J. Inst. Math. Jussieu (2025), **24**(4), 1463–1519
doi:10.1017/S1474748025000076

POLYNOMIALS REPRESENTED BY NORM FORMS VIA THE BETA SIEVE

ALEC SHUTE

School of Mathematics, University of Bristol, Bristol, BS8 1UG
(alec.shute@bristol.ac.uk)

(Received 17 January 2023; revised 23 January 2025; accepted 24 January 2025; first published online 25 February 2025)

Abstract A central question in Arithmetic geometry is to determine for which polynomials $f \in \mathbb{Z}[t]$ and which number fields K the Hasse principle holds for the affine equation $f(t) = \mathbf{N}_{K/\mathbb{Q}}(\mathbf{x}) \neq 0$. Whilst extensively studied in the literature, current results are largely limited to polynomials and number fields of low degree. In this paper, we establish the Hasse principle for a wide family of polynomials and number fields, including polynomials that are products of arbitrarily many linear, quadratic or cubic factors. The proof generalises an argument of Irving [27], which makes use of the beta sieve of Rosser and Iwaniec. As a further application of our sieve results, we prove new cases of a conjecture of Harpaz and Wittenberg on locally split values of polynomials over number fields, and discuss consequences for rational points in fibrations.

Contents

1 Introduction 1464

2 Main sieve results for binary forms 1469

3 Levels of distribution 1471

4 Application of the beta sieve 1477

5 Application to the Hasse principle 1495

6 Application to the Harpaz–Wittenberg conjecture 1504

2020 Mathematics subject classification: Primary 11D57, 14G05, 14G12

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

A The Brauer group for the equation $f(t) = \mathbf{N}(\mathbf{x}) \neq 0$ 1512

References 1518

1. Introduction

Let K be a number field of degree n, and let $f \in \mathbb{Z}[t]$ be a polynomial. A central problem in arithmetic geometry is to determine under what conditions f can take values equal to a norm of an element of K. In order to address this question, we take an integral basis $\omega_1,\dots,\omega_n$ for K, viewed as a vector space over $\mathbb{Q}$, and define the *norm form* as $\mathbf{N}(\mathbf{x}) = N_{K/\mathbb{Q}}(\omega_1 x_1 + \cdots + \omega_n x_n)$, where $N_{K/\mathbb{Q}}(\cdot)$ is the field norm. We then seek to understand when the equation

$$f(t) = \mathbf{N}(\mathbf{x}) \neq 0 \tag{1.1}$$

has a solution with $(t,x_1,\dots,x_n) \in \mathbb{Q}^{n+1}$. A necessary condition for solubility over $\mathbb{Q}$ is that (1.1) must have solutions in $\mathbb{R}^{n+1}$ and in $\mathbb{Q}_p^{n+1}$ for every prime p. We say that the *Hasse principle* holds if this condition alone is sufficient to guarantee existence of a solution to (1.1) over $\mathbb{Q}$.

Local to global questions for (1.1) have received much attention over the years. The first case to consider is when f is a nonzero constant polynomial. Here, the Hasse principle for (1.1) is known as the *Hasse norm principle*. More precisely, we say that the *Hasse norm principle* holds for the extension $K/\mathbb{Q}$ if $\mathbb{Q}^\times \cap N_{K/\mathbb{Q}}(I_K) = N_{K/\mathbb{Q}}(K^\times)$, where $I_K = \mathbb{A}_K^\times$ is the group of ideles of K. The Hasse norm principle has been extensively studied, beginning with the work of Hasse himself, who established that it holds for cyclic extensions $K/\mathbb{Q}$ (a result known as the *Hasse norm theorem*), but does not hold for certain biquadratic extensions, such as $K = \mathbb{Q}(\sqrt{13},\sqrt{-3})$. The Hasse norm principle is also known to hold if the degree of K is prime (Bartels [1]) or if the normal closure of K has Galois group S_n (Kunyavskiĭ and Voskresenskiĭ [32]) or A_n for $n \neq 4$ (Macedo [35]).

When $[K : \mathbb{Q}] = 2$ and f is irreducible of degree 3 or 4, (1.1) defines a Châtelet surface. There are now many known counterexamples to the Hasse principle for Châtelet surfaces, including one by Iskovskikh [28], which we discuss in more detail in Example 5.4. However, Colliot-Thélène, Sansuc and Swinnerton-Dyer [14] prove that the Brauer–Manin obstruction accounts for all failures of the Hasse principle. A similar result holds when f is an irreducible polynomial of degree at most 3 and $[K : \mathbb{Q}] = 3$, as proved by Colliot-Thélène and Salberger [8]. Both of these results make use of fibration and descent methods.

In the case when f is an irreducible quadratic and K is a quartic extension containing a root of f, the Hasse principle and weak approximation are known to hold for (1.1) thanks to the work of Browning and Heath-Brown [4]. This result was generalised by Derenthal, Smeets and Wei [16, Theorem 2] to prove that the Brauer–Manin obstruction is the only obstruction to the Hasse principle and weak approximation for irreducible quadratics f and arbitrary number fields K. Moreover, in [16, Theorem 4], they give an explicit description of the Brauer groups that can be obtained in this family.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Results when f is not irreducible have so far been limited to products of linear polynomials. Suppose that f takes the form

$$f(t) = c\prod_{i=1}^{r}(t-e_i)^{m_i}, \tag{1.2}$$

for some $c \in \mathbb{Q}^*, e_1, \dots, e_r \in \mathbb{Q}$ and $m_1, \dots, m_r \in \mathbb{N}$. When $r = 1$, the Brauer–Manin obstruction is the only obstruction to the Hasse principle and weak approximation for any smooth projective model of (1.1). This is a special case of the work of Colliot-Thélène and Sansuc [9] on principal homogeneous spaces under algebraic tori. Heath-Brown and Skorobogatov [26] treat the case $r = 2$ by combining descent methods with the Hardy–Littlewood circle method, under the assumption that $\gcd(m_1, m_2, \deg K) = 1$. This assumption was later removed by Colliot-Thélène, Harari and Skorobogatov [13]. Thanks to the work of Browning and Matthiesen [5], it is now settled that for any number field K and polynomial f of the form (1.2) (for arbitrary $r \geqslant 1$), the Brauer–Manin obstruction is the only obstruction to the Hasse principle and weak approximation for any smooth projective model of (1.1). Their result is inspired by additive combinatorics results of Green, Tao and Ziegler [20], [21], combined with vertical torsors introduced by Schindler and Skorobogatov [40].

In general, it has been conjectured by Colliot-Thélène [7] that all failures of the Hasse principle for any smooth projective model of (1.1) are explained by the Brauer–Manin obstruction. Assuming Schinzel's hypothesis, this holds true for f an arbitrary polynomial and $K/\mathbb{Q}$ a cyclic extension, as demonstrated by work of Colliot-Thélène and Swinnerton-Dyer on pencils of Severi–Brauer varieties [12]. Recently, Skorobogatov and Sofos also establish unconditionally that when $K/\mathbb{Q}$ is cyclic, (1.1) satisfies the Hasse principle for a positive proportion of polynomials f of degree d, when their coefficients are ordered by height [42, Theorem 1.3].

In [27], Irving introduces an entirely new approach to studying the Hasse principle for (1.1), which rests on sieve methods. Irving's main result [27, Theorem 1.1] states that if $f \in \mathbb{Z}[t]$ is an irreducible cubic, then the Hasse principle holds for (1.1) under the following assumptions:

(1) K satisfies the Hasse norm principle.

(2) There exists a prime $q \geqslant 7$ and a finite set of primes S, such that for all $p \notin S$, either $p \equiv 1 \pmod{q}$ or the inertia degrees of p in $K/\mathbb{Q}$ are coprime.

(3) The number field generated by f is not contained in the cyclotomic field $\mathbb{Q}(\zeta_q)$.

An example provided by Irving in [27] is the number field $\mathbb{Q}(\alpha)$, where α is a root of $x^q - 2$ and $q \geqslant 7$ is prime. We shall comment on this further in Example 5.9.

In this paper, we generalize Irving's arguments to establish the Hasse principle for a wide new family of polynomials and number fields. Our results cover for the first time polynomials of arbitrarily large degree which are not a product of linear factors. In fact, under suitable assumptions on K, we can deal with polynomials that are products of arbitrarily many linear, quadratic and cubic factors.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Throughout this paper, we let $\widehat{K}$ denote the Galois closure of K, and we let $G = \mathrm{Gal}(\widehat{K}/\mathbb{Q})$, viewed as a permutation group on n letters. We define

$$T(G) = \frac{1}{\#G}\#\{\sigma \in G : \text{ the cycle lengths of } \sigma \text{ are not coprime}\}. \tag{1.3}$$

We now state our main results.

Theorem 1.1. *Let K be a number field satisfying the Hasse norm principle. Let $f \in \mathbb{Z}[t]$ be a polynomial, all of whose irreducible factors have degree at most 2. Let k denote the number of distinct irreducible factors of f, and let j denote the number of distinct irreducible quadratic factors of f which generate a quadratic field contained in $\widehat{K}$. Suppose that $T(G) \leqslant \frac{0.39}{k+j+1}$. Then the Hasse principle holds for (1.1).*

In practice, the constant 0.39 can be improved slightly, particularly when the majority of the factors of f are linear, although it will always be less than $1/2$. The precise optimal constant is obtained by finding the maximal value of κ such that (4.55) holds.

When $G = S_n$, we shall see in Lemma 5.10 that $T(S_n) \to 0$ as $n \to \infty$, so in the setting of Theorem 1.1, we can establish the Hasse principle provided n is sufficiently large in terms of the degree of f. We illustrate this by treating the case when f is a product of two irreducible quadratics.

Corollary 1.2. *Let $f \in \mathbb{Z}[t]$ be a product of two quadratic polynomials such that the number field L generated by f is a biquadratic extension of $\mathbb{Q}$. Let K be a number field of degree n with $G = S_n$. Suppose that $L \cap \widehat{K} = \mathbb{Q}$. Then the Hasse principle holds for (1.1), provided that*

$$n \notin \{2,3,\dots,10,12,14,15,16,18,20,22,24,26,28,30,36,42,48\}.$$

We remark that without the assumption $L \cap \widehat{K} = \mathbb{Q}$, a similar result to Corollary 1.2 still holds, although a larger list of degrees n would need to be excluded. For example, if $L \cap \widehat{K}$ is quadratic, then the Hasse principle holds for (1.1) for all primes $n \geqslant 11$ and all integers $n > 90$, whilst if $L \cap \widehat{K} = L$, then the Hasse prinicple holds for all primes $n \geqslant 13$ and all integers $n > 150$.

We cannot hope to deal with all small values of n in Corollary 1.2. For example, the work of Iskovskikh [28] shows that the Hasse principle can fail when $n = 2$ (see Example 5.4). However, as we shall discuss in Appendix A, in the case $n \geqslant 3$, there is no Brauer–Manin obstruction to the Hasse principle, and so according to the conjecture of Colliot-Thélène mentioned above, we should expect the Hasse principle to hold.

Our second main result allows f to contain irreducible cubic factors but requires more restrictive assumptions on the number field K, more similar to Irving's setup in [27].

Theorem 1.3. *Let $f \in \mathbb{Z}[t]$ be a polynomial, all of whose irreducible factors have degree at most 3. Then the Hasse principle holds for (1.1) under the following assumptions for K.*

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

(1) *K satisfies the Hasse norm principle.*

(2) *The set $\mathscr{P}$ of primes p for which the inertia degrees of p in $K/\mathbb{Q}$ are not coprime satisfies Assumption 2.2.*

As an example, Assumption 2.2 is satisfied if there exists a prime q such that $\frac{\deg f+1}{q-1} \leqslant 0.32380$, and such that for all but finitely many primes $p \not\equiv 1 \pmod{q}$, the inertia degrees of p in $K/\mathbb{Q}$ are coprime. The constant 0.32380 appearing in Assumption 2.2 could likely be improved with more work, and in specific examples, the required bounds could be computed more precisely using (4.59). We remark that we have also dropped the assumption made in [27] that the number field generated by f is not contained in $\mathbb{Q}(\zeta_q)$. This assumption is not essential to Irving's argument, but allows for the treatment of smaller values of q. Reinserting this assumption and optimising (4.59), we could recover Irving's result from our work.

We prove Theorems 1.1 and 1.3 by applying the beta sieve of Rosser and Iwaniec [18, Theorem 11.13]. The main sieve results we obtain are stated in Theorems 2.1 and 2.3. These results in fact prove the existence of a solution to (1.1) with t arbitrarily close to a given adelic solution. Consequently, the above results could be extended to prove weak approximation for (1.1), provided that weak approximation holds for the norm one torus $\mathbf{N}(\mathbf{x}) = 1$. For example, the work of Kunyavskiĭ and Voskresenskiĭ [32] and Macedo [35] demonstrates that weak approximation for the norm one torus holds when $G = S_n$ or $G = A_n$ and $n \neq 4$, and so weak approximation holds in the setting of Corollary 1.2.

In Section 6, we find a second application of Theorem 2.1 to a conjecture of Harpaz and Wittenberg [23, Conjecture 9.1], which we restate in Conjecture 6.1 and henceforth refer to as the *Harpaz–Wittenberg conjecture*. The conjecture concerns a collection of number field extensions $L_i/k_i/k$, $i \in \{1,\dots,n\}$, where $k_i \cong k[t]/(P_i(t))$ for monic irreducible polynomials $P_i \in k[t]$. Roughly speaking, the conjecture predicts, under certain hypotheses, the existence of an element $t_0 \in k$ such that $P_1(t_0),\dots,P_n(t_0)$ are locally split (i.e., each place in k_i dividing $P_i(t_0)$ has a degree 1 place of L_i above it).

A major motivation for the conjecture is the development of the theory of rational points in fibrations. Given a fibration $\pi : X \to \mathbb{P}^1_k$, a natural question is to what extent we can deduce arithmetic information about X from arithmetic information about the fibres of π. A famous conjecture of Colliot-Thélène [7, p.174] predicts that for any smooth, proper, geometrically irreducible, rationally connected variety X over a number field k, the rational points $X(k)$ are dense in the Brauer–Manin set $X(\mathbb{A}_k)^{\mathrm{Br}}$. (In other words, the Brauer–Manin obstruction is the only obstruction to weak approximation.) Applied to this conjecture, the above question becomes whether density of $X(k)$ in $X(\mathbb{A}_k)^{\mathrm{Br}}$ follows from density of $X_c(k)$ in $X_c(\mathbb{A}_k)^{\mathrm{Br}}$ for a general fibre $X_c := \pi^{-1}(c)$ of π (see [24, Question 1.2]). Applications of the Harpaz–Wittenberg conjecture to this question are studied in [23] and [24].

Harpaz and Wittenberg [23, Section 9.2] demonstrate that their conjecture follows from the homogeneous version of Schinzel's hypothesis (commonly reffered to as (HH_1)) in the case of abelian extensions L_i/k_i, or more generally, almost abelian extensions (see [23, Definition 9.4]). Examples of almost abelian extensions include cubic extensions, and extensions of the form $k(c^{1/p})/k$ for $c \in k$ and p prime. The work of Heath-Brown

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

and Moroz [25] establishes (HH_1) for primes represented by binary cubic forms, from which the Harpaz–Wittenberg conjecture can be deduced in the case $k=\mathbb{Q}, n=1$ and $\deg P_1=3$. Using a geometric reformulation of [23, Conjecture 9.1], the authors establish their conjecture in low-degree cases – namely, when $\sum_{i=1}^n [k_i:k]\leqslant 2$ or $\sum_{i=1}^n [k_i:k]=3$ and $[L_i:k_i]=2$ for all i.

The Harpaz–Wittenberg conjecture is related to the study of polynomials represented by norm forms. As a consequence of the work of Matthiesen [36] on norms as products of linear polynomials, the Harpaz–Wittenberg conjecture holds in the case $k_1=\cdots=k_n=k=\mathbb{Q}$ [23, Theorem 9.14]. Similarly, we can deduce from [27, Theorem 1.1] that the Harpaz–Wittenberg conjecture holds in the case $n=2, k=\mathbb{Q}, k_1=K, k_2=\mathbb{Q}, L_1=K(2^{1/q})$ and $L_2=\mathbb{Q}(2^{1/q})$, where $q\geqslant 7$ is a prime such that $K\not\subseteq\mathbb{Q}(\zeta_q)$ [23, Theorem 9.15].

Besides the work of Matthiesen [36] for $k_1=\cdots=k_n=k=\mathbb{Q}$, the aforementioned results apply only to the case $n\leqslant 2$. In Section 6, we prove the following theorem, which establishes the Harpaz–Wittenberg conjecture in a new family of extensions $k_1/\mathbb{Q},\dots,k_n/\mathbb{Q}$, where n may be arbitrarily large, and each extension $k_i/\mathbb{Q}$ may have degree up to 3.

Theorem 1.4. *Let $n\geqslant 1$. Let $k=\mathbb{Q}$, and for $i\in\{1,\dots,n\}$, let k_i, M_i be linearly disjoint number fields over $\mathbb{Q}$. Let $L_i=M_ik_i$ be the compositum of k_i and M_i. Define*

$$T_i=\frac{1}{\#\mathrm{Gal}(\widehat{M_i}/\mathbb{Q})}\#\{\sigma\in\mathrm{Gal}(\widehat{M_i}/\mathbb{Q}):\sigma \text{ has no fixed point}\}. \tag{1.4}$$

Let $d=\sum_{i=1}^n[k_i:\mathbb{Q}]$. Then the Harpaz–Wittenberg conjecture holds in the following cases.

(1) *$[k_i:\mathbb{Q}]\leqslant 2$ for all $i\in\{1,\dots,n\}$ and $\sum_{i=1}^n T_i\leqslant 0.39/d$.*

(2) *$[k_i:\mathbb{Q}]\leqslant 3$ for all $i\in\{1,\dots,n\}$, and there exist primes q_i satisfying $\sum_{i=1}^n 1/(q_i-1)\leqslant 0.32380/d$, and integers t_i coprime to q_i, such that for all but finitely many primes $p\not\equiv t_i \pmod{q_i}$, there is a place of degree 1 in M_i above p.*

Corollary 1.5. *Let $q_1,\dots,q_n$ be distinct primes, and let $r_1,\dots,r_n\in\mathbb{N}$ be such that $g_i(x)=x^{q_i}-r_i$ is irreducible for all i. Let $M_i=\mathbb{Q}[x]/(g_i)$ and let k_i, L_i and d be as in Theorem 1.4. Suppose that one of the following holds:*

(1) *$[k_i:\mathbb{Q}]\leqslant 2$ for all $i\in\{1,\dots,n\}$ and $\sum_{i=1}^n 1/q_i\leqslant 0.39/d$,*

(2) *$[k_i:\mathbb{Q}]\leqslant 3$ for all $i\in\{1,\dots,n\}$ and $\sum_{i=1}^n 1/(q_i-1)\leqslant 0.32380/d$.*

Then the Harpaz–Wittenberg conjecture holds for $k=\mathbb{Q}$ and for such choices of k_i and L_i.

We remark that when applied to the setting of [23, Theorem 9.15], the above result requires a stronger bound on q. However, with a more careful optimisation of (4.58), it should be possible to recover [23, Theorem 9.15] from our approach.

By combining Theorem 1.4 with [23, Theorem 9.17] (with the choice $B=0, M''=\emptyset$ and $M'=\mathbb{P}^1_k\backslash U$), we obtain the following result about rational points in fibrations.

Theorem 1.6. *Let X be a smooth, proper, geometrically irreducible variety over $\mathbb{Q}$. Let $\pi: X\to\mathbb{P}^1_{\mathbb{Q}}$ be a dominant morphism whose general fibre is rationally connected. Let*

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$k_1,\dots,k_n$ denote the residue fields of the closed points of $\mathbb{P}^1_{\mathbb{Q}}$ above which π has nonsplit fibres, and let L_i/k_i be finite extensions which split these nonsplit fibres. Assume that

(1) *The smooth fibres of π satisfy the Hasse principle and weak approximation.*

(2) *The hypotheses of Theorem 1.4 hold.*

Then $X(\mathbb{Q})$ is dense in $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}(X)}$.

It would be interesting to investigate whether Condition (1) in Theorem 1.6 could be relaxed to the assumption that the smooth fibres $X_c(\mathbb{Q})$ are dense in $X_c(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}(X_c)}$, as in the setting of [24, Question 1.2] discussed above. This would require an extension of Theorem 1.4 to cover a stronger version of the Harpaz–Wittenberg conjecture, involving strong approximation of an auxiliary variety W off a finite set of places [24, Proposition 6.1]. Strong approximation of W was studied by Browning and Schindler [6], for example, who established [24, Question 1.2] in the case when the rank of π is at most 3, and at least one of its nonsplit fibres lies above a rational of $\mathbb{P}^1_{\mathbb{Q}}$.

2. Main sieve results for binary forms

Let $f \in \mathbb{Z}[x,y]$ be a non-constant binary form with nonzero discriminant. We write $f(x,y)$ as a product of distinct irreducible factors

$$f(x,y) = \prod_{i=0}^{m} f_i(x,y) \prod_{i=m+1}^{k} f_i(x,y), \tag{2.1}$$

where $f_i(x,y)$ are linear forms for $1 \leqslant i \leqslant m$, and forms of degree $k_i \geqslant 2$ for $m+1 \leqslant i \leqslant k$. If $y \mid f(x,y)$, then we define $f_0(x,y) = y$, and otherwise, we let $f_0(x,y) = 1$. Hence, we have $y \nmid f_i(x,y)$ for all $i \geqslant 1$.

For $i \in \{0,\dots,k\}$, we define

$$\nu_i(p) = \#\{[x:y] \in \mathbb{P}^1(\mathbb{F}_p) : f_i(x,y) \equiv 0 \pmod p\}, \tag{2.2}$$

$$\nu(p) = \#\{[x:y] \in \mathbb{P}^1(\mathbb{F}_p) : f(x,y) \equiv 0 \pmod p\}. \tag{2.3}$$

Let $\mathscr{P}$ be a set of primes, and let $\mathscr{P}_{\leqslant x} = \{p \in \mathscr{P} : p \leqslant x\}$. We denote by $\pi(x)$ the number of primes less than x. For all $i \in \{0,\dots,k\}$, we need to assume $\mathscr{P}$ has the following properties, for some $\alpha, \theta_i > 0$ and any $A \geqslant 1$:

$$\sum_{p \in \mathscr{P}_{\leqslant x}} 1 = \alpha \pi(x) \left(1 + O_A\left((\log x)^{-A}\right)\right), \tag{2.4}$$

$$\sum_{p \in \mathscr{P}_{\leqslant x}} \nu_i(p) = \alpha \theta_i \pi(x) \left(1 + O_A\left((\log x)^{-A}\right)\right). \tag{2.5}$$

The reason we require explicit error terms in (2.4) and (2.5) is so that the *sieve dimensions*, introduced in Section 4.2, exist. We note that for $i = 0$, we have $\theta_0 = 1$ if $f_0(x,y) = y$, and $\theta_0 = 0$ if $f_0(x,y) = 1$. Additionally, from (2.5), it follows that

$$\sum_{p \in \mathscr{P}_{\leqslant x}} \nu(p) = \alpha \theta \pi(x) \left(1 + O_A\left((\log x)^{-A}\right)\right), \tag{2.6}$$

where $\theta = \theta_0 + \cdots + \theta_k$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Let $\mathscr{B} \subseteq [-1,1]^2$ denote a bounded region whose boundary is a piecewise continuous simple closed curve of finite length. The perimeter of $\mathscr{B}$ will always be assumed to be bounded by some absolute constant C. In the applications in Sections 5 and 6, we shall make the choice

$$\mathscr{B} = \left\{(x,y) \in (0,1]^2 : \left|\frac{x}{y} - r\right| < \xi\right\}, \tag{2.7}$$

for a fixed real number $r > 0$ and a small parameter $\xi > 0$, and so we may choose $C = 4$, for example. We also define $\mathscr{B}N = \{(Nx,Ny) : (x,y) \in \mathscr{B}\}$.

Let Δ be an integer and let $a_0,b_0 \in \mathbb{Z}/\Delta\mathbb{Z}$. We now state the main sieve results which will be used in the proof of Theorem 1.1 and Theorem 1.3. They concern the sifting function

$$S(\mathscr{P},\mathscr{B},N) = \#\left\{(a,b) \in \mathscr{B}N \cap \mathbb{Z}^2 : \begin{array}{l} a \equiv a_0, b \equiv b_0 \pmod{\Delta} \\ p \mid f(a,b) \implies p \notin \mathscr{P} \end{array}\right\}. \tag{2.8}$$

Theorem 2.1. *Let $f(x,y)$ be a binary form consisting of distinct irreducible factors, all of degree at most 2. Then there exists a finite set of primes S_0, depending on f, such that the following holds:*

Let S be a finite set of primes containing S_0. Let Δ be an integer with only prime factors in S, and let $a_0,b_0 \in \mathbb{Z}/\Delta\mathbb{Z}$. Let $\mathscr{P}$ be a set of primes disjoint from S and satisfying (2.4) and (2.5) for some $\alpha,\theta_i > 0$. Assume that $\alpha\theta \leqslant 0.39$. Then $S(\mathscr{P},\mathscr{B},N) > 0$ for N sufficiently large.

We also have a similar sieve result when f may contain irreducible factors of degree up to 3, but with a less general sifting set $\mathscr{P}$, satisfying the following assumption.

Assumption 2.2. There exists a positive integer n, a finite set of primes S_0, primes $q_1,\dots,q_n$, and integers $t_1,\dots,t_n$ with $q_j \nmid t_j$ for all $j \in \{1,\dots,n\}$, such that

$$\mathscr{P}\backslash S_0 \subseteq \bigcup_{j=1}^{n} \{p \equiv t_j \pmod{q_j}\} \tag{2.9}$$

and

$$\deg f \sum_{j=1}^{n} \frac{1}{q_j - 1} \leqslant 0.32380.$$

Theorem 2.3. *Let $f(x,y)$ be a binary form consisting of distinct irreducible factors, all of degree at most 3. Then there exists a finite set of primes S_0, depending on f, such that the following holds:*

Let $\mathscr{P}$ be a set of primes satisfying Assumption 2.2 with the above choice of S_0. Let S be a finite set of primes containing S_0. Let Δ be an integer with only prime factors in S, and let $a_0,b_0 \in \mathbb{Z}/\Delta\mathbb{Z}$. Then $S(\mathscr{P},\mathscr{B},N) > 0$ for N sufficiently large.

For brevity, in the remainder of the paper, we shall denote the condition $a \equiv a_0 \pmod{\Delta}, b \equiv b_0 \pmod{\Delta}$ by $C(a,b)$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

3. Levels of distribution

Crucial to the success of the beta sieve in proving Theorems 2.1 and 2.3 is a good level of distribution result, which provides an approximation of the quantities

$$\#\{(a,b) \in \mathscr{B}N \cap \mathbb{Z}^2 : p \mid f_i(a,b), d \mid f(a,b)\}$$

by multiplicative functions, at least on average over p and d. (Here, and throughout this section, we keep the notation from Section 2.) In this section, we provide such an estimate, following similar arguments developed by Daniel [15, Lemma 3.3]. We slightly generalise the setup as follows:

Let g_1, g_2 be binary forms with nonzero discriminants. Throughout this section, we fix S, Δ, and $C(a,b)$, and assume that S contains all primes dividing the discriminants of g_1 and g_2. We allow all implied constants to depend only the degrees of g_1 and g_2 and a small positive constant ϵ, which for convenience, we allow to take different values at different points in the argument.

Let $\mathscr{R}$ be a bounded region of $\mathbb{R}^2$ whose boundary is a piecewise continuous simple closed curve of finite length. We denote by $\mathrm{Vol}(\mathscr{R})$ and $P(\mathscr{R})$ the volume and perimeter of $\mathscr{R}$, respectively. Let

$$A(d_1,d_2) = \#\{(a,b) \in \mathscr{R} \cap \mathbb{Z}^2 : C(a,b), d_1 \mid g_1(a,b), d_2 \mid g_2(a,b)\}, \tag{3.1}$$

$$\varrho(d_1,d_2) = \#\{(a,b) \pmod{d_1 d_2} : d_1 \mid g_1(a,b), d_2 \mid g_2(a,b)\}. \tag{3.2}$$

We define

$$r(d_1,d_2) = A(d_1,d_2) - \frac{\varrho(d_1,d_2)\,\mathrm{Vol}(\mathscr{R})}{d^2\Delta^2}. \tag{3.3}$$

In what follows, we let $d = d_1 d_2$, and we assume that $\gcd(d_1,d_2) = \gcd(d,\Delta) = 1$. The main aim of this section is to prove the following proposition.

Proposition 3.1. *Suppose that g_1 does not contain any linear factors. Then for any $D_1, D_2 > 0$ and any $\epsilon > 0$, we have*

$$\sum_{\substack{d_1 \leqslant D_1, d_2 \leqslant D_2 \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1}} \sup_{P(\mathscr{R}) \leqslant N} |r(d_1,d_2)| \ll (D_1 D_2)^{\epsilon} (D_1 D_2 + N(D_1 D_2)^{1/2} + N D_2).$$

As a corollary, we obtain the following level of distribution result.

Corollary 3.2. *Suppose that g_1 does not contain any linear factors. Let $\mathscr{B} \subseteq [-1,1]^2$ be as in Section 2, and let $\mathscr{R} = \mathscr{B}N$. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that for any $D_1, D_2 > 0$ with $D_2 \ll N^{1-\epsilon}$ and $D_1 D_2 \ll N^{2-\epsilon}$, we have*

$$\sum_{\substack{d_1 \leqslant D_1, d_2 \leqslant D_2 \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1}} \left| A(d_1,d_2) - \frac{N^2 \varrho(d_1,d_2)\,\mathrm{Vol}(\mathscr{B})}{d^2\Delta^2} \right| \ll N^{2-\delta}. \tag{3.4}$$

Proposition 3.1 and Corollary 3.2 are generalisations of Irving's results from [27, Section 3], which can be recovered by taking $g_1(x,y) = f(x,y)$ to be the cubic form Irving

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

considered, and $g_2(x,y) = yf(x,y)$. The method of proof is inspired by the pioneering work of Daniel on the divisor-sum problem for binary forms, which requires a similar level of distribution result (see [15, Lemma 3.3]). Daniel's argument is more delicate, keeping track of powers of $\log N$ in place of factors of N^{ϵ}, and Corollary 3.2 could be similarly refined, but this yields no advantage for our applications.

Before proceeding with the proof of Proposition 3.1, we recall the following standard lattice point counting result.

Lemma 3.3. *Let $\Lambda \subseteq \mathbb{R}^2$ be a full-rank lattice, and let $\mathscr{R} \subseteq \mathbb{R}^2$ be as defined before Proposition 3.1. Then*

$$\#(\mathscr{R} \cap \Lambda) = \frac{\mathrm{Vol}(\mathscr{R})}{\det \Lambda} + O\left(\frac{P(\mathscr{R})}{\lambda_1} + 1\right),$$

where λ_1 is the length of a shortest nonzero vector in Λ.

Proof. Let $\mathscr{F}$ be a fundamental domain of Λ. The translates $v + \mathscr{F}$ for $v \in \Lambda$ tile $\mathbb{R}^2$. Define sets

$$S^- = \{v \in \Lambda : (v + \mathscr{F}) \subseteq \mathscr{R}\}, \qquad S^+ = \{v \in \Lambda : (v + \mathscr{F}) \cap \mathscr{R} \neq \emptyset\}.$$

Then

$$\frac{\mathrm{Vol}(S^- + \Lambda)}{\det(\Lambda)} = \#S^- \leqslant \#(\mathscr{R} \cap \Lambda) \leqslant \#S^+ = \frac{\mathrm{Vol}(S^+ + \Lambda)}{\det(\Lambda)}.$$

Moreover, $S^- + \Lambda \subseteq \mathscr{R} \subseteq S^+ + \Lambda$, so $\mathrm{Vol}(S^- + \Lambda) \leqslant \mathrm{Vol}(\mathscr{R}) \leqslant \mathrm{Vol}(S^+ + \Lambda)$. Therefore,

$$\left|\#(\mathscr{R} \cap \Lambda) - \frac{\mathrm{Vol}(\mathscr{R})}{\det(\Lambda)}\right| \leqslant \#S^+ - \#S^-.$$

However, $S^+ - S^- = \{v \in \Lambda : (v + \mathscr{F}) \cap \partial\mathscr{R} \neq \emptyset\}$, where $\partial\mathscr{R}$ denotes the boundary of $\mathscr{R}$. Each segment of $\partial\mathscr{R}$ of length λ_1 can intersect at most four translates of $\mathscr{F}$. Therefore, $S^+ - S^- \ll P(\mathscr{R})/\lambda_1 + 1$, as required. □

We now commence with the proof of Proposition 3.1. We introduce the quantities $R^*(d_1,d_2), \varrho^*(d_1,d_2)$ and $r^*(d_1,d_2)$ which are defined similarly to $A(d_1,d_2), \varrho(d_1,d_2)$ and $r(d_1,d_2)$ but with the added condition $\gcd(a,b,d) = 1$. We write $(a_1,b_1) \sim (a_2,b_2)$ if there exists an integer λ such that $(a_1,b_1) \equiv \lambda(a_2,b_2) \pmod{d}$. This forms an equivalence relation on points $(a,b) \in \mathbb{Z}^2$ with $\gcd(a,b,d) = 1$. Moreover, the properties $g_1(a,b) \equiv 0 \pmod{d_1}$ and $g_2(a,b) \equiv 0 \pmod{d_2}$ are preserved under this equivalence. We may therefore define

$$\mathscr{U}(d_1,d_2) = \left\{ a,b \pmod{d} : \begin{array}{l} \gcd(a,b,d) = 1 \\ d_1 \mid g_1(a,b), d_2 \mid g_2(a,b) \end{array} \right\} \Big/ \sim .$$

For $\mathscr{C} \in \mathscr{U}(d_1,d_2)$, we define

$$\Lambda(\mathscr{C}) = \{y \in \mathbb{Z}^2 : y \equiv \lambda(a,b) \pmod{d} \text{ for some } (a,b) \in \mathscr{C} \text{ and some } \lambda \in \mathbb{Z}\}.$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

It is easy to check that $\Lambda(\mathscr{C})$ is a lattice in $\mathbb{Z}^2$, and its set of primitive points is $\mathscr{C}$. For $e \in \mathbb{Z}$, we define

$$\Lambda(\mathscr{C},e) = \{(a,b) \in \Lambda(\mathscr{C}) : e \mid \gcd(a,b)\}.$$

By Möbius inversion, we have

$$R^*(d_1,d_2) = \sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \sum_{e \mid d} \mu(e) \#\{(a,b) \in \mathscr{R} \cap \Lambda(\mathscr{C},e) : C(a,b)\}.$$

Since $\gcd(d,\Delta) = 1$, the set $\{(a,b) \in \Lambda(\mathscr{C},e) : C(a,b)\}$ is a coset of the lattice $\Lambda(\mathscr{C},e\Delta)$, which has determinant $de\Delta^2$. Therefore, by Lemma 3.3,

$$R^*(d_1,d_2) = \sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \sum_{e \mid d} \mu(e) \left(\frac{\mathrm{Vol}(\mathscr{R})}{de\Delta^2} + O\left(1 + \frac{P(\mathscr{R})}{\lambda_1(\mathscr{C})}\right)\right), \tag{3.5}$$

where $\lambda_1(\mathscr{C})$ denotes the length of the shortest nonzero vector in $\Lambda(\mathscr{C})$.

Each equivalence class $\mathscr{C} \in \mathscr{U}(d_1,d_2)$ consists of $\varphi(d)$ elements, and so

$$\sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \sum_{e \mid d} \frac{\mu(e)}{e} = \sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \frac{\varphi(d)}{d} = \frac{\varrho^*(d_1,d_2)}{d}.$$

Moreover, we have $\#\mathscr{U}(d_1,d_2) \ll d^\epsilon$, as we now explain. We observe that $\#\mathscr{U}(d_1,d_2) = \varrho^*(d_1,d_2)/\varphi(d)$, and ϱ^* is multiplicative by the Chinese remainder theorem. For primes $p \notin S$, we may apply Hensel's lemma to show that $\varrho^*(p^e,1), \varrho^*(1,p^e) = O(p^e)$ for any integer $e \geqslant 1$. Therefore, by the trivial bound for the divisor function [22, Section 18.1], we conclude that

$$\#\mathscr{U}(d_1,d_2) = \frac{\varrho^*(d_1,d_2)}{\varphi(d)} \ll \frac{d^{1+\epsilon}}{\varphi(d)} \ll d^\epsilon. \tag{3.6}$$

Applying (3.5), and (3.6), we obtain

$$\begin{aligned} &\sum_{\substack{d_1 \leqslant D_1, d_2 \leqslant D_2 \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1}} \sup_{P(\mathscr{R}) \leqslant N} |r^*(d_1,d_2)| \\ &\ll_\epsilon (D_1D_2)^\epsilon \left(D_1D_2 + N \sum_{\substack{d_1 \leqslant D_1, d_2 \leqslant D_2 \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1}} \sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \lambda_1(\mathscr{C})^{-1} \right). \end{aligned} \tag{3.7}$$

Let $v_1(\mathscr{C})$ denote a shortest nonzero vector of $\Lambda(\mathscr{C})$, and let $\|\cdot\|$ be the usual Euclidean norm. Then $\|v_1(\mathscr{C})\|^2 \ll |\det \Lambda(\mathscr{C})| = d \leqslant D_1D_2$. Therefore,

$$\sum_{\substack{d_1 \leqslant D_1, d_2 \leqslant D_2 \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1}} \sum_{\mathscr{C} \in \mathscr{U}(d_1,d_2)} \lambda_1(\mathscr{C})^{-1} \ll \sum_{0 < a^2+b^2 \ll D_1D_2} \frac{M(a,b)}{\sqrt{a^2+b^2}}, \tag{3.8}$$

where

$$M(a,b) = \#\left\{ \begin{matrix} d_1 \leqslant D_1, d_2 \leqslant D_2, \mathscr{C} \in \mathscr{U}(d_1,d_2) : \\ \gcd(d_1,d_2) = \gcd(d,\Delta) = 1, v_1(\mathscr{C}) = (a,b) \end{matrix} \right\}.$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

For any d_1,d_2 enumerated by $M(a,b)$, we have $d_1 \mid g_1(a,b)$ and $d_2 \mid g_2(a,b)$, so

$$M(a,b) \leqslant \#\{d_1 \leqslant D_1, d_2 \leqslant D_2 : d_1 \mid g_1(a,b), d_2 \mid g_2(a,b)\}.$$

Since g_1 contains no linear factors, we know that $g_1(a,b) \neq 0$ whenever $(a,b) \neq (0,0)$. Suppose in addition that $g_2(a,b) \neq 0$. Then by the trivial bound for the divisor function, we have $M(a,b) \ll (D_1D_2)^{\epsilon}$. We deduce that

$$\begin{aligned}\sum_{\substack{0<a^2+b^2\ll D_1D_2\\ g_2(a,b)\neq 0}} \frac{M(a,b)}{\sqrt{a^2+b^2}} &\ll (D_1D_2)^{\epsilon} \sum_{0<a^2+b^2\ll D_1D_2} \frac{1}{\sqrt{a^2+b^2}}\\ &\ll (D_1D_2)^{1/2+\epsilon}.\end{aligned}$$

Now suppose that $g_2(a,b) = 0$. Then as above, we have $O(D_1^{\epsilon})$ choices for d_1, but now D_2 choices for d_2, so that $M(a,b) \ll D_1^{\epsilon}D_2$. We obtain

$$\sum_{\substack{0<a^2+b^2\ll D_1D_2\\ g_2(a,b)=0}} \frac{M(a,b)}{\sqrt{a^2+b^2}} \ll D_1^{\epsilon}D_2 \sum_{\substack{0<a^2+b^2\ll D_1D_2\\ g_2(a,b)=0}} \frac{1}{\sqrt{a^2+b^2}}.$$

For a fixed $b \neq 0$, $g_2(a,b)$ is a nonzero polynomial in a, and so has $O(1)$ roots. Therefore,

$$\begin{aligned}\sum_{\substack{0<a^2+b^2\ll D_1D_2\\ g_2(a,b)=0}} \frac{1}{\sqrt{a^2+b^2}} &= \sum_{\substack{0<a^2+b^2\ll D_1D_2\\ b\neq 0\\ g_2(a,b)=0}} \frac{1}{\sqrt{a^2+b^2}} + \sum_{\substack{0<a^2\ll D_1D_2\\ g_2(a,0)=0}} \frac{1}{a}\\ &\ll \sum_{b\ll\sqrt{D_1D_2}} \frac{1}{b} + \sum_{a\ll\sqrt{D_1D_2}} \frac{1}{a}\\ &\ll (D_1D_2)^{\epsilon}.\end{aligned}$$

To summarize, we have established the following generalisation of [27, Lemma 3.2].

Lemma 3.4. *Suppose that g_1 does not contain any linear factors. Then for any $D_1,D_2 > 0$ and any $\epsilon > 0$, we have*

$$\sum_{\substack{d_1\leqslant D_1, d_2\leqslant D_2\\ \gcd(d_1,d_2)=\gcd(d,\Delta)=1}} \sup_{P(\mathscr{R})\leqslant N} |r^*(d_1,d_2)| \ll (D_1D_2)^{\epsilon}(D_1D_2 + N(D_1D_2)^{1/2} + ND_2).$$

Now we remove the restriction $\gcd(a,b,d) = 1$. Below, we write $A(d_1,d_2) = A(\mathscr{R},d_1,d_2;C(a,b))$ in order to make the dependence on $\mathscr{R}$ and $C(a,b)$ clear. Let $k_1 = \deg g_1$ and $k_2 = \deg g_2$. We work with multiplicative functions ψ_k for $k = k_1$ and $k = k_2$, which map prime powers p^r to $p^{\lceil r/k\rceil}$. We follow the same argument as Irving, but with ψ_{k_1},ψ_{k_2} in place of ψ_3,ψ_4. The motivation for this definition of ψ_k comes from the fact that for any integers $d,e,k \geqslant 1$ with $e \mid \psi_k(d)$, and for any prime p, we have

$$p \mid \frac{\psi_k(d)}{e} \iff p \mid \frac{d}{\gcd(d,e^k)}. \tag{3.9}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Since $\gcd(d_1,d_2)=1$, we have

$$A(\mathscr{R},d_1,d_2;C(a,b)) = \sum_{\substack{e_1\mid\psi_{k_1}(d_1)\\ e_2\mid\psi_{k_2}(d_2)}} N(d_1,d_2,e_1,e_2), \tag{3.10}$$

where

$$N(d_1,d_2,e_1,e_2) = \#\left\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2 : \begin{array}{l} C(a,b), d_1 \mid g_1(a,b), d_2 \mid g_2(a,b), \\ \gcd(a,b,\psi_{k_1}(d_1)\psi_{k_2}(d_2)) = e_1e_2 \end{array}\right\}. \tag{3.11}$$

We make a change of variables $a' = a/e_1e_2, b' = b/e_1e_2$ in (3.11). Let $\overline{e_1e_2}$ denote the multiplicative inverse of e_1e_2 modulo Δ, which exists due to the assumption $\gcd(d_1d_2,\Delta)=1$. The congruence condition $C(a,b)$ is equivalent to the congruence condition $a' \equiv \overline{e_1e_2}a_0 \pmod{\Delta}$ and $b' \equiv \overline{e_1e_2}b_0 \pmod{\Delta}$, which we denote by $C'(a',b')$. We have

$$\begin{aligned} d_1 \mid g_1(a,b) &\iff d_1 \mid (e_1e_2)^{k_1} g_1(a',b') \\ &\iff d_1 \mid e_1^{k_1} g_1(a',b') \\ &\iff \frac{d_1}{\gcd(d_1,e_1^{k_1})} \mid g_1(a',b'), \end{aligned}$$

and similarly for $d_2 \mid g_2(a,b)$. For convenience, we define

$$f_1 = \frac{d_1}{\gcd(d_1,e_1^{k_1})}, \quad f_2 = \frac{d_2}{\gcd(d_2,e_2^{k_2})}.$$

Changing notation from a',b' back to a,b, we deduce that $N(d_1,d_2,e_1,e_2)$ can be rewritten as

$$\begin{aligned} &\#\left\{(a,b)\in\mathscr{R}/(e_1e_2)\cap\mathbb{Z}^2 : \begin{array}{l} C'(a,b), f_1 \mid g_1(a,b), f_2 \mid g_2(a,b), \\ \gcd(a,b,\psi_{k_1}(d_1)\psi_{k_2}(d_2)/e_1e_2) = 1 \end{array}\right\} \\ &= \#\left\{(a,b)\in\mathscr{R}/(e_1e_2)\cap\mathbb{Z}^2 : \begin{array}{l} C'(a,b), f_1 \mid g_1(a,b), f_2 \mid g_2(a,b), \\ \gcd(a,b,f_1f_2) = 1 \end{array}\right\}. \\ &= R^*(\mathscr{R}/(e_1e_2),f_1,f_2;C'(a,b)). \end{aligned} \tag{3.12}$$

The above arguments, but with the congruence conditions removed, and with the specific choice $\mathscr{R}=[0,d_1d_2]^2$ also demonstrate that

$$\begin{aligned} \varrho(d_1,d_2) &= \sum_{\substack{e_1\mid\psi_{k_1}(d_1)\\ e_2\mid\psi_{k_2}(d_2)}} \#\left\{(a,b)\in\mathscr{R}/(e_1e_2)\cap\mathbb{Z}^2 : \begin{array}{l} f_1 \mid g_1(a,b), f_2 \mid g_2(a,b), \\ \gcd(a,b,\psi_{k_1}(d_1)\psi_{k_2}(d_2)/e_1e_2) = 1 \end{array}\right\} \\ &= \sum_{\substack{e_1\mid\psi_{k_1}(d_1)\\ e_2\mid\psi_{k_2}(d_2)}} \left(\frac{d_1d_2}{e_1e_2f_1f_2}\right)^2 \varrho^*(f_1,f_2). \end{aligned} \tag{3.13}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

We denote the quantity

$$R^*(\mathscr{R}/(e_1e_2),f_1,f_2;C'(a,b))-\frac{\mathrm{Vol}(\mathscr{R}/(e_1e_2))\varrho^*(f_1,f_2)}{(f_1f_2\Delta)^2}$$

by $E(e_1,e_2,f_1,f_2)$. Combining (3.10), (3.12) and (3.13), we have

$$\sum_{\substack{d_1\leqslant D_1,d_2\leqslant D_2\\(d_1,d_2)=(d_1d_2,\Delta)=1}}\sup_{P(\mathscr{R})\leqslant N}|r(d_1,d_2)|=\sum_{\substack{d_1\leqslant D_1,d_2\leqslant D_2\\(d_1,d_2)=(d_1d_2,\Delta)=1}}\sup_{P(\mathscr{R})\leqslant N}\sum_{\substack{e_1|\psi_{k_1}(d_1)\\e_2|\psi_{k_2}(d_2)}}|E(e_1,e_2,f_1,f_2)| \tag{3.14}$$

$$\leqslant\sum_{e_1\leqslant D_1,e_2\leqslant D_2}\sum_{\substack{f_1\leqslant D_1/e_1,f_2\leqslant D_2/e_2\\(f_1,f_2)=(f_1f_2,\Delta)=1}}\delta(e_1,f_1)\delta(e_2,f_2)\sup_{P(\mathscr{R})\leqslant N}|E(e_1,e_2,f_1,f_2)|, \tag{3.15}$$

where for integers $e,f,k,D\geqslant 1$, we have defined

$$\delta(e,f)=\#\left\{d\leqslant D:e\mid\psi_k(d),f=\frac{d}{\gcd(d,e^k)}\right\}.$$

We claim that $\delta(e,f)\ll e^\epsilon$. To see this, suppose that p is a prime and let $r=\nu_p(d),s=\nu_p(e)$ and $t=\nu_p(f)$. There is a unique choice of r for a given k,s and t provided that $t>0$ – namely, $r=ks+t$. If $t=0$, then we deduce from $f=d/\gcd(d,e^k)$ that $r\leqslant ks$. Taking a product over primes, we conclude that each d enumerated by $\delta(e,f)$ is a divisor of e^k multiplied by a quantity that is uniquely determined by e and f. The claim follows since the number of divisors of e^k is $O(e^\epsilon)$. In our situation, where $e_1\leqslant D_1$ and $e_2\leqslant D_2$, we obtain $\delta(e_1,f_1)\delta(e_2,f_2)\ll(D_1D_2)^\epsilon$. Therefore, applying Lemma 3.4 for each choice of e_1,e_2 in (3.15), we conclude that

$$\sum_{\substack{d_1\leqslant D_1,d_2\leqslant D_2\\(d_1,d_2)=(d_1d_2,\Delta)=1}}\sup_{P(\mathscr{R}\leqslant N)}|r(d_1,d_2)|\ll(D_1D_2)^\epsilon\sum_{\substack{e_1\leqslant D_1\\e_2\leqslant D_2}}\left(\frac{D_1D_2}{e_1e_2}+N\left(\frac{D_1D_2}{e_1e_2}\right)^{1/2}+\frac{ND_2}{e_2}\right)$$
$$\ll(D_1D_2)^\epsilon\left(D_1D_2+N(D_1D_2)^{1/2}+ND_2\right),$$

which completes the proof of Proposition 3.1.

Remark 3.5. If $g_2(a,b)\neq 0$ for all $(a,b)\neq(0,0)$, then we do not need to consider the case $g_2(a,b)=0$ in the analysis of the sum in (3.8), and so in our final level of distribution result, we do not require the assumption $D_2\ll N^{1-\epsilon}$.

When $g_1(a,b)$ does contain linear factors, we can still obtain a basic level of distribution result from the above argument using the trivial estimate $\lambda_1(\mathscr{C})^{-1}\leqslant 1$ in (3.7). This establishes the following lemma.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Lemma 3.6. *Let g_1,g_2 be arbitrary binary forms with nonzero discriminant. Then for any $\epsilon>0$, there exists $\delta>0$ such that for any $D_1,D_2>0$ with $D_1D_2\ll N^{1-\epsilon}$, we have*

$$\sum_{\substack{d_1\leqslant D_1,d_2\leqslant D_2\\ \gcd(d_1,d_2)=\gcd(d,\Delta)=1}}\sup_{P(\mathscr{R})\leqslant N}|r(d_1,d_2)|\ll N^{2-\delta}.$$

4. Application of the beta sieve

In this section, we prove Theorem 2.1 and Theorem 2.3 by combining the level of distribution results from Section 3 with the beta sieve of Rosser and Iwaniec [18, Theorem 11.12]. We state the precise version of this theorem we need in Theorem 4.1.

We recall some of the notation from Section 2. We fix a region $\mathscr{R}=\mathscr{B}N$ for some $\mathscr{B}\subseteq[-1,1]^2$ as in Section 2. Then $\mathscr{R}$ has volume $\gg N^2$ and perimeter $\ll N$. Let d denote the largest degree among the irreducible factors of f. (We specialise to the cases $d=2$ and $d=3$ later. The reason our methods are unable to deal with larger values of d is explained in Remark 4.5.) Then there exists $x\ll N^d$ such that the largest prime factor of $f(a,b)$ for $(a,b)\in\mathscr{R}\cap\mathbb{Z}^2$ is strictly less than x. Let S be a finite set of primes, including all primes dividing the discriminant of $f(x,y)$. In Section 4.6, we also append to S all primes bounded by some constant P_1. Let Δ be an integer with only prime factors in S. Without loss of generality, we may assume every prime in S divides Δ because taking a multiple of Δ can only decrease the sifting function $S(\mathscr{P},\mathscr{B},N)$ from (4.2). Additionally, by taking an appropriate multiple of Δ and appropriate lifts of a_0,b_0, we may assume that $\nu_p(f(a_0,b_0))<\nu_p(\Delta)$ for all $p\in S$.

All implied constants in this section are allowed to depend on Δ. Let $\mathscr{P}$ be a set of primes disjoint from S satisfying (2.4) and (2.5). We also define $\mathscr{P}'$ to be the set of primes not in $\mathscr{P}\cup S$. Let $P(x)$ denote the product of primes in $\mathscr{P}_{<x}$, and similarly for $P'(x)$. We also define $X=\mathrm{Vol}(\mathscr{R})/\Delta^2$.

For a sequence of non-negative real numbers $\mathscr{A}=(a_n)$, and a parameter $z\geqslant 1$, we define the sifting function

$$S(\mathscr{A},\mathscr{P},z)=\sum_{\gcd(n,P(z))=1}a_n.$$

We make the choice

$$a_n=\#\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2:C(a,b),f(a,b)=n\},\tag{4.1}$$

so that

$$\begin{aligned}S(\mathscr{A},\mathscr{P},x)&=\#\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2:C(a,b),\gcd(f(a,b),P(x))=1\}\\&=S(\mathscr{B},\mathscr{P},N),\end{aligned}\tag{4.2}$$

where $S(\mathscr{B},\mathscr{P},N)$ is as defined in (2.8). Our aim is to prove that $S(\mathscr{A},\mathscr{P},x)>0$ for sufficiently large N (which may depend on Δ). For a prime $p\in\mathscr{P}$ and for any $i\in\{0,\dots,k\}$, we also consider the sequences $\mathscr{A}_p,\mathscr{A}_p^{(i)}$ defined similarly to (4.1) but with the additional conditions $p\mid f(a,b),p\mid f_i(a,b)$, respectively, so that

$$S(\mathscr{A}_p,\mathscr{P},p)=\#\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2:C(a,b),p\mid f(a,b),\gcd(f(a,b),P(p))=1\},$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$$S(\mathscr{A}_p^{(i)},\mathscr{P},p) = \#\{(a,b) \in \mathscr{R} \cap \mathbb{Z}^2 : C(a,b), p \mid f_i(a,b), \gcd(f(a,b),P(p)) = 1\}.$$

Using the Buchstab identity, we have

$$S(\mathscr{A},\mathscr{P},x) = S(\mathscr{A},\mathscr{P},N^\gamma) - \sum_{\substack{N^\gamma \leqslant p < x \\ p \in \mathscr{P}}} S(\mathscr{A}_p,\mathscr{P},p),$$

for a parameter $\gamma \in (0,1)$ to be chosen later. We denote $S(\mathscr{A},\mathscr{P},N^\gamma)$ by S_1. If $p \mid f(a,b)$, then $p \mid f_i(a,b)$ for some i. Therefore, we have the decomposition

$$S(\mathscr{A},\mathscr{P},x) \geqslant S_1 - \sum_{i=0}^{m} S_2^{(i)} - \sum_{i=m+1}^{k} \left(S_3^{(i)} + S_4^{(i)}\right),$$

where

$$\begin{aligned} S_2^{(i)} &= \sum_{\substack{N^\gamma \leqslant p \ll N \\ p \in \mathscr{P}}} S(\mathscr{A}_p^{(i)},\mathscr{P},p), \\ S_3^{(i)} &= \sum_{\substack{N^\gamma \leqslant p < N^{\beta_i} \\ p \in \mathscr{P}}} S(\mathscr{A}_p^{(i)},\mathscr{P},p), \\ S_4^{(i)} &= \sum_{\substack{N^{\beta_i} \leqslant p < x \\ p \in \mathscr{P}}} S(\mathscr{A}_p^{(i)},\mathscr{P},p), \end{aligned} \tag{4.3}$$

for parameters $\beta_i \geqslant \gamma$ to be chosen later.

4.1. The beta sieve

Like most combinatorial sieves, the beta sieve provides a mechanism to estimate sifting functions of the form $S(\mathscr{A},\mathscr{P},z)$ given arithmetic information about the related quantities $|\mathscr{A}_d| := \sum_{d|n} a_n$ for squarefree integers d. More specifically, we require an approximation $|\mathscr{A}_d| = |\mathscr{A}_1| g(d) + r(d)$, where $g(d)$ is a multiplicative function supported on squarefree integers and

$$R(z) := \sum_{\substack{d \leqslant z \\ d \text{ squarefree}}} |r(d)|$$

is small. Define

$$V(z) = \sum_{d|P(z)} \mu(d) g(d) = \prod_{p \in \mathscr{P}_{\leqslant z}} (1 - g(p)).$$

We shall assume that for some $\kappa, L \geqslant 0$, we have

$$V(w) \leqslant \left(\frac{\log z}{\log w}\right)^\kappa \left(1 + \frac{L}{\log w}\right) V(z) \tag{4.4}$$

for all $2 \leqslant w \leqslant z$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

For some choice of sieve weights $\Lambda^{\pm} = (\lambda_d^{\pm})_{d \geqslant 1}$, define

$$V^{\pm}(z) = \sum_{d \mid P(z)} \lambda_d^{\pm} g(d).$$

We recall that if $\Lambda^{\pm}$ are upper and lower bound sieves of level z (i.e., if $\lambda_d^{\pm}$ are supported on squarefree integers $d < z$ and

$$\sum_{d \mid m} \lambda_d^{-} \leqslant \sum_{d \mid m} \mu(d) \leqslant \sum_{d \mid m} \lambda_d^{+}$$

for all integers m), then we have

$$\begin{aligned} S(\mathscr{A}, \mathscr{P}, z) &\leqslant |\mathscr{A}_1| V^{+}(z) + R(z), \\ S(\mathscr{A}, \mathscr{P}, z) &\geqslant |\mathscr{A}_1| V^{-}(z) - R(z). \end{aligned} \tag{4.5}$$

The main theorem of the beta sieve we apply is given in [18, Theorem 11.12]. We record this theorem here for convenience, in the special case $s = 1$.

Theorem 4.1. *Suppose κ, L are such that the assumption (4.4) holds. Then there is a choice of upper and lower bound sieve weights $\Lambda^{\pm}$ (taking values in $\{-1,0,1\}$), and explicit constants $A(\kappa), B(\kappa) \geqslant 0$ such that, as $z \to \infty$, we have*

$$\begin{aligned} V^{+}(z) &\leqslant (A(\kappa) + o(1)) V(z), \\ V^{-}(z) &\geqslant (B(\kappa) + o(1)) V(z). \end{aligned} \tag{4.6}$$

Notation 4.2. Throughout the remainder of this chapter, for a sequence $\mathscr{A}$, a set of primes $\mathscr{P}$, a multiplicative function g, and a sifting level $z \geqslant 1$, we define $\Lambda^{\pm}(\mathscr{A}, \mathscr{P}, g, z)$ to be the corresponding upper and lower bound beta sieves with these parameters. We sometimes apply (4.6) directly without reference to a sequence, in which case the parameter $\mathscr{A}$ is omitted from the notation.

In our applications of the beta sieve, the required bounds on $R(z)$ are provided by Corollary 3.2 and Lemma 3.6. For $i \in \{0, \dots, k\}$, we define multiplicative functions

$$\begin{aligned} \varrho_i(d_1, d_2) &= \#\{a, b \pmod{d_1 d_2} : d_1 \mid f_i(a,b), d_2 \mid f(a,b)\}, \\ \varrho_i(d) &= \#\{a, b \pmod{d} : f_i(a,b) \equiv 0 \pmod{d}\}, \\ \varrho(d) &= \#\{a, b \pmod{d} : f(a,b) \equiv 0 \pmod{d}\}. \end{aligned}$$

We note that the function $\varrho_i(d_1, d_2)$ is the same as the function $\varrho(d_1, d_2)$ from (3.2) with $g_1(x,y) = f_i(x,y)$ and $g_2(x,y) = f(x,y)$, but in this section, we add a subscript to keep track of the dependence on i. When $\gcd(d_1, d_2) = 1$, we have $\varrho_i(d_1, d_2) = \varrho_i(d_1)\varrho(d_2)$. Moreover, for any $i \in \{1, \dots, k\}$ and any prime $p \notin S$, we have

$$\varrho_i(p) = \nu_i(p)(p-1) + 1, \tag{4.7}$$

$$\varrho(p) = \nu(p)(p-1) + 1, \tag{4.8}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

where $\nu_i(p)$ and $\nu(p)$ are as defined in (2.2) and (2.3). We define multiplicative functions

$$g(p) := \frac{\varrho(p)}{p^2}, \qquad g_i(p) := \frac{\varrho_i(p)}{p^2}$$

and define

$$V(x) = \prod_{p \in \mathscr{P}_{\leqslant z}} (1 - g(p)), \qquad V_i(x) = \prod_{p \in \mathscr{P}'_{\leqslant z}} (1 - g_i(p)) \tag{4.9}$$

for $i \in \{1, \dots, k\}$.

4.2. Sieve dimensions

We prove in Lemma 4.3 that the functions V, V_i defined above satisfy the hypothesis (4.4) with the sieve dimensions

$$\kappa := \alpha\theta, \qquad \kappa_i := 1 - \alpha\theta_i,$$

where α, θ_i and θ are as defined in (2.4), (2.5) and (2.6).

In the notation of Theorem 4.1, we write $A = A(\kappa), B = B(\kappa)$ and $A_i = A(\kappa_i)$. We assume throughout this section that $\kappa < 1/2$, and so $\kappa_i > 1/2$. Then A and B are defined in [18, Equations (11.62, 11.63)] (see also Section 4.9), and A_i is defined in [18, Equations (11.42), (11.57)]. A table of numerical values of these constants can be found in [18, Section 11.19].

Lemma 4.3. *Let $x \geqslant 1$. For $i \in \{1, \dots, k\}$, and V, V_i as in (4.9), there exist constants $c, c_i > 0$ such that*

$$V(x) = \frac{c}{(\log x)^{\kappa}} \left(1 + O((\log x)^{-1})\right), \tag{4.10}$$

$$V_i(x) = \frac{c_i}{(\log x)^{\kappa_i}} \left(1 + O((\log x)^{-1})\right). \tag{4.11}$$

The asymptotic in (4.11) also holds for $i = 0$ when $f_0 \not\equiv 1$.

Proof. We follow a similar approach to [27, Lemma 4.2]. Below, we denote by C a constant which is allowed to vary from line to line. We have

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$$\log V(x) = -\sum_{p\in\mathscr{P}_{\leqslant x}}\left(\sum_{m=1}^{\infty}\frac{\varrho(p)^m}{mp^{2m}}\right) \tag{4.12}$$

$$= -\sum_{p\in\mathscr{P}_{\leqslant x}}\frac{\nu(p)(p-1)+1}{p^2} + C + O((\log x)^{-1}). \tag{4.13}$$

$$= -\sum_{p\in\mathscr{P}_{\leqslant x}}\frac{\nu(p)}{p} + C + O((\log x)^{-1}), \tag{4.14}$$

where in (4.14) we have used that $\nu(p)\leqslant \deg f$ for all but finitely many primes p. To estimate the sum in (4.14), we apply partial summation, together with our assumption (2.6). For $t\geqslant 2$, we define

$$A_t = \sum_{p\in\mathscr{P}_{\leqslant t}}\nu(p). \tag{4.15}$$

Then

$$\begin{aligned}\sum_{p\in\mathscr{P}_{\leqslant x}}\frac{\nu(p)}{p} &= \frac{A_x}{x} + \int_2^x \frac{A_t}{t^2}\mathrm{d}t\\ &= \kappa\int_2^x \frac{\pi(t)\left(1+O\left((\log t)^{-1}\right)\right)}{t^2}\mathrm{d}t + O((\log x)^{-1})\\ &= \kappa\int_2^x \frac{\mathrm{d}t}{t\log t} + C + O((\log x)^{-1})\\ &= \kappa\log\log x + C + O((\log x)^{-1}).\end{aligned} \tag{4.16}$$

We deduce (4.10) by taking the exponential of (4.16).

We can prove (4.11) in a similar way. When $i=0$ and $f_0\not\equiv 1$, we have $\varrho_0(p)=p$, and so the result is a consequence of Mertens' theorem [29, Equation (2.16)]. For any $i\in\{1,\dots,k\}$, we have

$$\begin{aligned}\log V_i(x) &= \sum_{p\in\mathscr{P}'_{\leqslant x}}\frac{\nu_i(p)}{p} + C + O((\log x)^{-1})\\ &= \sum_{p\leqslant x \text{ prime}}\frac{\nu_i(p)}{p} - \sum_{p\in\mathscr{P}_{\leqslant x}}\frac{\nu_i(p)}{p} + C + O((\log x)^{-1}).\end{aligned} \tag{4.17}$$

Similarly to above, using partial summation and (2.5), we have

$$\sum_{p\in\mathscr{P}_{\leqslant x}}\frac{\nu_i(p)}{p} = \alpha\theta_i\log\log x + C + O((\log x)^{-1}). \tag{4.18}$$

To treat the first sum in (4.17), we define L_i to be the number field generated by f_i. For all but finitely many primes p, the quantity $\nu_i(p)$ is equal to the number of degree one prime ideals $\mathfrak{p}$ in L_i above p. Let $\pi_{L_i}(x)$ denote the number of prime ideals $\mathfrak{p}$ in L of norm

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

at most x. This count is dominated by degree one ideals. In fact, the number of prime ideals of degree at least 2 enumerated by $\pi_{L_i}(x)$ is $O(x^{1/2})$ because such an ideal must lie over a rational prime $p \leqslant x^{1/2}$ and each rational prime p has at most $[L_i : \mathbb{Q}]$ prime ideals in L_i lying above it. Therefore,

$$\sum_{p \leqslant x \text{ prime}} \nu_i(p) = \pi_{L_i}(x) + O(x^{1/2}).$$

Using partial summation, as above, together with the Prime ideal theorem [38], we deduce that

$$\sum_{p \leqslant x \text{ prime}} \frac{\nu_i(p)}{p} = \log\log x + C + O((\log x)^{-1}). \tag{4.19}$$

Combining this with (4.18) and taking exponentials, we deduce the asymptotic in (4.11). □

In the following lemma, we record three more useful estimates following similar arguments to Lemma 4.3.

Lemma 4.4. *There exists constants $C, C_i > 0$ such that*

$$\sum_{p \in \mathscr{P}_{\leqslant x}} \frac{\varrho_i(p)}{p^2} = (1-\kappa_i)\log\log x + C + O((\log x)^{-1}), \tag{4.20}$$

$$\sum_{p \in \mathscr{P}'_{\leqslant x}} \frac{\varrho_i(p)}{p^2} = \kappa_i \log\log x + C_i + O((\log x)^{-1}), \tag{4.21}$$

$$\sum_{p \in \mathscr{P}'_{\leqslant x}} \frac{\varrho_i(p)}{p^2} \log p = \kappa_i \log x + O(1). \tag{4.22}$$

Proof. The estimates (4.20) and (4.21) are immediate consequences of (4.16), (4.18) and (4.19), together with fact that

$$\frac{\varrho_i(p)}{p^2} = \frac{(p-1)\nu_i(p)+1}{p^2} = \frac{\nu_i(p)}{p} + O(p^{-2}).$$

To prove (4.22), we proceed via partial summation in a very similar manner to (4.16). We recall from the Prime number theorem that

$$\pi(t) = \frac{t}{\log t} + \frac{t}{(\log t)^2} + O\left(\frac{t}{(\log t)^3}\right). \tag{4.23}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

For A_t as defined in (4.15), we have

$$\begin{aligned}\sum_{p\in\mathscr{P}'_{\leqslant x}}\frac{\varrho_i(p)}{p^2}\log p &= \sum_{p\in\mathscr{P}'_{\leqslant x}}\frac{\nu_i(p)}{p}\log p+O(1)\\ &= \frac{A_x\log x}{x}-\int_2^x A_t\left(\frac{\log t}{t}\right)'\mathrm{d}t+O(1)\\ &= \kappa_i\int_2^x\frac{(\log t-1)\pi(t)(1+O((\log t)^{-A})}{t^2}\mathrm{d}t+O(1)\\ &= \kappa_i\int_2^x\frac{1}{t}+O\left(\frac{1}{t(\log t)^2}\right)\mathrm{d}t+O(1)\quad\text{(from (4.3))}\\ &= \kappa_i\log x+O(1),\end{aligned}$$

as required. □

4.3. The sum S_1

We apply the lower bound sieve $\Lambda^-(\mathscr{A},\mathscr{P},g,N^\gamma)$ and the level of distribution result from Corollary 3.2 with $g_1(x,y)=1, g_2(x,y)=f(x,y), D_1=1$ and $D_2=N^\gamma$. The hypotheses of Corollary 3.2 require that $\gamma<1$. We obtain

$$S_1\geqslant(B+o(1))XV(N^\gamma). \tag{4.24}$$

By Lemma 4.3, we have

$$V(N^\gamma)\sim\frac{c}{(\log N^\gamma)^\kappa}.$$

For any $\epsilon>0$, taking γ sufficiently close to 1, we obtain

$$S_1\geqslant\frac{(cB-\epsilon+o(1))X}{(\log N)^\kappa}. \tag{4.25}$$

4.4. The sums $S_2^{(i)}$

We write $f_i(a,b)=pr$, for $p\in\mathscr{P}$ and $N^\gamma<p\ll N$. We apply the *switching principle*, which transforms the sum over p defining $S_2^{(i)}$ into a much shorter sum over the variable r.

Let $R=N^{1-\gamma}$. The sums $S_2^{(i)}$ only involve linear factors $f_i(x,y)$ since we assume $i\in\{0,\dots,m\}$. Therefore, for $(a,b)\in\mathscr{R}\cap\mathbb{Z}^2$, we have $f_i(a,b)\ll N$, and so $|r|\ll R$. Let $z=N^{1/3}$. We shall take γ arbitrarily close to 1; for now, we assume that $\gamma>2/3$. Then by definition of $S(\mathscr{A}_p^{(i)},\mathscr{P},p)$, we know that $\gcd(r,P(R))=1$ and $\gcd(f(a,b),P(z))=1$.

Let $r'=|r|/\gcd(r,\Delta)$. We now explain why r' only has prime factors in $\mathscr{P}'$. Fix $l\in S$. Recall the assumption that $\nu_l(f(a_0,b_0))<\nu_l(\Delta)$. Since $f_i(a_0,b_0)\mid f(a_0,b_0)$ and $f_i(a,b)\equiv f_i(a_0,b_0)\pmod{\Delta}$ (by the condition $C(a,b)$), we have

$$\nu_l(r)=\nu_l(f_i(a,b))\leqslant\nu_l(f(a,b))=\nu_l(f(a_0,b_0))<\nu_l(\Delta). \tag{4.26}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Therefore, $\gcd(r',\Delta)=1$, and by the assumption that every prime in S divides Δ, this implies r' has no prime factors in S. Moreover, p is the smallest prime in $\mathscr{P}$ dividing $f_i(a,b)$. Since $r<p$, this means that r' has no prime factors in $\mathscr{P}$, and hence only has prime factors in $\mathscr{P}'$.

However, $f_i(a,b)/r'$ has no prime factors in $\mathscr{P}'$. Therefore,

$$S_2^{(i)} \leqslant \sum_{\substack{r' \ll R \\ \gcd(r', P(R)\Delta)=1}} S_2^{(i)}(r'), \tag{4.27}$$

where

$$S_2^{(i)}(r') = \#\left\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2 : \begin{array}{l} C(a,b), r' \mid f_i(a,b), \\ \gcd(f_i(a,b)/r', P'(z))=1 \\ \gcd(f(a,b),P(z))=1 \end{array}\right\}. \tag{4.28}$$

Below, for convenience, we change notation from r' back to r.

Let Λ_1^+,Λ_2^+ be upper bound sieves of level z. Defining $A(dr,e),r(dr,e)$ as in Section 3 with $g_1(x,y)=f_i(x,y)$ and $g_2(x,y)=f(x,y)$, and writing

$$m_1 = \gcd\left(\frac{f_i(a,b)}{r}, P'(z)\right), \qquad m_2 = \gcd(f(a,b),P(z)),$$

we obtain

$$\begin{aligned} S_2^{(i)}(r) &\leqslant \sum_{\substack{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2 \\ C(a,b)}} \sum_{d\mid m_1} \mu(d) \sum_{e\mid m_2} \mu(e) \\ &\leqslant \sum_{\substack{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2 \\ C(a,b)}} \sum_{d\mid m_1} \lambda_1^+(d) \sum_{e\mid m_2} \lambda_2^+(e) \\ &\leqslant \sum_{\substack{d\mid P'(z) \\ \gcd(d,r)=1}} \lambda_1^+(d) \sum_{e\mid P(z)} A(dr,e)\lambda_2^+(e) \\ &= \sum_{\substack{d\mid P'(z) \\ \gcd(d,r)=1}} \lambda_1^+(d) \sum_{e\mid P(z)} \lambda_2^+(e)\left(\frac{X\varrho(dr,e)}{(dre)^2} + r(dr,e)\right) \\ &\leqslant \left(\frac{X\varrho_i(r)}{r^2} \sum_{\substack{d\mid P'(z) \\ \gcd(d,r)=1}} \lambda_1^+(d) g_i(d) \sum_{e\mid P(z)} \lambda_2^+(e) g(d)\right) + \sum_{\substack{d,e\leqslant z \\ \gcd(dr,e)=1 \\ \gcd(dre,\Delta)=1}} |r(dr,e)|. \end{aligned} \tag{4.29}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Choose λ_1^+,λ_2^+ to be the beta sieves $\Lambda^+(\mathscr{P}'\backslash\{p \mid r\},g_i,N^{1/3}),\Lambda^+(\mathscr{P},g,N^{1/3})$, respectively. Then the remainder term in (4.29) can be bounded by

$$\sum_{\substack{d,e\leqslant N^{1/3}\\ \gcd(dr,e)=1\\ \gcd(dre,\Delta)=1}} |r(dr,e)| \leqslant \sum_{\substack{d\ll N^{1/3}R\\ e\ll N^{1/3}\\ \gcd(dr,e)=1\\ \gcd(dre,\Delta)=1}} |r(d,e)|,$$

which is negligible by Lemma 3.6 because $N^{2/3}R = N^{5/3-\gamma} \ll N^{1-\epsilon}$.

Define a multiplicative function $h_i(r)$ supported on squarefree integers r, which is zero unless all prime factors of r are in $\mathscr{P}'$ and

$$h_i(r) = \frac{\varrho_i(r)}{r^2}\prod_{p\mid r}\left(1-\frac{\varrho_i(p)}{p^2}\right)^{-1} \tag{4.30}$$

otherwise. Applying Theorem 4.1, we conclude that

$$S_2^{(i)}(r) \leqslant Xh_i(r)V(N^{1/3})V_i(N^{1/3})(AA_i+o(1)).$$

From Lemma 4.3, we obtain

$$S_2^{(i)} \leqslant \frac{(cc_iAA_i+o(1))X}{(\log N^{1/3})^{\kappa+\kappa_i}}\sum_{r\ll R} h_i(r).$$

To deal with the sum over $h_i(r)$, we note that

$$\sum_{r\ll R} h_i(r) \leqslant \prod_{\substack{p\in\mathscr{P}'\\ p\ll R}}\left(1+\sum_{m=1}^{\infty} h_i(p^m)\right) = \prod_{\substack{p\in\mathscr{P}'\\ p\ll R}} (1-g_i(p))^{-1}(1+O(p^{-2})). \tag{4.31}$$

Applying Lemma 4.3 to the product, we obtain

$$\sum_{r\ll R} h_i(r) \ll (\log R)^{\kappa_i}.$$

Since $R = N^{1-\gamma}$, we deduce that

$$S_2^{(i)} \ll \frac{cc_iAA_iX}{(\log N)^{\kappa}}\left(\frac{(1-\gamma)^{\kappa_i}}{(1/3)^{\kappa_i+\kappa}}\right).$$

Therefore, $S_2^{(i)}$ can be made negligible compared to S_1 by taking γ arbitrarily close to 1.

4.5. The sums $S_3^{(i)}$

For a fixed prime p and an upper bound sieve λ^+ of level z, we have

$$S(\mathscr{A}_p^{(i)},\mathscr{P},z) \leqslant \sum_{d\mid P(z)} \lambda^+(d)A(p,d),$$

where $A(p,d)$ is as in Section 3 with $g_1(x,y) = f_i(x,y)$ and $g_2(x,y) = f(x,y)$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

We first deal with the primes p in the interval $I := (N^{\gamma}, N^{2-\gamma-\epsilon}]$. For any $p \in I$, let λ^+ be the beta sieve $\Lambda^+(\mathscr{A}_p^{(i)}, \mathscr{P}, g_i(p)g, N^{\gamma})$. We obtain

$$\begin{aligned} S(\mathscr{A}_p^{(i)}, \mathscr{P}, p) &\leqslant S(\mathscr{A}_p^{(i)}, \mathscr{P}, N^{\gamma}) \\ &\leqslant (A + o(1))XV(N^{\gamma})g_i(p) + \sum_{\substack{d \leqslant N^{\gamma} \\ \gcd(d, p\Delta)=1}} |r(p, N^{\gamma})|. \end{aligned} \tag{4.32}$$

We apply Corollary 3.2 with $D_1 = N^{2-\gamma-\epsilon}$ and $D_2 = N^{\gamma}$. These choices of D_1, D_2 satisfy the hypotheses of Corollary 3.2. Taking a sum over $p \in I$, the contribution from the remainder term in (4.32) is negligible. We obtain

$$\sum_{p \in I \cap \mathscr{P}} S(\mathscr{A}_p^{(i)}, \mathscr{P}, p) \leqslant (A + o(1))XV(N^{\gamma}) \sum_{p \in I \cap \mathscr{P}} g_i(p). \tag{4.33}$$

It follows from Lemma 4.4 that

$$\begin{aligned} \sum_{N^{\gamma} < p \leqslant N^{2-\gamma-\epsilon}} g_i(p) &= \log\log N^{2-\gamma-\epsilon} - \log\log N^{\gamma} + o(1) \\ &= \log(2 - \gamma - \epsilon) - \log\gamma + o(1). \end{aligned}$$

Therefore, the contribution to $S_3^{(i)}$ from this range is negligible if we take γ arbitrarily close to 1.

In the remaining range $N^{2-\gamma-\epsilon} < p < N^{\beta_i}$, we split into dyadic intervals $(R, 2R]$. Note that for $p \in (R, 2R]$, the assumption $\gamma < 1$ implies that $N^{2-\epsilon}/R < p$, so $S(\mathscr{A}_p^{(i)}, \mathscr{P}, p) \leqslant S(\mathscr{A}_p^{(i)}, \mathscr{P}, N^{2-\epsilon}/R)$. For each dyadic interval, we apply the beta sieve $\Lambda^+(\mathscr{A}_p^{(i)}, \mathscr{P}, g_i(p)g, N^{2-\epsilon}/R)$ and the level of distribution result from Corollary 3.2 with $D_1 = 2R$ and $D_2 = N^{2-\epsilon}/R$. At this point, we need to assume that $\beta_i < 2$ for all i, so that $D_2 \geqslant 1$. We obtain

$$\sum_{\substack{N^{2-\gamma-\epsilon} < p < N^{\beta_i} \\ p \in \mathscr{P}}} S(\mathscr{A}_p^{(i)}, \mathscr{P}, p) \tag{4.34}$$

$$\begin{aligned} &\leqslant \sum_{\substack{R \text{ dyadic} \\ N^{2-\gamma-\epsilon} < R < N^{\beta_i}}} \sum_{\substack{p \in (R, 2R] \\ p \in \mathscr{P}}} S(\mathscr{A}_p^{(i)}, \mathscr{P}, N^{2-\epsilon}/R) \\ &\leqslant (A + o(1))X \sum_{\substack{R \text{ dyadic} \\ N^{2-\gamma-\epsilon} < R < N^{\beta_i}}} V(N^{2-\epsilon}/R) \sum_{\substack{p \in (R, 2R] \\ p \in \mathscr{P}}} g_i(p) \\ &\leqslant \frac{(cA + o(1))X}{(\log N)^{\kappa}} \sum_{\substack{N^{2-\gamma-\epsilon} \leqslant p < N^{\beta_i} \\ p \in \mathscr{P}}} \frac{g_i(p)}{(2 - \epsilon - \frac{\log p}{\log N})^{\kappa}}, \end{aligned} \tag{4.35}$$

where the last line follows from Lemma 4.3 and the fact that $V(N^{2-\epsilon}/R) < V(N^{2-\epsilon-\log p/\log N})$ for all $p \in (R, 2R]$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

We denote the sum in (4.35) by $T(\beta_i,\kappa)$. Since $\gamma<1$, we have $2-\gamma-\epsilon>1$ for sufficiently small ϵ, and so we may upper bound $T(\beta_i,\kappa)$ by enlarging its range of summation to $N<p<N^{\beta_i}$. Define

$$A(t)=\sum_{p\in\mathscr{P}_{\leqslant t}} g_i(p), \qquad h(t)=\left(2-\epsilon-\frac{\log t}{\log N}\right)^{-\kappa}.$$

From Lemma 4.4, we have $A(t)=(1-\kappa_i)\log\log t+C+o(1)$ for some constant C. In particular, for any $r>0$, we have $A(N^r)-A(N)=(1-\kappa_i)\log r+o(1)$. Applying summation by parts, followed by the substitution $t=N^s$, we obtain

$$\begin{aligned}
T(\beta_i,\kappa) &\leqslant (A(N^{\beta_i})-A(N))h(N^{\beta_i})-\int_N^{N^{\beta_i}}(A(t)-A(N))h'(t)\mathrm{d}t\\
&=(A(N^{\beta_i})-A(N))h(N^{\beta_i})-\int_1^{\beta_i}(A(N^s)-A(N))\frac{\partial h(N^s)}{\partial s}\mathrm{d}s\\
&=(1-\kappa_i)\log\beta_i h(N^{\beta_i})-(1-\kappa_i)\int_1^{\beta_i}(\log s)\frac{\partial h(N^s)}{\partial s}\mathrm{d}s+o(1)\\
&=(1-\kappa_i)\int_1^{\beta_i}\frac{h(N^s)}{s}\mathrm{d}s+o(1). \qquad (4.36)
\end{aligned}$$

Taking ϵ sufficiently small and redefining it appropriately, we may replace $h(N^s)$ by $(2-s)^{-\kappa}$ at the cost of adding ϵ in (4.36). Combining with (4.35), we conclude that for any $\epsilon>0$,

$$S_3^{(i)}\leqslant\frac{(cA+\epsilon+o(1))X}{(\log N)^{\kappa}}\cdot(1-\kappa_i)\int_1^{\beta_i}(2-s)^{-\kappa}\frac{\mathrm{d}s}{s}. \qquad (4.37)$$

Due to the factor $1-\kappa_i=\alpha\theta_i$ appearing in the above estimate, $S_3^{(i)}$ becomes negligible compared to S_1 as $\alpha\to 0$. We perform a more precise quantitative comparison in Section 4.7.

4.6. The sums $S_4^{(i)}$

We begin in a similar manner to the sums $S_2^{(i)}$, by writing $f_i(a,b)=pr$, for $p\in\mathscr{P}$, where now $N^{\beta_i}\leqslant p<x$ and $R=x/N^{\beta_i}$. Let $D_1=N^{\eta_1}$ and $D_2=N^{\eta_2}$ for parameters $\eta_1,\eta_2>0$ (which may depend on r and i) to be chosen later. We assume $\eta_2<\beta_i$ so that the condition $\gcd(f(a,b),P(p))=1$ can be replaced by the weaker condition $\gcd(f(a,b),P(N^{\eta_2}))=1$. Proceeding as in Section 4.4, we have

$$S_4^{(i)}\leqslant\sum_{\substack{r\ll R\\ \gcd(r,P(N^{\beta_i})\Delta)=1}} S_4^{(i)}(r), \qquad (4.38)$$

where

$$S_4^{(i)}(r)=\#\left\{(a,b)\in\mathscr{R}\cap\mathbb{Z}^2:\ \begin{array}{l} C(a,b),r\mid f_i(a,b),\\ \gcd(f_i(a,b)/r,P'(N^{\eta_1}))=1\\ \gcd(f(a,b),P(N^{\eta_2}))=1\end{array}\right\}. \qquad (4.39)$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Similarly to (4.29), for any upper bound sieves λ_1^+,λ_2^+ of levels N^{η_1},N^{η_2}, the quantity $S_4^{(i)}(r)$ is bounded above by

$$\left(\frac{X\varrho_i(r)}{r^2}\sum_{\substack{d\mid P'(N^{\eta_1})\\ \gcd(d,r)=1}}\lambda_1^+(d)g_i(d)\sum_{e\mid P(N^{\eta_2})}\lambda_2^+(e)g(d)\right)+\sum_{\substack{d\leqslant N^{\eta_1},e\leqslant N^{\eta_2}\\ \gcd(dr,e)=1\\ \gcd(dre,\Delta)=1}}|r(dr,e)|.$$

In order to ensure the error terms from applying the level of distribution result from Corollary 3.2 are negligible after summing over r, we need η_1,η_2 to satisfy

$$\eta_1,\eta_2>0,\quad \eta_2\leqslant 1-\delta\qquad(\text{for all } r\leqslant R),\tag{4.40}$$

$$\sum_{r\leqslant R}N^{\eta_1+\eta_2}\leqslant N^{2-\delta}.\tag{4.41}$$

Remark 4.5. There are no η_1,η_2 satisfying (4.40) and (4.41) unless $R\leqslant N^{2-\delta}$, i.e., $x\leqslant N^{2-\delta+\beta_i}$. Since we had to assume $\beta_i<2$ in the treatment of the sums $S_3^{(i)}$ in Section 4.5, this means our approach cannot handle the case $d\geqslant 4$, in which $f(x,y)$ has an irreducible factor of degree $\geqslant 4$. Therefore, we proceed with the additional assumption that $d\leqslant 3$.

Choose λ_1^+,λ_2^+ to be the beta sieves $\Lambda^+(\mathscr{P}'\setminus\{p\mid r\},g_i,N^{\eta_1}),\Lambda^+(\mathscr{P},g,N^{\eta_2})$, respectively. Recalling the definition of $h_i(r)$ from (4.30), we obtain

$$S_4^{(i)}(r)\leqslant(AA_i+o(1))Xh_i(r)V_i(N^{\eta_1})V(N^{\eta_2}).\tag{4.42}$$

Using Lemma 4.3 to estimate the products in (4.42), and taking a sum over r, we obtain

$$S_4^{(i)}\leqslant\frac{(cc_iAA_i+o(1))X}{(\log N)^{\kappa_i+\kappa}}\sum_{r\leqslant R}\frac{h_i(r)}{\eta_1^{\kappa_i}\eta_2^{\kappa}}.\tag{4.43}$$

We divide the sum over r into dyadic intervals $r\in(R_1,2R_1]$, and take η_1,η_2 depending only on R_1 and i. To obtain a good estimate for (4.43), we maximise $\eta_1^{\kappa_i}\eta_2^{\kappa}$ subject to the constraints

$$\eta_1,\eta_2>0,\qquad \eta_2\leqslant 1-\delta,\qquad \eta_1+\eta_2\leqslant 2-\delta-\frac{\log R_1}{\log N}.$$

By a similar computation to [27, Section 6.5], the optimal solution is

$$\eta_1=\frac{\kappa_i}{\kappa+\kappa_i}\left(2-\delta-\frac{\log R_1}{\log N}\right),$$
$$\eta_2=\frac{\kappa}{\kappa+\kappa_i}\left(2-\delta-\frac{\log R_1}{\log N}\right).$$

We note that for $\delta>0$ sufficiently small, this solution satisfies $\eta_2\leqslant 1-\delta$ due to the assumption $\kappa<1/2$. Substituting this choice of η_1,η_2 into (4.43), we obtain

$$\sum_{r\leqslant R}\frac{h_i(r)}{\eta_1^{\kappa_i}\eta_2^{\kappa}}\leqslant\sum_{r\leqslant R}w(r,\delta)h_i(r),\tag{4.44}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

where

$$w(r,\delta) = \left(\frac{\kappa}{\kappa+\kappa_i}\right)^{-\kappa} \left(\frac{\kappa_i}{\kappa+\kappa_i}\right)^{-\kappa_i} \left(2-\delta-\frac{\log r}{\log N}\right)^{-(\kappa+\kappa_i)}. \tag{4.45}$$

We treat the sum in (4.44) using partial summation, for which we require estimates for $\sum_{r\leqslant t} h_i(r)$. For the case $d=2$, the estimate already found in (4.31) is sufficient. Below, we find a more refined estimate which we use in the case $d=3$.

We first consider the contribution to (4.44) from squarefree values of r. We would like to apply [18, Theorem A.5], which states that under certain hypothesis on the function $h_i(r)$, we have

$$\sum_{\substack{m\leqslant x\\ \mu^2(m)=1}} h_i(m) = c_{h_i}(\log x)^{\kappa_i} + O((\log x)^{\kappa_i-1}), \tag{4.46}$$

where

$$c_{h_i} = \frac{1}{\Gamma(\kappa_i+1)} \prod_p \left(1-\frac{1}{p}\right)^{\kappa_i} (1+h_i(p)).$$

In the following lemma, we verify that the function $h_i(r)$ satisfies the required hypotheses for [18, Theorem A.5].

Lemma 4.6. *For any $x\geqslant 1$ and any $2\leqslant w<z$, the function $h_i(r)$ satisfies the following estimates:*

$$\prod_{w\leqslant p<z} (1+h_i(p)) \ll \left(\frac{\log z}{\log w}\right)^{\kappa_i}, \tag{4.47}$$

$$\sum_p h_i(p)^2 \log p < \infty, \tag{4.48}$$

$$\sum_{p\leqslant x} h_i(p)\log p = \kappa_i \log x + O(1). \tag{4.49}$$

Proof. To prove (4.47), we note that $1+h_i(p) = \left(1-\frac{\varrho_i(p)}{p^2}\right)^{-1}$ for all $p\in\mathscr{P}'$. The result is then immediate from Lemma 4.3. To prove (4.45), we recall that $\varrho_i(p)\ll p$, and so $h_i(p)\ll p^{-1}$. Therefore,

$$\sum_p h_i(p)^2\log p \ll \sum_p \frac{\log p}{p^2} < \infty.$$

Finally, we note that

$$\sum_{p\leqslant x} h_i(p)\log p = \sum_{p\in\mathscr{P}'_{\leqslant x}} \frac{\varrho_i(p)}{p^2}\log p + O(1),$$

so that (4.49) follows by applying Lemma 4.4. □

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

We can now evaluate the sum in (4.44) using partial summation. We obtain

$$\begin{aligned}\sum_{\substack{r\leqslant R\\ \mu^2(r)=1}} w(r,\delta)h_i(r) &= w(R,\delta)\sum_{\substack{r\leqslant R\\ \mu^2(r)=1}} h_i(r) - \int_1^R \left(\sum_{\substack{r\leqslant t\\ \mu^2(r)=1}} h_i(r)\right) w'(t,\delta)\mathrm{d}t\\ &= c_{h_i}\left[(\log R)^{\kappa_i} w(R,\delta) - \int_1^R w'(t,\delta)(\log t)^{\kappa_i}\mathrm{d}t\right]\\ &\quad + o(1)\left[(\log R)^{\kappa_i} w(R,\delta) + \int_1^R w'(t,\delta)(\log t)^{\kappa_i}\mathrm{d}t\right]\\ &= c_{h_i}\kappa_i \int_1^R w(t,\delta)(\log t)^{\kappa_i-1}t^{-1}\mathrm{d}t + o((\log R)^{\kappa_i})\\ &= c_{h_i}\kappa_i(\log N)^{\kappa_i}\int_0^{\log R/\log N} w(N^s,\delta)s^{\kappa_i-1}\mathrm{d}s + o((\log R)^{\kappa_i})\\ &\leqslant (c_{h_i}\kappa_i + o(1))(\log N)^{\kappa_i}\int_0^{d-\beta_i} W(s)\mathrm{d}s,\end{aligned}$$

where $W(s) = w(N^s,0)s^{\kappa_i-1}$.

We now consider the contribution to (4.30) from those r which are not squarefree. Let d_i denote the degree of f_i. From [15, Lemma 3.1], we have $\varrho_i(p^\alpha) \ll p^{2\alpha(1-1/d_i)}$ for all primes $p \notin S$ and any positive integer α. By the multiplicativity of $h_i(r)$, it follows that $h_i(r) \ll r^{-2/d_i+\epsilon}$.

We recall that a positive integer n is *squareful* if for any prime $p \mid n$, we also have $p^2 \mid n$. Since $d_i \leqslant 3$, we have

$$\sum_{r \text{ squareful}} h_i(r) \ll \sum_{r \text{ squareful}} r^{-2/d_i+\epsilon} \leqslant \sum_{r \text{ squareful}} r^{-2/3+\epsilon} < \infty,$$

where the last inequality follows from partial summation together with the fact that there are $O(M^{1/2})$ squareful positive integers less than M. Since $h_i(r)$ is supported on integers with no prime factors in S, for any $\epsilon > 0$, there exists a set of primes S_0, depending only on ϵ and f, such that for any $S \supseteq S_0$, we have

$$\sum_{\substack{r \text{ squareful}\\ r>1}} h_i(r) < \epsilon.$$

For the remainder of this section, we assume that $S \supseteq S_0$. Proceeding as in [27, Lemma 6.2], we use that $w(r,\delta) \ll 1$, and decompose each non-squarefree r into $r = r_1r_2$, where r_1 is squarefree and $r_2 > 1$ is squareful. We have

$$\sum_{\substack{r\leqslant R\\ \mu^2(r)=0}} w(r,\delta)h_i(r) \ll \sum_{\substack{r_1\leqslant R\\ \mu^2(r_1)=1}} h_i(r_1) \sum_{\substack{r_2 \text{ squareful}\\ right_2>1}} h_i(r_2).$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Combining with (4.46), we deduce that for any $\epsilon > 0$, we have

$$\sum_{\substack{r \leqslant R \\ \mu^2(r)=0}} w(r,\delta)h_i(r) \leqslant (\epsilon + o(1))(\log N)^{\kappa_i}.$$

In conclusion, we have the upper bound

$$S_4^{(i)} \leqslant \frac{(cc_iAA_ic_{h_i}\kappa_i + \epsilon + o(1))X}{(\log N)^{\kappa}} \int_0^{d-\beta_i} W(s)\mathrm{d}s, \tag{4.50}$$

where

$$W(s) = \left(\frac{\kappa}{\kappa+\kappa_i}\right)^{-\kappa} \left(\frac{\kappa_i}{\kappa+\kappa_i}\right)^{-\kappa_i} (2-s)^{-(\kappa+\kappa_i)} s^{\kappa_i - 1}. \tag{4.51}$$

Remark 4.7. Whilst c, c_i, c_{h_i} all depend on S, the ratio $\kappa_i c_i c_{h_i} AA_i/B$ of the constants in the bounds for $S_4^{(i)}$ and S_1 is independent of S, as seen below in (4.52). Therefore, we may indeed take ϵ to be arbitrarily small in (4.50) without sacrificing anything in our comparison of S_1 and $S_4^{(i)}$. The $o(1)$ terms in (4.50) and in the bounds for the other sums S_1, $S_2^{(i)}, S_3^{(i)}$ also depend on S (via the quantity L in (4.4)), but this does not matter because we may take N to be sufficiently large in terms of S.

4.7. Proof of Theorem 2.1

We now suppose that $d = 2$ (i.e., $f(x,y)$ consists of irreducible factors of degree at most 2). We first obtain a qualitative result by considering the limit as $\alpha \to 0$. As $\alpha \to 0$, we have $\kappa \to 0$ and $\kappa_i \to 1$. Therefore, we have $A_i \to A(1)$, which is equal to $2e^{\gamma}$, where $\gamma = 0.57721\ldots$ is the Euler–Mascheroni constant. Moreover, by [18, Equation (11.62)], we have $A(\kappa), B(\kappa) \to 1$ as $\kappa \to 0$. By a very similar computation to [27, p. 248], we have

$$c_i c_{h_i} = \frac{e^{-\gamma\kappa_i}}{\Gamma(1+\kappa_i)}. \tag{4.52}$$

Therefore, the ratio of the constants in the bounds for S_1 and $S_4^{(i)}$ is $\kappa_i c_i c c_{h_i} AA_i/cB$, which is independent of S and tends to zero as $\alpha \to 0$. Also,

$$\lim_{\alpha \to 0} W(s) = (2-s)^{-1}.$$

Therefore,

$$\lim_{\alpha \to 0} S_4^{(i)} \ll \frac{(B + o(1))X}{(\log N)^{\kappa}} (\log 2 - \log \beta_i). \tag{4.53}$$

For all $\epsilon > 0$, by choosing β_i sufficiently close to 2, we have the bound

$$\lim_{\alpha \to 0} S_4^{(i)} \leqslant \frac{(\epsilon B + o(1))X}{(\log N)^{\kappa}} \ll \epsilon S_1. \tag{4.54}$$

We recall from Sections 4.4 and 4.5 that $S_2^{(i)}$ and $S_3^{(i)}$ are also negligible compared to S_1 as $\alpha \to 0$. Therefore, we see that $S(\mathscr{A}, \mathscr{P}, x) > 0$ for sufficiently small α and sufficiently large N.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

To obtain the best quantitative bounds, the choices of $\beta_i \in (1,2)$ should be optimised so as to minimise $S_3^{(i)} + S_4^{(i)}$. However, in practice, numerical computations suggest that the optimal choices for β_i are extremely close to 2, and little is lost in taking them arbitrarily close to 2, as above. In this case, the contributions from the sums $S_4^{(i)}$ are negligible. Taking a sum over i of the estimates from (4.37) and combining with (4.25), for any $\epsilon > 0$, we have

$$S(\mathscr{A},\mathscr{P},x) \geqslant \left(\frac{(c-\epsilon+o(1))X}{(\log N)^{\kappa}}\right)\left(B - A\alpha \sum_{i=m+1}^{k} \theta_i \int_1^2 (2-s)^{-\kappa}\frac{\mathrm{d}s}{s}\right).$$

Let $r(\kappa) = B/A$. Then we have established that $S(\mathscr{A},\mathscr{P},x) > 0$ provided that

$$r(\kappa) - \alpha \sum_{i=m+1}^{k} \theta_i \int_1^2 (2-s)^{-\kappa}\frac{\mathrm{d}s}{s} > 0. \tag{4.55}$$

We recall that $\theta = \theta_0 + \cdots + \theta_k$ and $\kappa = \alpha\theta$. Therefore, we may replace $\alpha\sum_{i=m+1}^{k}\theta_i$ with the trivial upper bound κ, after which we find by numerical computations (see Section 4.9 for details) that the largest value of κ we can take in (4.55) is $\kappa = 0.39000\ldots$. Thus the condition $\alpha\theta \leqslant 0.39$ is enough to ensure that $S(\mathscr{A},\mathscr{P},x) > 0$ for sufficiently large N. This completes the proof of Theorem 2.1.

4.8. Proof of Theorem 2.3

We now discuss the case $d = 3$, where $f(x,y)$ may contain irreducible factors of degree up to 3. We recall in the case $d = 2$, the sums $S_4^{(i)}$ could be made negligible compared to S_1 by choosing β_i arbitrarily close to 2, due to the factor $\log 2 - \log\beta_i$ appearing in (4.53). When $d = 3$, we obtain the same bound as in (4.53), but with $\log 2 - \log\beta_i$ replaced with $\log 2 - \log(\beta_i - 1)$. Consequently, $S_4^{(i)}$ is no longer negligible, even when $\alpha \to 0$ and β_i is arbitrarily close to 2. In fact, it can be checked that its limit as $\alpha \to 0$ and $\beta_i \to 2$ is larger than S_1, and so the above methods break down in this case.

However, we recall the additional hypothesis in Theorem 2.3 that Assumption 2.2 holds. By enlarging $\mathscr{P}$ if necessary, we may assume that equality holds in (2.9) – namely,

$$\mathscr{P}\backslash S = \bigcup_{j=1}^{n} \{p \equiv t_i \pmod{q_i}\}.$$

By Dirichlet's theorem on primes in arithmetic progressions, (2.4) holds with $\alpha \leqslant \sum_{j=1}^{n}\frac{1}{q_j - 1} \leqslant 0.32380$. Moreover, it follows from Lemma 5.8 (a version of the Chebotarev density theorem) that (2.5) holds for some $\theta_i \leqslant 3$. We now explain why this choice of $\mathscr{P}$ is easier to handle than arbitrary choices of $\mathscr{P}$ of the same density α.

When applying the switching principle for $S_4^{(i)}$, we wrote $f_i(a,b) = pr$ for $p \in \mathscr{P}$. Since we may assume $q_1,\ldots,q_n \in S_0$, the congruence condition $C(a,b)$ forces $f_i(a,b)$ to lie in a particular congruence class modulo $q_1,\ldots,q_n$. Combined with the fact that $p \equiv t_j \pmod{q_j}$ for some j, we see that r lies in a particular congruence class modulo q_j for some $j \in \{1,\ldots,n\}$, and this congruence class depends only on $t_j, C(a,b)$ and f. Moreover, by (4.26),

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

we have that $\gcd(r,\Delta)=\gcd(f_i(a_0,b_0),\Delta)$ depends only on $C(a,b)$ and f. We deduce that there exist $r_1,\dots,r_n$, depending only on $t_1,\dots,t_n,C(a,b)$ and f, such that $r':=|r|/\gcd(r,\Delta)$ lies in the set

$$\mathscr{T}:=\{r'\in\mathbb{N}:r'\equiv r_j \pmod{q_j} \text{ for some } j\in\{1,\dots,n\}\}.$$

Adding this condition into (4.43) and (4.44) and changing notation from r' back to r, we obtain

$$S_4^{(i)}\leqslant\frac{(cc_iAA_i+o(1))X}{(\log N)^{\kappa_i+\kappa}}\sum_{r\in\mathscr{T}_{\leqslant x}}w(r,\delta)h_i(r). \tag{4.56}$$

As demonstrated by Irving [27, Lemma 6.1], the argument based on [18, Theorem A.5] we used to obtain (4.46) can be generalised to give

$$\sum_{\substack{r\in\mathscr{T}_{\leqslant x}\\ \mu^2(r)=1}}h_i(r)\leqslant\sum_{j=1}^{n}\sum_{\substack{r\leqslant x\\ r\equiv r_j \pmod{q_j}}}h_i(r)=c_{h_i}(\log x)^{\kappa_i}\sum_{j=1}^{n}\frac{1}{q_j-1}+O((\log x)^{\kappa_i-1}). \tag{4.57}$$

Proceeding as before, we deduce the same estimate for $S_4^{(i)}$ as in (4.50), but with an additional factor $\alpha_0:=\sum_{j=1}^{n}\frac{1}{q_j-1}$. It is now clear qualitatively that for sufficiently large $q_1,\dots,q_n$ (i.e., as $\alpha_0\to 0$), the sums $S_4^{(i)}$ are once again negligible compared to S_1.

We now make the above discussion more quantitative in order to complete the proof of Theorem 2.3. Combining everything, we see that $S(\mathscr{A},\mathscr{P},x)>0$ for sufficiently large N provided that

$$r(\kappa)-\sum_{i:\deg f_i=2}\alpha\theta_i\int_1^2(2-s)^{-\kappa}\frac{\mathrm{d}s}{s}-\sum_{i:\deg f_i=3}\alpha_0A_i\kappa_ic_ic_{h_i}\int_0^1W(s)\mathrm{d}s>0. \tag{4.58}$$

We denote the left-hand side of (4.58) by $F(\boldsymbol{\theta})$. Recalling (4.51) and (4.52), we have

$$A_i\kappa_ic_ic_{h_i}\int_0^1W(s)\mathrm{d}s=\frac{A_i\kappa_ie^{-\gamma\kappa_i}\kappa^{-\kappa}\kappa_i^{-\kappa_i}(\kappa+\kappa_i)^{\kappa+\kappa_i}}{\Gamma(1+\kappa_i)}\int_0^1(2-s)^{-(\kappa+\kappa_i)}s^{\kappa_i-1}\mathrm{d}s. \tag{4.59}$$

For a fixed choice of $\kappa<1/2$, the integrand is a decreasing function of κ_i because $0\leqslant\frac{s}{(2-s)}\leqslant 1$ for any $s\in[0,1]$. The functions $\Gamma(1+\kappa_i)^{-1}$, $\kappa_i^{-\kappa_i}$ and $e^{-\gamma\kappa_i}$ are also decreasing in κ_i in the range $\kappa_i\in(1/2,1)$. Let $t=\alpha_0\deg f$. Since $\kappa_i\geqslant 1-\kappa\geqslant 1-t$, we therefore obtain an upper bound by replacing κ_i with $1-t$ in all these terms. The remaining terms in (4.59) are all increasing in κ_i, and we apply the trivial bound $\kappa_i\leqslant 1$. Finally, we note that $\kappa^{-\kappa}(\kappa+\kappa_i)^{\kappa+\kappa_i}$ is an increasing function in κ for $\kappa<1/2$, and so we may replace κ by t in this expression. Therefore, (4.59) can be bounded by

$$\frac{A(1)e^{\gamma(t-1)}t^{-t}(t+1)^{t+1}}{\Gamma(2-t)}\int_0^1(2-s)^{-1}s^{-\kappa}\mathrm{d}s. \tag{4.60}$$

We denote the factor outside the integral in (4.60) by $H(t)$. The integral in (4.60) is equal to the first integral in (4.58), as can be seen by making the substitution $u=2-s$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Therefore,

$$F(\boldsymbol{\theta}) \geqslant r(\kappa) - \left(\sum_{i:\deg f_i=2} \alpha\theta_i + \sum_{i:\deg f_i=3} \alpha_0 H(t) \right) \int_0^1 (2-s)^{-1} s^{-\kappa} \mathrm{d}s.$$

Recalling that $A(1) = 2e^{\gamma}$, it can be checked that $H(t) > 4$ for all $t > 0.2$, and so in particular, $H(t) > 2\theta_i$ whenever $\deg f_i = 2$. Moreover, r is a decreasing function of κ, so $r(\kappa) \geqslant r(t)$. We obtain

$$\begin{aligned} F(\boldsymbol{\theta}) &\geqslant r(t) - H(t) \int_0^1 (2-s)^{-1} s^{-t} \mathrm{d}s \left(\sum_{i:\deg f_i=2} \alpha/2 + \sum_{i:\deg f_i=3} \alpha_0 \right) \\ &\geqslant r(t) - \frac{tH(t)}{3} \int_0^1 (2-s)^{-1} s^{-t} \mathrm{d}s \end{aligned} \tag{4.61}$$

for any $t > 0.2$. We find by numerical computations that the above expression is positive provided that $t \leqslant 0.32380..$ Since $t = \alpha_0 \deg f$, this is implied by the condition (2.9) assumed in Theorem 2.3.

4.9. Details of the numerical computations

We now explain how we obtained numerical values with guaranteed error bounds in the proofs of Theorem 2.1 and Theorem 2.3. From [18, Equations (11.62), (11.63)], we have for $0 < \kappa < 1/2$ that

$$A(\kappa) = \frac{2e^{\gamma\kappa}}{\Gamma(1-\kappa)} \frac{q(\kappa)}{p(\kappa)+q(\kappa)}, \qquad B(\kappa) = \frac{2e^{\gamma\kappa}}{p(\kappa)+q(\kappa)},$$

where

$$\begin{aligned} p(\kappa) &= \int_0^{\infty} z^{-\kappa} \exp(-z + \kappa \operatorname{Ei}(-z)) \, \mathrm{d}z \\ q(\kappa) &= \int_0^{\infty} z^{-\kappa} \exp(-z - \kappa \operatorname{Ei}(-z)) \, \mathrm{d}z \end{aligned}$$

and where $\operatorname{Ei}(-z) = -\int_z^{\infty} \frac{e^{-u}}{u} \, \mathrm{d}u$ is the standard exponential integral. For any $0 < \kappa < 1/2$, the integrand defining $p(\kappa)$ is monotonically decreasing, decays exponentially as $z \to \infty$, and is bounded above by 2. Similarly, the integrand defining $q(\kappa)$ is monotonically decreasing and decays exponentially as $z \to \infty$, but diverges as $z \to 0$. To analyse its behaviour near 0, we note that for $\delta \in (0,1)$ and $\kappa < 0.4$, we have

$$\begin{aligned} &\int_0^{\delta} z^{-\kappa} e^{-z} \exp(-\kappa \operatorname{Ei}(-z)) \, \mathrm{d}z \\ &\leqslant \int_0^{\delta} z^{-0.4} \exp\left(0.4 \int_z^{\infty} \frac{e^{-u}}{u} \, \mathrm{d}u \right) \mathrm{d}z \\ &\leqslant \int_0^{\delta} z^{-0.4} \exp\left(0.4 \int_z^{1} \frac{\mathrm{d}u}{u} + 0.4 \int_1^{\infty} e^{-u} \, \mathrm{d}u \right) \mathrm{d}z \end{aligned}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$$\begin{aligned}&\leqslant \int_0^{\delta} z^{-0.4}\exp(-0.4\log(z)+0.4e^{-1})\,\mathrm{d}\,z\\&\leqslant 2\int_0^{\delta} z^{-0.8}\,\mathrm{d}\,z\\&\leqslant 10\delta^{0.2}.\end{aligned}$$

Therefore, we can obtain estimates for $p(\kappa),q(\kappa)$ (and hence also $A(\kappa),B(\kappa)$ with provable error bounds using standard numerical integration methods such as the trapezoid rule. The functions from (4.55) and (4.61) are bounded, continuous, and monotonically decreasing in κ, and so it is straightforward to find their roots to the desired precision. This was implemented in Sage, which produced the values 0.39000,0.32380 correct to five decimal places in 12.3 seconds on a standard laptop with 2 cores.

5. Application to the Hasse principle

In this section, we apply the main sieve results (Theorem 2.1 and Theorem 2.3) obtained in Section 4 in order to prove Theorem 1.1 and Theorem 1.3.

5.1. Algebraic reduction of the problem

Let K be a number field of degree n satisfying the Hasse norm principle. In [27, Lemma 2.6], Irving turns the problem of establishing the Hasse principle for

$$f(t)=\mathbf{N}(x_1,\dots,x_n)\neq 0 \tag{5.1}$$

into a sieve problem. Irving assumes that $f(t)\in\mathbb{Z}[t]$ is an irreducible cubic polynomial. However, in the following result, we demonstrate that Irving's strategy can be applied to establish a similar result for an arbitrary polynomial $f\in\mathbb{Z}[t]$. We recall that $f(x,y)$ denotes the homogenisation of f. Throughout this section, we make the choice

$$\mathscr{P}=\{p\notin S:\ \text{the inertia degrees of } p \text{ in } K/\mathbb{Q} \text{ are not coprime}\} \tag{5.2}$$

for a finite set of primes S containing all ramified primes in $K/\mathbb{Q}$.

Proposition 5.1. *Suppose that (5.1) has solutions over $\mathbb{Q}_p$ for every p and over $\mathbb{R}$. Let $\mathscr{P}$ and S be as in (5.2). Then there exists $\Delta\in\mathbb{N}$, divisible only by primes in S, integers a_0,b_0, and real numbers $r,\xi>0$ such that the following implication holds:*

Suppose that a,b are integers for which

(1) $a\equiv a_0 \pmod{\Delta}$ *and* $b\equiv b_0 \pmod{\Delta}$,

(2) $|a/b-r|<\xi$,

(3) $bf(a,b)$ *has no prime factors in* $\mathscr{P}$.

Then (5.1) has a solution over $\mathbb{Q}$.

By multiplicativity of norms, it suffices to find integers a,b such that b and $f(a,b)$ are in $N_{K/\mathbb{Q}}(K^*)$. Since K satisfies the Hasse norm principle, we have $\mathbb{Q}^*\cap N_{K/\mathbb{Q}}(I_K)=$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$N_{K/\mathbb{Q}}(K^*)$, where I_K denotes the ideles of K. Consequently, to show that $c \in \mathbb{Q}^*$ is a norm from K, it suffices to find elements $x_v \in K_v^*$ for each place v of K, such that

(1) For all but finitely many places v of K, we have $x_v \in \mathscr{O}_v^*$ (this ensures that $(x_v) \in I_K$).

(2) For all places w of $\mathbb{Q}$, we have

$$\prod_{v|w} N_{K_v/\mathbb{Q}_w}(x_v) = c. \tag{5.3}$$

The arguments in [27, Lemma 2.2, Lemma 2.3, Lemma 2.4] go through without changes. We summarise them below.

Lemma 5.2. *Suppose $c \neq 0$ is an integer. Then*

(1) *If $p \nmid c$ and $K/\mathbb{Q}$ is unramified at p, then there exist $x_v \in \mathscr{O}_v^*$ for each place v of K above p, such that $\prod_{v|p} N_{K_v/\mathbb{Q}_p}(x_v) = c$.*

(2) *Suppose that $K/\mathbb{Q}$ is unramified at p, and that the inertia degrees above p in K are coprime. Then there exist $x_v \in K_v^*$ for each place v of K above p, such that $\prod_{v|p} N_{K_v/\mathbb{Q}_p}(x_v) = c$.*

(3) *Let p be a place of $\mathbb{Q}$. Suppose that there exists $x_1,\dots,x_n \in \mathbb{Q}_p$ such that $c = \mathbf{N}(x_1,\dots,x_n)$. Then there exists $x_v \in K_v^*$ for each place v of K above p, such that $\prod_{v|p} N_{K_v/\mathbb{Q}_p}(x_v) = c$.*

We now give a slight generalisation of [27, Lemma 2.5] to the case of an arbitrary polynomial f.

Lemma 5.3. *Let p be a prime for which $f(t) = \mathbf{N}(x_1,\dots,x_n) \neq 0$ has a solution over $\mathbb{Q}_p$. Then there exists $a_0,b_0 \in \mathbb{Z}$ and $l \in \mathbb{N}$ (all depending on p) such that for any $a,b \in \mathbb{Z}$ satisfying $a \equiv a_0 \pmod{p^l}, b \equiv b_0 \pmod{p^l}$, we have $b, f(a,b) \in \mathbf{N}(\mathbb{Q}_p^n)\backslash\{0\}$.*

Proof. We define $N = \mathbf{N}(\mathbb{Q}_p^n)\backslash\{0\}$. Let $t_1 \in \mathbb{Q}_p$ be such that $f(t_1) = \mathbf{N}(x_1,\dots,x_n) \neq 0$ has a solution over $\mathbb{Q}_p$. Choose $a_1,b_1 \in \mathbb{Z}_p$ such that $\nu_p(b_1)$ is a multiple of n, and $a_1/b_1 = t_1$. Then $b_1 \in N$ and $f(a_1,b_1) \in N$.

The set $N \subseteq \mathbb{Q}_p$ is open, and so $N \times N \subseteq \mathbb{Q}_p^2$ is open. Moreover, the map $\varphi : \mathbb{Q}_p^2 \to \mathbb{Q}_p^2$ sending (a,b) to $(f(a,b),b)$ is continuous in the p-adic topology. Therefore, the set $\varphi^{-1}(N \times N)$ is open, and contains the element (a_1,b_1). Hence, there is a small p-adic ball with centre (a_1,b_1), all of whose elements (a,b) satisfy $b, f(a,b) \in N$. This ball can be described by congruence conditions $a \equiv a_0 \pmod{p^l}, b \equiv b_0 \pmod{p^l}$ for a sufficiently large integer l, as claimed in the lemma. □

Proof of Proposition 5.1. The proof closely follows the argument in [27, Lemma 2.6]. By the Hasse norm principle, to find solutions to (5.1), it suffices to find integers a,b such that properties (1) and (2) stated before Lemma 5.2 hold with $c = b$ and $c = f(a,b)$. We divide the places of $\mathbb{Q}$ into four sets:

(1) $p \in S$. Here, Lemma 5.3 gives congruence conditions $a \equiv a_{0,p} \pmod{p^l}$, $b \equiv b_{0,p} \pmod{p^l}$ which ensure that $b, f(a,b) \in N(\mathbb{Q}_p^n)\backslash\{0\}$. By part (3) of Lemma 5.2, we deduce that property (2) stated before Lemma 5.2 holds with $c = b$ and $c = f(a,b)$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

The congruence conditions at each prime $p \in S$ can be merged into one congruence condition $a \equiv a_0 \pmod{\Delta}, b \equiv b_0 \pmod{\Delta}$ using the Chinese remainder theorem.

(2) $p \notin S$ and $p \notin \mathscr{P}$. If $p \nmid b$, we apply part (1) of Lemma 5.2 with $c = b$. If $p \mid b$, we apply part (2) of Lemma 5.2 with $c = b$. The same argument works for $f(a,b)$ by choosing $c = f(a,b)$.

(3) $p \in \mathscr{P}$. Since we are assuming that $bf(a,b)$ has no prime factors in $\mathscr{P}$, for these primes, part (1) of Lemma 5.2 applies with $c = b$ and $c = f(a,b)$.

(4) $p = \infty$. We follow a similar argument to Lemma 5.3. We may assume that (5.1) is everywhere locally soluble, so in particular, there exists $r \in \mathbb{R}$ such that $f(r) \in \mathbf{N}(\mathbb{R}^n)\backslash\{0\}$. Since f is continuous and $\mathbf{N}(\mathbb{R}^n)\backslash\{0\}$ is open in the Euclidean topology, we can find $\xi > 0$ such that $f(t) \in \mathbf{N}(\mathbb{R}^n)\backslash\{0\}$ whenever $|t-r| < \xi$. Clearly, solubility of $f(t) = \mathbf{N}(\mathbf{x}) \neq 0$ and $f(-t) = \mathbf{N}(\mathbf{x}) \neq 0$ over $\mathbb{Q}$ are equivalent; consequently, we may assume $r > 0$. Suppose in addition that $t \in \mathbb{Q}$, and write $t = a/b$ for $a, b \in \mathbb{N}$. Since b is positive, it is automatically in $\mathbf{N}(\mathbb{R}^n)\backslash\{0\}$. By multiplicativity of norms, we conclude that $b, f(a,b) \in \mathbf{N}(\mathbb{R}^n)\backslash\{0\}$. The condition (5.3) now follows from part (3) of Lemma 5.2.

□

Example 5.4. Let $f(t) = (t^2-2)(-t^2+3)$ and $K = \mathbb{Q}(i)$, so that

$$f(x,y) = (x^2-2y^2)(-x^2+3y^2)$$
$$N_{K/\mathbb{Q}}(u,v) = u^2+v^2.$$

It is known that there is a Brauer–Manin obstruction to the Hasse principle for the equation $(t^2-2)(-t^2+3) = u^2+v^2 \neq 0$ by the work of Iskovskikh [28]. However, Proposition 5.1 still applies.

When $p \equiv 1 \pmod 4$, the prime p splits in $K/\mathbb{Q}$, and so the inertia degrees of p in $K/\mathbb{Q}$ are coprime. However, when $p \equiv 3 \pmod 4$, the prime p is inert in $K/\mathbb{Q}$ and has degree 2, so the inertia degrees are not coprime. Therefore, we have $\mathscr{P} = \{p : p \equiv 3 \pmod 4\}$. We choose $S = \{2\}$. In this example, it can be checked that the congruence conditions $a \equiv 8 \pmod{16}$ and $b \equiv 1 \pmod{16}$ are sufficient. Finally, for the infinite place, we just have the condition $f(a,b) > 0$. The sieve problem we obtain is to find integers a,b such that

(1) $a \equiv 8 \pmod{16}, b \equiv 1 \pmod{16}$,

(2) $f(a,b) > 0$,

(3) $f(a,b)$ has no prime factors $p \equiv 3 \pmod 4$.

We remark that $f(a/b) = b^{-4}f(a,b)$, and since $2 = [K:\mathbb{Q}]$ divides 4, b^{-4} is automatically a norm from K. This explains why in (3) above, we can consider prime factors of $f(a,b)$ rather than $bf(a,b)$.

An integer is the sum of two squares if and only if it is non-negative and all prime factors $p \equiv 3 \pmod 4$ occur with an even exponent. The above conditions are stronger than this, so the algebraic reduction performed in Proposition 5.1 is consistent with what we already knew for this example.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Since the Hasse principle fails for this example, we know that the above sieve problem cannot have a solution. This is indeed the case since condition (1) implies that $-a^2+3b^2 \equiv 3 \pmod 4$, and so $f(a,b)$ must contain a prime factor $p \equiv 3 \pmod 4$.

The fact that the above sieve problem has no solutions also does not contradict Theorem 2.1. For a prime $p > 3$, we have that $\nu_1(p)$ is 2 if $p \equiv \pm 1 \pmod 8$ and zero otherwise, and $\nu_2(p)$ is 2 if $p \equiv \pm 1 \pmod{12}$ and zero otherwise. Consequently, by Dirichlet's theorem on primes in arithmetic progressions, for $i \in \{1,2\}$, asymptotically one half of the primes $p \in \mathscr{P}$ have $\nu_i(p) = 2$ and half have $\nu_i(p) = 0$. We conclude that $\theta_1 = \theta_2 = 1$, so that $\theta = 2$. Theorem 2.1 therefore requires the density of $\mathscr{P}$ to be less than $0.39000\ldots/\theta = 0.19503\ldots$. However, the density of $\mathscr{P}$ here is $1/2$, and so Theorem 2.1 does not apply to this example.

5.2. The Chebotarev density theorem

In order to apply the sieve results from Section 2 to prove Theorem 1.1 and Theorem 1.3, we need to compute the densities (2.4) and (2.5) for the choice of $\mathscr{P}$ given in (5.2). For more complicated number fields K, such as when $K/\mathbb{Q}$ is non-abelian, it is not possible to write down the set $\mathscr{P}$ explicitly in terms of congruence conditions. However, we can still compute the densities (2.4) and (2.5) by appealing to the *Chebotarev density theorem*, which can be viewed as a vast generalisation of Dirichlet's theorem on primes in arithmetic progressions.

Let $K/\mathbb{Q}$ be a number field of degree n. Let p be a prime, unramified in $K/\mathbb{Q}$. We can factorise the ideal (p) as $(p) = \mathfrak{p}_1 \cdots \mathfrak{p}_r$, where $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are distinct prime ideals in $\mathscr{O}_K$. The *splitting type* of p in $K/\mathbb{Q}$ is the partition $(a_1, \ldots, a_r)$ of n, where a_i is the inertia degree of $\mathfrak{p}_i$ (i.e., $N(\mathfrak{p}_i) = p^{a_i}$). Equivalently, the splitting type is the list of degrees of the irreducible factors of the minimum polynomial of $K/\mathbb{Q}$, when factorised modulo p.

Suppose first that $K/\mathbb{Q}$ is Galois, with Galois group G. Then G acts transitively on $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_r\}$. Fix $i \in \{1, \ldots, r\}$. The *Decomposition* group $D_{\mathfrak{p}_i}$ is the stabilizer of $\mathfrak{p}_i$ under this action. Note that $D_{\mathfrak{p}_i}$ is cyclic, and there is an isomorphism

$$\psi_i : D_{\mathfrak{p}_i} \to \mathrm{Gal}((\mathscr{O}_K/\mathfrak{p}_i)/(\mathbb{Z}/(p))).$$

The group $\mathrm{Gal}((\mathscr{O}_K/\mathfrak{p}_i)/(\mathbb{Z}/(p)))$ is generated by the *Frobenius element* defined by $x \mapsto x^p$, which has order a_i. Let σ_i denote the preimage of the Frobenius element under ψ_i. We define the *Artin symbol*

$$\left[\frac{K/\mathbb{Q}}{p}\right] = \{\sigma_1, \ldots, \sigma_r\}.$$

The Artin symbol is a conjugacy class of G. Indeed, all the $\mathfrak{p}_i$'s lie in the same orbit of G (there is only one orbit as G acts transitively). Stabilisers of points in the same orbit of an action are conjugate, and so all the $D_{\mathfrak{p}_i}$ are conjugate.

We now come to the statement of the Chebotarev density theorem. For a conjugacy class C of G, we let $\pi_C(x)$ denote the number of primes $p \leqslant x$ whose Artin symbol is

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

equal to C. We define the (natural) *density* of a set of primes $\mathscr{P}$ to be

$$\lim_{x\to\infty}\left(\frac{\#\{p\in\mathscr{P}:p\leqslant x\}}{\pi(x)}\right),$$

if such a limit exists.

Theorem 5.5 (Chebotarev density theorem). *Let C be a conjugacy class of G. The density of primes p for which the Artin symbol is equal to C is $\#C/\#G$.*

In order to obtain the explicit error terms from (2.4) and (2.5), we need an effective version of the Chebotarev density theorem. The following lemma is a straightforward consequence of a more refined result due to Lagarias and Odlyzko [33, Theorem 1].

Theorem 5.6 (Effective Chebotarev density theorem). *For any $A\geqslant 1$, we have*

$$\pi_C(x)=\pi(x)\left(\frac{\#C}{\#G}+O_A((\log x)^{-A})\right),$$

where the implied constant may depend on K,C and A.

We now consider the non-Galois case. As above, let $(p)=\mathfrak{p}_1\cdots\mathfrak{p}_r$ be the factorisation of (p) in $\mathscr{O}_K$. Let $\widehat{K}$ denote the Galois closure of K, and let $G=\mathrm{Gal}(\widehat{K}/\mathbb{Q})$. This time, there is no action of G on $\{\mathfrak{p}_1,\dots,\mathfrak{p}_r\}$ because the $\mathfrak{p}_i$'s could split further in $\widehat{K}$, and elements of G could permute prime factors which are above different $\mathfrak{p}_i$'s.

To get around this, we define $H=\mathrm{Gal}(\widehat{K}/K)$, and instead consider the action of G on the set X of left cosets of H in G. For an element $\sigma\in G$, the cyclic group $\langle\sigma\rangle$ generated by σ acts by left multiplication on X. The sizes of the orbits of this action form a partition of $[G:H]=[K:\mathbb{Q}]=n$. Moreover, it can be checked that conjugate elements of G give the same orbit sizes, so we can associate a single partition of n with each Artin symbol $\left(\left[\frac{\widehat{K}/\mathbb{Q}}{p}\right]\right.$. The following fact relating this partition with the splitting type of p can be found in [30, Ch. 3, Proposition 2.8].

Lemma 5.7. *Let $\sigma\in\left[\frac{\widehat{K}/\mathbb{Q}}{p}\right]$. Then p has splitting type $(a_1,\dots,a_r)$ in $K/\mathbb{Q}$ if and only if the action of $\langle\sigma\rangle$ on X has orbit sizes $(a_1,\dots,a_r)$.*

Let $K=\mathbb{Q}(\alpha)$, and let g be the minimum polynomial of α. We can also view $\langle\sigma\rangle$ as acting on the set of n roots of g in $\widehat{K}$. By definition of H, we have that $\sigma\alpha=\sigma'\alpha$ if and only if $\sigma H=\sigma' H$. It follows that the orbit sizes of $\langle\sigma\rangle$ acting on X are the same as the orbit sizes of $\langle\sigma\rangle$ acting on the roots of g, which in turn are the cycle lengths of σ viewed as a permutation on the n roots of g in $\widehat{K}$. The set of $\sigma\in G$ with cycle lengths $(a_1,\dots,a_r)$ is a union $\bigcup_{i=1}^{s}C_i$ of conjugacy classes C_i. We may now apply Theorem 5.6 to each of these conjugacy classes separately. Putting everything together, we have the following result on densities of splitting types in non-Galois extensions.

Lemma 5.8. *Let K be a number field of degree n over $\mathbb{Q}$, and let $\widehat{K}$ denote its Galois closure. Let $G=\mathrm{Gal}(\widehat{K}/\mathbb{Q})$, viewed as a permutation group on the n roots of the minimum polynomial of K in $\widehat{K}$. For a partition $\mathbf{a}=(a_1,\dots,a_r)$ of n, let $\mathscr{P}(\mathbf{a})$ denote the set of*

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

primes with splitting type $\mathbf{a}$ *in* $K/\mathbb{Q}$*, and let* $T(\mathbf{a})$ *denote the proportion of elements of* G *with cycle shape* $\mathbf{a}$*. Then for any* $A \geqslant 1$*,*

$$\#\{p \in \mathscr{P}(\mathbf{a}) : p \leqslant x\} = \pi(x)\left(T(\mathbf{a}) + O_A((\log x)^{-A})\right).$$

5.3. Proof of Theorem 1.1 and Theorem 1.3

Proof of Theorem 1.1. Since K satisfies the Hasse norm principle, we may apply the algebraic reduction from Proposition 5.1. Therefore, it suffices to show that conditions (1)–(3) from Proposition 5.1 hold for $\mathscr{P}$ as defined in (5.2), and for some choice of S containing all the ramified primes in $K/\mathbb{Q}$. Let $F(x,y)$ be the binary form obtained from $yf(x,y)$ after removing any repeated factors. Clearly, to prove $bf(a,b)$ is free from prime factors in $\mathscr{P}$, it suffices to prove $F(a,b)$ is, and so we may replace $bf(a,b)$ with $F(a,b)$ in Condition (3) of Proposition 5.1. We apply Theorem 2.1 to the binary form $F(x,y)$, which has nonzero discriminant. We choose S to be the union of the ramified primes in $K/\mathbb{Q}$ and the set S_0 from Theorem 2.1, and we make the choice $\mathscr{R} = \mathscr{B}N$, where $\mathscr{B}$ takes the form (2.7) for the parameters $r,\xi > 0$ coming from the application of Proposition 5.1.

It remains only to check that (2.4) and (2.5) hold with $\alpha\theta \leqslant 0.39$. By Lemma 5.8, (2.4) holds with $\alpha = T(G)$. We now compute the quantity θ. We claim that $\theta \leqslant k + j + 1$. We recall that θ is a sum over the quantities θ_i associated to each irreducible factor of f, plus an additional term $\theta_0 = 1$ coming from the homogenising factor $f_0(x,y) = y$. We may therefore reduce to the case where f is itself irreducible of degree at most 2, with the goal of proving that

$$\theta \leqslant \begin{cases} 3, & \text{if } f \text{ is quadratic and } L \subseteq \widehat{K}, \\ 2, & \text{if } f \text{ otherwise,} \end{cases}$$

where L denotes the number field generated by f.

If f is linear, then $\nu_f(p) = 1$ for all $p \notin S$, and so $\theta = \theta_0 + 1 = 2$, as required. We now consider the case where f is an irreducible quadratic. Let

$$\nu_f(p) = \#\{t \pmod p : f(t) \equiv 0 \pmod p\}.$$

If $L \subseteq \widehat{K}$, then Lemma 5.8 could be applied to compute θ, with the desired error terms from (2.6). However, we apply the trivial bound $\nu_f(p) \leqslant 2$ for $p \notin S$ since it is not possible to improve on this in general. We therefore obtain $\theta \leqslant \theta_0 + 2 = 1 + 2 = 3$, which is satisfactory.

We now assume that $L \not\subseteq \widehat{K}$. We want to show that $\theta = 2$, or equivalently that $\nu_f(p) = 1$ on average over $p \in \mathscr{P}$. Let $M = \widehat{K}L$ be the compositum of $\widehat{K}$ and L. Since $\widehat{K} \cap L = \mathbb{Q}$, we have by [34, Ch. VI, Theorem 1.14] that $M/\mathbb{Q}$ is Galois, and

$$\mathrm{Gal}(M/\mathbb{Q}) \cong \mathrm{Gal}(L/\mathbb{Q}) \times \mathrm{Gal}(\widehat{K}/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathrm{Gal}(\widehat{K}/\mathbb{Q}).$$

We have $\nu_f(p) = 2$ if p is split in L, and $\nu_f(p) = 0$ if p is inert in L, and so

$$\theta = 1 + 2 \lim_{x\to\infty} \left(\frac{\#\{p \leqslant x : p \in \mathscr{P}, p \text{ split in } L\}}{\#\{p \leqslant x : p \in \mathscr{P}\}} \right). \tag{5.4}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Let $\sigma' = (\tau,\sigma)$ be an element of $\mathrm{Gal}(M/\mathbb{Q})$, where $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ and $\sigma \in \mathrm{Gal}(\widehat{K}/\mathbb{Q})$. Applying Lemma 5.7, the primes $p \in \mathscr{P}$ correspond to the σ' for which σ has non-coprime cycle lengths (so these primes have density $T(G)$ as mentioned above). If in addition, the prime p is split in L, then we require that $\tau = \mathrm{id}$. Therefore, by Lemma 5.8, asymptotically as $x \to \infty$, one half of the primes in $\mathscr{P}$ are also split in L. We conclude from (5.4) that (2.6) holds with $\theta = 1 + 2(1/2) = 2$. □

Proof of Theorem 1.3. We begin in the same manner as in the proof of Theorem 1.1, by appealing to Proposition 5.1 to reduce to a sieve problem. The binary form $F(x,y)$ has degree at most one higher than the degree of f. Therefore, we deduce from Theorem 2.3 that the Hasse principle holds for (1.1) provided that $(\deg f + 1)\sum_{j=1}^{n} \frac{1}{q_j - 1} \leqslant 0.32380$. □

Example 5.9. We consider the example $K = \mathbb{Q}(2^{1/q})$ discussed by Irving [27], where q is prime. Since K has prime degree, it satisfies the Hasse norm principle by work of Bartels [1]. We now compute $G = \mathrm{Gal}(\widehat{K}/\mathbb{Q})$. The minimum polynomial of $K/\mathbb{Q}$ is $x^q - 2$, which has roots $\{\beta, \beta\omega, \dots, \beta\omega^{q-1}\}$, where ω is a primitive qth root of unity and β is the real root of $x^q - 2$. We identify these roots with $\{0, \dots, q-1\}$ in the obvious way. An element $\sigma \in G$ is determined by the image of 0 and 1 since $\beta, \beta\omega$ multiplicatively generate all the other roots. Therefore, σ takes the form $\sigma_{a,b} : x \mapsto ax + b$ for some $a \in \mathbb{F}_q^{\times}, b \in \mathbb{F}_q$. Conversely, the maps $\sigma_{1,b}$ correspond to the q different embeddings K into $\widehat{K}$, and the maps $\sigma_{a,0}$ for $a \in \mathbb{F}_q^{\times}$ are elements of $\mathrm{Gal}(\widehat{K}/K) \leqslant G$. Combining these, we see that $\sigma_{a,b} \in G$ for any $a \in \mathbb{F}_q^{\times}, b \in \mathbb{F}_q$. We conclude that $G \cong \mathrm{AGL}(1,q)$, the group of affine linear transformations on $\mathbb{F}_q$.

When $a = 1$ and $b \neq 0$, $\sigma_{a,b}$ is a q-cycle. For all other choices of a,b, the equation $ax + b = x$ has a solution $x \in \mathbb{F}_q$, and so $\sigma_{a,b}$ has a fixed point. Therefore, $T(G) = (q-1)/\#G = 1/q$. From this, we see that Irving's choice of $\mathscr{P} = \{p \notin S : p \equiv 1 \pmod q\}$ is not quite optimal,because it has density $\alpha = 1/(q-1)$, whereas the set of primes we actually need to sift out has density $T(G) = 1/q$. In fact, we can see directly that even when $p \equiv 1 \pmod q$, there is sometimes a solution to $x^q - 2 \equiv 0 \pmod p$ (e.g., $q = 3$, $p = 31$, $x = 4$). However, it can be checked that even with this smaller sieve dimension, we are still not able to handle the cases $q = 5$ or $q = 3$ when f is an irreducible cubic.

We remark that we can replace the number 2 by any positive integer r such that $x^q - r$ is irreducible in the above example. (A necessary and sufficient condition for irreducibility of $x^q - r$ is given in [31, Theorem 8.16].) We can still take a sifting set $\mathscr{P}$ contained in $\{p \notin S : p \equiv 1 \pmod q\}$ and with density $1/q$, and the Galois group is still $\mathrm{AGL}(1,q)$, and so generalising to $x^q - r$ does not affect the analysis.

5.4. Proof of Corollary 1.2

We now consider the case when $[K : \mathbb{Q}] = n$ and $G = S_n$, with a view to proving Corollary 1.2. Such number fields automatically satisfy the Hasse principle by the work of Kunyavskiĭ and Voskresenskiĭ [32]. To ease notation, we shall write $T(n)$ in place of $T(S_n)$. In the following lemma, we find an estimate for $T(n)$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Lemma 5.10. *For all $n \geqslant 1$, we have*

$$T(n) = 1 - \sum_{k|n} \frac{\mu(k)\Gamma((n+1)/k)}{\Gamma(1/k)\Gamma(n/k+1)}, \tag{5.5}$$

$$T(n) < \frac{2}{\sqrt{\pi}} n^{1/r-1}\omega(n), \tag{5.6}$$

where r is the smallest prime factor of n and $\omega(n)$ is the number of prime factors of n.

Proof. Define

$$T_k(n) = \frac{1}{n!}\#\{\sigma \in S_n : \text{ the cycle lengths of } \sigma \text{ are all divisible by } k\}.$$

By Möbius inversion, we have $T(n) = 1 - \sum_{k|n} \mu(k) T_k(n)$. We now find an explicit formula for $T_k(n)$. For $j \geqslant 1$, let a_{jk} denote the number of cycles of length jk in σ. The cycle lengths of σ are all a multiple of k if and only if $\sum_{j=1}^{n/k} jka_{jk} = n$. We apply the well-known formula for the number of permutations of S_n with a given cycle shape to obtain

$$\begin{aligned} T_k(n) &= \frac{1}{n!} \sum_{\substack{a_k, a_{2k}, \ldots, a_n \\ \sum_{j=1}^{n/k} jka_{jk} = n}} \frac{n!}{\prod_{j=1}^{n/k} (jk)^{a_{jk}} a_{jk}!} \\ &= \sum_{\substack{b_1, \ldots, b_{n/k} \\ \sum_{j=1}^{n/k} jb_j = n/k}} \frac{1}{\prod_{j=1}^{n/k} (jk)^{b_j} b_j!} \\ &= \sum_{i=1}^{n/k} k^{-i} \sum_{\substack{b_1, \ldots, b_{n/k} \\ \sum_{j=1}^{n/k} jb_j = n/k \\ \sum_{j=1}^{n/k} b_j = i}} \frac{1}{\prod_{j=1}^{n/k} j^{b_j} b_j!} \\ &= \frac{1}{m!} \sum_{i=1}^{m} k^{-i} c(m,i), \end{aligned} \tag{5.7}$$

where $m = n/k$ and $c(m,i)$ is the number of $\sigma' \in S_m$ with exactly i cycles. The quantity $c(m,i)$ is called the *Stirling number of the first kind*. In order to evaluate (5.7), we follow the argument from [17, Example II.12]. We define a bivariate generating function

$$P(w,z) := \sum_{i=0}^{\infty} w^i \sum_{m=0}^{\infty} \frac{z^m}{m!} c(m,i). \tag{5.8}$$

By [17, Proposition II.4], we have

$$\sum_{m=0}^{\infty} \frac{z^m}{m!} c(m,i) = \frac{1}{i!} \left(\log \left(\frac{1}{1-z} \right) \right)^i.$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Therefore,

$$P(w,z)=\sum_{i=0}^{\infty}\frac{w^i}{i!}\left(\log\left(\frac{1}{1-z}\right)\right)^i=\exp\left(w\log\left(\frac{1}{1-z}\right)\right)=(1-z)^{-w}.$$

Applying the Binomial theorem, we find that the z^m coefficient of $P(w,z)$ is equal to $w(w+1)\cdots(w+m-1)/m!$. However, if we substitute $w=1/k$, the z^m coefficient of (5.8) is precisely (5.7). We conclude that

$$\begin{aligned}\frac{1}{m!}\sum_{i=1}^{m}k^{-i}c(m,i)&=(1/k)(1+1/k)\cdots(m-1+1/k)/m!\\&=\frac{\Gamma(m+1/k)}{\Gamma(1/k)\Gamma(m+1)},\end{aligned}$$

which completes the proof of (5.5).

We now establish the upper bound in (5.6). A basic bound on the gamma function is that for any real $s\in(0,1)$ and any positive real number x, we have

$$x^{1-s}<\frac{\Gamma(x+1)}{\Gamma(x+s)}<(1+x)^{1-s}.$$

Applying this with $x=m$ and $s=1/k$, we have

$$\frac{\Gamma(m+1/k)}{\Gamma(m+1)}<m^{1/k-1}.$$

Moreover, we have

$$\frac{1}{\Gamma(1/k)}=\frac{1}{\Gamma(1+1/k)}\frac{\Gamma(1+1/k)}{\Gamma(1/k)}\leqslant\frac{2}{\sqrt{\pi}k},$$

since $\Gamma(1+1/k)$ for integers $k\geqslant 2$ achieves its minimum at $k=2$, where we have $\Gamma(1+1/k)=\Gamma(3/2)=\sqrt{\pi}/2$. We conclude that

$$T_k(n)<\frac{2}{\sqrt{\pi}k}(n/k)^{1/k-1}=\frac{2}{\sqrt{\pi}}n^{1/k-1}k^{-1/k}<\frac{2}{\sqrt{\pi}}n^{1/k-1}.$$

Taking a sum over $k=p$ prime, we obtain

$$T(n)<\sum_{p\mid n}T_p(n)\leqslant\frac{2}{\sqrt{\pi}}n^{1/r-1}\omega(n),$$

as required. □

Proof of Corollary 1.2. We recall the setting of Corollary 1.2. We assume that $G=S_n$, and f is a product of two quadratics generating a biquadratic extension L of $\mathbb{Q}$. We apply Theorem 2.1 to the binary form $\prod_{i=0}^{2}f_i(x,y)$, where $f_0(x,y)=y$ and $f_1(x,y),f_2(x,y)$ are the homogenisations of the two quadratic factors of f. We also assume that $L\cap\widehat{K}=\mathbb{Q}$, and so by the proof of Theorem 1.1, we have $\theta_0=\theta_1=\theta_2=1$, and $\theta=3$. By maximising

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

the value of κ in (4.55) directly, we find by numerical computations that the largest value of κ we can take here is $0.42214\ldots$. (The slight improvement over $\kappa \leqslant 0.39$ for the general case comes from computing $\alpha \sum_{i=m+1}^{k} \theta_i = 2\alpha$ in our example, whilst in the proof of Theorem 1.1, we applied the trivial bound $\alpha \sum_{i=m+1}^{k} \theta_i \leqslant \kappa = 3\alpha$.) Hence, the Hasse principle holds for $f(t) = \mathbf{N}(\mathbf{x}) \neq 0$ provided that $T(n) \leqslant 0.42214\ldots/3 = 0.14071\ldots$.

We use the upper bound (5.6) from Lemma 5.10 to reduce the n for which $T(n) \geqslant 0.14071\ldots$ to finitely many cases, and then the exact formula (5.5) to find precisely which n satisfy $T(n) \geqslant 0.14071\ldots$. We find that $T(n) \leqslant 0.14071\ldots$ unless $n \in \{2,3,\ldots,10,12,14,15,16,18,20,22,24,26,28,30,36,42,48\}$. □

6. Application to the Harpaz–Wittenberg conjecture

In this section, we apply the sieve result from Theorem 2.1 to prove Theorem 1.4. We recall the statement of [23, Conjecture 9.1] (we shall only work with the ground field $\mathbb{Q}$).

Conjecture 6.1 (Harpaz, Wittenberg). *Let $P_1,\ldots,P_n \in \mathbb{Q}[t]$ be pairwise distinct irreducible monic polynomials. Let $k_i = \mathbb{Q}[t]/(P_i(t))$ be the corresponding number fields. Let $a_i \in k_i$ denote the class of t. For each $i \in \{1,\ldots,n\}$, let L_i/k_i be a finite extension, and let $b_i \in k_i^*$. Let S_0 be a finite set of places of $\mathbb{Q}$ including the archimedean place, and all finite places above which, for some i, either b_i is not a unit or L_i/k_i is ramified. For each $v \in S_0$, fix an element $t_v \in \mathbb{Q}_v$. Suppose that for every $i \in \{1,\ldots,n\}$ and every $v \in S_0$, there exists $x_{i,v} \in (L_i \otimes_{\mathbb{Q}} \mathbb{Q}_v)^*$ such that $b_i(t_v - a_i) = N_{L_i\otimes_{\mathbb{Q}}\mathbb{Q}_v/k_i\otimes_{\mathbb{Q}}\mathbb{Q}_v}(x_{i,v})$ in $k_i \otimes_{\mathbb{Q}} \mathbb{Q}_v$. Then there exists $t_0 \in \mathbb{Q}$ satisfying the following conditions.*

(1) *t_0 is arbitrarily close to t_v for all $v \in S_0$.*
(2) *For every $i \in \{1,\ldots,n\}$ and every place $\mathfrak{p}$ of k_i with $\mathrm{ord}_{\mathfrak{p}}(t_0 - a_i) > 0$, either $\mathfrak{p}$ lies above a place of S_0 or the field L_i possesses a place of degree 1 over $\mathfrak{p}$.*

We remark that below, the b_i and $x_{i,v}$ appearing in Conjecture 6.1 do not play a role, and so in the cases that Theorem 1.4 applies, it establishes a stronger version of Conjecture 6.1, where the assumption on the existence of the elements $x_{i,v}$ is removed. We discuss this further in Section 6.1.

We can reduce Conjecture 6.1 to a sieve problem as follows. Let $f_i(x,y) = c_i N_{k_i/\mathbb{Q}}(x - a_i y)$, where $c_i \in \mathbb{Q}$ is chosen such that the coefficients of $f_i(x,y)$ are coprime integers. Then $f_i(x,y)$ is an irreducible polynomial in $\mathbb{Z}[x,y]$.

Below, we suppose that S is a finite set of primes containing all primes in S_0 and all primes dividing any of the denominators $c_1,\ldots,c_n$. For $i \in \{1,\ldots,n\}$, we define $\mathscr{P}_i$ to be the set of primes $p \notin S$, such that for some place $\mathfrak{p}$ of k_i above p, L_i does not possess a place of degree 1 above $\mathfrak{p}$.

Lemma 6.2. *Let $k_1,\ldots,k_n$ and $L_1,\ldots,L_n$ and S_0 be as in Conjecture 6.1, and let $\mathscr{P}_i$ and $f_i(x,y)$ be as defined above. Suppose that there exists a finite set of primes $S \supset S_0 \backslash \{\infty\}$*

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

such that for any congruence condition $C(x,y)$ on x,y modulo an integer Δ with only prime factors in S, and any real numbers $r,\xi>0$, there exists $x_0,y_0\in\mathbb{N}$ such that

(i) *$C(x_0,y_0)$ holds,*

(ii) *$|x_0/y_0-r|<\xi$,*

(iii) *$f_i(x_0,y_0)$ has no prime factors in $\mathscr{P}_i$ for all $i\in\{1,\dots,n\}$.*

Then Conjecture 6.1 holds for this choice of $k_1,\dots,k_n,L_1,\dots,L_n$ and S_0.

Proof. From [23, Remark 9.3 (iii)], we are free to adjoin to S_0 a finite number of places, and so we may assume that $S_0=S\cup\{\infty\}$. We may assume that the value t_∞ prescribed in Conjecture 6.1 is non-negative, and let $t_0=x_0/y_0$. (If alternatively $t_\infty<0$, then we instead choose $t_0=-x_0/y_0$ and proceed similarly.) Then property (1) of Conjecture 6.1 immediately follows from (i) and (ii) by appropriate choices of $C(x,y),r$ and ξ. Let $\mathfrak{p}$ be a place of k_i above a prime $p\notin S$, satisfying $\mathrm{ord}_{\mathfrak{p}}(t_0-a_i)>0$. Then

$$f_i(x_0,y_0)=y_0^{\deg f_i}f_i(x_0/y_0,1)=y_0^{\deg f_i}c_iN_{k_i/\mathbb{Q}}(t_0-a_i).$$

Now $\mathrm{ord}_{\mathfrak{p}}(t_0-a_i)>0$ implies that $\mathrm{ord}_p(N_{k_i/\mathbb{Q}}(t_0-a_i))>0$. Since $p\notin S$, we have $\mathrm{ord}_p(y_0c_i)\geqslant 0$, and so $p\mid f_i(x_0,y_0)$. By (iii), we have $p\notin\mathscr{P}_i$, and so by construction of $\mathscr{P}_i$, we deduce that property (2) of Conjecture 6.1 holds. □

In view of Lemma 6.2, we let

$$\boldsymbol{\mathcal{P}}=(\mathscr{P}_1,\dots,\mathscr{P}_n),$$
$$\mathscr{R}=\{(x_0,y_0)\in[0,N]^2:|x_0/y_0-r|<\xi\},$$

and we aim to show that the sifting function

$$S(\mathscr{A},\boldsymbol{\mathcal{P}},x)=\#\{(x_0,y_0)\in\mathscr{R}\cap\mathbb{Z}^2:C(x_0,y_0),\gcd(f_i(x_0,y_0),P_i(x))=1\forall i\}\tag{6.1}$$

is positive for sufficiently large N. We do not attempt here to generalise Theorem 2.1 to deal with different sifting sets $\mathscr{P}_i$ for each i, but instead define below in (6.14) a set $\mathscr{P}\supseteq\bigcup_{i=1}^n\mathscr{P}_i$ and replace each of the conditions $\gcd(f_i(x_0,y_0),P_i(x))=1$ with $\gcd(f_i(x_0,y_0),P(x))=1$.

Proof of Theorem 1.4. We recall that L_i is the compositum k_iM_i, for a number field M_i which is linearly disjoint to k_i over $\mathbb{Q}$. Consequently, $[L_i:k_i]=[M_i:\mathbb{Q}]$. Writing $M_i=\mathbb{Q}(\beta_i)$ using the primitive element theorem, we therefore have that the minimum polynomial of β_i over $\mathbb{Q}$ and over k_i coincide. We denote this minimum polynomial by g_i.

Let $\mathfrak{p}$ denote a place of k_i. For all but finitely many places, the inertia degrees of the places of L_i above $\mathfrak{p}$ are the degrees of g_i when factored modulo $\mathfrak{p}$. If g_i has a root modulo p, then it has a root modulo every $\mathfrak{p}\mid p$, and so $p\notin\mathscr{P}_i$. Therefore, for suitably chosen S, we have

$$\mathscr{P}_i\subseteq\widetilde{\mathscr{P}_i}:=\{p\notin S:M_i\text{ does not possess a place of degree 1 above }p\}.\tag{6.2}$$

Define $\mathscr{P}=\bigcup_{i=1}^n\widetilde{\mathscr{P}_i}$. Clearly, to show that the sifting function $S(\mathscr{A},\boldsymbol{\mathcal{P}},x)$ from (6.1) is positive, it suffices to show the sifting function $S(\mathscr{A},\mathscr{P},x)$ (in the notation of (4.2)) is positive.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

By the Chebotarev density theorem (Lemma 5.8), the sets $\widetilde{\mathscr{P}_i}$ have density $\alpha_i = T_i$, where T_i is defined in (1.4). The sets $\widetilde{\mathscr{P}_i}$ are examples of *Frobenian sets* as defined by Serre [41, Section 3.3.1], which roughly means that away from S their membership is determined Artin symbols of some Galois extension (here $\widehat{M_i}$). It follows from [41, Proposition 3.7 b)] that the intersection of Frobenian sets is Frobenian. Since $\mathscr{P}$ is a disjoint union of such intersections, we conclude from Theorem 5.5 that the set $\mathscr{P}$ does indeed have a density. We shall bound its density trivially by $\sum_{i=1}^n \alpha_i$.

We now bound the value of θ defined in (2.6). Here, we have already defined f_i to be a binary form, and so no additional term θ_0 coming from homogenisation is required. We apply the trivial estimate $\nu_i(p) \leqslant \deg f_i = [k_i : \mathbb{Q}]$ for all $p \notin S$. We conclude that $\theta \leqslant \sum_{i=1}^n [k_i : \mathbb{Q}] = d$. Combining Lemma 6.2 and Theorem 2.1 completes the proof of part (1) of Theorem 1.4.

We now turn to the cubic case. Assumption 2.2 holds for our choice $\mathscr{P}$ and f. As in Section 4.8, we conclude that $S(\mathscr{A}, \boldsymbol{\mathcal{P}}, x) > 0$ for sufficiently large N, provided that $t \leqslant 0.32380$, where $t = \deg f \sum_{i=1}^n 1/(q_i - 1)$. Rearranging, and recalling $\deg f = d$, we complete the proof of part (2) of Theorem 1.4. □

Remark 6.3. In contrast to Section 5, now $\mathrm{Gal}(\widehat{M_i}/\mathbb{Q}) = S_n$ is not a case we can handle because there the proportion of fixed point free elements is $1 - 1/e$ as $n \to \infty$ (where $e = 2.718\ldots$ is Euler's constant), which is much too large.

For a permutation group G acting on $X = \{1, \ldots, k\}$, we define $h(G)$ to be the proportion of elements of G with no fixed point. The family of of groups G for which $h(G)$ is smallest are the *Frobenius groups*. These are the groups where G has a nontrivial element fixing one point of X, but no nontrivial elements fixing more than one point of X. We state two known results about Frobenius groups.

Lemma 6.4 [43, Theorem 1]. *Any Frobenius group can be realised as a Galois group over* $\mathbb{Q}$.

Lemma 6.5 [3, Theorem 3.1]. *Let G be a transitive permutation group on k letters.*

(1) *We have $h(G) \geqslant 1/k$, with equality if and only if G is a Frobenius group of order $k(k-1)$ and k is a prime power.*

(2) *In all other cases, $h(G) \geqslant 2/k$.*

Proof of Corollary 1.5. As computed in Example 5.9, we have that $G_i := \mathrm{Gal}(\widehat{M_i}/\mathbb{Q})$ is isomorphic to the group $\mathrm{AGL}(1, q_i)$ of affine linear transformations on $\mathbb{F}_{q_i}$. This is a Frobenius group of order $q_i(q_i - 1)$. By Lemma 6.5, we have $T_i = h(G_i) = 1/q_i$. (This also agrees with our computation in Example 5.9.) If $[k_i : \mathbb{Q}] \leqslant 2$ for all i, we can therefore apply part (1) of Theorem 1.4 provided that $\sum_{i=1}^n 1/q_i \leqslant 0.39/d$. Moreover, for all $i \in \{1, \ldots, n\}$, the sifting sets $\mathscr{P}_i$ are contained in $\{p \notin S : p \equiv 1 \pmod{q_i}\}$. Indeed, when $p \not\equiv 1 \pmod{q_i}$, the qth power map on $\mathbb{F}_p^\times$ is a bijection, and so $x^{q_i} - r_i$ has a root modulo p.

The minimum polynomial $x^{q_i} - r_i$ has a root modulo p for all but finitely many $p \not\equiv 1 \pmod{q_i}$, and these finitely many exceptional primes can be included in S_0. Therefore, we can apply part (2) of Theorem 1.4 provided that $\sum_{i=1}^n 1/(q_i - 1) \leqslant 0.32380/d$. □

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

6.1. The hypothesis on b_i

Given that the quantities b_i play no role in Theorem 1.4, it is natural to ask under what circumstances we should expect a stronger version of Conjecture 6.1 to hold, without the hypothesis on the b_i. The following proposition demonstrates that the hypothesis remains unchanged after passing to maximal abelian subextensions of each L_i/k_i.

Proposition 6.6. *Let L/k be an extension of number fields. Let L'/k be the maximal abelian subextension of L/k. Let S be a sufficiently large set of places of k, including all archimedean places. For each $v \in S$, fix an element $b_v \in k_v^*$. Then the following are equivalent:*

(1) *There exists an S-unit $b \in k$ such that for all $v \in S$, b/b_v is in the image of the norm map $(L \otimes_k k_v)^* \to k_v^*$.*

(2) *There exists an S-unit $b \in k$ such that for all $v \in S$, b/b_v is in the image of the norm map $(L' \otimes_k k_v)^* \to k_v^*$.*

We prove Proposition 6.6 at the end of this section. We now demonstrate how the lack of the hypothesis on b_i in Theorem 1.4 is explained by Proposition 6.6. We shall choose S to consist of all places that lie above a finite set of places S_0 of k, which corresponds to the set S_0 from Conjecture 6.1. We recall that due to [23, Remark 9.3 (iii)], we are free to make S_0, and hence S, large enough that Proposition 6.6 applies. We apply Proposition 6.6 with $L = L_i$ and $k = k_i$ for each extension L_i/k_i from Conjecture 6.1, and with $b_v = 1/(t_v - a_i)$. Clearly, if L/k contains no nontrivial abelian subextensions (so that $L' = k$), then Condition (2) above is trivially satisfied, and so Proposition 6.6 implies that the hypothesis on b_i in Conjecture 6.1 is vacuous. To complete the argument, it suffices to show that in the setting of Theorem 1.4, we have $L_i' = k_i$ for all i. In fact, in the following lemma, we show that the hypotheses of Theorem 1.4 force the stronger property that the L_i/k_i contain no nontrivial Galois subextensions.

Lemma 6.7. *Suppose that L/k is one of the extensions L_i/k_i from Theorem 1.4, and let $T = T_i$ be as in (1.4). Suppose that $T[k : \mathbb{Q}] < 1/2$. Then L/k has no nontrivial Galois subextensions.*

Proof. We recall from the proof of Theorem 1.4 that $T \geqslant \alpha$, where α is the natural density of the set $\mathscr{P}$ of primes $p \notin S$ such that there is some place $\mathfrak{p}$ of k above p for which L does not possess a place of degree 1 above $\mathfrak{p}$. For $x \geqslant 1$, we have the trivial bound

$$\begin{aligned} T\pi(x) &\geqslant \#(\mathscr{P}_{\leqslant x}) \\ &\geqslant \frac{1}{[k : \mathbb{Q}]} \#\{\mathfrak{p} \subseteq \mathscr{O}_k : N(\mathfrak{p}) \leqslant x, L \text{ has no degree one place above } \mathfrak{p}\}, \end{aligned}$$

since there are at most $[k : \mathbb{Q}]$ prime ideals $\mathfrak{p}$ above each p, and $N(\mathfrak{p}) \geqslant p$.

Suppose that N/k is a Galois subextension of L/k. In order for L to possess a degree one place above $\mathfrak{p}$, so must N. However, since N/k is Galois, N possesses a place of degree one above $\mathfrak{p}$ if and only in $\mathfrak{p}$ splits completely in N, and by the Chebotarev density theorem,

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

this occurs with density $1/[N:k]$. We deduce that

$$\begin{aligned} T\pi(x) &\geqslant \frac{1}{[k:\mathbb{Q}]}\#\{\mathfrak{p}\subseteq\mathscr{O}_k : N(\mathfrak{p})\leqslant x, N \text{ has no degree one place above } \mathfrak{p}\} \\ &\geqslant \frac{\#\{\mathfrak{p}\subseteq\mathscr{O}_k : N(\mathfrak{p})\leqslant x\}}{[k:\mathbb{Q}]}\left(1-\frac{1}{[N:k]}\right). \end{aligned}$$

Taking a limit at $x\to\infty$ and applying the prime ideal theorem, we conclude that

$$T\geqslant\frac{1}{[k:\mathbb{Q}]}\left(1-\frac{1}{[N:k]}\right).$$

The assumption $T[k:\mathbb{Q}]<1/2$ therefore implies that $N=k$. □

Let T denote the norm one torus associated to L/k, which is the algebraic group over k defined by the the equation $\mathbf{N}_{L/k}(\mathbf{x})=1$. We have a short exact sequence

$$1\to T\to R_{L/k}\mathbb{G}_m\xrightarrow{N_{L/k}}\mathbb{G}_m\to 1, \tag{6.3}$$

where $R_{L/k}$ denotes the Weil restriction. Let $T_{\overline{k}}=T\times_k\overline{k}$ and $\mathbb{G}_{m,\overline{k}}\cong\overline{k}^{\times}$. We define the character group of T to be $\widehat{T}=\mathrm{Hom}(T_{\overline{k}},\mathbb{G}_{m,\overline{k}})$, viewed as a Galois module via the natural action of $\mathrm{Gal}(\overline{k}/k)$.

Let $\mathscr{O}_S$ denote the ring of S-integers of k, and $\mathscr{O}_{L,S_L}$ the ring of S_L-integers of L, where S_L consists of all places of L above a place in S. We include in S all places of k which ramify in L. Then $\mathscr{O}_{L,S_L}/\mathscr{O}_S$ is étale, and so the equation $\mathbf{N}_{\mathscr{O}_{L,S_L}/\mathscr{O}_S}(\mathbf{z})=1$ defines a model $\mathscr{T}$ of T over $\mathscr{O}_S$. Similarly to (6.3), we have a short exact sequence

$$1\to\mathscr{T}\to R_{\mathscr{O}_{L,S_L}/\mathscr{O}_S}\mathbb{G}_{m,\mathscr{O}_{L,S_L}}\xrightarrow{N_{L/k}}\mathbb{G}_{m,\mathscr{O}_S}\to 1. \tag{6.4}$$

Let k_S denote the maximal subextension of $\overline{k}/k$ which is unramified at all places not contained in S. Below, we shall work with profinite group cohomology of $G_S:=\mathrm{Gal}(k_S/k)$. We note that L is a subextension of k_S since S is assumed to contain all ramified places of L/k. We may therefore define $G_{L,S}=\mathrm{Gal}(k_S/L)$. Let A_S denote the integral closure of $\mathscr{O}_S$ in k_S. The natural action of $\mathrm{Gal}(\overline{k}/k)$ on $\widehat{T}$ factors through G_S, so $\widehat{T}$ can be viewed as a G_S-module. The character group $\mathrm{Hom}(\mathscr{T}_{A_S},\mathbb{G}_{m,A_S})$ is nothing more than $\widehat{T}$ as a G_S-module, so we shall henceforth denote it by $\widehat{T}$.

Lemma 6.8. *Let $\varphi: H^1(G_S,\mathbb{Q}/\mathbb{Z})\to H^1(G_{L,S},\mathbb{Q}/\mathbb{Z})$ be the restriction map induced by the inclusion $G_{L,S}\hookrightarrow G_S$. Then there is an exact sequence*

$$0\to H^1(G_S,\widehat{T})\to H^1(G_S,\mathbb{Q}/\mathbb{Z})\xrightarrow{\varphi}H^1(G_{L,S},\mathbb{Q}/\mathbb{Z}). \tag{6.5}$$

Proof. We begin by taking character groups of the short exact sequence (6.4), or in other words, applying the contravariant functor $\mathrm{Hom}(-,\mathbb{G}_{m,A_S})$. we obtain a short exact sequence

$$0\to\mathbb{Z}\to\mathbb{Z}[L/k]\to\widehat{T}\to 0, \tag{6.6}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

where $\mathbb{Z}[L/k] \cong \mathbb{Z}[\mathscr{O}_{L,S_L}/\mathscr{O}_S]$ denotes the free abelian group generated by the k-linear embeddings $L \hookrightarrow \overline{k}$. We now take group cohomology of (6.6) to obtain a long exact sequence

$$\cdots \to H^1(G_S,\mathbb{Z}[L/k]) \to H^1(G_S,\widehat{T}) \to H^2(G_S,\mathbb{Z}) \to H^2(G_S,\mathbb{Z}[L/k]) \to \cdots. \tag{6.7}$$

For any groups $H \leqslant G$, any H-module N and any integer $i \geqslant 0$, Shapiro's lemma [37, Proposition 1.11] states that $H^i(H,N) \cong H^i(G, \mathrm{Coind}_H^G(N))$, where $\mathrm{Coind}_H^G(N) = \mathrm{Hom}_{\mathbb{Z}[H]}(\mathbb{Z}[G],N)$ denotes the coinduced module. We apply Shapiro's lemma to the first and last terms in the exact sequence (6.7) by choosing $G = G_S, H = G_{L,S}$ and $N = \mathbb{Z}$. Then

$$\mathrm{Coind}_H^G(N) = \mathrm{Hom}_{\mathbb{Z}[G_{L,S}]}(\mathbb{Z}[G_S],\mathbb{Z}) \cong \mathbb{Z}[G_S/G_{L,S}] \cong \mathbb{Z}[L/k],$$

and so we conclude that $H^i(G_S,\mathbb{Z}[L/k]) \cong H^i(G_{L,S},\mathbb{Z})$ for all $i \geqslant 0$. Moreover, $H^1(G_{L,S},\mathbb{Z})$ consists of all continuous group homomorphisms $G_{L,S} \to \mathbb{Z}$. Since $G_{L,S}$ is compact, and $\mathbb{Z}$ is discrete, we have $H^1(G_{L,S},\mathbb{Z}) = 0$. Therefore, we obtain from (6.7) the exact sequence

$$0 \to H^1(G_S,\widehat{T}) \to H^2(G_S,\mathbb{Z}) \to H^2(G_{L,S},\mathbb{Z}). \tag{6.8}$$

Now, for any subextension E of k_S/k, we have $H^2(G_{E,S},\mathbb{Z}) \cong H^1(G_{E,S},\mathbb{Q}/\mathbb{Z})$. To see this, we take group cohomology of the exact sequence $0 \to \mathbb{Z} \to \mathbb{Q} \to \mathbb{Q}/\mathbb{Z} \to 0$ to obtain a long exact sequence

$$\cdots \to H^1(G_{E,S},\mathbb{Q}) \to H^1(G_{E,S},\mathbb{Q}/\mathbb{Z}) \to H^2(G_{E,S},\mathbb{Z}) \to H^2(G_{E,S},\mathbb{Q}) \to \cdots,$$

and note that $H^1(G_{E,S},\mathbb{Q}) = H^2(G_{E,S},\mathbb{Q}) = 0$ since $G_{E,S}$ is a profinite group [39, Proposition 1.6.2 (c)]. Therefore, applying this fact with $E = k$ and $E = L$, we have an exact sequence

$$0 \to H^1(G_S,\widehat{T}) \to H^1(G_S,\mathbb{Q}/\mathbb{Z}) \xrightarrow{\psi} H^1(G_{L,S},\mathbb{Q}/\mathbb{Z}). \tag{6.9}$$

To complete the proof, we need to show that the map ψ we have obtained from the above argument is equal to the map φ defined in the lemma. We consider the diagram

$$\begin{array}{ccc}
H^2(G_S,\mathbb{Z}) & \longrightarrow & H^2(G_S,\mathbb{Z}[L/k]) \\
\| & & \downarrow{\scriptstyle \mathrm{sh}} \\
H^2(G_S,\mathbb{Z}) & \longrightarrow & H^2(G_{L,S},\mathbb{Z}) \\
\downarrow{\scriptstyle \delta^{-1}} & & \downarrow{\scriptstyle \delta^{-1}} \\
H^1(G_S,\mathbb{Q}/\mathbb{Z}) & \xrightarrow{\psi} & H^1(G_{L,S},\mathbb{Q}/\mathbb{Z})
\end{array}$$

Here, δ^{-1} denotes the inverse of the connecting homomorphism, and sh denotes the Shapiro map, (i.e., the isomorphism from the above application of Shapiro's lemma). By definition, ψ comes from applying these isomorphisms to the map $H^2(G_S,\mathbb{Z}) \to H^2(G_S,\mathbb{Z}[L/k])$ from (6.7), which is the map i_* in the notation of [39, Proposition 1.6.5]. Therefore, by [39, Proposition 1.6.5], the middle horizontal arrow is just the restriction

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

map. Finally, restriction maps commute with connecting homomorphisms [39, Proposition 1.5.2], and so ψ is the restriction homomorphism induced by the inclusion $G_{L,S} \hookrightarrow G_S$, as required. □

In what follows, we denote by $\mathrm{Cl}_{S_L}(L)$ the S_L-ideal class group (i.e., the quotient of the usual class group $\mathrm{Cl}(L)$ by the classes of all prime ideals in S_L). Since $\mathrm{Cl}(L)$ is finite, by adjoining finitely many primes to S, we may assume that $\mathrm{Cl}_{S_L}(L) = 0$.

Lemma 6.9. *Let S be as above. If $a \in k^*$ lies in the image of the norm map $(k_v \otimes_k L)^* \to k_v^*$ for all $v \notin S$, then there exists $y \in N_{L/k}(L^*)$ such that $ay \in \mathscr{O}_S^\times$.*

Proof. Fix a place $v \notin S$ at which a is not a unit. Let $w_1 \cdots w_r$ be the factorisation of v into prime ideals in $\mathscr{O}_L$, and let $c_i = [L_{w_i} : k_v]$ be the corresponding inertia degrees. (Since L/k is unramified outside S, the ideals $w_1, \ldots, w_r$ are distinct.) We define $c = \gcd(c_1, \ldots, c_r)$. The image of the norm map $(k_v \otimes_k L)^* \to k_v^*$ consists of the elements of k_v^* whose valuation is divisible by c. In particular, we may write $v(a) = \sum_{i=1}^r n_i c_i$ for some integers $n_1, \ldots, n_r$. Since $\mathrm{Cl}_{S_L}(L) = 0$, we can find an element $z_v \in L$ such that $w_i(z_v) = n_i$ for all $i \in \{1, \ldots, r\}$, and such that z_v is a unit at all other places not in S_L. Let $y_v = N_{L/k}(z_v)$. Then $v(y_v) = \sum_{i=1}^r n_i c_i = v(a)$, and y_v is a unit at all other places outside S. The result now follows by taking y to be the product of the elements y_v^{-1} over the (finitely many) places $v \notin S$ at which a is not a unit. □

Proof of Proposition 6.6. The implication (1) $\implies$ (2) is trivial, so we only need to prove (2) $\implies$ (1). Taking Galois cohomology of (6.3), we obtain a long exact sequence

$$1 \to T(k) \to L^\times \xrightarrow{N_{L/k}} k^\times \xrightarrow{\delta_k} H^1(k, T) \to H^1(k, R_{L/k}\mathbb{G}_m) \to \cdots. \tag{6.10}$$

Hilbert's Theorem 90 [11, Theorem 1.3.2] implies that $H^1(k, R_{L/k}\mathbb{G}_m)$ is trivial, and so we can naturally identify $H^1(k,T)$ with $\frac{k^\times}{N_{L/k}(L^\times)}$. In a similar way, we can identify $H^1(k_v, T)$ with the quotient of k_v^* by the image of the norm map $(L \otimes_k k_v)^* \to k_v^*$. We can also take group cohomology of the short exact sequence (6.4) to obtain a long exact sequence

$$1 \to \mathscr{T}(\mathscr{O}_S) \to \mathscr{O}_{L,S_L}^\times \to \mathscr{O}_S^\times \xrightarrow{\delta_{\mathscr{O}_S}} H^1(G_S, \mathscr{T}_{A_S}) \to H^1(G_S, A_{S_L}^\times) \to \cdots, \tag{6.11}$$

where $A_{S_L} = A_S \otimes_{\mathscr{O}_S} \mathscr{O}_{L,S_L}$ is the integral closure of $\mathscr{O}_{L,S_L}$ in k_S. (We remark that the ring A_S is denoted by $\mathscr{O}_S$ in [39].) By [39, Proposition 8.3.11 (ii)], we have $H^1(G_S, A_{S_L}^\times) \cong \mathrm{Cl}_{S_L}(L)$, which we recall is trivial due to our choice of S. Therefore, we may identify $H^1(G_S, \mathscr{T}_{A_S})$ with $\frac{\mathscr{O}_S^\times}{\mathscr{O}_S^\times \cap N_{L/k}(L^\times)}$.

Let $v \in S$, and let $\overline{k}_v$ denote an algebraic closure of k_v. Fixing an embedding $k_S \hookrightarrow \overline{k_v}$ determines a surjection $\mathrm{Gal}(\overline{k_v}/k_v) \to G_S$, which induces restriction maps $H^1(G_S, \mathscr{T}_{A_S}) \to H^1(k_v, T)$ and $H^1(k,T) \to H^1(k_v, T)$ called *localisation maps*. Below, we denote these maps by res_v. Under the above identifications, we obtain a commutative diagram

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

$$\begin{array}{ccccc} & & \overset{\mathrm{res}_v}{\frown} & & \\ H^1(G_S,\mathscr{T}_{A_S}) & \hookrightarrow & H^1(k,T) & \longrightarrow & H^1(k_v,T) \\ \cong\downarrow & & \cong\downarrow & & \cong\downarrow \\ \dfrac{\mathscr{O}_S^\times}{\mathscr{O}_S^\times \cap N_{L/k}(L^\times)} & \hookrightarrow & \dfrac{k^\times}{N_{L/k}(L^\times)} & \longrightarrow & \dfrac{k_v^*}{\mathrm{im}((L\otimes_k k_v)^* \to k_v^*)} \end{array} \tag{6.12}$$

where $(L\otimes_k k_v)^* \to k_v^*$ denotes the norm map, and the bottom arrows are induced by the inclusions $\mathscr{O}_S^\times \hookrightarrow k \hookrightarrow k_v^*$. Consequently, we can reformulate Condition (1) from Proposition 6.6 as

(3) The class of $(b_v)_{v\in S}$ in $\prod_{v\in S} H^1(k_v,T)$ belongs to the image of the map $\prod_{v\in S}\mathrm{res}_v$.

We assume that S is also large enough that $\mathrm{Cl}_{S_{L'}}(L') = 0$, so that the above argument gives a similar reformulation of (2), but with the torus T' associated to L'/k in place of T.

Let $(-)^\vee = \mathrm{Hom}(-,\mathbb{Q}/\mathbb{Z})$. Poitou–Tate duality [39, Theorem 4.20 b)] gives an exact sequence

$$H^1(k,T) \xrightarrow{\mathrm{res}} \prod{}' H^1(k_v,T) \xrightarrow{\xi} H^1(k,\widehat{T})^\vee. \tag{6.13}$$

Here, the restricted product is over all places v of k, with the added assumption that the specified element of $H^1(k_v,T)$ comes from $H^1(\mathscr{O}_v,T)$ for all but finitely many v, and the first map is induced by the residue maps res_v for each place v of k.

Let $\iota : \prod_{v\in S} H^1(k_v,T) \hookrightarrow \prod{}' H^1(k_v,T)$ be defined by adding trivial classes at the places $v\notin S$, and let $\xi_S = \xi\circ\iota$. We now explain how to deduce from (6.13) an exact sequence of the form

$$H^1(G_S,\mathscr{T}_{A_S}) \xrightarrow{\prod_{v\in S}\mathrm{res}_v} \prod_{v\in S} H^1(k_v,T) \xrightarrow{\xi_S} H^1(G_S,\widehat{T})^\vee. \tag{6.14}$$

Suppose that $b = (b_v)_{v\in S} \in \ker(\xi_S)$. Then $\iota(b)\in\ker\xi$. By the exactness of (6.13), this implies $\iota(b) = \mathrm{res}(a)$ for some $a\in H^1(k,T)$. Moreover, we know that $\mathrm{res}_v(a)$ is trivial at all places $v\notin S$. Recalling the identifications from (6.12), this means that a lies in the image of the norm map $(k_v\otimes_L k)^* \to k_v^*$ for all $v\notin S$. By Lemma 6.9, there is a representative of the class $a\in H^1(k,T)$ which is contained in $\mathscr{O}_S^\times$. Therefore, a in fact lies in $H^1(G_S,\mathscr{T}_{A_S})$. It follows that $\prod_{v\in S}\mathrm{res}_v(a) = b$, establishing the exactness of (6.14). Hence, condition (1) from the proposition is equivalent to the condition $(b_v)_{v\in S}\in\ker\xi_S$.

The map ξ_S is induced by the localisation maps $\prod_{v\in S}\mathrm{res}_v : H^1(G_S,\widehat{T}) \to \prod_{v\in S} H^1(k_v,\widehat{T})$ and the pairing

$$\prod_{v\in S}\Big(H^1(k_v,T)\times H^1(k_v,\widehat{T})\Big) \xrightarrow{\cup} \prod_{v\in S} H^2(k_v,\mathbb{G}_m) \xrightarrow{\sum_{v\in S}\mathrm{inv}_v} \mathbb{Q}/\mathbb{Z},$$

where $\cup$ denotes the cup product applied at each place $v\in S$ and inv_v denotes the local invariant map, as defined in [39, pp. 156].

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Consider the diagram

$$\begin{array}{ccc}
\prod_{v\in S} H^1(k_v,T) & \longrightarrow & \prod_{v\in S} H^1(k_v,T') \\
\downarrow \cup & & \cup \downarrow \\
\prod_{v\in S} \mathrm{Hom}\left(H^1(k_v,\widehat{T}),H^2(k_v,\mathbb{G}_m)\right) & \longrightarrow & \prod_{v\in S} \mathrm{Hom}\left(H^1(k_v,\widehat{T'}),H^2(k_v,\mathbb{G}_m)\right) \\
\downarrow \sum_{v\in S} \mathrm{inv}_v & & \sum_{v\in S} \mathrm{inv}_v \downarrow \\
\prod_{v\in S} H^1(k_v,\widehat{T})^\vee & \longrightarrow & \prod_{v\in S} H^1(k_v,\widehat{T'})^\vee \\
\downarrow \prod_{v\in S} \mathrm{res}_v & & \prod_{v\in S} \mathrm{res}_v \downarrow \\
H^1(G_S,\widehat{T})^\vee & \xrightarrow{\theta} & H^1(G_S,\widehat{T'})^\vee
\end{array}$$

where all the horizontal arrows are induced by the map $T \xrightarrow{N_{L/L'}} T'$. (The norm map $N_{L/L'} : R_{L/k}\mathbb{G}_m \to R_{L'/k}\mathbb{G}_m$ restricts to a map $T \to T'$ because $N_{L/k}(z) = N_{L/L'}(N_{L'/k}(z))$ for any $z \in L$.) This diagram commutes, thanks to the functorality properties of the localisation maps res_v [39, Proposition 1.5.2], and the projection formula for the cup product [39, Proposition 1.4.2]. The composition of the vertical arrows are the map ξ_S and the corresponding map ξ'_S for T'. Therefore, we obtain a commutative diagram

$$\begin{array}{ccc}
\prod_{v\in S} H^1(k_v,T) & \longrightarrow & \prod_{v\in S} H^1(k_v,T') \\
\downarrow \xi_S & & \xi'_S \downarrow \\
H^1(G_S,\widehat{T})^\vee & \xrightarrow{\theta} & H^1(G_S,\widehat{T'})^\vee
\end{array}$$

Applying Lemma 6.8, we have $H^1(G_S,\widehat{T}) \cong \ker\varphi$, where φ is as defined in the exact sequence (6.5). As discussed in the paragraphs preceding Lemma A.5, φ sends cyclic subextensions M/k of k_S (with a given choice of generator for $\mathrm{Gal}(M/k)$) to their compositum LM/L, and so elements of $\ker\varphi$ correspond to cyclic subextensions of L/k. However, since L'/k is the maximal abelian subextension of L/k, it contains all cyclic subextensions of L/k, and so the map $H^1(G_S,\widehat{T'}) \to H^1(G_S,\widehat{T})$ is surjective. Since $\mathbb{Q}/\mathbb{Z}$ is a divisible group, it is an injective object in the category of abelian groups, and so the contravariant functor $\mathrm{Hom}(-,\mathbb{Q}/\mathbb{Z})$ is exact. We deduce that θ is an injection. Therefore, if $\xi'_S((b_v)_{v\in S}) = 0$, then $\xi_S((b_v)_{v\in S}) = 0$, so $(b_v)_{v\in S} \in \ker\xi'_S \implies (b_v)_{v\in S} \in \ker\xi_S$. Recalling (6.14), this means that (2) $\implies$ (1). □

A. The Brauer group for the equation $f(t) = \mathbf{N}(\mathbf{x}) \neq 0$

This appendix will be concerned with the Brauer group of a smooth projective model X of the equation $f(t) = \mathbf{N}(\mathbf{x}) \neq 0$. In particular, we prove that in the setting of Corollary 1.2, we have $\mathrm{Br}(X) = \mathrm{Br}(\mathbb{Q})$ whenever $n \geqslant 3$. We are grateful to Colliot-Thélène for providing the arguments presented in this appendix.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

A.1. Main results

Theorem A.1. *Let k be a field of characteristic zero. Let K/k be an extension of degree n, and let L/k be the Galois closure. Suppose that $\mathrm{Gal}(L/k) = S_n$. Let $f(t) \in k[t]$ be a squarefree polynomial. Let Y/k be the affine variety given by the equation $f(t) = \mathbf{N}(x_1,\dots,x_n) \neq 0$, and $Y \to \mathbb{A}^1_k$ its projection onto t. Let $\pi : X \to \mathbb{P}^1_k$ be a smooth projective birational model of $Y \to \mathbb{A}^1_k$. Suppose that L and the number field generated by f are linearly disjoint over k. Then $\mathrm{Br}(k) = \mathrm{Br}(X)$.*

Theorem A.2. *Let k be a field of characteristic zero. Let K/k be a finite extension of degree $n \geqslant 3$, such that the Galois closure L/k satisfies $\mathrm{Gal}(L/k) = S_n$. Let $c \in k^\times$, and let Z be a smooth projective model of $\mathbf{N}(x_1,\dots,x_n) = c$. Then $\mathrm{Br}(k) \to \mathrm{Br}(Z)$ is surjective.*

A.2. Proof of Theorem A.2

Proof. The key ideas of the proof are discussed in detail by Bayer-Fluckiger and Parimala [2], and so here we just give a sketch. We would like to show that $\mathrm{Br}(Z)/\mathrm{Im}(\mathrm{Br}(k)) = 0$. We begin by reducing to the case $c = 1$. Suppose that T is the norm one torus given by $\mathbf{N}(x_1,\dots,x_n) = 1$, and let T^c denote a smooth compactification of T. Let k_s denote the separable closure of k, and let $\overline{Z} = Z \times_k k_s$, and $\overline{T} = T \times_k k_s$. By [13, Lemme 2.1], we have an isomorphism $H^1(k, \mathrm{Pic}\,\overline{Z}) \cong H^1(k, \mathrm{Pic}\,\overline{T}^c)$. Combining this with [2, Theorem 2.4], we have

$$\mathrm{Br}(Z)/\mathrm{Im}(\mathrm{Br}(k)) \hookrightarrow H^1(k, \mathrm{Pic}\,\overline{Z}) \cong H^1(k, \mathrm{Pic}\,\overline{T^c}) \cong \mathrm{Br}(T^c)/\mathrm{Br}(k), \tag{A.1}$$

and hence, it suffices to show that $\mathrm{Br}(T^c)/\mathrm{Br}(k) = 0$.

Let $G = \mathrm{Gal}(L/k)$. The character group $\widehat{T} = \mathrm{Hom}(T_{\overline{k}}, \mathbb{G}_{m,\overline{k}})$ can be viewed as a G-lattice. By [10, Proposition 9.5 (ii)], we have an isomorphism

$$\mathrm{Br}(T^c)/\mathrm{Br}(k) \cong Ш^2_{\mathrm{cycl}}(G, \widehat{T}),$$

where

$$Ш^2_{\mathrm{cycl}}(G, \widehat{T}) = \ker\left[H^2(G, \widehat{T}) \to \prod_{g \in G} H^2(\langle g \rangle, \widehat{T}) \right].$$

Let M/k be a finite extension with M linearly disjoint from L, and let $L' = LM, K' = KM, k' = kM$. Then the extension K'/k' has degree n and Galois closure L', with $\mathrm{Gal}(L'/k') = G$. Moreover, by a construction of Frölich [19], we may choose M in such a way that L'/k' is unramified.

Using [2, Proposition 4.1], we have $Ш^2_{\mathrm{cycl}}(G, \widehat{T}) \cong Ш^2(k', \widehat{T})$, where $\widehat{T}$ is regarded as a $\mathrm{Gal}(k'_s/k')$-module via the surjection $\mathrm{Gal}(k'_s/k') \to G$. In turn, this is isomorphic to $Ш^1(k', T)^\vee$ by Poitou–Tate duality [2, Corollary 4.5]. To summarise, we have isomorphisms

$$\mathrm{Br}(T^c)/\mathrm{Br}(k) \cong Ш^2_{\mathrm{cycl}}(G, \widehat{T}) \cong Ш^2(k', \widehat{T}) \cong Ш^1(k', T)^\vee, \tag{A.2}$$

and so it suffices to show that $Ш^1(k', T) = 0$. However, $Ш^1(k', T)$ is isomorhpic to the knot group $\kappa(K'/k') = \frac{k'^\times \cap N_{K'/k'}(\mathbb{A}^\times_{K'})}{N_{K'/k'}(k'^\times)}$. Due to the assumption $\mathrm{Gal}(L'/k') = G = S_n$, we

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

may apply the result of Kunyavskiĭ and Voskresenskiĭ [32] to deduce that the Hasse norm principle holds for the extension K'/k', and hence $\kappa(K'/k') = 0$. □

A.3. Proof of Theorem A.1

Before commencing with the proof, we require one more fact about the smooth projective model X from the statement of Theorem A.1.

Lemma A.3. *In the notation of Theorem A.1, the base change $X_K = X \times_k K$ is a K-rational variety.*

Proof. Since X is a smooth projective model of Y, it suffices to show that Y_K is K-rational. Let $K = k[x]/(p(x))$, where $p(x)$ is an irreducible polynomial over k. Let a denote the class of x. Over K, the polynomial $p(x)$ factorises as $p(x) = \prod_{i=0}^{r} q_i(x)$, where $q_0(x), \dots, q_r(x) \in K[x]$ are distinct and irreducible, and $q_0(x) = x - a$. Let $K_i = K[x]/(q_i(x))$. We shall construct a birational map of the form

$$\begin{aligned} \varphi : Y_K &\to \mathbb{A}^1_K \times \prod_{i=1}^{r} R_{K_i/K}\mathbb{A}^1 \\ (t, x_0, \dots, x_{n-1}) &\mapsto (t, z_1, \dots, z_r), \end{aligned} \tag{A.3}$$

where $R_{K_i/K}$ denotes the Weil restriction. Since $R_{K_i/K}\mathbb{A}^1 \cong \mathbb{A}^{\deg q_i}_K$, the right-hand side of (A.3) is isomorphic to $\mathbb{A}^n_K$, which is a Zariski open subset of $\mathbb{P}^n_K$. Therefore, φ induces a birational map $Y_K \dashrightarrow \mathbb{P}^n_K$, as desired.

Let $\overline{k}$ denote an algebraic closure of k. We denote by $\mathrm{Emb}_k(K, \overline{k})$ the embeddings $K \hookrightarrow \overline{k}$ fixing k, or in other words, the conjugates of K/k in $\overline{k}$. Over $\overline{k}$, the polynomials $p(x), q_0(x), \dots, q_r(x)$ split as

$$p(x) = \prod_{\sigma \in \mathrm{Emb}_k(K, \overline{k})} (x - \sigma(a)), \qquad q_i(x) = \prod_{\substack{\sigma \in \mathrm{Emb}_k(K, \overline{k}) \\ q_i(\sigma(a)) = 0}} (x - \sigma(a)). \tag{A.4}$$

For each i, we fix an isomorphism $K_i \cong K(\sigma_i(a))$ for some $\sigma_i \in \mathrm{Emb}_k(K, \overline{k})$ satisfying $q_i(\sigma_i(a)) = 0$, and view $\sigma_i(a)$ as the class of x in K_i/K. (The particular choice of representative σ_i does not matter.) Since $q_i(x)$ is the minimum polynomial of $\sigma_i(a)$ over K, it splits over $\overline{k}$ as the product of the conjugates of $\sigma_i(a)$, and so

$$q_i(x) = \prod_{\sigma' \in \mathrm{Emb}_K(K_i, \overline{k})} (x - \sigma' \sigma_i(a)).$$

For $i \in \{0, \dots, r\}$, we define $z_i \in R_{K_i/K}\mathbb{A}^1$ as

$$z_i = x_0 + \sigma_i(a) x_1 + \cdots + \sigma_i(a)^{n-1} x_{n-1}. \tag{A.5}$$

The polynomial $\sum_{j=0}^{\deg q_i - 1} z_i^{(j)} x^j$ representing z_i is the reduction of $x_0 + x_1 x + \cdots + x_{n-1} x^{n-1}$ modulo q_i. Consequently, by the Chinese remainder theorem, $z_0, \dots, z_r \in \prod_{i=0}^{r} R_{K_i/K}\mathbb{A}^1$ uniquely determine $x_0, \dots, x_{n-1} \in \mathbb{A}^1_K$.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

For any number field extension E/M, and any $y \in E$, we have

$$N_{E/M}(y) = \prod_{\sigma \in \mathrm{Emb}_M(E, \overline{M})} \sigma(y).$$

Therefore,

$$N_{K/k}(y) = \prod_{\sigma \in \mathrm{Emb}_k(K, \overline{k})} \sigma(y) = \prod_{i=0}^{r} \prod_{\sigma' \in \mathrm{Emb}_K(K_i, \overline{k})} \sigma' \sigma_i(y) = \prod_{i=0}^{r} N_{K_i/K}(\sigma_i(y)),$$

and so

$$\mathbf{N}(x_0, \dots, x_{n-1}) = \prod_{i=0}^{r} N_{K_i/K}(z_i). \tag{A.6}$$

We deduce that the equations (A.5) define an isomorphism from Y_K to the variety $V \subseteq \mathbb{A}^1_K \times \prod_{i=0}^{r} R_{K_i/K}\mathbb{A}^1$ given by $z_0 \prod_{i=1}^{r} N_{K_i/K}(z_i) = f(t) \neq 0$. For $t, z_1, \dots, z_r$ satisfying the Zariski open condition $\prod_{i=1}^{r} N_{K_i/K}(z_i) \neq 0$, we have $z_0 = f(t)/\prod_{i=1}^{n} N_{K_i/K}(z_i)$. Therefore, the projection of V onto $\mathbb{A}^1_K \times \prod_{i=1}^{r} R_{K_i/K}\mathbb{A}^1$ is birational. We conclude that the map φ from (A.3) is birational. □

We now commence with the proof of Theorem A.1. Let k be a field of characteristic zero. For a smooth irreducible variety X/k with function field $\kappa(x)$, we recall that $\mathrm{Br}(X)$ consists of all elements of $\mathrm{Br}(\kappa(X))$ which are unramified everywhere on X. Constant classes are unramified, and so we have $\mathrm{Br}(k) \subseteq \mathrm{Br}(X) \subseteq \mathrm{Br}(\kappa(X))$. By the purity theorem [11, Theorem 3.7.1], the ramification locus of $\mathscr{A} \in \mathrm{Br}(\kappa(X))$ is pure of codimension one. Consequently, to check $\mathscr{A} \in \mathrm{Br}(X)$, it suffices to check it is unramified at all codimension one points of X.

Let C be a codimension one point. We recall from [11, Section 1.4.3] the *residue map*

$$\partial_C : \mathrm{Br}(\kappa(X)) \to H^1(\kappa(C), \mathbb{Q}/\mathbb{Z})$$

is such that $\mathscr{A}$ is unramified at C if and only if $\partial_C(\mathscr{A})$ is trivial.

In our setting, codimension one points of X come in two types:

(1) Irreducible components of fibres $X_c = \pi^{-1}(c)$ above codimension one points c of $\mathbb{P}^1_k$,
(2) Codimension one points on the generic fibre X_η of $\pi : X \to \mathbb{P}^1_k$.

We recall that $\kappa(X) = \kappa(X_\eta)$. The codimension one points of X_η are a subset of the codimension one points of X, and so we have an inclusion $\mathrm{Br}(X) \hookrightarrow \mathrm{Br}(X_\eta)$. Since X_η is a smooth projective model of $\mathbf{N}_{K(t)/k(t)}(x_1, \dots, x_n) = f(t)$ over $k(t)$, it follows from Theorem A.2, applied to the extension $K(t)/k(t)$ and with $Z = X_\eta$, that $\mathrm{Br}(k(t)) \to \mathrm{Br}(X_\eta)$ is surjective. Putting everything together, we obtain a commutative diagram

$$\begin{array}{ccccc} \mathrm{Br}(X) & \hookrightarrow & \mathrm{Br}(X_\eta) & \hookrightarrow & \mathrm{Br}(\kappa(X_\eta)) = \mathrm{Br}(\kappa(X)) \\ \uparrow & & \uparrow\!\uparrow & & \\ \mathrm{Br}(k) & \hookrightarrow & \mathrm{Br}(k(t)) & & \end{array}$$

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

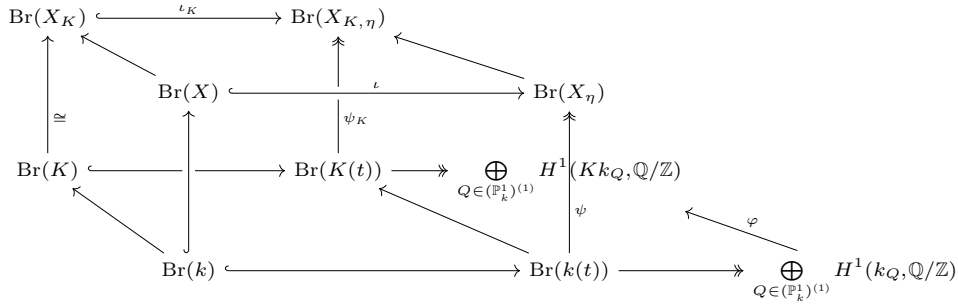
Let $\alpha \in \mathrm{Br}(X)$. By the above diagram, we can find $\beta \in \mathrm{Br}(k(t))$ whose image in $\mathrm{Br}(X_\eta)$ is equal to the image of α in $\mathrm{Br}(X_\eta)$. We want to show that β is the image of an element of $\mathrm{Br}(k)$ because then it follows from commutativity of the diagram that α is the image of an element of $\mathrm{Br}(k)$.

For any $n \geqslant 1$ and any field k, we have $\mathrm{Br}(\mathbb{P}^n_k) = \mathrm{Br}(k)$ [11, Theorem 6.1.3]. In particular, we have $\mathrm{Br}(k) = \mathrm{Br}(\mathbb{P}^1_k)$. Also, $k(t) = \kappa(\mathbb{P}^1_k)$, so $\mathrm{Br}(k(t)) = \mathrm{Br}(\kappa(\mathbb{P}^1_k))$. Therefore, as discussed above, to prove that β is in the image of $\mathrm{Br}(k)$, it suffices to show β is unramified at every codimension one point of $\mathbb{P}^1_k$. This is formalised by the *Faddeev exact sequence* [11, Theorem 1.5.2], which is the exact sequence

$$0 \to \mathrm{Br}(k) \hookrightarrow \mathrm{Br}(k(t)) \to \bigoplus_{Q \in (\mathbb{P}^1_k)^{(1)}} H^1(k_Q, \mathbb{Q}/\mathbb{Z}) \twoheadrightarrow H^1(k, \mathbb{Q}/\mathbb{Z}) \to 0, \tag{A.7}$$

where $(\mathbb{P}^1_k)^{(1)}$ denotes the codimension one points of $\mathbb{P}^1_k$ and the third map is the direct sum of the residue maps ∂_Q. In other words, to show that $\beta \in \mathrm{Br}(k(t))$ is actually in $\mathrm{Br}(k)$, it suffices to show that $\partial_Q(\beta) = 0$ for all $Q \in (\mathbb{P}^1_k)^{(1)}$. We have $\partial_Q(\beta) = 0$ unless Q is an irreducible factor of $f(t)$ by [11, Proposition 11.1.5], so we suppose from now on that Q is an irreducible factor of $f(t)$.

By Lemma A.3, the base change $X_K = X \times_k K$ is birational to $\mathbb{P}^n_K$. Since the Brauer group is a birational invariant on smooth projective varieties [11, Corollary 6.2.11], it follows that $\mathrm{Br}(X_K) = \mathrm{Br}(\mathbb{P}^n_K) = \mathrm{Br}(K)$. Therefore, we obtain the following commutative diagram:



The map φ is a direct sum over the restriction maps $\varphi_Q : H^1(k_Q, \mathbb{Q}/\mathbb{Z}) \to H^1(Kk_Q, \mathbb{Q}/\mathbb{Z})$ for each $Q \in (\mathbb{P}^1_k)^{(1)}$.

Lemma A.4. *We have $\partial_Q(\beta) \in \ker \varphi_Q$, where φ_Q is as defined above.*

Proof. Let β_K denote the image of β in $\mathrm{Br}(K(t))$, and $\partial_{K,Q}$ the residue map at Q on $\mathrm{Br}(K(t))$. We want to show that $\partial_{K,Q}(\beta_K) = 0$. By exactness of the Faddeev exact sequence over K, for this it suffices to show β_K is in the image $\mathrm{Br}(K) \to \mathrm{Br}(K(t))$. We know that $\psi(\beta) = \iota(\alpha)$. Applying a base change to K, we see that $\psi_K(\beta_K) = \iota_K(\alpha_K)$, where α_K, β_K are the images of α, β under base change. Since $\mathrm{Br}(K) \cong \mathrm{Br}(X_K)$, we have that α_K is in the image of $\mathrm{Br}(K) \to \mathrm{Br}(X_K)$, and hence by commutativity of the diagram, β_K is in the image of $\mathrm{Br}(K) \to \mathrm{Br}(K(t))$, as required. □

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

Let M be a number field, and let $G_M = \mathrm{Gal}(\overline{M}/M)$. For a finite Galois extension M'/M, we consider $\mathrm{Gal}(M'/M)$ as a topological space with the discrete topology. We can then put the profinite topology on G_M, which is defined as the inverse limit

$$G_M = \varprojlim_{M'/M \text{ Galois}} \mathrm{Gal}(M'/M).$$

We recall that $H^1(M,\mathbb{Q}/\mathbb{Z}) = \mathrm{Hom}_{\mathrm{cont}}(G_M,\mathbb{Q}/\mathbb{Z})$, the continuous group homomorphisms $G_M \to \mathbb{Q}/\mathbb{Z}$ [11, pp.16]. Suppose that $\theta \in \mathrm{Hom}_{\mathrm{cont}}(G_M,\mathbb{Q}/\mathbb{Z})$. Then $\ker\theta$ is an open subgroup of G_M. Since G_M is a profinite group, this implies that $\ker\theta$ has finite index in G_M. By the fundamental theorem of Galois theory, $\mathrm{im}\,\theta \cong G_M/\ker\theta \cong \mathrm{Gal}(M'/M)$, for some finite Galois extension M'/M. Moreover, $\mathrm{im}\,\theta$ is a finite subgroup of $\mathbb{Q}/\mathbb{Z}$. All finite subgroups of $\mathbb{Q}/\mathbb{Z}$ are cyclic groups of the form $\frac{1}{n}\mathbb{Z}/\mathbb{Z}$ for some positive integer n. Consequently, $\ker\theta = \mathrm{Gal}(M'/M)$ for a cyclic extension M'/M. To summarise, we have the identification

$$H^1(M,\mathbb{Q}/\mathbb{Z}) = \{M'/M \text{ cyclic, with a given map } \gamma : \mathrm{Gal}(M'/M) \hookrightarrow \mathbb{Q}/\mathbb{Z}\}.$$

We now describe $\varphi_Q : H^1(k_Q,\mathbb{Q}/\mathbb{Z}) \to H^1(Kk_Q,\mathbb{Q}/\mathbb{Z})$ explicitly. Using the above identification, we view an element $\theta \in H^1(k_Q,\mathbb{Q}/\mathbb{Z})$ as a pair $(M'/k_Q,\gamma)$. The map φ_Q is given by taking the compositum with K. More precisely, it sends the above pair to $(KM'/Kk_Q,\gamma)$, where now γ is viewed as a map $\mathrm{Gal}(KM'/KK_Q) \hookrightarrow \mathbb{Q}/\mathbb{Z}$ via the natural identification of $\mathrm{Gal}(KM'/Kk_Q)$ as a subgroup of $\mathrm{Gal}(M'/k_Q)$. Therefore, $(M'/k_Q,\gamma) \in \ker\varphi_Q$ if and only if M'/k_Q is a cyclic subextension of Kk_Q/k_Q.

Due to the assumption that k_Q and L are linearly disjoint over k, we have $\mathrm{Gal}(Kk_Q/k_Q) \cong \mathrm{Gal}(K/k) \cong S_n$. We now complete the proof of Theorem A.1 with the following elementary group theory fact.

Lemma A.5. *Suppose that K/k is a finite extension of degree $n \geqslant 3$, and the Galois group of the Galois closure $\mathrm{Gal}(L/k)$ is isomorphic to S_n. Then there are no nontrivial cyclic extensions M/k with $M \subseteq K$.*

Proof. By the fundamental theorem of Galois theory, if M/k is a subextension of K/k, then $\mathrm{Gal}(L/K) \leqslant \mathrm{Gal}(L/M) \leqslant \mathrm{Gal}(L/k) = S_n$. However, $\mathrm{Gal}(L/K) \cong S_{n-1}$. (More explicitly, if $K = k(\alpha_1)$ and $\alpha_1,\ldots,\alpha_n$ are the roots of the minimum polynomial of α_1 over k, then $\mathrm{Gal}(L/K)$ consists of all permutations of $\{\alpha_1,\ldots,\alpha_n\}$ which fix α_1.) However, S_{n-1} is a maximal subgroup of S_n, and so $M = k$ or $M = K$. Since $n \geqslant 3$, the extension K/k is not cyclic. Therefore, $M = k$. □

In conclusion, the map φ_Q is injective by Lemma A.5, and $\partial_Q(\beta) \in \ker\varphi_Q$ by Lemma A.4, and hence $\partial_Q(\beta) = 0$. This means that all the residue maps of β are trivial, so β is in the image of $\mathrm{Br}(k) \to \mathrm{Br}(k(t))$. Hence, α is in the image of $\mathrm{Br}(k) \to \mathrm{Br}(X)$, and so $\mathrm{Br}(X) = \mathrm{Br}(k)$. This completes the proof of Theorem A.1.

Acknowledgements. The author is grateful to Jean-Louis Colliot-Thélène for providing the statements and proofs in Appendix A, and to Julian Lyczak and Alexei Skorobogatov for helpful discussions on Brauer groups. The author would like to thank Olivier Wittenberg for many useful comments on the Harpaz–Wittenberg conjecture,

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

including suggesting the statement and proof of Proposition 6.6, and to Florian Wilsch for helpful discussions and feedback on an earlier version of Section 6.1. The author would like to thank Tim Browning for valuable feedback and guidance during the development of this work. Finally, the author is grateful to the anonymous referee for many useful comments on an earlier version of this paper. The author is supported by the University of Bristol and the Heilbronn Institute for Mathematical Research.

Competing interests. The authors have no competing interest to declare.

Data availability. No new data were created or analysed in this study. Data sharing is not applicable to this article.

References

[1] Bartels H-J (1981) Zur Arithmetik von Konjugationsklassen in algebraischen Gruppen. *J. Algebra.* **70**, 179–199.
[2] Bayer-Fluckiger E and Parimala R (2020) *On unramified Brauer groups of torsors over tori*. Doc. Math. 25.
[3] Boston N, Dabrowski W, Foguel T, Gies P, Jackson D, Ose D and Walker J (1993) The proportion of fixed-point-free elements of a transitive permutation group. *Comm. Algebra* **21**, 3259–3275.
[4] Browning T and Heath-Brown DR (2012) Quadratic polynomials represented by norm forms. *Geom. Funct. Anal.* **22**(5), 1124–1190.
[5] Browning T and Matthiesen L (2017) Norm forms for arbitrary number fields as products of linear polynomials. *Ann. Sci. Éc. Norm. Supér.* **50**, 1383–1446.
[6] Browning T and Schindler D (2019) Strong approximation and a conjecture of Harpaz and Wittenberg. *Int. Math. Res. Not. IMRN*, 4340–4369.
[7] Colliot-Thélène J-L (2003) Points rationnels sur les fibrations. In *Higher Dimensional Varieties and Rational Points*. Springer, 171–221.
[8] Colliot-Thélène J-L and Salberger P (1989) Arithmetic on some singular cubic hyper- surfaces. *Proc. Lond. Math. Soc.* **3**, 519–549.
[9] Colliot-Thélène J-L and Sansuc J-J (1977) La R-équivalence sur les tores. *Ann. Sci. Éc. Norm. Supér.* **10**, 175–229.
[10] Colliot-Thélène J-L and Sansuc J-J (1987) Principal homogeneous spaces under flasque tori: Applications. *J. Algebra* **106**, 148–205.
[11] Colliot-Thélène J-L and Skorobogatov A (2021) *The Brauer–Grothendieck Group*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics. Cham: Springer Nature.
[12] Colliot-Thélène J-L and Swinnerton-Dyer P (1994) Hasse principle and weak approximation for pencils of Severi–Brauer and similar varieties. *J. Reine Angew. Math.* **1994**, 49–112.
[13] Colliot-Thélène J-L, Harari D and Skorobogatov A (2003) *Valeurs d'un polynôme à une variable représentées par une norme. London Math. Soc. Lecture Note Ser.*, 69–90.
[14] Colliot-Thélène J-L, Sansuc J-J and Swinnerton-Dyer P (1987) Intersections of two quadrics and Châtelet surfaces. *I. J. Reine Angew. Math.* **373**, 37–107.
[15] Daniel S (1999) On the divisor-sum problem for binary forms. *J. Reine Angew. Math.* **507**, 107–129.
[16] Derenthal U, Smeets A and Wei D (2015) Universal torsors and values of quadratic polynomials represented by norms. *Math. Ann.* **361**, 1021–1042.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press

[17] Flajolet P and Sedgewick R (2009) *Analytic Combinatorics*. Cambridge : Cambridge University Press.
[18] Friedlander J and Iwaniec H (2010) *Opera de Cribro*, vol. 57. American Mathematical Society.
[19] Fröhlich A (1962) On non ramified extensions with prescribed Galois group. *Mathematika* **9**.
[20] Green B and Tao T (2010) Linear equations in primes. *Ann. of Math.*, 1753–1850.
[21] Green B, Tao T and Ziegler T (2012) An inverse theorem for the Gowers U s+1[N]-norm. *Ann. of Math.*, 1231–1372.
[22] Hardy G and Wright E (2008) *An Introduction to the Theory of Numbers*, 6th edn. Oxford University Press.
[23] Harpaz Y and Wittenberg O (2016) On the fibration method for zero-cycles and rational points. *Ann. of Math.*, 229–295.
[24] Harpaz Y, Wei D and Wittenberg O (2022) Rational points on fibrations with few non-split fibres. *J. Reine Angew. Math.* **2022**, 89–133.
[25] Heath-Brown DR and Moroz B (2002) Primes represented by binary cubic forms. *Proc. Lond. Math. Soc.* **84**, 257–288.
[26] Heath-Brown DR and Skorobogatov A (2002) Rational solutions of certain equations involving norms. *Acta Math.* **189**, 161–177.
[27] Irving A (2017) Cubic polynomials represented by norm forms. *J. Reine Angew. Math.* **2017**, 217–250.
[28] Iskovskikh V (1971) A counterexample to the Hasse principle for a system of two quadratic forms in five variables. *Mat. Zametki* **10**, 253–257.
[29] Iwaniec H and Kowalski E (2004) *Analytic Number Theory*, vol. 53. American Mathematical Society.
[30] Janusz J (1996) *Algebraic Number Fields*, vol. 7. American Mathematical Society.
[31] Karpilovsky G (1989 *Topics in Field Theory*. North-Holland Mathematics Studies . Elsevier.
[32] Kunyavskiĭ B and Voskresenskiĭ V (1984) Maximal tori in semisimple algebraic groups. *VINITI* **15**, 1269–1284, preprint.
[33] Lagarias J and Odlyzko A (1977) Effective versions of the Chebotarev density theorem. *Algebraic Number Fields* (A. Frölich edit.), 409–464.
[34] Lang S (2002) *Algebra*, revised third edn. Graduate Texts in Mathematics. Springer.
[35] Macedo A (2020) The Hasse norm principle for An-extensions. *J. Number Theory* **211**, 500–512.
[36] Matthiesen L (2018) On the square-free representation function of a norm form and nilsequences. *J. Inst. Math. Jussieu* **17**, 107–135.
[37] Milne J (2020) Class field theory (v4.03). Available at www.jmilne.org/math/.
[38] Mitsui T (1968) on the prime ideal theorem, dedicated to Professor S. Iyanaga on his 60th birthday. *J. Math. Soc. Japan* **20**, 233–247.
[39] Neukirch J, Schmidt A and Wingberg K (2013) *Cohomology of Number Fields*, vol. 323. Springer Science & Business Media.
[40] Schindler D and Skorobogatov A (2014) Norms as products of linear polynomials. *J. Lond. Math. Soc.* **89**, 559–580.
[41] Serre J-P (2012) *Lectures on NX (p)*. Boca Raton, FL : CRC Press.
[42] Skorobogatov A and Sofos E (2023) Schinzel hypothesis on average and rational points. *Invent. Math.* **231**, 673–739.
[43] Sonn J (1980) SL(2, 5) and Frobenius Galois groups over Q. *Canad. J. Math.* **32**, 281–293.

https://doi.org/10.1017/S1474748025000076 Published online by Cambridge University Press