathfrak{p}_2^2\mathfrak{p}_1>\mathfrak{z}_6}} T(\mathfrak{p}_1\mathfrak{p}_2, \mathfrak{p}_2). \tag{A.8}$$

Thus, to get a lower bound for $T((1), \mathfrak{z}_4)$, it suffices to get lower bounds for the sums $S_1, \dots, S_6$. Recalling the definition of $T(\mathfrak{d}, \mathfrak{z})$, we see that we have the lower bound (valid for $N(\mathfrak{d}\mathfrak{z}) \leqslant X^{n(1-\epsilon)}$)

$$\begin{aligned} T(\mathfrak{d}, \mathfrak{z}) &= S(\mathcal{A}_\mathfrak{d}, \mathfrak{z}) - \tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}} S(\mathcal{B}_\mathfrak{d}, \mathfrak{z}) \\ &\geqslant -\tilde{\mathfrak{S}}\frac{\#\mathcal{A}}{\#\mathcal{B}} S(\mathcal{B}_\mathfrak{d}, \mathfrak{z}) \\ &= -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{N(\mathfrak{d})\log N(\mathfrak{z})}\omega\left(\frac{\log(X^n/N(\mathfrak{d})}{\log N(\mathfrak{z})}\right), \end{aligned}$$

where $\omega(\cdot)$ is the Buchstab function defined by

$$\begin{aligned} \omega(u) &= \frac{1}{u}, && 1 \leqslant u \leqslant 2, \\ u\frac{\partial\omega}{\partial u}(u) &= \omega(u-1) - \omega(u), && 2 \leqslant u. \end{aligned}$$

This lower bound allows us to obtain an explicit integral expression as a lower bound for $T((1), \mathfrak{z}_4)$. Moreover, we can restrict the summation in each of the S_i so that no subproduct of $\mathfrak{p}_1, \dots, \mathfrak{p}_4$ lies between $\mathfrak{z}_2$ and $\mathfrak{z}_3$ or between $\mathfrak{z}_5$ and $\mathfrak{z}_6$ since these parts are negligible by our Type II estimate. For example,

$$S_1 = \sum_{\substack{\mathfrak{z}_3<\mathfrak{p}_1\leqslant\mathfrak{z}_4 \\ \mathfrak{z}_1<\mathfrak{p}_4\leqslant\mathfrak{p}_3\leqslant\mathfrak{p}_2\leqslant\mathfrak{p}_1 \\ \mathfrak{p}_2^2\mathfrak{p}_1\leqslant\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4, \mathfrak{p}_4)$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\geqslant \sum_{\substack{\mathfrak{z}_3<\mathfrak{p}_1\leqslant\mathfrak{z}_4\\ \mathfrak{z}_1<\mathfrak{p}_4\leqslant\mathfrak{p}_3\leqslant\mathfrak{p}_2\leqslant\mathfrak{p}_1\\ \mathfrak{p}_2^2\mathfrak{p}_1\leqslant\mathfrak{z}_6}}{}' T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)+o(\#\mathcal{A}/\log X)$$

$$\geqslant (-1+o(1))\mathfrak{S}\#\mathcal{A}\sum_{\substack{\mathfrak{z}_3<\mathfrak{p}_1<\mathfrak{z}_4\\ \mathfrak{z}_1<\mathfrak{p}_4\leqslant\mathfrak{p}_3\leqslant\mathfrak{p}_2\leqslant\mathfrak{p}_1\\ \mathfrak{p}_2^2\mathfrak{p}_1\leqslant\mathfrak{z}_6}}{}' \frac{\omega\left(\frac{\log(X^n/N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4))}{\log N(\mathfrak{p}_4)}\right)}{N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4)\log N(\mathfrak{p}_4)}$$

$$=(-1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\int\cdots\int{}'\omega\left(\frac{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}{\alpha_4}\right)\frac{d\alpha_1\,d\alpha_2\,d\alpha_3\,d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4^2}.$$

Here $\sum'$ indicates that we have restricted the summation so that no subproduct of $\mathfrak{p}_1,\dots,\mathfrak{p}_4$ lies between $\mathfrak{z}_2$ and $\mathfrak{z}_3$ or between $\mathfrak{z}_5$ and $\mathfrak{z}_6$, and in the final line, we used partial summation with the change of variables $N(\mathfrak{p}_i)=X^{n\alpha_i}$, and the integration is over the region defined by

$$1-2\varpi\leqslant\alpha_1\leqslant 1/2+\epsilon,\quad 1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1,$$
$$\sum_{i\in J}\alpha_i\notin[\varpi,1-2\varpi]\cup[2\varpi,1-\varpi]\,\forall J\subseteq\{1,2,3,4\},$$
$$\alpha_1+\cdots+\alpha_{j-1}+2\alpha_j\leqslant 1\,\forall j\in\{2,3,4\}.$$

In principle, this should already give us a reasonable lower bound for $T((1),\mathfrak{z}_4)$. Unfortunately, it appears difficult to get a good numerical approximation to integrals over regions similar to the above one, presenting a practical difficulty. To get around this difficulty, we split the sums $S_1,\dots,S_6$ further into various subsums, and on these subsums, we relax some of the constraints (corresponding to obtaining an upper bound for the integrals appearing) so that we have explicit integrals which are amenable to numerical integration. The remainder of the appendix is spent performing such a decomposition explicitly and obtaining the corresponding numerical estimates.

From now on, we use the notation $\sum^*$ and $\int\cdots\int^*$ to denote the fact that we are summing or integrating over variables with various size constraints, which we only explicitly write down later. The constraints implied by the asterisk will remain the same within each display but may be different in different displays.

A.1. The sum S_1. We first split the summation according to whether $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_5$ or $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_6$ or $\mathfrak{z}_5<\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_6$. The final range makes a negligible contribution by our Type II estimate. This gives

$$S_1=S_{1,1}+S_{1,2}+o(\#\mathcal{A}/\log X).$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

We first concentrate on $S_{1,1}$ where $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_5$. We split this into two further sums $S_{1,1,1}$ and $S_{1,1,2}$ depending on whether $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2 \leqslant \mathfrak{z}_6$ or not. If $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2 \leqslant \mathfrak{z}_6$, then we can perform two further Buchstab decompositions. Thus, we find

$$\begin{aligned} S_{1,1,1} &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) \\ &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\dots,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{z}_1) - \sideset{}{^*}\sum_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}\sum_{\mathfrak{z}_1<\mathfrak{p}_5\leqslant\mathfrak{p}_4} T(\mathfrak{p}_1,\dots,\mathfrak{p}_5,\mathfrak{z}_1) \\ &\quad + \sideset{}{^*}\sum_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}\sum_{\mathfrak{z}_1<\mathfrak{p}_6\leqslant\mathfrak{p}_5\leqslant\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_6,\mathfrak{p}_6). \end{aligned}$$

By Proposition 6.2, the first two terms make a negligible contribution, and lower bounding the final term, we find that $S_{1,1,1}$ is

$$\begin{aligned} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\sideset{}{^*}\sum_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}\frac{1}{N(\mathfrak{p}_1\dots\mathfrak{p}_4)} \\ &\quad\times\int_{1-3\varpi}^{\alpha_5}\int_{1-3\varpi}^{\alpha_4}\omega\Big(\frac{1-\alpha_1-\dots-\alpha_6}{\alpha_6}\Big)\frac{d\alpha_5\,d\alpha_6}{\alpha_5\alpha_6^2} \\ &\quad+O(\epsilon\#\mathcal{A}/\log X) \\ &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} \\ &\quad\times\int\cdots\int^*\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\bigg(\frac{1}{1-3\varpi}\bigg(\log\Big(\frac{\alpha_4}{1-3\varpi}\Big)-1\bigg)+\frac{1}{\alpha_4}\bigg) \\ &=: -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{1,1,1}. \end{aligned}$$

Here we trivially bounded the Buchstab function by 1, and the integral $I_{1,1,1}$ is over the region defined by

$$\begin{aligned} &1-2\varpi\leqslant\alpha_1\leqslant 1/2+\epsilon, && 1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1, \\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4\leqslant 1-\varpi, && \alpha_1+2\alpha_2\leqslant 1-\varpi, \\ &\alpha_1+\alpha_2+2\alpha_3\leqslant 1, \end{aligned}$$

and this can be evaluated numerically in a feasible manner.

We will not perform further decompositions for the remaining parts of S_1, simply splitting the summation according to size conditions. The remaining part

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$S_{1,1,2}$ of $S_{1,1}$ with $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2 > \mathfrak{z}_6$ we lower bound directly, giving

$$\begin{aligned} S_{1,1,2} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\int\cdots\int^* \omega\Big(\frac{1-\alpha_1-\cdots-\alpha_4}{\alpha_4}\Big)\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4^2},\\ &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\int\cdots\int^* \frac{4}{7}\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4^2},\\ &=: -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{1,1,2}. \end{aligned}$$

Here we used the fact that $\alpha_1+\alpha_2+\alpha_3+\alpha_4 < 2\varpi$ so $1-\alpha_1-\alpha_2-\alpha_3-\alpha_4 > 2\alpha_4$ to note that the value of the Buchstab function is always at most $4/7$ since $\sup_{u>7/4}\omega(u) = 4/7$. The integration in $I_{1,1,2}$ is over the region defined by the conditions

$$\begin{aligned} &1-2\varpi \leqslant \alpha_1 \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1,\\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \geqslant 1-\varpi, && \alpha_1+2\alpha_2 \leqslant 1-\varpi,\\ &\alpha_1+\alpha_2+\alpha_3+\alpha_4 \leqslant 2\varpi, && \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1,\\ &\alpha_1+\alpha_2+2\alpha_3 \leqslant 1. \end{aligned}$$

This gives our lower bound for $S_{1,1}$. We now consider $S_{1,2}$. We note that we have the constraints $\mathfrak{p}_1\mathfrak{p}_2^2 \leqslant \mathfrak{z}_6$ and $\mathfrak{p}_3 \leqslant \mathfrak{p}_2$ so $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_6$. Thus, by our Type II estimate, we can restrict to $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_5$ at the cost of a negligible error term. We now split the summation depending on whether $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_2$ or $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 > \mathfrak{z}_3$ (the intermediate range being negligible by our Type II estimate). This gives

$$\begin{aligned} S_{1,2} &= \sum_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}^* T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)\\ &= \sum_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2}}^* T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3}}^* T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\ &=: S_{1,2,1} + S_{1,2,2} + o(\#\mathcal{A}/\log X). \end{aligned}$$

We first consider $S_{1,2,1}$. Here we have the constraints $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_5$ and $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_2$, so we see that $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^3) \leqslant N(\mathfrak{z}_5\mathfrak{z}_2) < X^n$ for all terms in consideration. We then obtain the lower bound for $S_{1,2,1}$

$$\begin{aligned} S_{1,2,1} &\geqslant -(1+o(1))\frac{\#\mathcal{A}}{n\log X}\int\cdots\int^* \omega\Big(\frac{1-\alpha_1-\cdots-\alpha_4}{\alpha_4}\Big)\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4^2}\\ &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\int\cdots\int^* \frac{4}{7}\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4^2}, \end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$=: -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{1,2,1}.$$

We bounded the Buchstab function above by $4/7$ since $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^3) \leqslant X^n$ and so $1-\alpha_1-\cdots-\alpha_4 \geqslant 2\alpha_4$. The integration in $I_{1,2,1}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \alpha_1 \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1,\\
&\alpha_2+\alpha_3+\alpha_4 \leqslant \varpi, && \alpha_1+2\alpha_2 \leqslant 1-\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+\alpha_4 \geqslant 1-\varpi, && \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1,\\
&\alpha_1+\alpha_2+2\alpha_3 \leqslant 1, && \alpha_1+\alpha_2+\alpha_3 \leqslant 2\varpi.
\end{aligned}$$

For $S_{1,2,2}$, we obtain the lower bound

$$\begin{aligned}
S_{1,2,2} &\geqslant -(1+o(1))\frac{\#\mathcal{A}}{n\log X}\int\cdots\int^{*}\omega\left(\frac{1-\alpha_1-\cdots-\alpha_4}{\alpha_4}\right)\frac{d\alpha_1\ldots d\alpha_4}{\alpha_1\ldots\alpha_4^2}\\
&\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\int\cdots\int^{*}\frac{d\alpha_1\ldots d\alpha_4}{\alpha_1\ldots\alpha_4^2},\\
&=: -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{1,2,2}.
\end{aligned}$$

Here we bounded the Buchstab function above by 1 and the integration in $I_{1,2,2}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \alpha_1 \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1,\\
&\alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi, && \alpha_1+2\alpha_2 \leqslant 1-\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+\alpha_4 \geqslant 1-\varpi, && \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1,\\
&\alpha_1+\alpha_2+2\alpha_3 \leqslant 1, && \alpha_1+\alpha_2+\alpha_3 \leqslant 2\varpi.
\end{aligned}$$

This completes our lower bound for S_1.

A.2. The sum S_2. We now consider the sum S_2. There is a negligible contribution whenever any product of three of $\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \mathfrak{p}_4$ lies between $\mathfrak{z}_2$ and $\mathfrak{z}_3$. We therefore split the summation according to the range of each of these triple products.

$$\begin{aligned}
S_2 &= \sum_{\mathfrak{p}_1,\ldots,\mathfrak{p}_4}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4)\\
&= \sum_{\substack{\mathfrak{p}_1,\ldots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\leqslant\mathfrak{z}_2}}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\ldots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3>\mathfrak{z}_3\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_4\leqslant\mathfrak{z}_2}}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4)
\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$+ \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_4>\mathfrak{z}_3\\ \mathfrak{p}_1\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)$$

$$+ \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)$$

$$=: S_{2,1}+S_{2,2}+S_{2,3}+S_{2,4}+S_{2,5}+o(\#\mathcal{A}/\log X).$$

We decompose $S_{2,1}$ once more, depending on the size of $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$, giving

$$S_{2,1} = \sum^{*}_{\mathfrak{p}_1,\dots,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)$$

$$= \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)$$

$$=: S_{2,1,1}+S_{2,1,2}+o(\#\mathcal{A}/\log X).$$

We also split $S_{2,5}$ according to the size of $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2$ and $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)N(\mathfrak{p}_4)^{19/4}$.

$$S_{2,5} = \sum^{*}_{\mathfrak{p}_1,\dots,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)$$

$$= \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant\mathfrak{z}_6\\ N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)N(\mathfrak{p}_4)^{19/4}\leqslant X^n}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant\mathfrak{z}_6\\ N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)N(\mathfrak{p}_4)^{19/4}>X^n}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)$$

$$+ \sum^{*}_{\substack{\mathfrak{p}_1,\dots,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2>\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)$$

$$=: S_{2,5,1}+S_{2,5,2}+S_{2,5,3}.$$

For each of $S_{2,1,1}$, $S_{2,1,2}$, $S_{2,2}$, $S_{2,3}$, $S_{2,4}$, $S_{2,5,1}$ and $S_{2,5,2}$, we obtain lower bounds in an analogous manner to $S_{1,1,1}$. We have $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant\mathfrak{z}_6$, and so we can perform two further Buchstab iterations. We have

$$S_{2,1,1} = \sum^{*}_{\mathfrak{p}_1,\dots,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{z}_1) - \sum^{*}_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}\sum_{\mathfrak{z}_1<\mathfrak{p}_5\leqslant\mathfrak{p}_4} T(\mathfrak{p}_1,\dots,\mathfrak{p}_5,\mathfrak{z}_1)$$

$$+ \sum^{*}_{\mathfrak{p}_1,\dots,\mathfrak{p}_4}\sum_{\mathfrak{z}_1<\mathfrak{p}_6\leqslant\mathfrak{p}_5\leqslant\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_6,\mathfrak{p}_6)$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\begin{aligned}
&\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}\\
&\qquad\times\int\cdots\int^{*}\frac{d\alpha_1\ldots d\alpha_4}{\alpha_1\ldots\alpha_4}\frac{4}{7}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right)\\
&=: -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,1,1}.
\end{aligned}$$

Here we used the fact that $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)N(\mathfrak{p}_4)^{19/4} < X^n$ to bound the Buchstab function by $4/7$ since it is only ever evaluated at arguments larger than $7/4$. The integral $I_{2,1,1}$ is over the region defined by

$$\begin{aligned}
&\alpha_1+\alpha_2+\alpha_3+\alpha_4<\varpi, \qquad 1-3\varpi<\alpha_4<\alpha_3<\alpha_2<\alpha_1<4\varpi-1,\\
&\alpha_2+\alpha_1<\varpi.
\end{aligned}$$

In an entirely analogous manner, we obtain

$$\begin{aligned}
S_{2,1,2} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,1,2},\\
S_{2,2} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,2},\\
S_{2,3} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,3},\\
S_{2,4} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,4},\\
S_{2,5,1} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,5,1},
\end{aligned}$$

where the integrals $I_{2,1,2}$, $I_{2,2}$, $I_{2,3}$, $I_{2,4}$, $I_{2,5,1}$ are all of the form

$$\int\cdots\int^{*}\frac{d\alpha_1\ldots d\alpha_4}{\alpha_1\ldots\alpha_4}\frac{4}{7}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right)$$

for some constrained region in $\mathbb{R}^4$. Explicitly, $I_{2,1,2}$ is over the region defined by

$$\begin{aligned}
&\alpha_1+\alpha_2+\alpha_3+\alpha_4\geqslant 1-2\varpi, \quad 1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1\leqslant 4\varpi-1,\\
&\alpha_2+\alpha_1\leqslant\varpi, \qquad\qquad\qquad \alpha_1+\alpha_2+\alpha_3\leqslant\varpi.
\end{aligned}$$

The integral $I_{2,2}$ is over the region defined by

$$\begin{aligned}
&\alpha_1+\alpha_2+\alpha_3\geqslant 1-2\varpi,\ \ 1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1\leqslant 4\varpi-1,\\
&\qquad\alpha_2+\alpha_1\leqslant\varpi,\ \ \alpha_1+\alpha_2+\alpha_4\leqslant\varpi.
\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The integral $I_{2,3}$ is over the region defined by

$$\alpha_1+\alpha_2+\alpha_4 \geqslant 1-2\varpi, \quad 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1 \leqslant 4\varpi-1,$$
$$\alpha_2+\alpha_1 \leqslant \varpi, \quad \alpha_1+\alpha_3+\alpha_4 \leqslant \varpi.$$

The integral $I_{2,4}$ is over the region defined by

$$\alpha_1+\alpha_3+\alpha_4 \geqslant 1-2\varpi, \quad 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1 \leqslant 4\varpi-1,$$
$$\alpha_2+\alpha_1 \leqslant \varpi, \quad \alpha_2+\alpha_3+\alpha_4 \leqslant \varpi.$$

The integral $I_{2,5,1}$ is over the region defined by

$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1-\varpi, \quad 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1 \leqslant 4\varpi-1,$$
$$\alpha_2+\alpha_1 \leqslant \varpi, \quad \alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_1+\alpha_2+\alpha_3+19\alpha_4/4 \leqslant 1.$$

When dealing with the sum $S_{2,5,2}$, we cannot bound the Buchstab function by $4/7$, so instead we bound it by 1. In this way, we obtain

$$S_{2,5,2} \geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,5,2},$$
$$I_{2,5,2} = \int\cdots\int^{*}\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right),$$

with the integral $I_{2,5,2}$ over the region defined by

$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1-\varpi, \quad 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1 \leqslant 4\varpi-1,$$
$$\alpha_2+\alpha_1 \leqslant \varpi, \quad \alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_1+\alpha_2+\alpha_3+19\alpha_4/4 \geqslant 1.$$

Finally, for the sum $S_{2,5,3}$, we cannot perform further Buchstab iterations, so we just bound it directly. This gives

$$S_{2,5,3} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{2,5,3},$$
$$I_{2,5,3} = \int\cdots\int^{*}\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4^2},$$

with the integral $I_{2,5,3}$ over the region defined by

$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \geqslant 1-\varpi, \quad 1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1 \leqslant 4\varpi-1,$$
$$\alpha_2+\alpha_1 \leqslant \varpi, \quad \alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi.$$

This completes our decomposition of S_2.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

A.3. The sum S_3. We now consider the sum S_3. By our Type II estimate, there is a negligible contribution when any product of two of $\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \mathfrak{p}_4$ lies between $\mathfrak{z}_2$ and $\mathfrak{z}_3$. We now split the summation according to the size of the pairwise products, noting that in all cases, we have $\mathfrak{p}_1\mathfrak{p}_2 > \mathfrak{z}_3$. This gives

$$\begin{aligned}
S_3 &= \sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) \\
&= \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_1\mathfrak{p}_3 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_1\mathfrak{p}_3 > \mathfrak{z}_3 \\ \mathfrak{p}_1\mathfrak{p}_4, \mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) \\
&+ \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_1\mathfrak{p}_4 > \mathfrak{z}_3 \\ \mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_2\mathfrak{p}_3 > \mathfrak{z}_3 \\ \mathfrak{p}_1\mathfrak{p}_4 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) \\
&+ \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_1\mathfrak{p}_4, \mathfrak{p}_2\mathfrak{p}_3 > \mathfrak{z}_3 \\ \mathfrak{p}_2\mathfrak{p}_4 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_2\mathfrak{p}_4 > \mathfrak{z}_3 \\ \mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) \\
&+ \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_3\mathfrak{p}_4 > \mathfrak{z}_3}}^{*} T(\mathfrak{p}_1 \dots \mathfrak{p}_4, \mathfrak{p}_4) + o(\#\mathcal{A}/\log X) \\
&= S_{3,1} + S_{3,2} + S_{3,3} + S_{3,4} + S_{3,5} + S_{3,6} + S_{3,7} + o(\#\mathcal{A}/\log X).
\end{aligned}$$

The final three sums we obtain lower bounds without further decompositions, giving

$$\begin{aligned}
S_{3,5} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,5}, \\
S_{3,6} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,6}, \\
S_{3,7} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,7},
\end{aligned}$$

where $I_{3,5}$, $I_{3,6}$, $I_{3,7}$ are all integrals of the form

$$\int \dots \int^{*} \frac{d\alpha_1 \dots d\alpha_4}{\alpha_1 \dots \alpha_4^2}$$

over some region in $\mathbb{R}^4$. Explicitly, $I_{3,5}$ is over the region defined by

$$1/2 - \varpi \leqslant \alpha_1 \leqslant \varpi, \qquad 1 - 2\varpi - \alpha_1 \leqslant \alpha_2 \leqslant (1 - \varpi - \alpha_1)/2,$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, \quad \alpha_1+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_2+\alpha_3 \geqslant 1-2\varpi, \quad \alpha_2+\alpha_4 \leqslant \varpi.$$

The integral $I_{3,6}$ is over the region defined by

$$1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, \quad 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,$$
$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, \quad \alpha_2+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_3+\alpha_4 \leqslant \varpi.$$

The integral $I_{3,7}$ is over the region defined by

$$1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, \quad 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,$$
$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, \quad \alpha_3+\alpha_4 \geqslant 1-2\varpi.$$

We now consider $S_{3,4}$. Since $\mathfrak{p}_1\mathfrak{p}_2^2 \leqslant \mathfrak{z}_6$, we have $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_6$, and so we can restrict to $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_5$ at the cost of a negligible error term. We split the summation according to the size of $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$.

$$\begin{aligned} S_{3,4} &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) \\ &= \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_5}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X) \\ &= S_{3,4,1}+S_{3,4,2}+o(\#\mathcal{A}/\log X). \end{aligned}$$

Since $\mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{p}_1^2$ and $\mathfrak{p}_1\mathfrak{p}_4 \leqslant \mathfrak{z}_2$, we have $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^3) \leqslant N(\mathfrak{z}_2)^3 \leqslant X^n$, and so $N(\log(X^n/N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4))/\log(N(\mathfrak{p}_4))) > 3 > 7/4$. Thus, we obtain lower bounds

$$S_{3,4,1} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{3,4,1},$$
$$S_{3,4,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{3,4,2},$$

where $I_{3,4,1}$, $I_{3,4,2}$ are integrals of the form

$$\int\cdots\int^* \frac{4d\alpha_1\dots d\alpha_4}{7\alpha_1\dots\alpha_4^2}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

over some region in $\mathbb{R}^4$. Explicitly, $I_{3,4,1}$ is over the region defined by

$$\begin{aligned}
&1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,\\
&1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+\alpha_3 \leqslant 2\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, && \alpha_1+\alpha_4 \leqslant \varpi,\\
&\alpha_2+\alpha_3 \geqslant 1-2\varpi, && \alpha_1+\alpha_2+\alpha_3+\alpha_4 \leqslant 2\varpi.
\end{aligned}$$

The integral $I_{3,4,2}$ is over the region defined by

$$\begin{aligned}
&1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,\\
&1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+\alpha_3 \leqslant 2\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, && \alpha_1+\alpha_4 \leqslant \varpi,\\
&\alpha_2+\alpha_3 \geqslant 1-2\varpi, && \alpha_1+\alpha_2+\alpha_3+\alpha_4 \geqslant 1-\varpi.
\end{aligned}$$

We now consider $S_{3,3}$. We split the summation according to whether we can perform further Buchstab iterations and according to the size $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$, noting that terms with $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$ between $\mathfrak{z}_5$ and $\mathfrak{z}_6$ make a negligible contribution.

$$\begin{aligned}
S_{3,3} &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)\\
&= \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_5\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_5\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2>\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)\\
&\quad + \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_6}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\
&=: S_{3,3,1}+S_{3,3,2}+S_{3,3,3}+o(\#\mathcal{A}/\log X).
\end{aligned}$$

With $S_{3,3,1}$, we can decompose using two further Buchstab iterations, as we did with $S_{1,1,1}$. This results in the lower bound

$$\begin{aligned}
S_{3,3,1} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{3,3,1},\\
I_{3,3,1} &= \int\cdots\int^* \frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right),
\end{aligned}$$

with the integral $I_{3,3,1}$ over the region defined by

$$1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, \qquad 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1-\varpi, \quad \alpha_1+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_2+\alpha_3 \leqslant \varpi, \quad \alpha_1+\alpha_2+\alpha_3+\alpha_4 \leqslant 2\varpi.$$

With $S_{3,3,2}$, we split further depending on the size of $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$, giving

$$\begin{aligned} S_{3,3,2} &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) \\ &= \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant \mathfrak{z}_2}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4> \mathfrak{z}_3}} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X) \\ &=: S_{3,3,2,1} + S_{3,3,2,2} + o(\#\mathcal{A}/\log X). \end{aligned}$$

We directly lower bound $S_{3,3,2,1}$ and $S_{3,3,2,2}$, noting that $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^3) < X^{3\varpi n}$ (since $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_5$), so occurrences of the Buchstab function can be bounded by $4/7$. This gives

$$S_{3,3,2,1} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,3,2,1},$$
$$S_{3,3,2,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,3,2,2},$$

where both $I_{3,3,2,1}$ and $I_{3,3,2,2}$ are of the form

$$\int\cdots\int^* \frac{4}{7}\frac{d\alpha_1\,d\alpha_2\,d\alpha_3\,d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4^2}.$$

The integral $I_{3,3,2,1}$ is over the region defined by

$$1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, \quad 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,$$
$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$1-\varpi \leqslant \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, \quad \alpha_1+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_2+\alpha_3 \leqslant \varpi, \quad \alpha_2+\alpha_3+\alpha_4 \leqslant \varpi.$$

The integral $I_{3,3,2,2}$ is over the region defined by

$$1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, \quad 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,$$
$$1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, \quad \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,$$
$$1-\varpi \leqslant \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, \quad \alpha_1+\alpha_4 \geqslant 1-2\varpi,$$
$$\alpha_2+\alpha_3 \leqslant \varpi, \quad \alpha_1+\alpha_2+\alpha_3+\alpha_4 \leqslant 2\varpi,$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$\alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi.$$

The sum $S_{3,3,3}$ we lower bound directly, after splitting according to whether we can bound the Buchstab function by $4/7$ or not. We obtain

$$\begin{aligned} S_{3,3,3} &\geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}(I_{3,3,3,1}+I_{3,3,3,2}),\\ I_{3,3,3,1} &= \int\cdots\int^{*}\frac{4}{7}\frac{d\alpha_1\,d\alpha_2\,d\alpha_3\,d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4^2},\\ I_{3,3,3,2} &= \int\cdots\int^{*}\frac{d\alpha_1\,d\alpha_2\,d\alpha_3\,d\alpha_4}{\alpha_1\alpha_2\alpha_3\alpha_4^2}, \end{aligned}$$

where the integral $I_{3,3,3,1}$ is over the region defined by

$$\begin{aligned} &1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,\\ &1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,\\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, && \alpha_1+\alpha_4 \geqslant 1-2\varpi,\\ &\alpha_2+\alpha_3 \leqslant \varpi, && \alpha_1+\alpha_2+\alpha_3+\alpha_4 \geqslant 1-\varpi,\\ &\alpha_1+\alpha_2+\alpha_3+11\alpha_4/4 \leqslant 1. \end{aligned}$$

The integral $I_{3,3,3,2}$ is over the region defined by

$$\begin{aligned} &1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,\\ &1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,\\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, && \alpha_1+\alpha_4 \geqslant 1-2\varpi,\\ &\alpha_2+\alpha_3 \leqslant \varpi, && \alpha_1+\alpha_2+\alpha_3+\alpha_4 \geqslant 1-\varpi,\\ &\alpha_1+\alpha_2+\alpha_3+11\alpha_4/4 \geqslant 1. \end{aligned}$$

We now consider the sum $S_{3,2}$. We split the sum according to whether we can do further Buchstab iterations or not. This gives

$$\begin{aligned} S_{3,2} &= \sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4)\\ &= \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant \mathfrak{z}6}}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2> \mathfrak{z}6}}^{*} T(\mathfrak{p}_1\ldots\mathfrak{p}_4,\mathfrak{p}_4)\\ &=: S_{3,2,1}+S_{3,2,2}. \end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The terms in $S_{3,2,1}$ can undergo two more Buchstab iterations. As with $S_{1,1,1}$, we obtain

$$S_{3,2,1} \geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{3,2,1},$$
$$I_{3,2,1} = \int\cdots\int^{*}\frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right),$$

with the integral $I_{3,2,1}$ over the region defined by

$$\begin{aligned}
&1/2-\varpi\leqslant\alpha_1\leqslant\varpi, && 1-2\varpi-\alpha_1\leqslant\alpha_2\leqslant(1-\varpi-\alpha_1)/2,\\
&1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1, && \alpha_1+\alpha_2+2\alpha_3\leqslant 1,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4\leqslant 1, && \alpha_1+\alpha_3\geqslant 1-2\varpi,\\
&\alpha_2+\alpha_3\leqslant\varpi, && \alpha_1+\alpha_4\leqslant\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4\leqslant 1-\varpi.
\end{aligned}$$

We apply a direct bound to $S_{3,2,2}$, and note that since $\mathfrak{p}_1\mathfrak{p}_4, \mathfrak{p}_2\mathfrak{p}_3 \leqslant \mathfrak{z}_2$, we can bound occurrences of the Buchstab function by $4/7$. This gives

$$S_{3,2,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{3,2,2},$$
$$I_{3,2,2} = \int\cdots\int^{*}\frac{4d\alpha_1\dots d\alpha_4}{7\alpha_1\dots\alpha_4^2},$$

with the integral $I_{3,2,1}$ over the region defined by

$$\begin{aligned}
&1/2-\varpi\leqslant\alpha_1\leqslant\varpi, && 1-2\varpi-\alpha_1\leqslant\alpha_2\leqslant(1-\varpi-\alpha_1)/2,\\
&1-3\varpi\leqslant\alpha_4\leqslant\alpha_3\leqslant\alpha_2\leqslant\alpha_1, && \alpha_1+\alpha_2+2\alpha_3\leqslant 1,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4\leqslant 1, && \alpha_1+\alpha_3\geqslant 1-2\varpi,\\
&\alpha_2+\alpha_3\leqslant\varpi, && \alpha_1+\alpha_4\leqslant\varpi,\\
&\alpha_1+\alpha_2+\alpha_3+2\alpha_4\geqslant 1-\varpi.
\end{aligned}$$

Finally, we consider $S_{3,1}$. We split the summation according to whether we can perform further Buchstab iterations

$$\begin{aligned}
S_{3,1} &= \sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4}^{*} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)\\
&= \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2\leqslant\mathfrak{z}_6}}^{*} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2>\mathfrak{z}_6}}^{*} T(\mathfrak{p}_1\dots\mathfrak{p}_4,\mathfrak{p}_4)
\end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$=: S_{3,1,1} + S_{3,1,2}.$$

We split $S_{3,1,1}$ further depending on the size of $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$.

$$\begin{aligned} S_{3,1,1} &= \sideset{}{^*}\sum_{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{p}_2\dots\mathfrak{p}_4, \mathfrak{p}_4) \\ &= \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_2}} T(\mathfrak{p}_1\dots\mathfrak{p}_4, \mathfrak{p}_4) + \sideset{}{^*}\sum_{\substack{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4 \\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 > \mathfrak{z}_3}} T(\mathfrak{p}_1\dots\mathfrak{p}_4, \mathfrak{p}_4) + o(\#\mathcal{A}/\log X) \\ &=: S_{3,1,1,1} + S_{3,1,1,2} + o(\#\mathcal{A}/\log X). \end{aligned}$$

In both $S_{3,1,1,1}$ and $S_{3,1,1,2}$, we can perform two further Buchstab iterations. In $S_{3,1,1,1}$, we have $\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^2 \leqslant \mathfrak{z}_6$ and $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4 \leqslant \mathfrak{z}_2$, so $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^5) \leqslant X^n$, and it follows that we can bound occurrences of the Buchstab function by $4/7$. In $S_{3,1,1,2}$, we just bound the Buchstab function by 1. This gives

$$\begin{aligned} S_{3,1,1,1} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,1,1,1}, \\ S_{3,1,1,2} &\geqslant -(1+O(\epsilon))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,1,1,2}, \\ I_{3,1,1,1} &= \int\cdots\int^* \frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\frac{4}{7}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right), \\ I_{3,1,1,2} &= \int\cdots\int^* \frac{d\alpha_1\dots d\alpha_4}{\alpha_1\dots\alpha_4}\left(\frac{1}{1-3\varpi}\left(\log\left(\frac{\alpha_4}{1-3\varpi}\right)-1\right)+\frac{1}{\alpha_4}\right). \end{aligned}$$

Here the integral $I_{3,1,1,1}$ is over the region defined by

$$\begin{aligned} &1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2, \\ &1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+2\alpha_3 \leqslant 1, \\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1-\varpi, && \alpha_1+\alpha_3 \leqslant \varpi, \\ &\alpha_2+\alpha_3+\alpha_4 \leqslant \varpi. \end{aligned}$$

The integral $I_{3,1,1,2}$ is over the region defined by

$$\begin{aligned} &1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2, \\ &1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+2\alpha_3 \leqslant 1, \\ &\alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1-\varpi, && \alpha_1+\alpha_3 \leqslant \varpi, \\ &\alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi. \end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

The sum $S_{3,1,2}$ we lower bound directly, noting that $\mathfrak{p}_2\mathfrak{p}_4 \leqslant \mathfrak{p}_1\mathfrak{p}_3 \leqslant \mathfrak{z}_2$, so $N(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4^3) < X^{3\varpi n}$ and we can bound occurrences of the Buchstab function by $4/7$. This gives

$$S_{3,1,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{3,1,2},$$
$$I_{3,1,2} = \int\cdots\int^* \frac{4d\alpha_1\dots d\alpha_4}{7\alpha_1\dots\alpha_4^2},$$

with the integral $I_{3,1,2}$ over the region defined by

$$\begin{aligned}
&1/2-\varpi \leqslant \alpha_1 \leqslant \varpi, && 1-2\varpi-\alpha_1 \leqslant \alpha_2 \leqslant (1-\varpi-\alpha_1)/2,\\
&1-3\varpi \leqslant \alpha_4 \leqslant \alpha_3 \leqslant \alpha_2 \leqslant \alpha_1, && \alpha_1+\alpha_2+2\alpha_3 \leqslant 1,\\
&1-\varpi \leqslant \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \leqslant 1, && \alpha_1+\alpha_3 \leqslant \varpi.
\end{aligned}$$

This completes our lower bound for S_3.

A.4. The sum S_4. We split the sum S_4 first according to the size of $\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4$, then according to the size of $\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4$ or $\mathfrak{p}_2\mathfrak{p}_4$. This gives

$$\begin{aligned}
S_4 &= \sum_{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)\\
&= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\
&= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4\leqslant\mathfrak{z}_5}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4\leqslant\mathfrak{z}_2\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4>\mathfrak{z}_6}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)\\
&\quad + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3\\ \mathfrak{p}_2\mathfrak{p}_4\leqslant\mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4>\mathfrak{z}_3\\ \mathfrak{p}_2\mathfrak{p}_4>\mathfrak{z}_3}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\
&=: S_{4,1}+S_{4,2}+S_{4,3}+S_{4,4}+o(\#\mathcal{A}/\log X).
\end{aligned}$$

We perform no further decompositions and directly obtain a lower bound for the sums $S_{4,1}$ and $S_{4,2}$. This gives

$$S_{4,1} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{4,1}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$S_{4,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{4,2},$$

where $I_{4,1}$ and $I_{4,2}$ are integrals of the form

$$\int\cdots\int^{*}\omega\Big(\frac{\beta-\alpha_3}{\alpha_3}\Big)\omega\Big(\frac{1-\beta-\alpha_2-\alpha_4}{\alpha_4}\Big)\frac{d\beta\,d\alpha_2\,d\alpha_3\,d\alpha_4}{\alpha_2\alpha_3^2\alpha_4^2}. \tag{A.9}$$

(This arises from putting $N(\mathfrak{q}) = X^{n\beta-n\alpha_3}$, $N(\mathfrak{p}_i) = X^{n\alpha_i}$.) The integral $I_{4,1}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \leqslant 2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&\alpha_2+\alpha_3+\alpha_4 \leqslant \varpi, && 1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2.
\end{aligned}$$

The integral $I_{4,2}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \geqslant 1-\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&\alpha_2+\alpha_3+\alpha_4 \leqslant \varpi, && 1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2.
\end{aligned}$$

We split $S_{4,3}$ up further depending on the size of $\mathfrak{p}_2\mathfrak{p}_3$ and $\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4$. This gives

$$\begin{aligned}
S_{4,3} &= \sum_{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)\\
&= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\leqslant\mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3>\mathfrak{z}_3}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\
&= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\leqslant\mathfrak{z}_2\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4\leqslant\mathfrak{z}_5}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\leqslant\mathfrak{z}_2\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4>\mathfrak{z}_6}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)\\
&\quad + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3>\mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4\leqslant\mathfrak{z}_5}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3>\mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4>\mathfrak{z}_6}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)\\
&=: S_{4,3,1}+S_{4,3,2}+S_{4,3,3}+S_{4,3,4}+o(\#\mathcal{A}/\log X).
\end{aligned}$$

We now obtain lower bounds for $S_{4,3,1},\dots,S_{4,3,4}$ exactly as before. This gives

$$S_{4,3,1} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{4,3,1},$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$S_{4,3,2} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{4,3,2},$$
$$S_{4,3,3} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{4,3,3},$$
$$S_{4,3,4} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_{4,3,4}.$$

Here the integrals $I_{4,3,1},\dots,I_{4,3,4}$ are of the form (A.9). The integral $I_{4,3,1}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \leqslant 2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \leqslant \varpi,\\
&\alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi.
\end{aligned}$$

The integral $I_{4,3,2}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \geqslant 1-\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \leqslant \varpi,\\
&\alpha_2+\alpha_3+\alpha_4 \geqslant 1-2\varpi.
\end{aligned}$$

The integral $I_{4,3,3}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \leqslant 2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \geqslant 1-2\varpi,\\
&\alpha_2+\alpha_4 \leqslant \varpi.
\end{aligned}$$

The integral $I_{4,3,4}$ is over the region defined by

$$\begin{aligned}
&1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\
&\beta-\alpha_3+\alpha_2+\alpha_4 \geqslant 1-\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\
&1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \geqslant 1-2\varpi,\\
&\alpha_2+\alpha_4 \leqslant \varpi.
\end{aligned}$$

Finally, we consider $S_{4,4}$. We split $S_{4,4}$ according to the size of $\mathfrak{p}_2\mathfrak{p}_3$, and then $\mathfrak{q}\mathfrak{p}_4$ and $\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4$. This gives

$$S_{4,4} = \sideset{}{^*}\sum_{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3> \mathfrak{z}_3}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)$$

$$= \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3\leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3> \mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_4\leqslant \mathfrak{z}_2}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4)$$

$$+ \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3> \mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_4> \mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4\leqslant \mathfrak{z}_5}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + \sum_{\substack{\mathfrak{q},\mathfrak{p}_2,\mathfrak{p}_3,\mathfrak{p}_4\\ \mathfrak{p}_2\mathfrak{p}_3> \mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_4> \mathfrak{z}_3\\ \mathfrak{q}\mathfrak{p}_2\mathfrak{p}_4> \mathfrak{z}_6}}^{*} T(\mathfrak{q}\mathfrak{p}_2\mathfrak{p}_3\mathfrak{p}_4,\mathfrak{p}_4) + o(\#\mathcal{A}/\log X)$$

$$=: S_{4,4,1} + S_{4,4,2} + S_{4,4,3} + S_{4,4,4} + o(\#\mathcal{A}/\log X).$$

We then obtain lower bounds of $S_{4,4,i}$ exactly as before. This gives for each $i \in \{1,2,3,4\}$

$$S_{4,4,i} \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X} I_{4,4,i},$$

where $I_{4,4,i}$ is an integral of the form (A.9). Explicitly, $I_{4,4,1}$ is over the region

$$\begin{aligned} &1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\ &\alpha_2+\alpha_4 \geqslant 1-2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\ &1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \leqslant \varpi. \end{aligned}$$

The integral $I_{4,4,2}$ is over the region

$$\begin{aligned} &1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\ &\alpha_2+\alpha_4 \geqslant 1-2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\ &1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \geqslant 1-2\varpi,\\ &\beta-\alpha_3+\alpha_4 \leqslant \varpi. \end{aligned}$$

The integral $I_{4,4,3}$ is over the region

$$\begin{aligned} &1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\ &\alpha_2+\alpha_4 \geqslant 1-2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2,\\ &1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, && \alpha_2+\alpha_3 \geqslant 1-2\varpi,\\ &\beta-\alpha_3+\alpha_4 \geqslant 1-2\varpi, && \beta-\alpha_3+\alpha_2+\alpha_4 \leqslant 2\varpi. \end{aligned}$$

The integral $I_{4,4,4}$ is over the region

$$\begin{aligned} &1-2\varpi \leqslant \beta \leqslant 1/2+\epsilon, && 1-3\varpi \leqslant \alpha_3 \leqslant \beta/2,\\ &\alpha_2+\alpha_4 \geqslant 1-2\varpi, && 1-3\varpi \leqslant \alpha_4 \leqslant (1-\beta-\alpha_2)/2, \end{aligned}$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

$$1-\varpi-\beta \leqslant \alpha_2 \leqslant (1-\beta)/2, \qquad \alpha_2+\alpha_3 \geqslant 1-2\varpi,$$
$$\beta-\alpha_3+\alpha_4 \geqslant 1-2\varpi, \qquad \beta-\alpha_3+\alpha_2+\alpha_4 \geqslant 1-\varpi.$$

This completes our decomposition of the sum S_4.

A.5. The sums S_5 and S_6. The sums S_5 and S_6 require no further decompositions, and we obtain the lower bounds

$$S_5 \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_5,$$
$$S_6 \geqslant -(1+o(1))\frac{\mathfrak{S}\#\mathcal{A}}{n\log X}I_6,$$

where

$$I_5 = \int_{1-2\varpi}^{1/2+\epsilon}\int_{1-\varpi-\alpha_1}^{(1-\alpha_1)/2}\frac{d\alpha_1\,d\alpha_2}{\alpha_1\alpha_2(1-\alpha_1-\alpha_2)}, \tag{A.10}$$
$$I_6 = \int_{(1-\varpi)/3}^{\varpi}\int_{(1-\varpi-\alpha_1)/2}^{\alpha_1}\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)\frac{d\alpha_1\,d\alpha_2}{\alpha_1\alpha_2^2}. \tag{A.11}$$

A.6. Numerical conclusion. Putting everything together, we find that the above manipulations give a decomposition of the form of Proposition 6.6, namely

$$S(\mathcal{A},\mathfrak{z}_4) = \sum_{\mathcal{R}\in\mathcal{S}_1}\sum_{\mathfrak{d}}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})S(\mathcal{A}_{\mathfrak{d}},\mathfrak{z}_1) - \sum_{\mathcal{R}\in\mathcal{S}_2}\sum_{\mathfrak{d}}\mathbf{1}_{\mathcal{R}}(\mathfrak{d})S(\mathcal{A}_{\mathfrak{d}},\mathfrak{z}_1)$$
$$+\sum_{\mathcal{R}\in\mathcal{S}_3}\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{R}}(\mathfrak{a}) - \sum_{\mathcal{R}\in\mathcal{S}_4}\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{R}}(\mathfrak{a}) + \sum_{\mathcal{R}\in\mathcal{S}_5}\sum_{\mathfrak{a}\in\mathcal{A}}\mathbf{1}_{\mathcal{R}}(\mathfrak{a}),$$

for certain sets of polytopes $\mathcal{S}_1,\dots,\mathcal{S}_5$ satisfying the properties claimed in the proposition. Specifically, all terms coming from $\mathcal{S}_1$ and $\mathcal{S}_2$ can be evaluated using Proposition 6.2, and all terms coming from $\mathcal{S}_3$ and $\mathcal{S}_4$ can be evaluated using Proposition 6.1. All the terms corresponding to $\mathcal{S}_5$ are terms which we discard for a lower bound by positivity, corresponding to the lower bounds we obtained for the subsums of $S_1,\dots,S_5$. All the terms we have considered throughout the appendix (including those we discard or deal with using Propositions 6.1 and 6.2) can be viewed as sums of the form $\mathbf{1}_{\mathcal{R}}(\mathfrak{a})$ (potentially summing over $O(1)$ polytopes) since all terms are sums of integers with at most $1/(3\varpi-1)$ prime factors, with the only restrictions being on the size of these prime factors.

We are left to check the final estimate, namely that

$$\sum_{\mathcal{R}\in\mathcal{S}_5} I_{\mathcal{R}} < 0.99.$$

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

From our previous work, we see that

$$\sum_{\mathcal{R}\in\mathcal{S}_5} I_{\mathcal{R}} = I_1 + I_2 + I_3 + I_4 + I_5,$$

where

$$\begin{aligned}
I_1 &= I_{1,1,1} + I_{1,1,2} + I_{1,2,1} + I_{1,2,2},\\
I_2 &= I_{2,1,1} + I_{2,1,2} + I_{2,2} + I_{2,3} + I_{2,4} + I_{2,5,1} + I_{2,5,2} + I_{2,5,3},\\
I_3 &= I_{3,1,1,1} + I_{3,1,1,2} + I_{3,1,2} + I_{3,2,1} + I_{3,2,2} + I_{3,3,1} + I_{3,3,2,1} + I_{3,3,2,2}\\
&\quad + I_{3,3,3,1} + I_{3,3,3,2} + I_{3,4,1} + I_{3,4,2} + I_{3,5} + I_{3,6} + I_{3,7},\\
I_4 &= I_{4,1} + I_{4,2} + I_{4,3,1} + I_{4,3,2} + I_{4,3,3} + I_{4,3,4} + I_{4,4,1} + I_{4,4,2} + I_{4,4,3} + I_{4,4,4},
\end{aligned}$$

and I_5, I_6 are given by (A.10) and (A.11). In particular, we obtain the required result, provided $I_1 + I_2 + I_3 + I_4 + I_5 + I_6 < 1$. All the integrals appearing are in a suitably explicit form that they can be calculated numerically. The following table gives the result of these numerical estimates. A Mathematica © file performing these computations is available along with this article at https://arxiv.org/abs/1507.05080.

Integral	Numerical upper bound	Integral	Numerical upper bound
$I_{1,1,1}$	0.00393	$I_{3,3,3,1}$	0.02824
$I_{1,1,2}$	0.03341	$I_{3,3,3,2}$	0.00045
$I_{1,2,1}$	0.05488	$I_{3,4,1}$	0.00350
$I_{1,2,2}$	0.00098	$I_{3,4,2}$	0.01194
$I_{2,1,1}$	0.00370	$I_{3,5}$	0.00615
$I_{2,1,2}$	0.00769	$I_{3,6}$	0.00038
$I_{2,2}$	0.00011	$I_{3,7}$	0.00158
$I_{2,3}$	0.00147	$I_{4,1}$	0.00001
$I_{2,4}$	0.00623	$I_{4,2}$	0.02744
$I_{2,5,1}$	0.00614	$I_{4,3,1}$	0.00161
$I_{2,5,2}$	0.00118	$I_{4,3,2}$	0.09657
$I_{2,5,3}$	0.00289	$I_{4,3,3}$	0.14092
$I_{3,1,1,1}$	0.00388	$I_{4,3,4}$	0.00054
$I_{3,1,1,2}$	0.00546	$I_{4,4,1}$	0.05416
$I_{3,1,2}$	0.00437	$I_{4,4,2}$	0.00736
$I_{3,2,1}$	0.00277	$I_{4,4,3}$	0.00499
$I_{3,2,2}$	0.00578	$I_{4,4,4}$	0.06736
$I_{3,3,1}$	0.01363	I_5	0.14018
$I_{3,3,2,1}$	0.01524	I_6	0.22180
$I_{3,3,2,2}$	0.00085		

This gives a total bound of 0.98977 for $I_1 + \cdots + I_6$ which is less than 0.99, as desired.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press

Conflict of Interest: There are no conflicts of interest.

References

[1] P. T. Bateman and R. A. Horn, 'A heuristic asymptotic formula concerning the distribution of prime numbers', *Math. Comp.* **16** (1962), 363–367.

[2] B. J. Birch, 'Forms in many variables', *Proc. R. Soc. Ser. A* **265** (1961/1962), 245–263.

[3] J. W. S. Cassels, *An Introduction to the Geometry of Numbers*, Classics in Mathematics (Springer, Berlin, 1997), Corrected reprint of the 1971 edition.

[4] M. D. Coleman, 'A zero-free region for the Hecke L-functions', *Mathematika* **37**(2) (1990), 287–304.

[5] H. Davenport, 'On a principle of Lipschitz', *J. Lond. Math. Soc. (2)* **26** (1951), 179–183.

[6] H. Davenport, 'Indefinite quadratic forms in many variables. II', *Proc. Lond. Math. Soc. (3)* **8** (1958), 109–126.

[7] W. Duke, 'Some problems in multidimensional analytic number theory', *Acta Arith.* **52**(3) (1989), 203–228.

[8] J. Friedlander and H. Iwaniec, 'The polynomial $X^2 + Y^4$ captures its primes', *Ann. of Math. (2)* **148**(3) (1998), 945–1040.

[9] H. Halberstam and H.E. Richert, *Sieve Methods*, L.M.S. monographs (Academic Press, 1974).

[10] G. Harman, 'On the distribution of αp modulo one. II', *Proc. Lond. Math. Soc. (3)* **72**(2) (1996), 241–260.

[11] G. Harman, *Prime-detecting Sieves*, London Mathematical Society Monographs Series, 33 (Princeton University Press, Princeton, NJ, 2007).

[12] D. R. Heath-Brown, 'Diophantine approximation with square-free numbers', *Math. Z.* **187**(3) (1984), 335–344.

[13] D. R. Heath-Brown, 'Primes represented by $x^3 + 2y^3$', *Acta Math.* **186**(1) (2001), 1–84.

[14] D. R. Heath-Brown and X. Li, 'Prime values of $a^2 + p^4$', *Invent. Math.* **208**(2) (2017), 441–499.

[15] D. R. Heath-Brown and B. Z. Moroz, 'Primes represented by binary cubic forms', *Proc. Lond. Math. Soc. (3)* **84**(2) (2002), 257–288.

[16] D. R. Heath-Brown and B. Z. Moroz, 'On the representation of primes by cubic polynomials in two variables', *Proc. Lond. Math. Soc. (3)* **88**(2) (2004), 289–312.

[17] H. Iwaniec, 'Primes represented by quadratic polynomials in two variables', *Acta Arith.* **24** (1973/74), 435–459. Collection of articles dedicated to Carl Ludwig Siegel on the occasion of his seventy-fifth birthday, V.

[18] S. Lang, *Diophantine Geometry*, Interscience Tracts in Pure and Applied Mathematics, 11 (Interscience Publishers (a division of John Wiley & Sons), New York–London, 1962).

[19] J. Neukirch, *Algebraic Number Theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 322 (Springer, Berlin, 1999), Translated from the 1992 German original and with a note by Norbert Schappacher, with a foreword by G. Harder.

[20] E. C. Titchmarsh, *The Theory of the Riemann Zeta-function*, 2nd edn (The Clarendon Press, Oxford University Press, New York, 1986), edited and with a preface by D. R. Heath-Brown.

[21] A. Weiss, 'The least prime ideal', *J. Reine Angew. Math.* **338** (1983), 56–94.

https://doi.org/10.1017/fmp.2019.8 Published online by Cambridge University Press