OPEN ACCESS

University of BRISTOL

Peer reviewed version

University of Bristol - Explore Bristol Research
General rights

POWER-FREE VALUES OF POLYNOMIALS ON SYMMETRIC VARIETIES

T.D. BROWNING AND A. GORODNIK

ABSTRACT. Given a symmetric variety Y defined over $\mathbb{Q}$ and a non-zero polynomial with integer coefficients, we use techniques from homogeneous dynamics to establish conditions under which the polynomial can be made r-free for a Zariski dense set of integral points on Y. We also establish an asymptotic counting formula for this set. In the special case that Y is a quadric hypersurface, we give explicit bounds on the size of r by combining the argument with a uniform upper bound for the density of integral points on general affine quadrics defined over $\mathbb{Q}$.

CONTENTS

1. INTRODUCTION

Given a polynomial with integer coefficients, the problem of determining whether or not it takes infinitely many square-free values has long been a central concern in analytic number theory. More generally, one can ask for r-free values, for any $r \geqslant 2$, where an integer is said to be r-free if it is not divisible by p^r for any prime p. In this paper we initiate an investigation of r-free values of polynomials whose arguments run over *thin sets*.

Let $Y \subset \mathbb{A}^n$ be an affine variety defined by a system of polynomial equations with integer coefficients, with $Y(\mathbb{Z}) \neq \emptyset$, and let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a polynomial. Nevo and Sarnak [21] define the *saturation number* $r(Y, f)$ to be the least positive integer r such that the set of $\mathbf{x} \in Y(\mathbb{Z})$, for which $f(\mathbf{x})$ has

at most r prime factors, is Zariski dense in Y. They show that $r(Y, f)$ is finite whenever Y is a principal homogeneous space of a semisimple algebraic group and f is "weakly primitive". In a similar spirit, we can define the *permeation number* $r^\square(Y, f)$ to be the least integer $r \geqslant 2$ such that the set

$$\{\mathbf{x} \in Y(\mathbb{Z}) : \ f(\mathbf{x}) \text{ is } r\text{-free}\} \tag{1.1}$$

is Zariski dense in Y. The following natural condition becomes relevant in this setting. We say that the polynomial f has an r-*power divisor on* Y if there is a prime p such that $p^r \mid f(\mathbf{x})$ for every $\mathbf{x} \in Y(\mathbb{Z}_p)$, where $\mathbb{Z}_p$ denotes the ring of p-adic integers. It is clear that when the polynomial f has an r-power divisor on Y, the set (1.1) is empty. On the other hand, in this paper we show that for some classes of varieties and sufficiently large r, the set (1.1) is Zariski dense provided that f has no r-power divisor on Y. We also establish an asymptotic counting formula describing the distribution of this set.

One of the earliest examples arises in work of Erdős [14], who showed that $r^\square(\mathbb{A}^1, f) \leqslant d - 1$, provided that f has degree d and contains no 2-power divisors. Assuming the truth of the *abc*-conjecture, Poonen [23] has established the equality $r^\square(\mathbb{A}^n, f) = 2$ for any polynomial $f \in \mathbb{Z}[X_1, \ldots, X_n]$ without 2-power divisors. Our main result establishes finiteness of the permeation number $r^\square(Y, f)$ for generic f and a general class of *symmetric varieties* $Y \subset \mathbb{A}^n$ over $\mathbb{Q}$. Thus, let G be a connected semisimple algebraic group defined over $\mathbb{Q}$ and let $\iota : G \to \mathrm{GL}_n$ be an almost faithful linear representation, also defined over $\mathbb{Q}$. We assume that G acts transitively on Y and that $Y(\mathbb{Z})$ is non-empty. Then $Y \simeq G/L$, where L is an algebraic subgroup of G defined over $\mathbb{Q}$. The symmetric varieties dealt with here are assumed to satisfy the following properties:

  (i) the group L is a symmetric subgroup of G (i.e. the Lie algebra of L is equal to the set of fixed points of a non-trivial involution defined over $\mathbb{Q}$);
  (ii) the connected component of L has no non-trivial $\mathbb{Q}$-rational characters;
  (iii) the group G is $\mathbb{Q}$-simple and simply connected; and
  (iv) the group $G(\mathbb{R})$ has no compact factors.

It is known that the set $Y(\mathbb{Z})$ of integral points can be parametrised by orbits of the arithmetic group $\Gamma = \iota^{-1}(\mathrm{GL}_n(\mathbb{Z}))$. According to Borel and Harish-Chandra [4], the set $Y(\mathbb{Z})$ is a union of finitely many Γ-orbits. This allows us to study the set of r-free points using techniques from homogeneous dynamics.

It is very natural to demand that $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be devoid of r-power divisors on Y. It turns out that our argument also requires knowledge of the arithmetic function

$$\varrho(\ell) = \# \left\{\mathbf{x} \in Y(\mathbb{Z}/\ell\mathbb{Z}) : f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\,\ell)\right\}, \tag{1.2}$$

for $\ell \in \mathbb{N}$. This function is multiplicative, by the Chinese remainder theorem, and we can only handle f for which the prime power constituents of $\varrho(\ell)$ satisfy the following assumption.

Hypothesis-ϱ. For any $r \geqslant 1$, there exists a constant $C_{Y,f,r} > 0$, depending on r and the coefficients of Y and f, such that $\varrho(p^r) \leqslant C_{Y,f,r} p^{r(\dim(Y)-1)}$, for any prime p.

Let Z denote the variety $Y \cap \{f = 0\}$. When $r = 1$ the upper bound for $\varrho(p)$ in Hypothesis-ϱ follows from the Lang–Weil estimate if Z has codimension 1 in Y. If we further assume that Z is a non-singular affine variety of codimension 1 in Y, then Hypothesis-ϱ follows from an application of Hensel's lemma. Since Y is non-singular, it is worth emphasising that Hypothesis-ϱ holds for generic choices of f. We shall see that Hypothesis-ϱ is also satisfied for quadric hypersurfaces (see Lemma 4.2 below).

Bearing this in mind, we may now record our first main result.

Theorem 1.1. *Let $Y \simeq G/L \subset \mathbb{A}^n$ be a symmetric variety over $\mathbb{Q}$ satisfying (i)–(iv), with $Y(\mathbb{Z}) \neq \emptyset$. Assume that $f \in \mathbb{Z}[X_1, \ldots, X_n]$ satisfies Hypothesis-ϱ. Then $r^\square(Y, f) < \infty$.*

More precisely, we show that there exists r_0 such that for $r \geqslant r_0$, if the set (1.1) is not empty, then it is Zariski dense in Y. Moreover, if L is additionally assumed to be semisimple and simply connected, then for $r \geqslant r_0$, the set (1.1) is Zariski dense provided only that f does not have r-power divisors on Y. The value of r_0 is not made explicit in this work, but it can be estimated using our method. It depends on $\dim(G), \deg(f)$ and on the uniform spectral gap property that was shown by Burger–Sarnak [9] and Clozel [10] to be enjoyed by the action of each non-compact simple factor of $G(\mathbb{R})$ on the congruence quotients $G(\mathbb{R})/\Gamma_\ell$, where

$$\Gamma_\ell = \{\gamma \in \Gamma : \iota(\gamma) = id \bmod \ell\}. \tag{1.3}$$

Although we shall not pursue it here, we note that the arguments in this paper could also be used to generalise the finite saturation results of Nevo and Sarnak [21] to a broader class of symmetric varieties.

Our argument also allows us to establish an asymptotic formula for the number of r-free points. For $r \geqslant 2$ and a polynomial $f \in \mathbb{Z}[X_1, \ldots, X_n]$, define

$$N_r(Y, f; H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) \text{ is } r\text{-free}\}, \tag{1.4}$$

where $|\mathbf{x}| = \max_{1 \leqslant i \leqslant n} |x_i|$. The main term in the asymptotic formula for $N_r(Y, f; H)$ will involve a product of local densities which we proceed to define here. To define the real density, we assume that the variety Y is the zero locus

of a family of polynomials $f_1, \ldots, f_\ell \in \mathbb{Z}[X_1, \ldots, X_n]$ that satisfy

$$\operatorname{rank}\left(\frac{\partial f_i}{\partial X_j}\right) = n - \dim(Y) \tag{1.5}$$

everywhere on Y. Then we define the real density by

$$\mu_\infty(Y; H) = \lim_{\epsilon \to 0} \frac{1}{\epsilon^\ell} \int_{\substack{|\mathbf{x}| \leqslant H \\ |f_1(\mathbf{x})|, \ldots, |f_\ell(\mathbf{x})| < \epsilon/2}} d\mathbf{x}. \tag{1.6}$$

For each prime p, the p-adic density is

$$\hat{\mu}_p(Y, f, r) = \lim_{t \to \infty} p^{-t \dim(Y)} \#\{\mathbf{x} \in Y(\mathbb{Z}/p^t\mathbb{Z}) : \ p^r \nmid f(\mathbf{x})\}. \tag{1.7}$$

We also define the Euler product

$$\mathfrak{S}(Y, f, r) = \prod_{p < \infty} \hat{\mu}_p(Y, f, r). \tag{1.8}$$

If L is semisimple, then under Hypotheses-ϱ, this product converges absolutely.

With this notation, we prove the following result.

Theorem 1.2. *Let $Y \simeq G/L \subset \mathbb{A}^n$ be a symmetric variety over $\mathbb{Q}$ satisfying (i)–(iv). We assume that L is semisimple and simply connected. Let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a polynomial satisfying Hypothesis-ϱ. Then for all sufficiently large r, there exists a constant $\delta > 0$ such that*

$$N_r(Y, f; H) = \mathfrak{S}(Y, f, r)\mu_\infty(Y; H) + O_r(\mu_\infty(Y; H)^{1-\delta}).$$

Moreover, $\mathfrak{S}(Y, f, r) > 0$ provided that f does not have r-power divisors on Y.

Throughout our work, unless stated otherwise, we will allow our implied constants to depend on the polynomial f and the variety Y, which are considered to be fixed once and for all. Any further dependence will be explicitly indicated by appropriate subscripts. In Theorem 1.2, for example, the implied constant in the error term is allowed to depend on r, on f and on the polynomials defining Y.

We also establish an asymptotic formula for $N_r(Y, f; H)$ when L is not assumed to be a semisimple simply connected group (see Theorem 3.4 and Remark 3.5 below). However, without this assumption, the variety $Y \simeq G/L$ may fail to satisfy the local-to-global principle. Moreover, the definition of the Euler product (1.8) requires the introduction of additional convergence factors, so that the main term in the asymptotic formula becomes significantly more involved.

Our remaining results are concerned with producing explicit upper bounds for $r^\square(Y, f)$ for quadric hypersurfaces. For $n \geqslant 3$, let $Q \in \mathbb{Z}[X_1, \ldots, X_n]$ be a non-singular indefinite quadratic form and let m be a non-zero integer.

We shall always assume that $-m\det(Q)$ is not the square of an integer when $n = 3$. We let $Y \subset \mathbb{A}^n$ denote the affine quadric

$$Q(X_1, \ldots, X_n) = m. \tag{1.9}$$

We observe that our general results (Theorem 1.1 and 1.2) are applicable in this setting:

Remark 1.3. The assumptions (i)–(iv) are satisfied in the setting of quadric hypersurfaces (1.9) (with a possible exception of G being $\mathbb{Q}$-simple when $n = 4$, which we discuss separately in Remark 2.4). In the case of quadric hypersurfaces, $G = \mathrm{Spin}(Q)$ is the spinor group of Q and $\iota : G \to \mathrm{GL}_n$ is the standard representation of the spinor group of Q. We let $\Gamma = \iota^{-1}(\mathrm{GL}_n(\mathbb{Z}))$ and $L = \mathrm{Stab}_G(\mathbf{x}_0)$, with $\mathbf{x}_0 \in Y(\mathbb{Q})$. Thus

$$\dim(G) = \tfrac{1}{2}n(n-1) \quad \text{and} \quad \dim(L) = \tfrac{1}{2}(n-1)(n-2).$$

Moreover, L is a symmetric subgroup of G and $L \simeq \mathrm{Spin}(Q|_V)$, where V is the orthogonal complement of $\mathbf{x}_0$. In particular, when $n \geqslant 4$, it follows that L is a semisimple simply connected algebraic group, and when $n = 3$, L is a one-dimensional torus. We observe that $\det(Q|_V) = \det(Q)/m$, so that when $n = 3$, $Q|_V$ is equivalent to the quadratic form $x^2 + m\det(Q)y^2$. Hence, if $-m\det(Q)$ is not a square, L is anisotropic over $\mathbb{Q}$, and the assumption (ii) is satisfied. The group $G = \mathrm{Spin}(Q)$ is simply connected, so that $G(\mathbb{R})$ is connected. Moreover, $G(\mathbb{R}) \simeq \mathrm{Spin}(r_1, r_2)$, where (r_1, r_2) is the signature of the quadratic form Q. Since Q is assumed to be isotropic over $\mathbb{Q}$, $G(\mathbb{R})$ is not compact. It is simple unless $(r_1, r_2) = (2, 2)$, in which case $G(\mathbb{R}) \simeq \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. Hence, $G(\mathbb{R})$ has no compact factors. It also follows that G is $\mathbb{Q}$-simple, unless $(r_1, r_2) = (2, 2)$. We discuss the case $(r_1, r_2) = (2, 2)$ in Remark 2.4.

Thus $r^\square(Y, f) < \infty$ for quadratic hypersurfaces (1.9) with $Y(\mathbb{Z}) \neq 0$ and any integral polynomial f satisfying Hypothesis-ϱ. When $n \geqslant 4$ and certain necessary conditions are met, Baker [1] has used a variant of the Hardy–Littlewood circle method to show that there exist infinitely many points $\mathbf{x} \in Y(\mathbb{Z})$ with all the coordinates x_i square-free, provided that the obvious local conditions are satisfied. A modification of Baker's argument would easily give $r^\square(Y, X_i) = 2$, for any $i \in \{1, \ldots, n\}$, provided that $n \geqslant 4$. In this paper, we give explicit bounds on $r^\square(Y, f)$ and establish an asymptotic formula for $N_r(Y, f; H)$ when f is an arbitrary non-singular form.

We define the Euler product as in (1.8). When $n \geqslant 4$, this product is absolutely convergent, and positive provided that f does not have r-power divisors on Y. It is only conditionally convergent when $n = 3$ and $-m\det(Q)$ is not a square.

Our first result specific to quadrics concerns the asymptotic behaviour of $N_r(Y, f; H)$ in the easier case $n \geqslant 4$.

Theorem 1.4. *Let $n \geqslant 4$ and let $Y \subset \mathbb{A}^n$ be the quadric hypersurface (1.9). Assume that f is a non-singular form of degree $d \geqslant 2$ and let $r \geqslant dn^2/(n-1)$. Then there exists a constant $\delta > 0$ such that*

$$N_r(Y, f; H) = \mathfrak{S}(Y, f, r)\mu_\infty(Y; H) + O_r(H^{n-2-\delta}).$$

Moreover, $\mathfrak{S}(Y, f, r) > 0$ provided that f does not have r-power divisors on Y.

Here, we note that $\mu_\infty(Y; H) \sim \mu_\infty(Y)H^{n-2}$, as $H \to \infty$, for some constant $\mu_\infty(Y) > 0$.

The case $n = 3$ is much harder because quadric surfaces may fail to satisfy the local-to-global principle. This phenomenon can be analysed using a cohomological invariant introduced by Borovoi and Rudnick [6, 5]. This invariant is a locally constant function

$$\delta : Y(\mathbf{A}) \to \{0, 2\},$$

defined on the adelic space $Y(\mathbf{A}) = Y(\mathbb{R}) \times \prod'_{p<\infty} Y(\mathbb{Q}_p)$. If $Y(\mathbb{Q}) = \emptyset$, then $\delta \equiv 0$. Otherwise, we fix $\mathbf{x}_0 \in Y(\mathbb{Q})$. Let $G = \mathrm{Spin}(Q)$ be the spinor group of Q. Then G acts transitively on Y and $G(\mathbf{A})$ acts on $Y(\mathbf{A})$, but the latter action is not transitive. Orbits $\mathscr{O}_{\mathbf{A}}$ for this action are open in $Y(\mathbf{A})$ and they are restricted direct products

$$\mathscr{O}_{\mathbf{A}} = \prod'_{p \leqslant \infty} \mathscr{O}_p, \tag{1.10}$$

where each $\mathscr{O}_p$ is an open orbits of $G(\mathbb{Q}_p)$ in $Y(\mathbb{Q}_p)$. We define

$$\nu_p(\mathscr{O}_p) = \begin{cases} +1 & \text{if } \mathscr{O}_p = G(\mathbb{Q}_p)\mathbf{x}_0, \\ -1 & \text{if } \mathscr{O}_p \neq G(\mathbb{Q}_p)\mathbf{x}_0. \end{cases}$$

We note that $\nu_p(\mathscr{O}_p) = 1$ for almost all p. Let

$$\nu(\mathscr{O}_{\mathbf{A}}) = \prod_{p \leqslant \infty} \nu_p(\mathscr{O}_p).$$

One can show that ν is independent of the choice of $\mathbf{x}_0 \in Y(\mathbb{Q})$. The function is extended to elements of $Y(\mathbf{A})$ by setting $\nu(\mathbf{x}) = \nu(G(\mathbf{A})\mathbf{x})$, for any $\mathbf{x} \in Y(\mathbf{A})$. Next, we set $\delta = 1+\nu$. This defines a locally constant function on $Y(\mathbf{A})$. It was shown in [6, 5] that $\delta(\mathscr{O}_{\mathbf{A}}) = 0$ if and only if $\mathscr{O}_{\mathbf{A}}$ contains no rational points. (This theory can be also interpreted in terms of the *integral Brauer–Manin obstruction*, as worked out by Colliot-Thélène and Xu [11].)

As in (1.6)–(1.7), we define local densities of adelic orbits (1.10). Since orbits of $G(\mathbb{R})$ in $Y(\mathbb{R})$ are open and connected, they are equal to connected

components of the quadratic surface $Y(\mathbb{R})$. The real density is defined by

$$\mu_\infty(\mathscr{O}_\mathbf{A}; H) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{\substack{\mathbf{x} \in \mathscr{O}_\infty^*, \, |\mathbf{x}| \leqslant H \\ |Q(\mathbf{x}) - m| < \epsilon/2}} d\mathbf{x}, \qquad (1.11)$$

where $\mathscr{O}_\infty^*$ is a fixed neighbourhood of $\mathscr{O}_\infty$ which does not intersect the other connected components of $Y(\mathbb{R})$. The p-adic densities are defined by

$$\hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r) = \lim_{t \to \infty} p^{-t(n-1)} \#\{\mathbf{x} \in \mathscr{O}_p \cap Y(\mathbb{Z}_p) \bmod p^t : \ p^r \nmid f(\mathbf{x})\}.$$

We note that $Y(\mathbb{Z}_p) \subset \mathscr{O}_p$ for almost all p. Hence, $\hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r) = \hat{\mu}_p(Y, f, r)$ for almost all p. We also define the Euler product

$$\mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) = \prod_{p < \infty} \hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r),$$

which differs from the Euler product (1.8) only at finitely many factors.

For $n = 3$ we have the following result.

Theorem 1.5. *Let $n = 3$ and let $Y \subset \mathbb{A}^3$ be the quadric surface (1.9). Assume that f is a non-singular form of degree $d \geqslant 2$ and let $r \geqslant 4d^2 + \frac{4}{3}d$. Then there exists $\delta > 0$ such that*

$$N_r(Y, f; H) = \sum_{\mathscr{O}_\mathbf{A} \subset Y(\mathbf{A})} \delta(\mathscr{O}_\mathbf{A}) \mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) \mu_\infty(\mathscr{O}_\mathbf{A}; H) + O_r(H^{1-\delta}),$$

where the sum is taken over finitely many orbits $\mathscr{O}_\mathbf{A}$ that have non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p < \infty} Y(\mathbb{Z}_p)$.

The following is an immediate consequence of Theorems 1.4 and 1.5, giving an explicit version of Theorem 1.1 in the setting of quadric hypersurfaces.

Corollary 1.6. *Let $n \geqslant 3$ and let $Y \subset \mathbb{A}^n$ be the quadric hypersurface (1.9). Assume that f is a non-singular form of degree $d \geqslant 2$. We set*

$$r_0(n, d) = \begin{cases} 4d^2 + \frac{4}{3}d & \text{if } n = 3, \\ dn^2/(n-1) & \text{if } n \geqslant 4. \end{cases}$$

When $n \geqslant 4$, we denote by $r_0(Y, f)$ the least r such that f has no r-power divisors on Y. When $n = 3$, we denote by $r_0(Y, f)$ the least r such that there exists $\mathbf{x} \in Y(\mathbb{Z})$ with $f(\mathbf{x})$ being r-free. Then

$$r^\square(Y, f) \leqslant \max\{r_0(n, d), r_0(Y, f)\}.$$

We illustrate Theorem 1.5 with some examples borrowed from the work of Borovoi and Rudnick [6, 5].

Example 1.7 (§6.4.1 in [6])**.** Let $Q(X_1, X_2, X_3) = -9X_1^2 + 2X_1 X_2 + 7X_2^2 + 2X_3^2$ and let $m = 1$. Then the equation defining Y can be rewritten

$$(X_2 - X_1)(9X_1 + 7X_2) = 1 - 2X_3^2.$$

One easily checks that $(-\frac{1}{2}, \frac{1}{2}, 1)$ and $(\frac{1}{3}, 0, 1)$ are points in $Y(\mathbb{Q})$, so that there are solutions over $\mathbb{Z}_p$ for every prime p with $X_3 = 1$. Moreover, in the asymptotic formula for $N_2(Y, X_3; H)$ one finds that $\mathfrak{S}(Y, X_3, 2) > 0$. On the other hand, $Y(\mathbb{Z}) = \emptyset$. From the point of view of Theorem 1.5 this means that $\delta(\mathscr{O}_\mathbf{A}) = 0$ for all adelic orbits $\mathscr{O}_\mathbf{A}$ that have non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p < \infty} Y(\mathbb{Z}_p)$.

Example 1.8 (§4 in [5])**.** Let us assume that the hyperboloid $Y(\mathbb{R})$ has two connected components. Consider the involution

$$\iota : Y(\mathbf{A}) \to Y(\mathbf{A}), \quad (\mathbf{y}_\infty, \mathbf{y}_f) \mapsto (-\mathbf{y}_\infty, \mathbf{y}_f).$$

It is clear that ι maps orbits $\mathscr{O}_\mathbf{A}$ to orbits, and it follows from the definition of the invariant ν that $\nu(\iota(\mathscr{O}_\mathbf{A})) = -\nu(\mathscr{O}_\mathbf{A})$. Hence, $\delta(\iota(\mathscr{O}_\mathbf{A})) + \delta(\mathscr{O}_\mathbf{A}) = 2$. Moreover,

$$\mathfrak{S}(\iota(\mathscr{O}_\mathbf{A}), f, r) = \mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) \quad \text{and} \quad \mu_\infty(\iota(\mathscr{O}_\mathbf{A}); H) = \mu_\infty(\mathscr{O}_\mathbf{A}; H).$$

Hence, Theorem 1.5 implies that

$$\begin{aligned}
N_r(Y, f; H) &= \sum_{\mathscr{O}_\mathbf{A} \subset Y(\mathbf{A})} \mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) \mu_\infty(\mathscr{O}_\mathbf{A}; H) + O_r(H^{1-\delta}) \\
&= \mathfrak{S}(Y, f, r) \mu_\infty(Y; H) + O_r(H^{1-\delta}).
\end{aligned}$$

In this case the main term happens to satisfy the Hardy–Littlewood prediction even though the integral points are far from being equidistributed with respect to the orbits $\mathscr{O}_\mathbf{A}$. Indeed, among $\iota(\mathscr{O}_\mathbf{A})$ and $\mathscr{O}_\mathbf{A}$, only one of the sets contains integral points.

Example 1.9 (§3 in [5])**.** Assume that $Y(\mathbb{Z}) \neq \emptyset$, but there exists a quadratic form in the genus of Q which does not represent m over $\mathbb{Z}$. In this case, Theorem 1.5 gives

$$N_r(Y, f; H) = 2\mathfrak{S}(Y, f, r) \mu_\infty(Y; H) + O_r(H^{1-\delta}).$$

Indeed, in this case it was was proved in [5] that $Y(\mathbb{R}) \times \prod_{p < \infty} Y(\mathbb{Z}_p)$ is contained in a single orbit $\mathscr{O}_\mathbf{A}$ with $\delta(\mathscr{O}_\mathbf{A}) = 2$.

We can do better than Theorems 1.4 and 1.5 when f is linear, in which case one can actually produce an asymptotic formula for $N_r(Y, f; H)$, for all $r \geqslant 2$. This has the following outcome.

Theorem 1.10. *Let $n \geqslant 3$ and let $Y \subset \mathbb{A}^n$ be the quadric hypersurface* (1.9), *with $Y(\mathbb{Z}) \neq \emptyset$. Assume that f is a linear form having no 2-power divisors on Y. When $n = 3$, we additionally assume that there exists $\mathbf{x} \in Y(\mathbb{Z})$ such that $f(\mathbf{x})$ is square-free. Then $r^\square(Y, f) = 2$.*

We now return to the setting of a general symmetric variety $Y \simeq G/L \subset \mathbb{A}^n$ defined over $\mathbb{Q}$, with G, L satisfying (i)–(iv). Let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a polynomial that satisfies Hypothesis-ϱ. Recall the counting function $N_r(Y, f; H)$ from (1.4). The igniting spark in its analysis is provided by the indicator function

$$\sum_{k^r | N} \mu(k) = \begin{cases} 1 & \text{if } N \text{ is } r\text{-free}, \\ 0 & \text{otherwise}, \end{cases}$$

where $N \in \mathbb{Z}$ is non-zero and μ is the Möbius function. Thus

$$N_r(Y, f; H) = \sum_{k=1}^{\infty} \mu(k) \# \left\{ \mathbf{x} \in Y(\mathbb{Z}) : \begin{array}{l} |\mathbf{x}| \leqslant H \\ 0 \neq f(\mathbf{x}) \equiv 0 \, (\mathrm{mod}\, k^r) \end{array} \right\}. \quad (1.12)$$

Since f has degree d it is clear that the summand vanishes unless $k \ll H^{d/r}$. Moreover, since $Y(\mathbb{Z})$ consists of finitely many Γ-orbits, we may break the sum into residue classes modulo k^r and find that estimating it reduces to estimating

$$\#\{\mathbf{x} \in \Gamma\mathbf{y} : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \, (\mathrm{mod}\, k^r)\}, \quad (1.13)$$

for given $\mathbf{y} \in Y(\mathbb{Z})$ and given $\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z})$ such that $f(\boldsymbol{\xi}) \equiv 0 \, (\mathrm{mod}\, k^r)$. The sets $\{\mathbf{x} \in \Gamma\mathbf{y} : \mathbf{x} \equiv \boldsymbol{\xi} \, (\mathrm{mod}\, k^r)\}$ are finite unions of Γ_{k^r}-orbits, where Γ_ℓ is given by (1.3) for $\ell \in \mathbb{N}$. Thus the investigation of (1.13) reduces to establishing an asymptotic formula for $\#\{\mathbf{x} \in \Gamma_\ell\mathbf{y} : |\mathbf{x}| \leqslant H\}$, as $H \to \infty$, which is uniform in ℓ. This estimate is the focus of §2 and lies at the heart of this paper (see Theorem 2.1). The error term involves a polynomial dependence on ℓ, meaning that it is only useful for handling the contribution to $N_r(Y, f; H)$ from sufficiently small values of k^r.

By taking r sufficiently large we can ensure that k is an arbitrarily small power of H. In this way, on observing that

$$\# \left\{ \mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ k^r \mid f(\mathbf{x}) \right\} \leqslant \# \left\{ \mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ k^2 \mid f(\mathbf{x}) \right\},$$

it is possible to reapply the results from §2 with $\ell = k^2$, in order to show that the larger values of k^r make a negligible contribution to $N_r(Y, f; H)$. This summarises our strategy behind the proof of Theorem 1.2. The proof of Theorem 1.1 requires a generalisation of Theorem 1.2, which gives an asymptotic formula for the number r-free points lying on a given adelic orbit (see §3.3).

Our proof of Theorems 1.4, 1.5 and 1.10 gets under way in §4 and relies on a more efficient method for handling the contribution from large values of k^r. Thus, when $Y \subset \mathbb{A}^n$ is given by (1.9), we will transform the problem into

one that involves counting integral points of bounded size on affine quadrics. Our bound needs to be uniform in the coefficients of the defining polynomial and, since it may be of general interest, we proceed to describe it here. Let $q \in \mathbb{Z}[T_1, \ldots, T_\nu]$ be a non-zero quadratic polynomial, for $\nu \geqslant 2$. Let

$$M(q; B) = \#\{\mathbf{t} \in \mathbb{Z}^\nu : |\mathbf{t}| \leqslant B, \ q(\mathbf{t}) = 0\},$$

for any $B \geqslant 1$. We will require an upper bound for $M(q; B)$ which is uniform in the coefficients of q and which is essentially as sharp and as general as possible. A trivial estimate is $M(q; B) = O_\nu(B^{\nu-1})$, which is optimal when q is reducible over $\mathbb{Q}$. Assuming that q is irreducible over $\mathbb{Q}$, a result of Pila [22] gives $M(q; B) = O_{\varepsilon, \nu}(B^{\nu-3/2+\varepsilon})$, for any $\varepsilon > 0$. Again, this is essentially best possible, as consideration of the polynomial $T_1 - T_2^2$ shows. Let q_0 denote the quadratic part of q, so that $q_0 = T_2^2$ in the previous example. One might hope for an improved bound when q_0 has rank at least 2. This is confirmed in the following result, which is a straightforward modification of ideas developed by Browning, Heath-Brown and Salberger [8, §§4–5].

Theorem 1.11. *Let $q \in \mathbb{Z}[T_1, \ldots, T_\nu]$ be quadratic, with $\nu \geqslant 2$. Let $\varepsilon > 0$. Assume that q is irreducible over $\mathbb{Q}$ and that $\mathrm{rank}(q_0) \geqslant 2$. Then*

$$M(q; B) = O_{\varepsilon, \nu}(B^{\nu-2+\varepsilon}).$$

The implied constant in this result depends only on the choice of ε and the number ν. This is the most important feature of Theorem 1.11, since it would be easy to prove a version of the theorem with an implied constant that is allowed to depend on q by first diagonalising q_0 and then completing the square where possible.

2. Counting on symmetric varieties with congruences

2.1. **The main estimate.** In this section we establish an asymptotic counting estimate for integral points on symmetric varieties that satisfy a congruence condition. Let

$$Y \simeq G/L \subset \mathbb{A}^n$$

be a symmetric variety satisfying the hypotheses (i)–(ii) from §1 and

   (iii′) the group G is $\mathbb{Q}$-simple; and
   (iv′) the group $G(\mathbb{R})$ is connected and has no compact factors.

When G is simply connected, $G(\mathbb{R})$ is connected, so that conditions (iii′)–(iv′) are weaker than conditions (iii)–(iv).

Recalling the definition (1.3) of Γ_ℓ, our aim is to estimate the cardinality of the sets

$$\{\mathbf{x} \in \Gamma_\ell \mathbf{y} : |\mathbf{x}| \leqslant H\},$$

as $H \to \infty$, uniformly in ℓ.

For $\mathbf{y} \in Y(\mathbb{Z})$, we set

$$L_{\mathbf{y}} = \mathrm{Stab}_G(\mathbf{y}) \quad \text{and} \quad B_H(\mathbf{y}) = \{\mathbf{x} \in G(\mathbb{R})\mathbf{y} : |\mathbf{x}| \leqslant H\}.$$

We fix compatible volume forms m_G, m_L, m_Y on $G(\mathbb{R})$, $L_{\mathbf{y}}(\mathbb{R})$, $G(\mathbb{R})\mathbf{y}$, respectively. Let

$$\mathscr{X}_\ell = G(\mathbb{R})/\Gamma_\ell \quad \text{and} \quad \mathscr{Z}_{\mathbf{y},\ell} = L_{\mathbf{y}}(\mathbb{R})/(\Gamma_\ell \cap L_{\mathbf{y}}(\mathbb{R})).$$

We consider $\mathscr{Z}_{\mathbf{y},\ell}$ as a submanifold of $\mathscr{X}_\ell$. We denote by $m_{\mathscr{X}_\ell}$ and $m_{\mathscr{Z}_{\mathbf{y},\ell}}$ the measures on $\mathscr{X}_\ell$ and $\mathscr{Z}_{\mathbf{y},\ell}$ induced by the corresponding measure on $G(\mathbb{R})$ and $L_{\mathbf{y}}(\mathbb{R})$. It follows from our assumptions that the spaces $\mathscr{X}_\ell$ and $\mathscr{Z}_{\mathbf{y},\ell}$ have finite measures.

With this notation, the main result of this section is the following.

Theorem 2.1. *Under assumptions (i)–(ii) and (iii′)–(iv′), there exists $\varrho > 0$ such that*

$$|\Gamma_\ell \mathbf{y} \cap B_H(\mathbf{y})| = \frac{m_{\mathscr{Z}_{\mathbf{y},\ell}}(\mathscr{Z}_{\mathbf{y},\ell})}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)} m_Y(B_H(\mathbf{y})) + O(\ell^{\dim(L)+\dim(G)} m_Y(B_H(\mathbf{y}))^{1-\varrho}).$$

The implied constant in the error term is uniform over $\mathbf{y} \in Y(\mathbb{Z})$ and $\ell \in \mathbb{N}$.

The proof of Theorem 2.1 follows the strategy developed by Duke–Rudnick–Sarnak [13] and Eskin–McMullen [15]. Quantitative estimates in this setting have also been obtained by Benoist–Oh [2]. The main novelty of our result is the uniformity over the congruence subgroups Γ_ℓ, which is pivotal for our application to power-free values of polynomials.

The following result shows that Theorem 2.1 always provides a non-trivial estimate.

Lemma 2.2. *The space $G(\mathbb{R})\mathbf{y} \simeq G(\mathbb{R})/L_{\mathbf{y}}(\mathbb{R})$ is not-compact, and we have $m_Y(B_H(\mathbf{y})) \to \infty$ as $H \to \infty$.*

Proof. To simplify notation in this proof, we write L for $L_{\mathbf{y}}(\mathbb{R})$. Let θ be a Cartan involution of $G(\mathbb{R})$ that commutes with σ, and K is the corresponding maximal compact subgroup of $G(\mathbb{R})$. Then we have the decompositions

$$\mathrm{Lie}(G(\mathbb{R})) = \mathrm{Lie}(K) \oplus \mathfrak{p} \quad \text{and} \quad \mathrm{Lie}(G(\mathbb{R})) = \mathrm{Lie}(L) \oplus \mathfrak{q}$$

defined by the (± 1)-eigenspaces of θ and σ respectively. By [3, §55], $G(\mathbb{R})/L$ is diffeomorphic to the vector bundle $K \times_{K \cap L} (\mathfrak{p} \cap \mathfrak{q})$. In particular, the

space $G(\mathbb{R})/L$ can only be compact if $\mathfrak{p} \cap \mathfrak{q} = 0$. We also have the Cartan decomposition $G(\mathbb{R}) = K \exp(\mathfrak{p})$, and its generalisation

$$G(\mathbb{R}) = K \exp(\mathfrak{p} \cap \mathfrak{q}) \exp(\mathfrak{p} \cap \mathrm{Lie}(L))$$

(see [24, Prop. 2.2]), that define diffeomorphisms

$$G(\mathbb{R}) \simeq K \times \mathfrak{p} \quad \text{and} \quad G(\mathbb{R}) \simeq K \times (\mathfrak{p} \cap \mathfrak{q}) \times (\mathfrak{p} \cap \mathrm{Lie}(L)).$$

Hence, it follows that if $\mathfrak{p} \cap \mathfrak{q} = 0$, then $\mathfrak{p} \subset \mathrm{Lie}(L)$. Since $[\mathrm{Lie}(K), \mathfrak{p}] \subset \mathfrak{p}$, the Lie algebra generated by $\mathfrak{p}$ is an ideal in $\mathrm{Lie}(G(\mathbb{R}))$ which corresponds to a connected normal subgroup of $G(\mathbb{R})$ contained in L. Moreover, it is clear that this subgroup is cocompact. Since $G(\mathbb{R})$ has no compact factors, it follows that $L = G(\mathbb{R})$, but the involution σ has been assumed to be non-trivial. This contradiction shows that $\mathfrak{p} \cap \mathfrak{q} \neq 0$, and the space $G(\mathbb{R})/L$ is not compact. The last assertion follows from explicit volume computations for symmetric spaces (see (2.4) below). $\qquad\square$

Theorem 2.1 is deduced from the following equidistribution result on the space $\mathscr{X}_\ell$, which will be established in §2.2. We denote by $\mu_{\mathscr{X}_\ell}$ and $\mu_{\mathscr{Z}_{\mathbf{y},\ell}}$ the normalised measures on $\mathscr{X}_\ell$ and $\mathscr{Z}_{\mathbf{y},\ell}$ respectively.

Proposition 2.3. *Under assumptions (i)–(ii) and (iii′)–(iv′), there exist $q \geqslant 1$ and $\varrho_1 > 0$ such that for every $\varphi \in C_c^\infty(\mathscr{X}_\ell)$, $\mathbf{y} \in Y(\mathbb{Z})$, and $g \in G(\mathbb{R})$,*

$$\int_{\mathscr{Z}_{\mathbf{y},\ell}} \varphi(gz)\, d\mu_{\mathscr{Z}_{\mathbf{y},\ell}}(z) = \int_{\mathscr{X}_\ell} \varphi\, d\mu_{\mathscr{X}_\ell} + O_{\mathbf{y}}\left(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)|g\mathbf{y}|^{-\varrho_1}\|\varphi\|_{C^q} \right),$$

where $\|\varphi\|_{C^q}$ denotes the C^q-norm of the function φ. The implied constant in the error term is uniform in ℓ.

We note that

$$m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) \ll |\Gamma : \Gamma_\ell| \leqslant G(\mathbb{Z}/\ell\mathbb{Z}) \ll \ell^{\dim(G)}, \tag{2.1}$$

in Proposition 2.3. Let us recall the definition of the C^q-norms. These norms are defined with respect to a fixed basis $D_1, \ldots, D_n$ of the Lie algebra of $G(\mathbb{R})$. For φ in $C_c^\infty(G(\mathbb{R}))$ (or $C_c^\infty(\mathscr{X}_\ell)$) we set

$$\|\varphi\|_{C^q} = \sum_D \|D\varphi\|_\infty, \tag{2.2}$$

where the sum is taken over all monomials in D_i's of degree at most q, and D_i's are right-invariant differential operators defined by

$$D_i\varphi(x) = \frac{d}{dt}\varphi(\exp(tD_i)x)|_{t=0}. \tag{2.3}$$

Proof of Theorem 2.1. We note that by [4] the set $Y(\mathbb{Z})$ consists of finitely many orbits of the arithmetic group Γ. Therefore, it suffices to prove the claim of the theorem for $\mathbf{y} = \gamma_0 \mathbf{y}_0$, for a fixed $\mathbf{y}_0 \in Y(\mathbb{Z})$, with estimates which are uniform over $\gamma_0 \in \Gamma$. Since $B_H(\mathbf{y}) = B_H(\mathbf{y}_0)$ in this case, in order to simplify notation, we denote this set by B_H in subsequent computations. We also write L for $L_{\mathbf{y}_0}$ and $\mathscr{Z}_\ell$ for $\mathscr{Z}_{\mathbf{y}_0,\ell}$.

Let

$$F_H(g) = \sum_{\gamma \in \Gamma_\ell/(\Gamma_\ell \cap L(\mathbb{R}))} \chi_{B_H}(g\gamma\mathbf{y}_0).$$

We note that this defines a function on $\mathscr{X}_\ell = G(\mathbb{R})/\Gamma_\ell$. Since Γ_ℓ is normal in Γ, we have $F_H(\gamma_0) = |\gamma_0\Gamma_\ell\mathbf{y}_0 \cap B_H| = |\Gamma_\ell\mathbf{y} \cap B_H|$. For a real-valued $\varphi \in C_c^\infty(\mathscr{X}_\ell)$, we consider the inner product $\langle F_H, \varphi \rangle_{L^2(m_{\mathscr{X}_\ell})}$ which can be unfolded as in [13, p. 151]:

$$\langle F_H, \varphi \rangle_{L^2(m_{\mathscr{X}_\ell})} = \int_{G(\mathbb{R})/\Gamma_\ell} \left(\sum_{\gamma \in \Gamma_\ell/(\Gamma_\ell \cap L(\mathbb{R}))} \chi_{B_H}(g\gamma\mathbf{y}_0) \right) \varphi(g\Gamma_\ell) \, dm_G(g)$$

$$= \int_{G(\mathbb{R})/(\Gamma_\ell \cap L(\mathbb{R}))} \chi_{B_H}(g\mathbf{y}_0)\varphi(g\Gamma_\ell) \, dm_G(g)$$

$$= \int_{G(\mathbb{R})/L(\mathbb{R})} \chi_{B_H}(g\mathbf{y}_0) \left(\int_{\mathscr{Z}_\ell} \varphi(gz) \, dm_{\mathscr{Z}_\ell}(z) \right) dm_Y(g\mathbf{y}_0)$$

$$= \int_{B_H} \left(\int_{\mathscr{Z}_\ell} \varphi(gz) \, dm_{\mathscr{Z}_\ell}(z) \right) dm_Y(g\mathbf{y}_0)$$

$$= m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell) \int_{B_H} \left(\int_{\mathscr{Z}_\ell} \varphi(gz) \, d\mu_{\mathscr{Z}_\ell}(z) \right) dm_Y(g\mathbf{y}_0).$$

Let $R \in (0, H)$. By Proposition 2.3, when $|g\mathbf{y}_0| \geqslant R$, we have

$$\int_{\mathscr{Z}_\ell} \varphi(gz) \, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell} + O(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)R^{-\varrho_1}\|\varphi\|_{C^q}).$$

Also, it is clear that

$$\int_{\mathscr{Z}_\ell} \varphi(gz) \, d\mu_{\mathscr{Z}_\ell}(z) \ll \|\varphi\|_\infty \leqslant \|\varphi\|_{C^q}.$$

Hence, it follows that $\langle F_H, \varphi \rangle_{L^2(m_{\mathscr{X}_\ell})}$ is

$$
\begin{aligned}
&= m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)\, m_Y(B_H \backslash B_R) \left(\int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell} + O(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) R^{-\varrho_1} \|\varphi\|_{C^q}) \right) \\
&\quad + O(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell) m_Y(B_R) \|\varphi\|_\infty) \\
&= m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)\, m_Y(B_H) \left(\int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell} \right) \\
&\quad + O(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) R^{-\varrho_1} m_Y(B_H) + m_Y(B_R)) \|\varphi\|_{C^q}) \\
&= \frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)}\, m_Y(B_H) \left(\int_{\mathscr{X}_\ell} \varphi \, dm_{\mathscr{X}_\ell} \right) \\
&\quad + O(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) R^{-\varrho_1} m_Y(B_H) + m_Y(B_R)) \|\varphi\|_{C^q}).
\end{aligned}
$$

It follows from [19, Cor. 6.10] that

$$
m_Y(B_H) \sim v\, H^a (\log H)^b, \quad \text{as } H \to \infty, \tag{2.4}
$$

for some $v, a > 0$ and $b \geqslant 0$. Hence, the last estimate with a suitable choice of the parameter R implies that for some $\varrho_2 > 0$

$$
\begin{aligned}
\langle F_H, \varphi \rangle_{L^2(m_{\mathscr{X}_\ell})} &= \frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)}\, m_Y(B_H) \left(\int_{\mathscr{X}_\ell} \varphi \, dm_{\mathscr{X}_\ell} \right) \\
&\quad + O(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell) m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) m_Y(B_H)^{1-\varrho_2} \|\varphi\|_{C^q}).
\end{aligned} \tag{2.5}
$$

We apply this estimate to a suitably chosen bump-function φ_ε on $\mathscr{X}_\ell$. We denote by O_ε^G the ε-neighbourhood of identity with respect to a Riemannian metric on $G(\mathbb{R})$. Let Φ_ε be a smooth non-negative function supported on O_ε^G such that

$$
\int_{G(\mathbb{R})} \Phi_\varepsilon \, dm_G = 1 \quad \text{and} \quad \|\Phi_\varepsilon\|_{C^q} \ll \varepsilon^{-\beta},
$$

with some $\beta > 0$ depending on $\dim(G)$. It follows from the definition of the C^q-norms (cf. (2.2)–(2.3)) that the functions $g \mapsto \Phi_\epsilon(g g_0)$, $g_0 \in G$, have the same norms as Φ_ε. Let

$$
\varphi_\varepsilon(g\Gamma_\ell) = \sum_{\gamma \in \Gamma_\ell} \Phi_\varepsilon(g\gamma_0^{-1}\gamma) = \sum_{\gamma \in \Gamma_\ell} \Phi_\varepsilon(g\gamma\gamma_0^{-1}).
$$

This defines a function on $\mathscr{X}_\ell = G(\mathbb{R})/\Gamma_\ell$ which also satisfies

$$
\int_{\mathscr{X}_\ell} \varphi_\varepsilon \, dm_{\mathscr{X}_\ell} = 1 \quad \text{and} \quad \|\varphi_\varepsilon\|_{C^q} \ll \varepsilon^{-\beta}. \tag{2.6}
$$

Our goal is to show that $|\Gamma_\ell \mathbf{y} \cap B_H|$ can be approximated by the inner products $\langle F_H, \varphi_\varepsilon \rangle_{L^2(m_{\mathscr{X}_\ell})}$. We observe that if for some $g \in G(\mathbb{R})$, we have

$$
\varphi_\varepsilon(g\Gamma_\ell) \neq 0, \tag{2.7}
$$

then $g\gamma_0^{-1} \in O_\varepsilon^G \Gamma_\ell$ and $g \in u\gamma_0\Gamma_\ell$ for some $u \in O_\varepsilon^G$. For such g,

$$F_H(g) = |g\Gamma_\ell \mathbf{y}_0 \cap B_H| = |\gamma_0\Gamma_\ell \mathbf{y}_0 \cap u^{-1}B_H|.$$

For $u \in O_\varepsilon^G$, we have $u = e + O(\varepsilon)$, so that there exists uniform $c > 0$ such that $B_{(1-c\varepsilon)H} \subset u^{-1}B_H \subset B_{(1+c\varepsilon)H}$. Hence, we deduce that for g satisfying (2.7),

$$|\gamma_0\Gamma_\ell \mathbf{y}_0 \cap B_{(1-c\varepsilon)H}| \leqslant |g\Gamma_\ell \mathbf{y}_0 \cap B_H| \leqslant |\gamma_0\Gamma_\ell \mathbf{y}_0 \cap B_{(1+c\varepsilon)H}|.$$

This implies that $F_H(\gamma_0) = |\Gamma_\ell \mathbf{y} \cap B_H|$ satisfies

$$F_{(1+c\varepsilon)^{-1}H}(g) \leqslant F_H(\gamma_0) \leqslant F_{(1-c\varepsilon)^{-1}H}(g).$$

Hence, it follows from (2.6) that

$$\left\langle F_{(1+c\varepsilon)^{-1}H}, \varphi_\varepsilon \right\rangle_{L^2(m_{\mathscr{X}_\ell})} \leqslant F_H(\gamma_0) \leqslant \left\langle F_{(1-c\varepsilon)^{-1}H}, \varphi_\varepsilon \right\rangle_{L^2(m_{\mathscr{X}_\ell})}. \tag{2.8}$$

Applying (2.5), we conclude that

$$\begin{aligned}
F_H(\gamma_0) \leqslant &\frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)} m_Y(B_{(1-c\varepsilon)^{-1}H}) \\
&+ O\left(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)m_Y(B_{(1-c\varepsilon)^{-1}H})^{1-\varrho_2}\varepsilon^{-\beta}\right).
\end{aligned}$$

The volumes of the sets B_H satisfy a regularity property. According to [16, Appendix], there exists $c_0 > 0$ such that for all $\delta \in (0, \delta_0)$ and $H \geqslant H_0$,

$$m_Y(B_{(1+\delta)H}) \leqslant (1 + c_0\delta)m_Y(B_H).$$

Therefore, it follows that

$$\begin{aligned}
F_H(\gamma_0) \leqslant &\frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)} m_Y(B_H) \\
&+ O\left(\frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)}\varepsilon m_Y(B_H) + m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)m_Y(B_H)^{1-\varrho_2}\varepsilon^{-\beta}\right).
\end{aligned}$$

We recall from (2.1) that $m_{\mathscr{X}_\ell}(\mathscr{X}_\ell) \ll \ell^{\dim(G)}$, and similarly

$$\begin{aligned}
m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell) = m_L(L(\mathbb{R})/(L(\mathbb{R}) \cap \Gamma_\ell)) &\ll |(\Gamma \cap L) : (\Gamma_\ell \cap L)| \\
&\leqslant |\iota(L)(\mathbb{Z}/\ell\mathbb{Z})| \\
&\ll \ell^{\dim(L)}.
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
F_H(\gamma_0) \leqslant &\frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)} m_Y(B_H) \\
&+ O\left(\ell^{\dim(L)+\dim(G)}(\varepsilon m_Y(B_H) + m_Y(B_H)^{1-\varrho_2}\varepsilon^{-\beta})\right).
\end{aligned}$$

Optimising in ε, we deduce that there exists $\varrho > 0$ such that

$$|\Gamma_\ell \mathbf{y} \cap B_H| = F_H(\gamma_0) \leqslant \frac{m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)} m_Y(B_H) + O(\ell^{\dim(L)+\dim(G)}m_Y(B_H)^{1-\varrho}).$$

This proves the required upper bound on $|\Gamma_\ell \mathbf{y} \cap B_H|$. The lower estimate on $|\Gamma_\ell \mathbf{y} \cap B_H|$ is proved similarly using the lower bound from (2.8). This completes proof of Theorem 2.1 assuming Proposition 2.3. $\square$

2.2. **Proof of the equidistribution result.** In this section we prove Proposition 2.3. To simplify notation, we write L for $L_\mathbf{y}$ and $\mathscr{Z}_\ell$ for $\mathscr{Z}_{\mathbf{y},\ell}$. We recall the Cartan decomposition

$$G(\mathbb{R}) = K\,A\,L(\mathbb{R}),$$

where K is a compact subgroup compatible with $L(\mathbb{R})$, and A is a suitable Cartan subgroup complementary to $L(\mathbb{R})$ (see, for instance, [24, Ch.7]). For $g \in G(\mathbb{R})$, we write $g = kah$ with $k \in K$, $a \in A$, and $h \in L(\mathbb{R})$. Then

$$\int_{\mathscr{Z}_\ell} \varphi(gz)\, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{Z}_\ell} \varphi_k(az)\, d\mu_{\mathscr{Z}_\ell}(z),$$

where $\varphi_k \in C_c^\infty(\mathscr{X}_\ell)$ is given by $\varphi_k(x) = \varphi(kx)$. Since K is compact,

$$\|\varphi_k\|_{C^q} \ll \|\varphi\|_{C^q} \quad \text{and} \quad |gy| = |ka\mathbf{y}| \ll |a\mathbf{y}|$$

uniformly for $k \in K$. Hence, the claim of the proposition will follow once we prove it for $g = a \in A$. Moreover, without loss of generality we may assume that a belongs to a fixed positive Weyl chamber A^+ in A for the action of A on the Lie algebra of $G(\mathbb{R})$.

In the proof we use parameters R, ε, η of the form

$$R = \eta_1 d(a,e), \quad \varepsilon = e^{-\eta_2 d(a,e)}, \quad \delta = e^{-\eta_3 d(a,e)},$$

with some $\eta_1, \eta_2, \eta_3 > 0$ that will be specified later. We equip the space $\mathscr{X}_\ell$ with an invariant Riemannian metric induced from a right-invariant Riemannian metric on $G(\mathbb{R})$ which is bi-invariant with respect to the maximal compact subgroup K. Fix $z_0 \in \mathscr{Z}_\ell$ and set

$$\mathscr{Z}_{\ell,R}^- = \{z \in \mathscr{Z}_\ell : d(z,z_0) < R\} \quad \text{and} \quad \mathscr{Z}_{\ell,R}^+ = \{z \in \mathscr{Z}_\ell : d(z,z_0) > R\}.$$

It follows from [20, Sec. 5] that

$$m_{\mathscr{Z}_\ell}(\mathscr{Z}_{\ell,R}^+) \ll m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)e^{-\theta R} \quad \text{and} \quad \mu_{\mathscr{Z}_\ell}(\mathscr{Z}_{\ell,R}^+) \ll e^{-\theta R}, \tag{2.9}$$

for some fixed $\theta > 0$.

We refine the open cover $\mathscr{Z}_\ell = \mathscr{Z}_{\ell,R}^- \cup \mathscr{Z}_{\ell,R-1}^+$ further. Let O_ε^G and O_ε^L denote the ε-neighbourhoods of identity in $G(\mathbb{R})$ and $L(\mathbb{R})$ respectively. Since these neighbourhoods are defined with respect to a invariant metric, $(O_\varepsilon^G)^{-1} = O_\varepsilon^G$ and $O_{\varepsilon_1}^G O_{\varepsilon_2}^G \subset O_{\varepsilon_1+\varepsilon_2}^G$, and similarly for O_ε^L. Let $\Omega = \{z_i : i \in I\}$ be a maximal subset of $\mathscr{Z}_{\ell,<R}$ such that $O_{\varepsilon/2}^L z_i$, $i \in I$, are disjoint. We observe that then

$$\mathscr{Z}_{\ell,R}^- \subset \bigcup_{i \in I} O_\varepsilon^L z_i. \tag{2.10}$$

Indeed, suppose that $z \in \mathscr{Z}_{\ell,R}^-$, but z does not belong to the union in (2.10). Then the set $\Omega \cup \{z\}$ satisfies the same disjointness property as Ω. Indeed, if $O_{\varepsilon/2}^L z \cap O_{\varepsilon/2}^L z_i \neq \emptyset$ for some i, then $z \in (O_{\varepsilon/2}^L)^{-1} O_{\varepsilon/2}^L z_i \subset O_\varepsilon^L z_i$ which is not the case. This contradicts maximality of Ω and proves (2.10).

For future reference, we prove some basic properties of the set Ω. First, we claim that for sufficiently small ε, the map $O_{3\varepsilon}^G \to O_{3\varepsilon}^G z_i$ is injective. Let $z_i = h_i \Gamma_\ell$ with $h_i \in L(\mathbb{R})$ satisfying $d(h_i, e) < R$. If $u_1 h_i \Gamma_\ell = u_2 h_i \Gamma_\ell$ for some $u_1, u_2 \in O_{3\varepsilon}^G$, then for some $\gamma \in \Gamma_\ell$, we obtain

$$\gamma = h_i^{-1} u_2^{-1} u_1 h_i \in h_i^{-1} O_{6\varepsilon}^G h_i \subset O_{6e^{cR}\varepsilon}^G,$$

with some fixed $c > 0$. Let us choose $\varepsilon \leqslant \varepsilon_0 e^{-cR}$ with sufficiently small $\varepsilon_0 > 0$. Then it follows from discreteness of Γ_ℓ that $\gamma = e$, so that $u_1 = u_2$. Hence, this shows that the map $O_{3\varepsilon}^G \to O_{3\varepsilon}^G z_i$ is injective.

We will also need an upper bound on $|\Omega|$ which is easy to deduce from the disjointness property. Since the map $O_{\varepsilon/2}^L \to O_{\varepsilon/2}^L z_i$ is injective, $m_{\mathscr{Z}_\ell}(O_{\varepsilon/2}^L z_i) \gg \varepsilon^d$ where $d = \dim(L)$. This implies that

$$|\Omega| \ll m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)\varepsilon^{-d}. \tag{2.11}$$

We choose a smooth function Ψ on $L(\mathbb{R})$ such that

$$0 \leqslant \Psi \leqslant 1, \quad \mathrm{supp}(\Psi) \subset O_{2\varepsilon}^L, \quad \Psi = 1 \text{ on } O_\varepsilon^L, \quad \|\Psi\|_{C^q} \ll \varepsilon^{-\beta},$$

for some $\beta > 0$ depending on $\dim(L)$. It follows from the definition of C^q-norms (cf. (2.2)–(2.3)) that the family of functions $h \mapsto \Psi(hh_0)$, $h_0 \in L(\mathbb{R})$, have the same Sobolev norms. Since the map $O_{2\varepsilon}^L \to O_{2\varepsilon}^L z_i$ is injective, we also obtain smooth functions χ_i on $\mathscr{Z}_\ell$ such that

$$\mathrm{supp}(\chi_i) \subset O_{2\varepsilon}^L z_i \quad \text{and} \quad \|\chi_i\|_{C^q} = \|\Psi\|_{C^q} \ll \varepsilon^{-\beta}.$$

We fix a total ordering on I and set

$$\psi_i = \chi_i \prod_{j>i}(1 - \chi_j) \quad \text{and} \quad \psi_\infty = 1 - \sum_i \psi_i = \prod_i (1 - \chi_i).$$

Then $\mathrm{supp}(\psi_i) \subset \mathrm{supp}(\chi_i) \subset O_{2\varepsilon}^L z_i$, and it follows from (2.10) that

$$\mathrm{supp}(\psi_\infty) \subset \mathscr{Z}_{\ell,R-1}^+.$$

It is also clear that

$$0 \leqslant \psi_i \leqslant 1 \quad \text{and} \quad \|\psi_i\|_{C^q} \ll \varepsilon^{-\beta_1}, \tag{2.12}$$

with some fixed $\beta_1 > 0$.

Let P be the non-expanding horospherical subgroup of $G(\mathbb{R})$ corresponding to A^+. This is the connected Lie subgroup of $G(\mathbb{R})$ whose Lie algebra consists of X such that $\frac{\|\mathrm{Ad}(a)X\|}{\|X\|}$ is uniformly bounded as $a \in A^+$. This property

implies, in particular, that for p in a neighbourhood of identity in P and all $a \in A^+$,

$$d(apa^{-1}, e) \ll d(p, e). \tag{2.13}$$

We note that since $L(\mathbb{R})$ is a symmetric subgroup in $G(\mathbb{R})$, it follows that

$$\mathrm{Lie}(G(\mathbb{R})) = \mathrm{Lie}(P) + \mathrm{Lie}(L(\mathbb{R}))$$

(see, for instance, [15, p. 199]). In particular, there is a subspace V of $\mathrm{Lie}(P)$ such that

$$\mathrm{Lie}(G(\mathbb{R})) = V \oplus \mathrm{Lie}(L(\mathbb{R})). \tag{2.14}$$

Let O_ε^V denote the ε-neighbourhood of identity in $\exp(V)$. It follows from (2.14) that the product map

$$\exp(V) \times L(\mathbb{R}) \to G(\mathbb{R}) \tag{2.15}$$

is a diffeomorphism in a neighbourhood of identity. For g in a neighbourhood of identity in $G(\mathbb{R})$, we write $g = \mathbf{v}(g)\mathbf{h}(g)$ where $\mathbf{v}$ and $\mathbf{h}$ are the smooth maps realising this diffeomorphism.

We observe that with respect to the decomposition (2.15), the Haar measure m_G restricted to $O_\varepsilon^V O_{2\varepsilon}^L$ decomposes as a product $m_V \otimes m_L'$ where m_V is a smooth measure on O_ε^V, and m_L' is the restriction of m_L to $O_{2\varepsilon}^L$. We also note that under the map $h \mapsto hz_i$, the measure m_L' projects to $m_{\mathscr{Z}_\ell}|_{O_{2\varepsilon}^L z_i}$. We choose a smooth non-negative function σ on $\exp(V)$ such that

$$\int_{\exp(V)} \sigma \, dm_V = 1, \quad \mathrm{supp}(\sigma) \subset O_\delta^V, \quad \|\sigma\|_{C^q} \ll \delta^{-\beta_2}, \tag{2.16}$$

with some $\beta_2 > 0$ depending on $\dim(V)$. Then

$$\int_{\mathscr{Z}_\ell} \varphi(az)\psi_i(z) \, dm_{\mathscr{Z}_\ell}(z) = \int_{\exp(V) \times O_{2\varepsilon}^L} \varphi(ahz_i)\sigma(v)\psi_i(hz_i) \, dm_V(v) dm_L'(h)$$

$$= \int_{\exp(V) \times O_{2\varepsilon}^L} \varphi(ahz_i)\Phi_i(vh) \, dm_V(v) dm_L'(h),$$

where Φ_i is the smooth function supported on $O_\delta^V O_{2\varepsilon}^L$, with $\delta \leqslant \varepsilon$, defined by

$$\Phi_i(g) = \sigma(\mathbf{v}(g))\tilde{\psi}_i(\mathbf{h}(g)),$$

where $\tilde{\psi}_i(h) = \psi_i(hz_i)$. Then $\|\tilde{\psi}_i\|_{C^q} = \|\psi_i\|_{C^q}$, and it follows from (2.12) and (2.16) that

$$\|\Phi_i\|_{C^q} \ll \delta^{-\beta_3},$$

with some $\beta_3 > 0$. By (2.13), for $v \in O_\delta^V$,

$$|\varphi(avz) - \varphi(az)| \ll \|\varphi\|_{C^1} d(ava^{-1}, e) \ll \|\varphi\|_{C^1}\delta.$$

Since

$$
\left| \int_{\exp(V)\times O_\varepsilon^L} \Phi_i(vh)\, dm_V(v) dm'_L(h) \right| = \left(\int_{\exp(V)} \sigma\, dm_V \right) \left(\int_{O_\varepsilon^L} \tilde{\psi}_i(h)\, dm'_L(h) \right)
$$
$$
\ll \varepsilon^d,
$$

where $d = \dim(L)$, we obtain that

$$
\int_{\mathscr{X}_\ell} \varphi(az)\psi_i(z)\, dm_{\mathscr{X}_\ell}(z) = \int_{\exp(V)\times O_{2\varepsilon}^L} \varphi(avhz_i)\Phi_i(vh)\, dm_V(v) dm'_L(h)
$$
$$
+ O(\|\varphi\|_{C^1}\delta\varepsilon^d)
$$
$$
= \int_{O_\delta^V O_{2\varepsilon}^L} \varphi(agz_i)\Phi_i(g)\, dm_G(g) + O(\|\varphi\|_{C^1}\delta\varepsilon^d).
$$

We recall that the map $g \mapsto gz_i$ is injective on $O_{3\varepsilon}^G$. Hence, under this map the measure $m_G|_{O_{3\varepsilon}^G}$ projects to the measure $m_{\mathscr{X}_\ell}|_{O_{3\varepsilon}^G z_i}$, and since $O_\delta^V O_{2\varepsilon}^L \subset O_{3\varepsilon}^G$, Φ_i defines a function φ_i supported on $O_{3\varepsilon}^G z_i$ such that

$$
\|\varphi_i\|_{C^q} = \|\Phi_i\|_{C^q} \ll \delta^{-\beta_3}, \tag{2.17}
$$

and

$$
\int_{O_{3\varepsilon}^G} \varphi(agz_i)\varphi_i(gz_i)\, dm_G(g) = \int_{\mathscr{X}_\ell} \varphi(ax)\varphi_i(x)\, dm_{\mathscr{X}_\ell}(x).
$$

Combining the above estimates, we conclude that

$$
\int_{\mathscr{X}_\ell} \varphi(az)\varphi_i(z)\, dm_{\mathscr{X}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi(ax)\varphi_i(x)\, dm_{\mathscr{X}_\ell}(x) + O(\|\varphi\|_{C^1}\delta\varepsilon^d).
$$

This formula allows us to use the exponential decay property of matrix coefficients for representations of $G(\mathbb{R})$ to estimate the original integral. It is known from the works [9, 10], which established bounds towards the generalised Ramanujan conjectures, that the action of each simple factor of $G(\mathbb{R})$ on the congruence quotients $\mathscr{X}_\ell = G(\mathbb{R})/\Gamma_\ell$ has the uniform spectral gap property. Namely, the unitary representation of non-compact simple factors of $G(\mathbb{R})$ on the orthogonal complement of the constant functions in $L^2(\mathscr{X}_\ell)$ are uniformly isolated from the trivial representations. Then, by [20, §3.4], there exists $\varrho > 0$ such that

$$
\int_{\mathscr{X}_\ell} \varphi(ax)\varphi_i(x)\, d\mu_{\mathscr{X}_\ell}(x) = \left(\int_{\mathscr{X}_\ell} \varphi\, d\mu_{\mathscr{X}_\ell} \right) \left(\int_{\mathscr{X}_\ell} \varphi_i\, d\mu_{\mathscr{X}_\ell} \right)
$$
$$
+ O\left(e^{-\varrho d(a,e)} \|\varphi\|_{C^q} \|\varphi_i\|_{C^q} \right). \tag{2.18}
$$

We note that the exponent ϱ is determined by the isolation property of the unitary representation, so that it is independent of ℓ.

Since $\sum_i \psi_i + \psi_\infty = 1$, we obtain

$$\int_{\mathscr{Z}_\ell} \varphi(az)\,dm_{\mathscr{Z}_\ell}(z) = \sum_i \int_{\mathscr{Z}_\ell} \varphi(az)\psi_i(z)\,dm_{\mathscr{Z}_\ell}(z) + \int_{\mathscr{Z}_\ell} \varphi(az)\psi_\infty(z)\,dm_{\mathscr{Z}_\ell}(z).$$

The last term can be estimated using the fact that $\mathrm{supp}(\psi_\infty) \subset \mathscr{Z}_{\ell,R-1}^+$. This gives

$$\int_{\mathscr{Z}_\ell} \varphi(az)\psi_\infty(z)\,dm_{\mathscr{Z}_\ell}(z) \leqslant m_{\mathscr{Z}_\ell}(\mathscr{Z}_{\ell,R-1}^+)\|\varphi\|_\infty \ll m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)e^{-\theta R}\|\varphi\|_\infty,$$

whence

$$\begin{aligned}
\int_{\mathscr{Z}_\ell} \varphi(az)\,dm_{\mathscr{Z}_\ell}(z) &= \sum_i \int_{\mathscr{X}_\ell} \varphi(ax)\varphi_i(x)\,dm_{\mathscr{X}_\ell}(x) \\
&\quad + O\left(|\Omega|\|\varphi\|_{C^1}\delta\varepsilon^d + e^{-\theta R}\|\varphi\|_\infty\right) \\
&= m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)\sum_i \int_{\mathscr{X}_\ell} \varphi(ax)\varphi_i(x)\,d\mu_{\mathscr{X}_\ell}(x) \\
&\quad + O\left(|\Omega|\|\varphi\|_{C^1}\delta\varepsilon^d + m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)e^{-\theta R}\|\varphi\|_\infty\right).
\end{aligned}$$

Next, we apply (2.18), combined with estimates (2.17) and (2.11), to deduce that the right hand side is

$$\begin{aligned}
&= m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)\sum_i \left(\int_{\mathscr{X}_\ell} \varphi\,d\mu_{\mathscr{X}_\ell}\right)\left(\int_{\mathscr{X}_\ell} \varphi_i\,d\mu_{\mathscr{X}_\ell}\right) \\
&\quad + O\left(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)|\Omega|e^{-\varrho d(a,e)}\|\varphi\|_{C^q}\delta^{-\beta_3} + |\Omega|\|\varphi\|_{C^1}\delta\varepsilon^d + m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)e^{-\theta R}\|\varphi\|_\infty\right) \\
&= \left(\int_{\mathscr{X}_\ell} \varphi\,d\mu_{\mathscr{X}_\ell}\right)\left(\sum_i \int_{\mathscr{X}_\ell} \varphi_i\,dm_{\mathscr{X}_\ell}\right) \\
&\quad + O\left(m_{\mathscr{Z}_\ell}(\mathscr{Z}_\ell)(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)\varepsilon^{-d}\delta^{-\beta_3}e^{-\varrho d(a,e)} + \delta + e^{-\theta R})\|\varphi\|_{C^q}\right).
\end{aligned}$$

Here we used that $m_{\mathscr{X}_\ell} = m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)\mu_{\mathscr{X}_\ell}$. Using (2.16), the sum above is

$$\begin{aligned}
\sum_i \int_{\mathscr{X}_\ell} \varphi_i\,dm_{\mathscr{X}_\ell} &= \sum_i \int_{O_{3\varepsilon}^G} \varphi_i(gz_i)\,dm_G(g) \\
&= \sum_i \int_{O_{3\varepsilon}^G} \Phi_i(g)\,dm_G(g) \\
&= \sum_i \int_{O_\delta^V \times O_{2\varepsilon}^L} \sigma(v)\psi_i(hz_i)\,dm_V(v)dm_L'(h) \\
&= \sum_i \int_{\mathscr{Z}_\ell} \psi_i\,dm_{\mathscr{Z}_\ell}.
\end{aligned}$$

Hence, it follows from (2.9) that

$$\sum_i \int_{\mathscr{X}_\ell} \varphi_i \, dm_{\mathscr{X}_\ell} = m_{\mathscr{X}_\ell}(\mathscr{Z}_\ell) \int_{\mathscr{Z}_\ell} (1 - \psi_\infty) \, d\mu_{\mathscr{Z}_\ell}$$

$$= m_{\mathscr{X}_\ell}(\mathscr{Z}_\ell) \left(1 + O\left(e^{-\theta R}\right)\right).$$

Since $m_{\mathscr{Z}_\ell} = m_{\mathscr{X}_\ell}(\mathscr{Z}_\ell)\mu_{\mathscr{Z}_\ell}$, we conclude that

$$\int_{\mathscr{Z}_\ell} \varphi(az) \, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell}$$

$$+ O\left((m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)\varepsilon^{-d}\delta^{-\beta_3}e^{-\varrho d(a,e)} + \delta + e^{-\theta R})\|\varphi\|_{C^q}\right).$$

We recall that this estimate holds under the previously made assumptions:

$$\varepsilon \leqslant \varepsilon_0 e^{-cR} \quad \text{and} \quad \delta \leqslant \varepsilon.$$

We take $\varepsilon = \varepsilon_0 e^{-cR}$ and $\delta = \varepsilon_0 e^{dcR/(\beta_3+1)}e^{-\varrho d(a,e)/(\beta_3+1)}$. This gives

$$\int_{\mathscr{Z}_\ell} \varphi(az) \, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell}$$

$$+ O\left(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)(e^{dcR/(\beta_3+1)}e^{-\varrho d(a,e)/(\beta_3+1)} + e^{-\theta R})\|\varphi\|_{C^q}\right).$$

We choose $R = \eta d(a,e)$ with sufficiently small $\eta > 0$. Then $\delta \leqslant \varepsilon$, and we deduce that

$$\int_{\mathscr{Z}_\ell} \varphi(az) \, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell} + O\left(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)e^{-\varrho' d(a,e)}\|\varphi\|_{C^q}\right), \qquad (2.19)$$

for some $\varrho' > 0$.

To finish the proof, it remains to compare $d(a,e)$ and $\|a\mathbf{y}\|$. First, we note that since d is an invariant Riemannian metric,

$$\|\log a\| \ll d(a,e) \ll \|\log a\|,$$

where $\|\cdot\|$ denotes the norm induced by the Riemannian metric on $\mathrm{Lie}(A)$. Since the A-action on $\mathbb{R}^n$ is diagonalisable, we can write $\mathbf{y} = \sum_i \mathbf{y}_i$ where $\mathbf{y}_i$'s are linearly independent eigenvectors of A. Then

$$a\mathbf{y} = \sum_i e^{\lambda_i(\log a)}\mathbf{y}_i$$

for some characters λ_i on $\mathrm{Lie}(A)$, and

$$|a\mathbf{y}| \ll \exp\left(\max_i \lambda_i(\log a)\right) \leqslant \exp(c\|\log a\|)$$

for some fixed $c > 0$. Hence, it follows from (2.19) that

$$\int_{\mathscr{Z}_\ell} \varphi(az) \, d\mu_{\mathscr{Z}_\ell}(z) = \int_{\mathscr{X}_\ell} \varphi \, d\mu_{\mathscr{X}_\ell} + O\left(m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)|a\mathbf{y}|^{-\varrho''}\|\varphi\|_{C^q}\right),$$

with $\varrho'' = c\varrho' > 0$. This therefore completes the proof of the Proposition 2.3.

Remark 2.4. We proved Theorem 2.1 under the assumption that the group G is $\mathbb{Q}$-simple, but the method of the proof sometimes works without this assumption. The only place where this assumption was used is the exponential mixing estimate (2.18). When G is not $\mathbb{Q}$-simple, the space $\mathscr{X}_\ell$ has a finite cover $\prod_{i=1}^r \mathscr{X}_\ell^{(i)}$, where $\mathscr{X}_\ell^{(i)}$ are the spaces corresponding to $\mathbb{Q}$-simple factors of G. In this case, we can generalise (2.18) to give

$$\int_{\mathscr{X}_\ell} \varphi(ax)\psi_i(x)\,d\mu_{\mathscr{X}_\ell}(x) = \left(\int_{\mathscr{X}_\ell} \varphi\,d\mu_{\mathscr{X}_\ell}\right)\left(\int_{\mathscr{X}_\ell} \psi_i\,d\mu_{\mathscr{X}_\ell}\right)$$
$$+ O\left(e^{-\varrho D(a)}\|\varphi\|_{C^q}\|\psi_i\|_{C^q}\right),$$

where $D(a) = \min_i d(a^{(i)}, e)$, and $a = a^{(1)} \cdots a^{(r)}$ is the decomposition of a with respect to the $\mathbb{Q}$-simple factors. Hence, if one shows that for every $a \in A$,

$$D(a) \gg d(a, e), \tag{2.20}$$

then the proof of Theorem 2.1 can be completed exactly as before.

We are particularly interested in quadric hypersurfaces $\{Q = m\}$ of signature $(2,2)$. After a suitable real change of variable, this quadratic surface can be reduced to the form $X_1 X_2 - X_3 X_4 = m$ with $m > 0$. Then after identifying $\mathbb{R}^4$ with the space $\mathrm{M}_2(\mathbb{R})$ of matrices, Q will be given by the determinant, and $G(\mathbb{R}) \simeq \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ with the action given

$$(g_1, g_2) \cdot X \mapsto g_1 X g_2^{-1}, \qquad (g_1, g_2) \in G(\mathbb{R}), \quad X \in \mathrm{M}_2(\mathbb{R}).$$

Then $Y(\mathbb{R}) \simeq G(\mathbb{R})/L(\mathbb{R})$, where $L(\mathbb{R})$ is the diagonal subgroup of $G(\mathbb{R})$. It is the symmetric subgroup with respect to the involution $(g_1, g_2) \mapsto (g_2, g_1)$. In this case the Cartan subgroup complementary to $L(\mathbb{R})$ is

$$A = \{(b, b^{-1}) : b \in B\},$$

where B denotes the diagonal subgroup of $\mathrm{SL}_2(\mathbb{R})$. It is clear that (2.20) holds in this case, so that Theorem 2.1 holds as well.

2.3. **Consequences.** Our next goal is to estimate

$$\#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}} : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi}\,(\mathrm{mod}\,\ell)\}$$

for a given orbit $\mathscr{O}_{\mathbf{A}}$ of $G(\mathbf{A})$ in $Y(\mathbf{A})$. To state this result we use a cohomological invariant $\delta : Y(\mathbf{A}) \to \{0, |\mathrm{Pic}(L)|\}$ introduced by Borovoi and Rudnick in [6]. This invariant is constant on orbits $\mathscr{O}_{\mathbf{A}}$ of $G(\mathbf{A})$ in $Y(\mathbf{A})$ and has the property that

$$\delta(\mathscr{O}_{\mathbf{A}}) = 0 \quad \Longleftrightarrow \quad Y(\mathbb{Q}) \cap \mathscr{O}_{\mathbf{A}} = \emptyset.$$

We note that $\delta \equiv 1$ when L is semisimple and simply connected. In particular, $\delta \equiv 1$ in the case of quadric hypersurfaces with $n \geqslant 4$.

Let $\mathscr{O}_{\mathbf{A}}$ be an orbit of $G(\mathbf{A})$ in $Y(\mathbf{A})$. This orbit is of the form $\mathscr{O}_{\mathbf{A}} = \mathscr{O}_\infty \times \mathscr{O}_f$ where $\mathscr{O}_\infty$ is an orbit of $G(\mathbb{R})$ in $Y(\mathbb{R})$, and $\mathscr{O}_f$ is an orbit of $G(\mathbf{A}_f)$ in $Y(\mathbf{A}_f)$.

For $\ell \in \mathbb{N}$ and $\boldsymbol{\xi} \in Y(\mathbb{Z}/\ell\mathbb{Z})$, we consider a family of open subsets $B_f(\boldsymbol{\xi}, \ell)$ of $Y(\mathbf{A}_f)$ defined by

$$B_f(\boldsymbol{\xi}, \ell) = \prod_{p < \infty} B_p(\boldsymbol{\xi}, \ell),$$

where

$$B_p(\boldsymbol{\xi}, \ell) = \left\{ \mathbf{y} \in Y(\mathbb{Z}_p) : \mathbf{y} \equiv \boldsymbol{\xi} \, (\mathrm{mod}\, p^{v_p(\ell)}) \right\}.$$

We also set

$$\mathscr{O}_\infty(H) = \{ \mathbf{y} \in \mathscr{O}_\infty : |\mathbf{y}| \leqslant H \}.$$

We fix a *gauge form* on Y; i.e., a nowhere zero regular differential form of top degree. Since Y is a homogeneous variety of a semisimple group, such a form exists and is unique up to a scalar multiple. This defines a measure m_Y on $Y(\mathbb{R})$, induced by the gauge form, and a measure $m_{Y,f}$ on $Y(\mathbf{A}_f)$. We refer to [6, §1] for a detailed discussion of gauge forms and corresponding measures. When L is semisimple, the measure $m_{Y,f}$ is the product of measures $m_{Y,p}$ on $Y(\mathbb{Q}_p)$ induced by the gauge form

$$m_{Y,f} = \prod_{p < \infty} m_{Y,p}.$$

To define $m_{Y,f}$ in general, we need to introduce suitable convergence factors. Let ϱ_L denote the representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the space $X^*(L) \otimes \mathbb{Q}$ of characters of G, let t_L be the rank of the group of $\mathbb{Q}$-characters of G, and let

$$L(s, \varrho_L) = \prod_{p < \infty} L_p(s, \varrho_L) \tag{2.21}$$

be the Artin L-function associated to ϱ_L. We recall that $L(s, \varrho_L)$ has a pole of order t_L at $s = 1$. The measure $m_{Y,f}$ on $Y(\mathbf{A}_f)$ is defined by

$$m_{Y,f} = \left(\lim_{s \to 1} (s-1)^{t_L} L(s, \varrho_L) \right) \prod_{p < \infty} L_p(1, \varrho_L)^{-1} m_{Y,p}.$$

Assumption (ii) implies that $t_L = 0$, so that we also have

$$m_{Y,f} = \prod_{p < \infty} m_{Y,p},$$

but this convergence is only conditional.

For our next results, we henceforth assume that $Y \simeq G/L$ satisfies the assumptions (i)–(iv).

Corollary 2.5. *Under assumptions (i)–(iv), there exists $\varrho > 0$ such that for every orbit $\mathscr{O}_{\mathbf{A}}$ of $G(\mathbf{A})$ in $Y(\mathbf{A})$, we have*

$$\#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}} : |\mathbf{x}| \leqslant H, \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\,\ell)\}$$
$$= \delta(\mathscr{O}_{\mathbf{A}}) m_Y(\mathscr{O}_{\infty}(H)) m_{Y,f}(\mathscr{O}_f \cap B_f(\boldsymbol{\xi}, \ell))$$
$$+ O\left(\ell^{\dim(L)+2\dim(G)} m_Y(\mathscr{O}_{\infty}(H))^{1-\varrho}\right).$$

The implied constant in the error term is uniform in ℓ and $\boldsymbol{\xi}$.

Proof. We introduce compact open subgroups

$$K_f(\ell) = \left\{g \in \prod_{p<\infty} G(\mathbb{Z}_p) : \iota(g_p) \equiv id \,(\mathrm{mod}\,p^{v_p(\ell)})\right\}$$

of $G(\mathbf{A}_f)$ and set

$$K(\ell) = G(\mathbb{R}) \times K_f(\ell).$$

Then $\Gamma_\ell = G(\mathbb{Q}) \cap K(\ell)$ are precisely the congruence subgroups defined in (1.3). The group Γ_ℓ acts on $Y(\mathbb{Q}) \cap B(\boldsymbol{\xi}, \ell)$. Given an orbit $\mathscr{O}$ of Γ_ℓ in $Y(\mathbb{Q}) \cap B(\boldsymbol{\xi}, \ell)$, following [6, §4] we define its *weight* $w(\mathscr{O})$ as follows. For $\mathbf{y} \in \mathscr{O}$, we fix gauge forms on G and $L_{\mathbf{y}}$ which are compatible with the chosen gauge form on Y. These forms define the corresponding measures m_G on $G(\mathbb{R})$ and $m_{L_{\mathbf{y}}}$ on $L_{\mathbf{y}}(\mathbb{R})$ which also induce measures $m_{\mathscr{X}_\ell}$ and $m_{\mathscr{Z}_{\mathbf{y},\ell}}$ on $\mathscr{X}_\ell = G(\mathbb{R})/\Gamma_\ell$ and $\mathscr{Z}_{\mathbf{y},\ell} = L_{\mathbf{y}}(\mathbb{R})/(\Gamma_\ell \cap L_{\mathbf{y}}(\mathbb{R}))$. The weight of the orbit $\mathscr{O}$ is defined by

$$w(\mathscr{O}) = \frac{m_{\mathscr{Z}_{\mathbf{y},\ell}}(\mathscr{Z}_{\mathbf{y},\ell})}{m_{\mathscr{X}_\ell}(\mathscr{X}_\ell)}.$$

One can check that this definition is independent of the choice of $\mathbf{y}$. Using the new notation, Theorem 2.1 can be restated as follows:

$$|\mathscr{O} \cap \mathscr{O}_{\infty}(H)| = w(\mathscr{O}) m_Y(\mathscr{O}_{\infty}(H)) + O(\ell^{\dim(L)+\dim(G)} m_Y(\mathscr{O}_{\infty}(H))^{1-\varrho}). \quad (2.22)$$

Let $B = \mathscr{O}_{\infty} \times (\mathscr{O}_f \cap B_f(\boldsymbol{\xi}, \ell))$. Then

$$Y(\mathbb{Q}) \cap B = \{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}} : \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\,\ell)\}.$$

We note that L is reductive by [4, Thm. 3.5] and, in particular, unimodular. Since G is simply connected and L is symmetric, it follows from [25, §8] that L is connected. Hence, all the assumptions of [6, §4] are satisfied, and according to [6, Thm. 4.2],

$$\sum_{\mathscr{O} \subset B} w(\mathscr{O}) = \delta(\mathscr{O}_{\mathbf{A}}) m_{Y,f}(\mathscr{O}_f \cap B_f(\boldsymbol{\xi}, \ell)),$$

where the sum is taken over the orbits $\mathscr{O}$ of Γ_ℓ contained in B. Hence, summing (2.22) over these orbits, we deduce the corollary. We note that the number orbits is at most $O(|\Gamma : \Gamma_\ell|) = O(\ell^{\dim(G)})$ which contributes an additional factor to the error term. $\qquad \square$

Finally, we deduce an estimate for the number of $\mathbf{x} \in Y(\mathbb{Z})$ with $|\mathbf{x}| \leqslant H$ and $\mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, \ell)$. Let

$$Y(\mathbb{R})_H = \{\mathbf{x} \in Y(\mathbb{R}) : |\mathbf{x}| \leqslant H\}.$$

Since the set $Y(\mathbb{Z})$ consist of finitely many orbits of Γ, by [4], we can sum the estimates from Corollary 2.5 to conclude as follows.

Corollary 2.6. *There exists $\varrho > 0$ such that*

$$\#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, \ell)\}$$

$$= \int_{Y(\mathbb{R})_H \times B_f(\boldsymbol{\xi}, \ell)} \delta \, d(m_Y \otimes m_{Y,f}) + O\left(\ell^{\dim(L) + 2\dim(G)} V(H)^{1-\varrho}\right),$$

where

$$V(H) = \max_{\mathcal{O}_\infty \subset Y(\mathbb{R})} m_Y(\mathcal{O}_\infty(H))$$

and where $\mathcal{O}_\infty$ runs over orbits of $G(\mathbb{R})$ in $Y(\mathbb{R})$. The implied constant in the error term is uniform on $\boldsymbol{\xi}$ and ℓ. If, in addition, L is assumed to be semisimple and simply connected, then

$$\int_{Y(\mathbb{R})_H \times B_f(\boldsymbol{\xi}, \ell)} \delta \, d(m_Y \otimes m_{Y,f}) = m_Y(Y(\mathbb{R})_H) m_{Y,f}(B_f(\boldsymbol{\xi}, \ell)).$$

3. Small moduli and the proof of Theorems 1.1 and 1.2

3.1. **Small moduli.** Let $Y \simeq G/L \subset \mathbb{A}^n$ be a symmetric variety satisfying (i)–(iv) and let $r \geqslant 2$. In §§3.1–3.2 we additionally assume that L is semisimple and simply connected and that the smoothness assumption (1.5) holds. The case when L is not a semisimple simply connected group will be discussed in §3.3.

We recall the expression (1.12) for $N_r(Y, f; H)$. The goal of this section is to estimate the contribution from small moduli

$$N^{(1)}(H) = \sum_{k \leqslant H^\Delta} \mu(k) \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ 0 \neq f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^r)\}, \quad (3.1)$$

as $H \to \infty$, for given $\Delta > 0$. We shall need to separate the contribution from $\mathbf{x}$ such that $f(\mathbf{x}) = 0$. Accordingly, we write

$$N^{(1)}(H) = \sum_{k \leqslant H^\Delta} \mu(k) \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^r)\}$$

$$+ O\left(H^\Delta E(f; H)\right),$$

where for any $g \in \mathbb{Z}[X_1, \ldots, X_n]$, we set

$$E(g; H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ g(\mathbf{x}) = 0\}. \quad (3.2)$$

Breaking the first cardinality into congruence classes modulo k^r, we conclude that

$$N^{(1)}(H) = \sum_{k \leqslant H^\Delta} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} V_{k^r}(H; \boldsymbol{\xi}) + O\left(H^\Delta E(f; H)\right),$$

where for any $\ell \in \mathbb{N}$ and $\boldsymbol{\xi} \in Y(\mathbb{Z}/\ell\mathbb{Z})$, we put

$$V_\ell(H; \boldsymbol{\xi}) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, \ell)\}.$$

We now shift our attention to estimating $V_\ell(H; \boldsymbol{\xi})$, as $H \to \infty$.

Proposition 3.1. *There exists $\delta > 0$ such that*

$$V_\ell(H; \boldsymbol{\xi}) = \mu_\infty(Y; H) \prod_{p < \infty} \hat{\mu}_p(Y; \boldsymbol{\xi}; \ell) + O(\ell^{\dim(L) + 2\dim(G)} \mu_\infty(Y; H)^{1-\delta}),$$

where $\mu_\infty(Y; H)$ is defined in (1.6) and

$$\hat{\mu}_p(Y; \boldsymbol{\xi}, \ell) = \lim_{t \to \infty} p^{-t\dim(Y)} \#\{\mathbf{x} \in Y(\mathbb{Z}/p^t\mathbb{Z}) : \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, p^{v_p(\ell)})\}.$$

The implied constant in this estimate depends only on Y and is independent of ℓ and $\boldsymbol{\xi}$.

Proof. This result is deduced from our work in §2. While there we stated the estimates in terms of the measures m_Y and $m_{Y,f} = \prod_{p < \infty} m_{Y,p}$, but they can also be interpreted using local densities. By [6, Lemma 1.8.2],

$$m_Y(Y(\mathbb{R})_H) = \mu_\infty(Y; H).$$

Also, the proof of [6, Lemma 1.8.2] gives

$$m_{Y,p}(U) = \lim_{t \to \infty} p^{-t\dim(Y)} \#(U \,\mathrm{mod}\, p^t), \tag{3.3}$$

for every open $U \subset Y(\mathbb{Z}_p)$. In particular,

$$m_{Y,p}(B_p(\boldsymbol{\xi}, \ell)) = \hat{\mu}_p(Y; \boldsymbol{\xi}, \ell).$$

The result now follows from Corollary 2.6. $\qquad\square$

Next, we claim that

$$\prod_{p < \infty} \hat{\mu}_p(Y; \boldsymbol{\xi}; \ell) \ll \frac{1}{\ell^{\dim(Y)}}, \tag{3.4}$$

for any $\ell \in \mathbb{N}$. Let $p \mid \ell$ and let $\mu = v_p(\ell)$. Recall that Y is non-singular and let p be a prime of good reduction for Y. We set

$$N(p^t) = \#\{\mathbf{x} \in Y(\mathbb{Z}/p^t\mathbb{Z}) : \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, p^{v_p(\ell)})\}.$$

It follows from Hensel's lemma that $N(p^{t+1}) = p^{\dim(Y)} N(p^t)$ for any $t \geqslant \mu$. Hence $\hat{\mu}_p(Y; \boldsymbol{\xi}; \ell) = p^{-\mu \dim(Y)} N(p^\mu) = p^{-\mu \dim(Y)}$, since $\boldsymbol{\xi} \in Y(\mathbb{Z}/p^\mu\mathbb{Z})$. The claim now easily follows.

Now there are $\varrho(k^r) \leqslant k^{rn}$ choices of $\boldsymbol{\xi}$ which we must consider. We substitute the estimate from Proposition 3.1 to obtain

$$N^{(1)}(H) = \mu_\infty(Y; H)S(H)$$
$$+ O\left(H^{\Delta(1+r\{n+\dim(L)+2\dim(G)\})}\mu_\infty(Y; H)^{1-\delta} + H^\Delta E(f; H)\right),$$

where $E(f; H)$ is given by (3.2) and

$$S(H) = \sum_{k \leqslant H^\Delta} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} \prod_{p < \infty} \hat{\mu}_p(Y; \boldsymbol{\xi}; k^r). \qquad (3.5)$$

Lemma 3.2. *Let $g \in \mathbb{Z}[X_1, \ldots, X_n]$ be such that $g \not\equiv 0$ on Y. Then there exists $\eta > 0$ such that $E(g; H) = O_g(\mu_\infty(Y; H)^{1-\eta})$, where $E(g; H)$ is given by (3.2).*

Proof. Pick a large prime p. It follows from Proposition 3.1 and (3.4) that

$$E(g; H) \leqslant \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \, p \mid g(\mathbf{x})\}$$
$$= \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/p\mathbb{Z}) \\ g(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, p)}} V_p(H; \boldsymbol{\xi})$$
$$\ll \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/p\mathbb{Z}) \\ g(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, p)}} \left(\frac{\mu_\infty(Y; H)}{p^{\dim(Y)}} + p^{\dim(L)+2\dim(G)}\mu_\infty(Y; H)^{1-\delta}\right),$$

for some $\delta > 0$. Since Y is irreducible, $\dim(Y \cap \{g = 0\}) < \dim(Y)$ and so it follows from the Lang-Weil estimates that

$$\#\{\boldsymbol{\xi} \in Y(\mathbb{Z}/p\mathbb{Z}) : \, g(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, p)\} \ll_g p^{\dim(Y)-1}.$$

Hence,

$$E(g; H) \ll_g \frac{\mu_\infty(Y; H)}{p} + p^{3\dim(G)-1}\mu_\infty(Y; H)^{1-\delta},$$

since $\dim(Y) = \dim(G) - \dim(L)$. This is satisfactory for the lemma, on choosing p appropriately. $\qquad\square$

We may now conclude that there exists $\delta > 0$ such that

$$N^{(1)}(H) = \mu_\infty(Y; H)S(H) + O\left(H^{\Delta(1+r\{n+\dim(L)+2\dim(G)\})}\mu_\infty(Y; H)^{1-\delta}\right),$$

where $S(H)$ is given by (3.5). Turning to an analysis of $S(H)$, we appeal to Hypothesis-ϱ, which gives

$$\varrho(k^r) \leqslant C_{Y,f,r}^{\omega(k)} k^{r(\dim(Y)-1)} \ll_{\varepsilon,r} k^{r(\dim(Y)-1)+\varepsilon},$$

for an appropriate constant $C_{Y,f,r} > 0$. Invoking (3.4) with $\ell = k^r$, and recalling that $r \geqslant 2$, we may therefore extend the sum over k to infinity, finding that

$$S(H) = \sum_{k=1}^{\infty} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} \prod_{p<\infty} \hat{\mu}_p(Y; \boldsymbol{\xi}; k^r) + O(H^{-\Delta/2}).$$

The main term here is equal to the Euler product $\mathfrak{S}(Y, f, r)$ that is defined in (1.8). Putting everything together, we have therefore established the following result, which completes our treatment of the small moduli.

Proposition 3.3. *Let* $Y \simeq G/L \subset \mathbb{A}^n$ *be a symmetric variety over $\mathbb{Q}$ satisfying (i)–(iv), with $Y(\mathbb{Z}) \neq \emptyset$, and with L semisimple and simply connected. Assume that $f \in \mathbb{Z}[X_1, \ldots, X_n]$ satisfies Hypothesis-ϱ. Then there exists $\delta > 0$ such that*

$$\begin{aligned} N^{(1)}(H) =& \mathfrak{S}(Y, f, r)\mu_{\infty}(Y; H) \\ &+ O\left(H^{\Delta(1+r\{n+\dim(L)+2\dim(G)\})}\mu_{\infty}(Y; H)^{1-\delta} + \mu_{\infty}(Y; H)H^{-\Delta/2}\right). \end{aligned}$$

Moreover, $\mathfrak{S}(Y, f, r) > 0$ provided that f has no r-power divisors on Y.

3.2. **Proof of Theorem 1.2.** Assuming $\Delta > 0$ is chosen to be sufficiently small in terms of $\delta, \dim(G)$ and r, the error terms in Proposition 3.3 can both be made smaller than the main term. It remains to show that the contribution

$$N^{(2)}(H) = \sum_{H^{\Delta}<k\ll H^{d/r}} |\mu(k)| \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ 0 \neq f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^r)\} \quad (3.6)$$

is negligible. Here, we have truncated the outer sum to $k \ll H^{d/r}$, on supposing that f has degree d. As remarked in §1, our key observation for handling large moduli is based on the inequality

$$\begin{aligned} N^{(2)}(H) &\leqslant \sum_{H^{\Delta}<k\ll H^{d/r}} |\mu(k)| \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^2)\} \\ &= \sum_{H^{\Delta}<k\ll H^{d/r}} |\mu(k)| \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^2\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^2)}} V_{k^2}(H; \boldsymbol{\xi}), \end{aligned}$$

in the notation of Proposition 3.1. Combining this result with (3.4) and Hypothesis-ϱ, we therefore conclude that

$$N^{(2)}(H) \ll \sum_{H^\Delta < k \ll H^{d/r}} |\mu(k)| \varrho(k^2) \left\{ \frac{\mu_\infty(Y;H)}{k^{2\dim(Y)}} + k^{2\dim(L)+4\dim(G)} \mu_\infty(Y;H)^{1-\delta} \right\}$$

$$\ll \sum_{H^\Delta < k \ll H^{d/r}} \left\{ \frac{\mu_\infty(Y;H)}{k^2} + k^{6\dim(G)} \mu_\infty(Y;H)^{1-\delta} \right\}$$

$$\ll \frac{\mu_\infty(Y;H)}{H^\Delta} + H^{\frac{d}{r}(1+6\dim(G))} \mu_\infty(Y;H)^{1-\delta}.$$

The first term is satisfactory and the second term is also satisfactory provided that r is taken to be sufficiently large in terms of $d, \dim(G)$ and δ. This completes the proof of Theorem 1.2.

3.3. **Generalisation of Theorem 1.2 and proof of Theorem 1.1.** In this section we discuss symmetric varieties $Y \simeq G/L$ when L is not necessarily a semisimple simply connected group. This requires a more delicate analysis because such varieties may fail to satisfy the Hardy–Littlewood asymptotic formula. Throughout this section we assume that conditions (i)–(iv) from §1 hold. Let $\mathscr{O}_\mathbf{A} = \prod'_{p \leqslant \infty} \mathscr{O}_p \subset Y(\mathbf{A})$ be an orbit of $G(\mathbf{A})$ in $Y(\mathbf{A})$. Our goal is to estimate the counting function

$$N_r(\mathscr{O}_\mathbf{A}; f; H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_\mathbf{A} : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) \text{ is } r\text{-free}\}.$$

We introduce local densities associated to the orbit $\mathscr{O}_\mathbf{A}$:

$$\hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r) = \lim_{t \to \infty} p^{-t\dim(Y)} \#\{\mathbf{x} \in Y(\mathbb{Z}_p) \cap \mathscr{O}_\mathbf{A} \bmod p^t : \ p^r \nmid f(\mathbf{x})\}.$$

For almost all p, $\mathscr{O}_p \supset Y(\mathbb{Z}_p)$ which implies that $\hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r) = \hat{\mu}_p(Y, f, r)$. We also define the corresponding Euler product

$$\mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) = L(1, \varrho_L) \prod_{p < \infty} L_p(1, \varrho_L)^{-1} \hat{\mu}_p(\mathscr{O}_\mathbf{A}, f, r),$$

where $L(s, \varrho_L)$ is given by (2.21). Letting $\hat{\mu}_p(Y) = m_{Y,p}(Y(\mathbb{Z}_p))$, in the notation of (3.3), it follows from Hypothesis-ϱ that the product

$$\prod_{p < \infty} \frac{\hat{\mu}_p(Y, f, r)}{\hat{\mu}_p(Y)}$$

converges absolutely, so that the Euler product $\mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r)$ also converges absolutely.

With this notation, we establish the following result.

Theorem 3.4. *Let $Y \simeq G/L \subset \mathbb{A}^n$ be a symmetric variety over $\mathbb{Q}$ satisfying (i)–(iv). Let $\mathscr{O}_\mathbf{A} \subset Y(\mathbf{A})$ be an orbit of $G(\mathbf{A})$. Assume that the polynomial*

$f \in \mathbb{Z}[X_1, \ldots, X_n]$ *satisfies Hypothesis-ϱ. Then for all sufficiently large r, there exists $\delta > 0$ such that*

$$N_r(\mathscr{O}_{\mathbf{A}}; f; H) = \delta(\mathscr{O}_{\mathbf{A}})\mathfrak{S}(\mathscr{O}_{\mathbf{A}}, f, r)\mu_Y(\mathscr{O}_{\infty}(H)) + O_r\left(\mu_Y(\mathscr{O}_{\infty}(H))^{1-\delta}\right).$$

Moreover, if we assume that there exists $\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}}$ such that $f(\mathbf{x})$ is r-free, then $\delta(\mathscr{O}_{\mathbf{A}}) > 0$ and $\mathfrak{S}(\mathscr{O}_{\mathbf{A}}, f, r) > 0$.

If the variety Y additionally satisfies the smoothness assumption (1.5), then it follow from the argument in [6, Lemma 1.8.2] that

$$\mu_Y(\mathscr{O}_{\infty}(H)) = \mu_{\infty}(\mathscr{O}_{\mathbf{A}}; H),$$

where the local density $\mu_{\infty}(\mathscr{O}_{\mathbf{A}}; H)$ is defined analogously to (1.11).

Remark 3.5. We note that

$$N_r(Y; f; H) = \sum_{\mathscr{O}_{\mathbf{A}} \subset Y(\mathbf{A})} N_r(\mathscr{O}_{\mathbf{A}}; f; H),$$

where the sum is taken over finitely many orbits $\mathscr{O}_{\mathbf{A}}$ that have non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p<\infty} Y(\mathbb{Z}_p)$. Hence, Theorem 3.4 implies an asymptotic formula for the counting function $N_r(Y; f; H)$. In fact, this asymptotic formula can be stated in terms of the Tamagawa volume of a suitable subset of $Y(\mathbf{A})$ defined by the integral Brauer–Manin obstruction, as introduced by Colliot-Thél'ene and Xu [11]. We denote by $Y(\mathbf{A})^{\mathrm{Br}(Y)}$ the kernel of the Brauer–Manin pairing. Since G is assumed to be simply connected, this kernel consists of orbits of $G(\mathbf{A})$ (see [11, Thm. 3.2]). If $\mathscr{O}_{\mathbf{A}} \cap Y(\mathbf{A})^{\mathrm{Br}(Y)} = \emptyset$, then the orbit $\mathscr{O}_{\mathbf{A}}$ contains no rational points, and $\delta(\mathscr{O}_{\mathbf{A}}) = 0$. On the other hand, if $\mathscr{O}_{\mathbf{A}} \subset Y(\mathbf{A})^{\mathrm{Br}(Y)}$, and $\mathscr{O}_{\mathbf{A}}$ has non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p<\infty} Y(\mathbb{Z}_p)$, then it follows from [11, Thm. 3.7] that $\mathscr{O}_{\mathbf{A}} \cap Y(\mathbb{Z}) \neq \emptyset$. In particular, we conclude that for these orbits $\delta(\mathscr{O}_{\mathbf{A}}) > 0$. Since $\delta : Y(\mathbf{A}) \to \{0, |\mathrm{Pic}(L)|\}$, we have $\delta(\mathscr{O}_{\mathbf{A}}) = |\mathrm{Pic}(L)|$ for these orbits. Thus, setting

$$Y(\mathbf{A}_f)^{(f,r)} = \left\{ (\mathbf{y}_p) \in \prod_{p<\infty} Y(\mathbb{Z}_p) : \ f(\mathbf{y}_p) \not\equiv 0 \ (\mathrm{mod}\, p^r) \right\},$$

we conclude that

$$\begin{aligned} N_r(Y; f; H) = &|\mathrm{Pic}(L)| \cdot (m_Y \otimes m_{Y,f}) \left((Y(\mathbb{R})_H \times Y(\mathbf{A}_f)^{(f,r)}) \cap Y(\mathbf{A})^{\mathrm{Br}(Y)}\right) \\ &+ O(V(H)^{1-\delta}), \end{aligned}$$

where $V(H) = \max_{\mathscr{O}_{\mathbf{A}} \subset Y(\mathbf{A})} \mu_{\infty}(\mathscr{O}_{\mathbf{A}}; H)$, with the maximum taken over the finitely many orbits having non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p<\infty} Y(\mathbb{Z}_p)$.

Proof of Theorem 3.4. The proof follows the same strategy as the proof of Theorem 1.2 presented in §3.1–3.2. As in the proof of Theorem 1.2, we start by estimating

$$N^{(1)}(\mathscr{O}_\mathbf{A}; H) = \sum_{k \leqslant H^\Delta} \mu(k) \#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_\mathbf{A} : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^r)\}$$
$$+ O\left(H^\Delta E_{\mathscr{O}_\mathbf{A}}(f; H)\right),$$

where $E_{\mathscr{O}_\mathbf{A}}(f; H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_\mathbf{A} : |\mathbf{x}| \leqslant H, \ f(\mathbf{x}) = 0\}$. Then

$$N^{(1)}(\mathscr{O}_\mathbf{A}; H) = \sum_{k \leqslant H^\Delta} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} V_{k^r}(\mathscr{O}_\mathbf{A}; H; \boldsymbol{\xi}) + O\left(H^\Delta E_{\mathscr{O}_\mathbf{A}}(f; H)\right),$$

where

$$V_\ell(\mathscr{O}_\mathbf{A}; H; \boldsymbol{\xi}) = \#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_\mathbf{A} : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, \ell)\},$$

for given ℓ and $\boldsymbol{\xi} \in Y(\mathbb{Z}/\ell\mathbb{Z})$. The quantity $V_\ell(\mathscr{O}_\mathbf{A}; H; \boldsymbol{\xi})$ can be estimated as in Proposition 3.1. We obtain that there exists $\delta > 0$ such that

$$\begin{aligned} V_\ell(\mathscr{O}_\mathbf{A}; H; \boldsymbol{\xi}) &= \delta(\mathscr{O}_\mathbf{A})\mu_Y(\mathscr{O}_\infty(H))\mathfrak{S}(\mathscr{O}_\mathbf{A}, \boldsymbol{\xi}; \ell) \\ &\quad + O(\ell^{\dim(L)+2\dim(G)}\mu_Y(\mathscr{O}_\infty(H))^{1-\delta}), \end{aligned} \tag{3.7}$$

where

$$\mathfrak{S}(\mathscr{O}_\mathbf{A}, \boldsymbol{\xi}; \ell) = L(1, \varrho_L) \prod_{p < \infty} L_p(1, \varrho_L)^{-1} \hat{\mu}_p(\mathscr{O}_\mathbf{A}, \boldsymbol{\xi}; \ell)$$

and

$$\hat{\mu}_p(\mathscr{O}_\mathbf{A}; \boldsymbol{\xi}, \ell) = \lim_{t \to \infty} p^{-t\dim(Y)} \#\{\mathbf{x} \in \mathscr{O}_p \cap Y(\mathbb{Z}_p) \bmod p^t : \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\, p^{v_p(\ell)})\}.$$

Indeed, this estimate can be directly deduced from Corollary 2.5 by observing that

$$m_{Y,p}(\mathscr{O}_p \cap B_p(\boldsymbol{\xi}, \ell)) = \hat{\mu}_p(\mathscr{O}_\mathbf{A}; \boldsymbol{\xi}, \ell),$$

which follows from (3.3). Next, we substitute (3.7) into our work above to deduce that

$$\begin{aligned} N^{(1)}(\mathscr{O}_\mathbf{A}; H) &= \delta(\mathscr{O}_\mathbf{A})\mu_Y(\mathscr{O}_\infty(H))S(\mathscr{O}_\mathbf{A}; H) + O\left(H^\Delta E_{\mathscr{O}_\mathbf{A}}(f; H)\right) \\ &\quad + O\left(H^{\Delta(1+r\{n+\dim(L)+2\dim(G)\})}\mu_Y(\mathscr{O}_\infty(H))^{1-\delta}\right), \end{aligned}$$

where

$$S(\mathscr{O}_\mathbf{A}; H) = \sum_{k \leqslant H^\Delta} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} \mathfrak{S}(\mathscr{O}_\mathbf{A}, \boldsymbol{\xi}; k^r).$$

Arguing as in the proof of Theorem 1.2, we find that

$$S(\mathscr{O}_{\mathbf{A}}; H) = \sum_{k=1}^{\infty} \mu(k) \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^r\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^r)}} \mathfrak{S}(\mathscr{O}_{\mathbf{A}}, \boldsymbol{\xi}; k^r) + O(H^{-\Delta/2}),$$

with the main term is equal to the Euler product $\mathfrak{S}(\mathscr{O}_{\mathbf{A}}, f, r)$.

Arguing as in Lemma 3.2, we easily use (3.7) to show that

$$E_{\mathscr{O}_{\mathbf{A}}}(f; H) = O_g(\mu_Y(\mathscr{O}_{\infty}(H))^{1-\eta}), \tag{3.8}$$

for some $\eta > 0$. Hence, we conclude that

$$\begin{aligned}
N^{(1)}(\mathscr{O}_{\mathbf{A}}; H) = {}& \delta(\mathscr{O}_{\mathbf{A}})\mathfrak{S}(\mathscr{O}_{\mathbf{A}}, f, r)\mu_Y(\mathscr{O}_{\infty}(H)) + O\left(\mu_Y(\mathscr{O}_{\infty}(H))H^{-\Delta/2}\right) \\
& + O\left(H^{\Delta(1+r\{n+\dim(L)+2\dim(G)\})}\mu_Y(\mathscr{O}_{\infty}(H))^{1-\delta}\right),
\end{aligned} \tag{3.9}$$

for every adelic orbit $\mathscr{O}_{\mathbf{A}} \subset Y(\mathbf{A})$. When $\Delta > 0$ is chosen sufficiently small, the error terms in this estimate can be made smaller than the main term.

Now it remains to estimate

$$\begin{aligned}
N^{(2)}(\mathscr{O}_{\mathbf{A}}; H) = {}& \sum_{H^\Delta < k \ll H^{d/r}} |\mu(k)| \#\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}} : |\mathbf{x}| \leqslant H,\ f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\, k^r)\} \\
& \leqslant \sum_{H^\Delta < k \ll H^{d/r}} |\mu(k)| \sum_{\substack{\boldsymbol{\xi} \in Y(\mathbb{Z}/k^2\mathbb{Z}) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\, k^2)}} V_{k^2}(\mathscr{O}_{\mathbf{A}}; H; \boldsymbol{\xi}).
\end{aligned}$$

Thus we can argue as in §3.2, using (3.7), combined with Hypotheses-ϱ and (3.4), to conclude that if r is taken sufficiently large, there exists $\delta' > 0$ such that

$$N^{(2)}(\mathscr{O}_{\mathbf{A}}; H) = O(\mu_Y(\mathscr{O}_{\infty}(H))^{1-\delta'}).$$

This completes the proof of the theorem. $\square$

Proof of Theorem 1.1. Since $Y(\mathbb{Z}) \neq \emptyset$, we may pick $\mathbf{x} \in Y(\mathbb{Z})$ with $f(\mathbf{x}) \neq 0$. Suppose that $|f(\mathbf{x})| = \prod_{1 \leqslant i \leqslant \ell} p_i^{r_i}$, for distinct primes $p_1, \ldots, p_\ell$. Then $f(\mathbf{x})$ is r_0-free, with $r_0 = 1 + \max_{1 \leqslant i \leqslant \ell}\{r_i\}$. We apply Theorem 3.4 to an orbit $\mathscr{O}_{\mathbb{A}}$ that contains this point $\mathbf{x}$. Then $\delta(\mathscr{O}_{\mathbf{A}}) > 0$ and $\mathfrak{S}(\mathscr{O}_{\mathbf{A}}, f, r) > 0$ if $r \geqslant r_0$. Comparing the asymptotic formula given by Theorem 3.4 with (3.8) we deduce that the set $\{\mathbf{x} \in Y(\mathbb{Z}) \cap \mathscr{O}_{\mathbf{A}} : f(\mathbf{x}) \text{ is } r\text{-free}\}$ is Zariski dense in Y. (We note that $\mu_Y(\mathscr{O}_{\infty}(H)) \to \infty$, as $H \to \infty$, by Lemma 2.2.) $\square$

4. THE ROADMAP FOR QUADRICS

It is now time to initiate the proof of Theorems 1.4, 1.5 and 1.10. From this point forwards, $n \geqslant 3$ and $Y \subset \mathbb{A}^n$ is the affine quadric (1.9), where

$Q \in \mathbb{Z}[X_1, \ldots, X_n]$ is a non-singular indefinite quadratic form and m is a non-zero integer such that $-m \det(Q) \neq \square$ when $n = 3$. We begin with a proof of Theorem 1.11 in §4.1. Next, in §4.2, we shall establish Hypothesis-ϱ for the polynomials $f \in \mathbb{Z}[X_1, \ldots, X_n]$ under consideration. Finally, in §4.3 we shall collect together the main steps in the proof of Theorems 1.4, 1.5 and 1.10. The primary ingredients in this endeavour here are the results in §3 and the treatment of large moduli in §5.

4.1. **Integral points on affine quadrics.** In this section we establish Theorem 1.11. We begin by noting that it suffices to assume that q is absolutely irreducible in the statement of the theorem, rather than merely irreducible over $\mathbb{Q}$. Indeed, if q factorises as $\ell_1 \ell_2$ for linear polynomials $\ell_1, \ell_2 \in \overline{\mathbb{Q}}[T_1, \ldots, T_\nu]$, neither one of which is proportional to a linear polynomial defined over $\mathbb{Z}$, then ℓ_1, ℓ_2 are not proportional to each other and ℓ_1 must be a conjugate of ℓ_2. Moreover, the integer points $\mathbf{t}$ in which we are interested must satisfy the pair of equations $\ell_1(\mathbf{t}) = \ell_2(\mathbf{t}) = 0$. Such points clearly contribute only $O_\nu(B^{\nu-2})$ to $M(q; B)$, which is satisfactory.

With the restriction to absolutely irreducible q in place, we will establish Theorem 1.11 by induction on $\nu \geqslant 2$, following the approach in [8, §4]. We henceforth set

$$R(X_0, X_1, \ldots, X_\nu) = X_0^2 q(X_1/X_0, \ldots, X_\nu/X_0) \qquad (4.1)$$

for the homogenised quadratic form associated to q. In particular q_0 is obtained by setting $X_0 = 0$ in R. Since q is absolutely irreducible it follows that R is absolutely irreducible and so has rank at least 3. Moreover, by hypothesis, the quadratic form $R(0, X_1, \ldots, X_\nu)$ has rank at least 2. We will need the following result, due to Browning, Heath-Brown and Salberger [8, Lemma 13].

Lemma 4.1. *Let $\varepsilon > 0$, let $B \geqslant 1$ and suppose that $R \in \mathbb{Z}[X_0, X_1, X_2]$ is a non-singular quadratic form such that the binary form $R(0, X_1, X_2)$ is also non-singular. Then for any $t \in \mathbb{Z} \cap [-B, B]$ we have*

$$\# \left\{ (x_1, x_2) \in \mathbb{Z}^2 : \begin{array}{l} |x_1|, |x_2| \leqslant B, \ R(t, x_1, x_2) = 0 \\ \gcd(t, x_1, x_2) = 1 \end{array} \right\} = O_\varepsilon(B^\varepsilon).$$

Following our convention, the implied constant in this estimate does not depend on t or on the coefficients of R. The case $\nu = 2$ of Theorem 1.11 is now a trivial consequence of Lemma 4.1 with $t = 1$. We will require a separate treatment of the case $\nu = 3$ when q is absolutely irreducible with $q = q_0$. In this case the statement of Theorem 1.11 follows from taking $d = 2$ in work of Heath-Brown [18, Thm. 3].

We now turn to the proof of Theorem 1.11 when $\nu \geqslant 3$, assuming that $q \neq q_0$ when $\nu = 3$. Our plan is to take hyperplane slices and apply the inductive hypothesis. We claim that there exists $\mathbf{a} \in \mathbb{Z}^\nu$, with $0 < |\mathbf{a}| = O_\nu(1)$, such

that the quadratic polynomial obtained by eliminating a variable from the pair of equations

$$q(T_1, \ldots, T_\nu) = a_1 T_1 + \cdots + a_\nu T_\nu = 0 \qquad (4.2)$$

is absolutely irreducible and has quadratic part with rank at least 2. Taking this claim on faith for the moment, we may assume after a possible change of variables that $\mathbf{a} = (0, \ldots, 0, 1)$. Thus, for any $k \in \mathbb{Z}$, the polynomial $q_k = q(T_1, \ldots, T_{\nu-1}, k)$ is both absolutely irreducible and has quadratic part with rank at least 2. In this way we obtain the bound

$$M(q; B) \leqslant \sum_{|k| \leqslant B} M(q_k; B) \ll_{\varepsilon, \nu} B^{\nu - 2 + \varepsilon},$$

by the inductive hypothesis. This completes the proof of the theorem subject to the claim.

Let us call a vector $\mathbf{a} \in \mathbb{C}^\nu$ *defective* if the polynomial induced by (4.2) fails to be absolutely irreducible or has quadratic part with rank at most 1. We will construct a proper subvariety $E \subset \mathbb{P}^{\nu-1}$ defined over $\mathbb{Z}$, with degree $O_\nu(1)$, such that $[\mathbf{a}] \in E$ whenever $\mathbf{a}$ is defective. Once this is achieved it is a simple matter to find a vector $\mathbf{a} \in \mathbb{Z}^\nu$ satisfying the claim. Indeed, for any $A > 1$, there are at least $c_1(\nu) A^\nu$ possible non-zero vectors $\mathbf{a} \in \mathbb{Z}^\nu$ for which $|\mathbf{a}| \leqslant A$, for an appropriate constant $c_1(\nu) > 0$. Moreover, it follows from the trivial estimate [8, Lemma 2] that there are at most $c_2(\nu) A^{\nu-1}$ defective vectors $\mathbf{a} \in \mathbb{Z}^\nu$ satisfying $|\mathbf{a}| \leqslant A$, for an appropriate constant $c_2(\nu) > 0$. The claim then follows on taking $A > c_2(\nu)/c_1(\nu)$.

It remains to construct the variety E. Let us begin by considering vectors $\mathbf{a} \in \mathbb{C}^\nu$ for which (4.2) is not absolutely irreducible. When q is homogeneous, so that $q = q_0$ and $\nu \geqslant 4$, then it is well-known (see [8, Lemma 7], for example) that there exists a non-zero form $F \in \mathbb{Z}[X_1, \ldots, X_\nu]$, with degree $O_\nu(1)$, such that $F(\mathbf{a}) = 0$ when the intersection is not absolutely irreducible. Alternatively, when $q \neq q_0$ and $\nu \geqslant 3$, we will work with the homogenised quadratic form (4.1). Let $U \subset \mathbb{P}^\nu$ denote the quadric $R = 0$. Using elimination theory we can construct a form $F \in \mathbb{Z}[X_1, \ldots, X_\nu]$, with degree $O_\nu(1)$, such that $F(\mathbf{a}) = 0$ whenever the intersection of U with the hyperplane $\sum_{i=1}^\nu a_i X_i = 0$ produces a reducible quadric. We need to show that F is non-zero. Let $x = [1, 0, \ldots, 0]$ and let Σ_x denote the set of hyperplanes in $\mathbb{P}^\nu$ containing x. Then the desired conclusion follows from the version of Bertini's theorem found in Fulton and Lazarsfeld [17, Thm. 1.1], which shows that $U \cap H$ is absolutely irreducible for generic $H \in \Sigma_x$. We let $E_1 \subset \mathbb{P}^{\nu-1}$ denote the projective hypersurface $F = 0$.

Shifting attention to the vectors $\mathbf{a} \in \mathbb{C}^\nu$ for which (4.2) has quadratic part with rank at most 1, the 2×2 minors of the underlying quadratic form give a system of six homogeneous quadratic equations whose simultaneous vanishing

at $\mathbf{a}$ encapsulates this property. We denote this variety by $E_2 \subset \mathbb{P}^{\nu-1}$. After verifying that E_2 is a proper subvariety, our construction is completed by taking $E = E_1 \cup E_2$.

Let $V \subset \mathbb{P}^{\nu-1}$ denote the quadric $q_0 = 0$. Then $\nu \geqslant 3$ and V has rank $r_V \geqslant 2$, on identifying the rank of a quadric with the rank of the underlying quadratic form. To prove that $E_2 \neq \mathbb{P}^{\nu-1}$ it suffices to show that for generic hyperplanes H in $\mathbb{P}^{\nu-1}$ the intersection $V \cap H$ produces a quadric with rank at least 2. Thus we need to know some elementary facts about how the ranks of quadratic forms diminish on linear subspaces. For a hyperplane H let $W = V \cap H$ and let r_W be the associated rank. It is well-known that $r_W \geqslant r_V - 2$. We will need slightly finer information (see Swinnerton-Dyer [27, p. 264], for example). If $r_V = \nu$ then $r_W = \nu - 1$ if H is not tangent to V. If $r_V < \nu$ then the singular points of V form a linear space L of dimension $\nu - r_V - 1$ and we have $r_W = r_V$ if $L \not\subset H$. In either case we deduce that $r_W \geqslant 2$ for generic H. This concludes the proof of Theorem 1.11.

4.2. Polynomial congruences modulo prime powers.

This section is concerned with counting solutions to certain systems of polynomial congruences modulo prime powers. Recalling the definition (1.2) of $\varrho(\ell)$, we begin by establishing Hypothesis-ϱ for the non-singular forms $f \in \mathbb{Z}[X_1, \ldots, X_n]$ that feature in Theorems 1.4, 1.5 and 1.10, when $Y \subset \mathbb{A}^n$ is the affine quadric (1.9). (Note that a linear form is automatically non-singular.)

Lemma 4.2. *Let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a non-singular form of degree $d \geqslant 1$. Then we have $\varrho(p^r) \ll_r p^{r(n-2)}$.*

Proof. The trivial bound is $\varrho(p^r) \leqslant p^{rn}$. Since we allow our implied constant to depend on r, we may henceforth assume that $p \nmid 2dm\Delta_f \det(Q)$, where Δ_f is the discriminant of f.

When $r = 1$ it follows from the Lang–Weil estimate that $\varrho(p) = O(p^{n-2})$. When $r \geqslant 1$ the statement of the lemma will follow provided we can show that

$$\varrho(p^{r+1}) = p^{n-2}\varrho(p^r).$$

To verify this we use an approach based on Hensel's lemma. Let $\mathbf{x} \, (\mathrm{mod}\, p^r)$ be counted by $\varrho(p^r)$ and consider the vectors $\mathbf{x} + p^r\mathbf{y}$ for $\mathbf{y} \, (\mathrm{mod}\, p)$. Such a vector runs modulo p^{r+1} and is counted by $\varrho(p^{r+1})$ if and only if

$$p^{-r}f(\mathbf{x}) + \mathbf{y}.\nabla f(\mathbf{x}) \equiv 0 \, (\mathrm{mod}\, p),$$
$$p^{-r}(Q(\mathbf{x}) - m) + \mathbf{y}.\nabla Q(\mathbf{x}) \equiv 0 \, (\mathrm{mod}\, p).$$

We claim that $\nabla f(\mathbf{x})$ and $\nabla Q(\mathbf{x})$ are not proportional modulo p, from which it will follow that there are p^{n-2} possibilities for $\mathbf{y} \, (\mathrm{mod}\, p)$, as required. Suppose for a contradiction that there exists $\lambda, \mu \in \mathbb{F}_p$, not both zero, such that $\lambda\nabla f(\mathbf{x}) \equiv \mu\nabla Q(\mathbf{x}) \, (\mathrm{mod}\, p)$. Since $p \nmid m$ we must have $p \nmid \mathbf{x}$. In particular,

$\lambda\mu \neq 0$ since f and Q are non-singular modulo p. It then follows from Euler's identity that

$$\begin{aligned} 0 &\equiv d\lambda f(\mathbf{x}) \,(\mathrm{mod}\,p) \\ &\equiv \lambda\mathbf{x}.\nabla f(\mathbf{x}) \,(\mathrm{mod}\,p) \\ &\equiv \mu\mathbf{x}.\nabla Q(\mathbf{x}) \,(\mathrm{mod}\,p) \\ &\equiv 2\mu m \,(\mathrm{mod}\,p), \end{aligned}$$

which is a contradiction.   □

When $n = 3$ it turns out that we shall also need a good bound for

$$\varrho(\ell; \mathbf{c}) = \#\left\{\mathbf{x} \in Y(\mathbb{Z}/\ell\mathbb{Z}) : f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\,\ell), \ \mathbf{c}.\mathbf{x} \equiv 0 \,(\mathrm{mod}\,\ell)\right\},$$

for any $\ell \in \mathbb{N}$ and $\mathbf{c} \in \mathbb{Z}^n$. Note that $\varrho(\ell) = \varrho(\ell; \mathbf{0})$. The quantity $\varrho(\ell; \mathbf{c})$ is a multiplicative function of ℓ and our next result is concerned with estimating it when $\ell = p^r$.

Lemma 4.3. *Let $n = 3$ and let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a non-singular form of degree d. Then we have $\varrho(p^r; \mathbf{c}) \ll_r p^{(1-1/d)r} \gcd(p^r, \mathbf{c})^3$.*

Proof. As in the proof of Lemma 4.2, we may proceed under the assumption that $p \nmid 2dm\Delta_f \det(Q)$. In particular f is non-singular modulo p and does not vanish identically on the linear form $\mathbf{c}.\mathbf{x}$ modulo p.

We begin by proving the result under the assumption that $p \nmid \mathbf{c}$, analysing $\varrho(p^r; \mathbf{c})$ via the non-singular change of variables $\mathbf{y} = (x_1, x_2, \mathbf{c}.\mathbf{x})$. Assuming without loss of generality that $p \nmid c_3$, we find that $\varrho(p^r; \mathbf{c})$ is equal to

$$\#\left\{(y_1, y_2) \,(\mathrm{mod}\,p^r) : \begin{array}{l} Q(c_3y_1, c_3y_2, -c_1y_1 - c_2y_2) \equiv c_3^2 m \,(\mathrm{mod}\,p^r) \\ f(c_3y_1, c_3y_2, -c_1y_1 - c_2y_2) \equiv 0 \,(\mathrm{mod}\,p^r) \end{array}\right\}.$$

There is no contribution from y_1, y_2 for which $p \mid (y_1, y_2)$. Suppose without loss of generality that $p \nmid y_2$. We make the further change of variables $y_1 = zy_2$, now finding that the contribution to $\varrho(p^r; \mathbf{c})$ is

$$\#\left\{(y_2, z) \,(\mathrm{mod}\,p^r) : p \nmid y_2, \ y_2^2 h(z) \equiv c_3^2 m \,(\mathrm{mod}\,p^r), \ g(z) \equiv 0 \,(\mathrm{mod}\,p^r)\right\},$$

where

$$h(z) = Q(c_3z, c_3, -c_1z - c_2) \quad \text{and} \quad g(z) = f(c_3z, c_3, -c_1z - c_2).$$

Here $g(z)$ is a polynomial of degree at most d which does not vanish identically modulo p. Moreover we are only interested in roots of $g(z)$ modulo p^r for which $p \nmid h(z)$. It follows from work of Stewart [26, Cor. 2] that the number of $z \,(\mathrm{mod}\,p^r)$ is $O(p^{(1-1/d)r})$. For given z there are then at most 2 available choices for y_2, which therefore completes the proof of the lemma when $p \nmid \mathbf{c}$.

Suppose now that $p^j \| \mathbf{c}$. If $j \geqslant r$ then we get a satisfactory bound for the lemma by taking the trivial bound $\varrho(p^r; \mathbf{c}) \leqslant p^{3r} = \gcd(p^r, \mathbf{c})^3$. Alternatively,

if $j < r$ we write $\mathbf{x} = \mathbf{u} + p^{r-j}\mathbf{v}$ for $\mathbf{u} \,(\mathrm{mod}\, p^{r-j})$ and $\mathbf{v} \,(\mathrm{mod}\, p^j)$. The number of $\mathbf{u}$ is precisely $\varrho(p^{r-j}; \tilde{\mathbf{c}})$, where $\tilde{\mathbf{c}} = p^{-j}\mathbf{c}$. The number of $\mathbf{v}$ is trivially at most p^{3j}. Hence we have $\varrho(p^r; \mathbf{c}) \leqslant p^{3j}\varrho(p^{r-j}; \tilde{\mathbf{c}})$. Applying our earlier bound for the case $p \nmid \mathbf{c}$, we therefore complete the proof of the lemma. $\qquad\square$

4.3. **Summary of the argument.** It is now time to survey the key steps in the proof of Theorems 1.4, 1.5 and 1.10. Let $r \geqslant 2$ and let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a non-singular form of degree $d \geqslant 1$. Recalling the counting function $N_r(Y, f; H)$ from (1.4), we let $\Delta > 0$ and consider the contributions $N^{(1)}(H)$ and $N^{(2)}(H)$ that were defined in (3.1) and (3.6), respectively. In view of Remark 1.3, we see that the treatment of $N^{(1)}(H)$ (i.e. the small moduli) is handled by Proposition 3.3 and (3.9), in which we can take $\mu_\infty(Y; H) \ll H^{n-2}$ and $\mu_\infty(\mathscr{O}_\infty(H)) \ll H^{n-2}$, $\dim(G) = n(n-1)/2$ and $\dim(L) = (n-1)(n-2)/2$.

When $n \geqslant 4$, L is simply connected and it follows from Proposition 3.3 that there exists $\delta > 0$ such that

$$N^{(1)}(H) = \mathfrak{S}(Y, f, r)\mu_\infty(Y; H)$$
$$+ O\left(H^{n-2}\left\{H^{-\Delta/2} + H^{\Delta(1+r(\frac{3}{2}n^2 - \frac{3}{2}n+1)-\delta)}\right\}\right).$$

Moreover, for $n = 3$, there exists $\delta > 0$ such that

$$N^{(1)}(H) = \sum_{\mathscr{O}_\mathbf{A} \subset Y(\mathbf{A})} \delta(\mathscr{O}_\mathbf{A})\mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r)\mu_\infty(\mathscr{O}_\mathbf{A}; H)$$
$$+ O\left(H^{1-\Delta/2} + H^{1+\Delta(1+10r)-\delta}\right),$$

where the sum is taken over finitely many orbits $\mathscr{O}_\mathbf{A}$ that have non-trivial intersection with $Y(\mathbb{R}) \times \prod_{p<\infty} Y(\mathbb{Z}_p)$. Our additional assumption on existence of r-free points (when $n = 3$) guarantees that for at least one of the orbits $\mathscr{O}_\mathbf{A}$ we have $\delta(\mathscr{O}_\mathbf{A}) > 0$ and $\mathfrak{S}(\mathscr{O}_\mathbf{A}, f, r) > 0$. On taking $\Delta > 0$ to be sufficiently small in terms of δ we can ensure that these error terms are all satisfactory from the point of view of Theorems 1.4, 1.5 and 1.10.

For given $\Delta > 0$, it remains to show that there exists $\eta > 0$, depending on Δ, r, d and n, such that

$$N^{(2)}(H) \ll H^{n-2-\eta}. \tag{4.3}$$

We shall do so provided that f is a non-singular form of degree $d \geqslant 1$, with r satisfying the lower bounds from Theorems 1.4, 1.5 or 1.10, which will thereby suffice to conclude their proof. For $\ell \in \mathbb{N}$, the estimation of $N^{(2)}(H)$ hinges upon good upper bounds for

$$U_\ell(H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ 0 \neq f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\,\ell)\}, \tag{4.4}$$

The following result summarises our treatment of $U_\ell(H)$ when $d = 1$.

Proposition 4.4. *Suppose that $f \in \mathbb{Z}[X_1, \ldots, X_n]$ is a linear form. Let $\varepsilon > 0$ and let $\ell \in \mathbb{N}$. Then $U_\ell(H) = O_\varepsilon(\ell^{-1} H^{n-2+\varepsilon})$.*

This result will be established in §5.1. Applying Proposition 4.4 with $\ell = k^r$ and $\varepsilon = \Delta/2$, we obtain

$$N^{(2)}(H) \ll H^{n-2+\Delta/2} \sum_{k > H^\Delta} \frac{|\mu(k)|}{k^2} \ll H^{n-2-\Delta/2}.$$

This is satisfactory for (4.3).

Estimating $U_\ell(H)$ for $d \geqslant 2$ is more difficult. For square-free $k \in \mathbb{N}$ we deal with this by noting that $U_{k^r}(H) \leqslant U_{k^j}(H)$, for any $j \leqslant r$, whence

$$N^{(2)}(H) \leqslant \sum_{H^\Delta < k \ll H^{d/r}} |\mu(k)| U_{k^j}(H). \tag{4.5}$$

We will establish the following result in §5.2.

Proposition 4.5. *Let $n \geqslant 4$ and suppose that $f \in \mathbb{Z}[X_1, \ldots, X_n]$ is a non-singular form of degree $d \geqslant 2$. Let $\varepsilon > 0$ and let $k \in \mathbb{N}$ be square-free such that $k^{jn/(n-1)} \leqslant H$. Then $U_{k^j}(H) = O_{\varepsilon,j}(k^{-j/(n-1)} H^{n-2+\varepsilon})$.*

This allows us to establish (4.3) when $n \geqslant 4$. We make the assumption that j and r are chosen so that

$$\frac{jdn}{(n-1)r} \leqslant 1. \tag{4.6}$$

Then it will follow that $k^{jn/(n-1)} \ll H$ in (4.5), since $k \ll H^{d/r}$. Hence Proposition 4.5 yields

$$N^{(2)}(H) \ll_{\varepsilon,j} H^{n-2+\varepsilon} \sum_{k > H^\Delta} \frac{|\mu(k)|}{k^{j/(n-1)}},$$

for any $\varepsilon > 0$. Here the exponent of k exceeds 1 if and only if $j > n - 1$. We choose $j = n$, with which choice we can conclude that (4.3) holds for any $\eta < \Delta/(n-1)$, provided that r satisfies (4.6) with $j = n$. But this is equivalent to $r \geqslant dn^2/(n - 1)$, which was one of the assumptions in Theorem 1.4.

When $n = 3$ we are not able to get such a good bound for $U_{k^j}(H)$. The following result will also be established in §5.2.

Proposition 4.6. *Let $n = 3$ and suppose that $f \in \mathbb{Z}[X_1, X_2, X_3]$ is a non-singular form of degree $d \geqslant 2$. Let $\varepsilon > 0$ and let $k \in \mathbb{N}$ be square-free such that $k^{4j/3} \leqslant H$. Then $U_{k^j}(H) = O_{\varepsilon,j}(k^{-j/(3d)} H^{1+\varepsilon})$.*

Let us see how this is sufficient to prove (4.3) when $n = 3$. We make the assumption that j and r are chosen so that

$$\frac{4jd}{3r} \leqslant 1. \tag{4.7}$$

Then, as before, it will follow that $k^{4j/3} \ll H$ in (4.5), since $k \ll H^{d/r}$. Hence we may apply Proposition 4.6 in (4.5), giving

$$N^{(2)}(H) \ll_{\varepsilon,j} H^{1+\varepsilon} \sum_{k>H^\Delta} \frac{|\mu(k)|}{k^{j/(3d)}},$$

for any $\varepsilon > 0$ and any j such that (4.7) holds. Here the exponent of k exceeds 1 if and only if $j > 3d$. We choose $j = 3d + 1$, with which choice we conclude that (4.3) holds with $n = 3$ and any $\eta < \Delta/(3d)$, provided that r satisfies the inequality $r \geqslant \frac{4}{3}d(3d + 1)$, which was one of the hypotheses of Theorem 1.5.

5. LARGE MODULI

5.1. **Linear polynomials.** In this section we establish Proposition 4.4. Let $Y \subset \mathbb{A}^n$ denote the quadric $Q = m$, for $n \geqslant 3$, and let $f \in \mathbb{Z}[X_1, \ldots, X_n]$ be a linear form. We are interested in the quantity

$$U_\ell(H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ 0 \neq f(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\,\ell)\},$$

for $\ell \in \mathbb{N}$. After a non-singular linear change of variables, we see that it suffices to prove Proposition 4.4 when $f = X_n$. In particular $U_\ell(H) = 0$ unless $\ell \ll H$, which we henceforth assume.

Suppose first that $n \geqslant 4$. Note that

$$U_\ell(H) = \sum_{\substack{h \in \mathbb{Z}, \ 0 < |h| \ll H \\ h \equiv 0 \,(\mathrm{mod}\,\ell)}} L_h(H),$$

where $L_h(H) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ x_n = h\}$, for any $h \in \mathbb{Z}$. For $n \geqslant 4$ we claim that

$$L_h(H) = O_\varepsilon(H^{n-3+\varepsilon}), \tag{5.1}$$

for any $\varepsilon > 0$, where the implied constant does not depend on h. This will suffice to establish Proposition 4.4 for $n \geqslant 4$, since there are $O(\ell^{-1}H)$ integer values of $h \ll H$ such that $\ell \mid h$.

To verify (5.1) we write

$$q(X_1, \ldots, X_{n-1}) = Q(X_1, \ldots, X_{n-1}, h) - m,$$

for the quadratic polynomial in $L_h(H)$. Then

$$q_0(X_1, \ldots, X_{n-1}) = Q(X_1, \ldots, X_{n-1}, 0)$$

and it follows that $\mathrm{rank}(q_0) \geqslant n - 2$, since Q is non-singular. In particular $\mathrm{rank}(q_0) \geqslant 2$ if $n \geqslant 4$. Moreover, if q were reducible, then the quadratic form

$$R_h(X_0, \ldots, X_{n-1}) = Q(X_1, \ldots, X_{n-1}, hX_0) - mX_0^2$$

would have rank at most 2. But this is impossible when $n \geqslant 4$. Indeed, if $h = 0$, then $\mathrm{rank}(R_0) \geqslant n - 1 \geqslant 3$. Equally, if $h \neq 0$, then $\mathrm{rank}(R_h) \geqslant n - 1 \geqslant 3$,

since $\operatorname{rank}(A + B) \geqslant |\operatorname{rank}(A) - \operatorname{rank}(B)|$ for any $n \times n$ matrices with integer coefficients. The estimate (5.1) is now a trivial consequence of Theorem 1.11.

It remains to establish Proposition 4.4 when $n = 3$ and $f = X_3$. We may write

$$Q(X_1, X_2, X_3) = P(X_1, X_2) + X_3(cX_1 + dX_2) + eX_3^2, \qquad (5.2)$$

where $P(X_1, X_2) = Q(X_1, X_2, 0)$ and $c, d, e \in \mathbb{Z}$. Since Q is non-singular, it follows that P has rank 1 or 2. We will need to deal with each of these cases separately.

Suppose first that $\operatorname{rank}(P) = 2$. Then after a non-singular change of variables in X_1 and X_2 alone it suffices to proceed under the assumption that $c = d = 0$ in (5.2). Hence

$$U_\ell(H) \leqslant \# \left\{ \mathbf{x} \in \mathbb{Z}^3 : \begin{array}{l} |\mathbf{x}| \ll H, \ 0 \neq x_3 \equiv 0 \, (\operatorname{mod} \ell) \\ P(x_1, x_2) = m - ex_3^2 \end{array} \right\}.$$

Let $h \in \mathbb{Z}$ such that $h \ll H$ and $0 \neq h \equiv 0 \, (\operatorname{mod} \ell)$. There are $O(\ell^{-1}H)$ such integers. If $m - eh^2 \neq 0$ then an application of Lemma 4.1 shows that there are $O_\varepsilon(H^\varepsilon)$ choices for $x_1, x_2 \ll H$ such that $P(x_1, x_2) = m - eh^2$. If $m - eh^2 = 0$, which can happen for at most 2 values of h, we deduce that $0 \neq m \equiv 0 \, (\operatorname{mod} \ell)$. Hence $\ell = O(1)$ and there are $O(H)$ choices for $x_1, x_2 \ll H$ such that $P(x_1, x_2) = 0$. This case therefore contributes $O(\ell^{-1}H)$ overall, which shows that Proposition 4.4 holds when $n = 3$ and $\operatorname{rank}(P) = 2$.

Next we suppose that $\operatorname{rank}(P) = 1$ in (5.2), still with $n = 3$ and $f = X_3$. Then Q takes the shape

$$Q(X_1, X_2, X_3) = aL_1(X_1, X_2)^2 + X_3 L_2(X_1, X_2) + eX_3^2,$$

for $a, e \in \mathbb{Z}$ and linear forms $L_1, L_2 \in \mathbb{Z}[X_1, X_2]$. Moreover aL_1 and L_2 are non-zero and non-proportional, since Q is non-singular. After a change of variables we obtain

$$U_\ell(H) \leqslant \# \left\{ \mathbf{x} \in \mathbb{Z}^3 : \begin{array}{l} |\mathbf{x}| \ll H, \ 0 \neq x_3 \equiv 0 \, (\operatorname{mod} \ell) \\ x_1^2 + x_2 x_3 + cx_3^2 = d \end{array} \right\},$$

for suitable integers $c, d = O(1)$ with $d \neq 0$. If $x_3 = 0$ then $\ell = O(1)$ and there are clearly $O(H)$ choices for x_1, x_2. Hence there is a contribution of $O(\ell^{-1}H)$ to $U_\ell(H)$ from this case. Alternatively, the contribution from non-zero x_3 is at most

$$\sum_{\substack{0 < |x_3| \ll H \\ x_3 \equiv 0 \, (\operatorname{mod} \ell)}} \#\{x_1 \ll H : x_1^2 \equiv d \, (\operatorname{mod} x_3)\} \ll H \sum_{\substack{0 < |x_3| \ll H \\ x_3 \equiv 0 \, (\operatorname{mod} \ell)}} \frac{\nu(x_3; d)}{|x_3|},$$

where

$$\nu(q; d) = \#\{n \, (\operatorname{mod} q) : n^2 \equiv d \, (\operatorname{mod} q)\}.$$

On noting that $\nu(q;d)$ is a multiplicative function of q, with $\nu(2^j;d) \leqslant 4$ and $\nu(q;d) = \sum_{k|q} |\mu(k)|(\frac{d}{q})$ when q is odd, it follows that $\nu(q;d) \leqslant 2^{\omega(q)+1} \ll_\varepsilon q^{\varepsilon/2}$, where $\omega(q)$ denotes the number of distinct prime factors of q. Next we observe that

$$\sum_{\substack{0<|x_3|\ll H \\ x_3 \equiv 0 \,(\mathrm{mod}\,\ell)}} \frac{1}{|x_3|} \ll \ell^{-1} \log H.$$

Thus the non-zero x_3 contribute $O_\varepsilon(\ell^{-1}H^{1+\varepsilon})$ overall. The same bound therefore holds for $U_\ell(H)$ when $n = 3$ and $\mathrm{rank}(P) = 1$, which thereby concludes the proof of Proposition 4.4.

5.2. **Higher degree polynomials.** We now place ourselves in the setting of Propositions 4.5 and 4.6. Let $f \in \mathbb{Z}[X_1,\ldots,X_n]$ be a non-singular form of degree $d \geqslant 2$. If Δ_f is the discriminant of f, then Δ_f and $m\det(Q)$ are both non-zero integers. Let $\varepsilon > 0$ be given once and for all. For square-free $k \in \mathbb{N}$ such that $k^{2j} \leqslant H$, we want to estimate the quantity $U_{k^j}(H)$ in (4.4).

Let us put $\ell = k^j$ for convenience. Our estimation of $U_\ell(H)$ is inspired by an argument of Browning and Munshi [7, Lemma 4]. We will require some elementary facts about integer sublattices, as established by Davenport [12, Lemma 5]. Suppose that $\Lambda \subset \mathbb{Z}^n$ is a lattice of rank r and determinant $\det(\Lambda)$. Then there exists a "minimal" basis $\mathbf{m}^{(1)},\ldots,\mathbf{m}^{(r)}$ of Λ with the property

$$\lambda_j \ll |\mathbf{u}|/|\mathbf{m}^{(j)}|, \quad (1 \leqslant j \leqslant r),$$

whenever $\mathbf{u} \in \Lambda$ is written as $\mathbf{u} = \sum_{j=1}^r \lambda_j \mathbf{m}^{(j)}$. Furthermore, the basis is constructed in such a way that $1 \leqslant |\mathbf{m}^{(1)}| \leqslant \cdots \leqslant |\mathbf{m}^{(r)}|$ and

$$\det(\Lambda) \leqslant \prod_{j=1}^r |\mathbf{m}^{(j)}| \ll \det(\Lambda). \tag{5.3}$$

Breaking into residue classes modulo ℓ, we obtain

$$U_\ell(H) \leqslant \sum_{\substack{\boldsymbol{\xi} \,(\mathrm{mod}\,\ell) \\ Q(\boldsymbol{\xi}) \equiv m \,(\mathrm{mod}\,\ell) \\ f(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\,\ell)}} \#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\,\ell)\}. \tag{5.4}$$

We denote the set whose cardinality appears in the inner sum by $S_\ell(H;\boldsymbol{\xi})$. If $S_\ell(H;\boldsymbol{\xi})$ is empty then there is nothing to prove. Alternatively, suppose we are given $\mathbf{x}_0 \in S_\ell(H;\boldsymbol{\xi})$. Then any other vector in the set must be congruent to $\mathbf{x}_0$ modulo ℓ. Making the change of variables $\mathbf{x} = \mathbf{x}_0 + \ell\mathbf{y}$ in $S_\ell(H;\boldsymbol{\xi})$, we have $|\mathbf{y}| < 2\ell^{-1}H$. Furthermore,

$$\mathbf{y}.\nabla Q(\mathbf{x}_0) + \ell Q(\mathbf{y}) = 0, \tag{5.5}$$

by Taylor's formula, since $Q(\mathbf{x}_0 + \ell\mathbf{y}) = m$ and $Q(\mathbf{x}_0) = m$. Note here that $\nabla Q(\mathbf{x}_0) \neq \mathbf{0}$ for any $\mathbf{x}_0 \in Y(\mathbb{Z})$.

When $n \geqslant 4$, it will be convenient to deal separately with the contribution from $\mathbf{y}$ for which $\mathbf{y}.\nabla Q(\mathbf{x}_0) = 0$. Using this linear equation to eliminate one of the variables, we arrive at a quadratic form in $n-1$ variables. We claim that this quadratic form is non-singular. When $n \geqslant 4$ this automatically implies that it is also absolutely irreducible. To see the claim, suppose that $\mathbf{B}$ is the underlying symmetric matrix associated to Q, so that $\nabla Q(\mathbf{y}) = 2\mathbf{B}\mathbf{y}$ and $\nabla Q(\mathbf{x}_0) = 2\mathbf{B}\mathbf{x}_0$. Then if the quadric obtained from $\mathbf{Y}.\nabla Q(\mathbf{x}_0) = Q(\mathbf{Y}) = 0$ is singular, there must exist $\lambda, \mu \in \overline{\mathbb{Q}}$ and $\mathbf{y} \neq \mathbf{0}$, such that $(\lambda, \mu) \neq (0,0)$ and

$$Q(\mathbf{y}) = 0, \quad 2\mathbf{B}(\lambda\mathbf{y} + \mu\mathbf{x}_0) = \mathbf{0}.$$

Since $\mathbf{B}$ is non-singular, this implies that $\lambda\mathbf{y} + \mu\mathbf{x}_0 = \mathbf{0}$, which is impossible since $0 \neq m = Q(\mathbf{x}_0)$. Hence the claim follows and Theorem 1.11 implies that for $n \geqslant 4$ the overall contribution to $S_\ell(H; \boldsymbol{\xi})$ from $\mathbf{y}$ for which $\mathbf{y}.\nabla Q(\mathbf{x}_0) = 0$ is $\ll_\varepsilon (H/\ell)^{n-3+\varepsilon}$.

Returning to general $n \geqslant 3$, (5.5) implies that the $\mathbf{y}$ under consideration satisfy the congruence $\mathbf{y}.\nabla Q(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\,\ell)$, since $\mathbf{x}_0 \equiv \boldsymbol{\xi} \,(\mathrm{mod}\,\ell)$. Let us write

$$\Lambda_{\boldsymbol{\xi}} = \{\mathbf{y} \in \mathbb{Z}^n : \mathbf{y}.\nabla Q(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\,\ell)\}.$$

Our work so far has shown that

$$\#S_\ell(H; \boldsymbol{\xi}) \leqslant 2\#\{\mathbf{y} \in \Lambda_{\boldsymbol{\xi}} : |\mathbf{y}| < 2\ell^{-1}H, \ (5.5) \text{ holds}\},$$

when $n = 3$, and

$$\#S_\ell(H; \boldsymbol{\xi}) \ll_\varepsilon \left(\frac{H}{\ell}\right)^{n-3+\varepsilon} + \#\left\{\mathbf{y} \in \Lambda_{\boldsymbol{\xi}} : \begin{array}{l} |\mathbf{y}| < 2\ell^{-1}H, \ (5.5) \text{ holds} \\ \mathbf{y}.\nabla Q(\mathbf{x}_0) \neq 0 \end{array}\right\},$$

when $n \geqslant 4$.

The set $\Lambda_{\boldsymbol{\xi}}$ defines an integer lattice of rank n. To calculate its determinant we write $\hat{\ell} = \ell/\gcd(\ell, \nabla Q(\boldsymbol{\xi}))$ and note that $\hat{\ell}\mathbb{Z}^n \subset \Lambda_{\boldsymbol{\xi}}$. Thus we have

$$\det(\Lambda_{\boldsymbol{\xi}}) = [\mathbb{Z}^n : \Lambda_{\boldsymbol{\xi}}] = \frac{[\mathbb{Z}^n : \hat{\ell}\mathbb{Z}^n]}{[\Lambda_{\boldsymbol{\xi}} : \hat{\ell}\mathbb{Z}^n]}.$$

The numerator here is clearly $\hat{\ell}^n$ and the denominator is seen to be $\hat{\ell}^{n-1}$, since there are $\hat{\ell}^{n-1}$ distinct values of $\mathbf{y} \,(\mathrm{mod}\,\hat{\ell})$ for which $\mathbf{y}.\nabla Q(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\,\ell)$. Thus $\det(\Lambda_{\boldsymbol{\xi}}) = \hat{\ell}$. We claim that in fact $\ell \ll \det(\Lambda_{\boldsymbol{\xi}}) \leqslant \ell$, the upper bound being trivial. For the lower bound, note that $\boldsymbol{\xi}.\nabla Q(\boldsymbol{\xi}) = 2Q(\boldsymbol{\xi}) \equiv 2m \,(\mathrm{mod}\,\ell)$ in (5.4), whence $\gcd(\ell, \nabla Q(\boldsymbol{\xi})) \ll 1$.

Let $\mathbf{M}$ denote the non-singular matrix formed from taking column vectors to be a minimal basis $\mathbf{m}_1, \ldots, \mathbf{m}_n$ for $\Lambda_{\boldsymbol{\xi}}$. Making the change of variables

$\mathbf{y} = \mathbf{M}\boldsymbol{\lambda}$, we arrive at the equation $q(\boldsymbol{\lambda}) = 0$, where if $b_i = \ell^{-1}\mathbf{m}_i.\nabla Q(\mathbf{x}_0)$ for $i = 1, \ldots, n$, then

$$q(\boldsymbol{\lambda}) = Q(\lambda_1\mathbf{m}_1 + \cdots + \lambda_n\mathbf{m}_n) + b_1\lambda_1 + \cdots + b_n\lambda_n. \tag{5.6}$$

This is obtained from (5.5) by substitution and dividing through by ℓ. It is defined over $\mathbb{Z}$ and the quadratic homogeneous part q_0 has underlying matrix $\mathbf{M}^T\mathbf{B}\mathbf{M}$ of full rank $n \geqslant 3$, where $\mathbf{B}$ is the matrix associated to Q. Our argument now diverges according to whether $n \geqslant 4$ or $n = 3$.

The case $n \geqslant 4$. Recalling the definition of $S_\ell(H; \boldsymbol{\xi})$, our main aim in this section is to establish the following result, which may prove to be of independent interest.

Proposition 5.1. *Assume that $n \geqslant 4$ and let $\varepsilon > 0$. Then we have*

$$\#\{\mathbf{x} \in Y(\mathbb{Z}) : |\mathbf{x}| \leqslant H, \ \mathbf{x} \equiv \boldsymbol{\xi} \,(\mathrm{mod}\,\ell)\} \ll_\varepsilon \left(\frac{H}{\ell}\right)^{n-3+\varepsilon} \left(1 + \frac{H}{\ell^{n/(n-1)}}\right),$$

for any $\ell \in \mathbb{N}$ and any $\boldsymbol{\xi} \in (\mathbb{Z}/\ell\mathbb{Z})^n$, where the implied constant is uniform in ℓ and $\boldsymbol{\xi}$.

Before proving this result let us see how it suffices to complete the proof of Proposition 4.5. Taking $\ell = k^j$ for square-free $k \in \mathbb{N}$, we have $\ell^{n/(n-1)} \leqslant H$, by hypothesis. Thus the second term dominates the first term in Proposition 5.1. Substituting this into (5.4) and applying Lemma 4.2, we finally arrive at the bound

$$U_\ell(H) \ll_\varepsilon \frac{\varrho(\ell)H^{n-2+\varepsilon}}{\ell^{n-2+1/(n-1)+\varepsilon}} \leqslant \frac{C_j^{\omega(\ell)}H^{n-2+\varepsilon}}{\ell^{1/(n-1)+\varepsilon}},$$

for an appropriate constant $C_j > 0$ depending on j. Taking $C_j^{\omega(\ell)} = O_{\varepsilon,j}(\ell^\varepsilon)$, we therefore conclude the proof of Proposition 4.5.

Proof of Proposition 5.1. We employ the properties of the minimal basis that were recorded above. This leads to the inequality

$$\#S_\ell(H; \boldsymbol{\xi}) \ll_\varepsilon \left(\frac{H}{\ell}\right)^{n-3+\varepsilon} + \#\left\{\boldsymbol{\lambda} \in \mathbb{Z}^n : \begin{array}{l} \lambda_i \ll (|\mathbf{m}_i|\ell)^{-1}H \text{ for } 1 \leqslant i \leqslant n \\ q(\boldsymbol{\lambda}) = 0, \ b_1\lambda_1 + \cdots + b_n\lambda_n \neq 0 \end{array}\right\}.$$

We begin by considering $q(\boldsymbol{\lambda})$ when $\lambda_n = h$ is a fixed integer, with $h \neq 0$ when $b_1 = \cdots = b_{n-1} = 0$. Put

$$r(X_1, \ldots, X_{n-1}) = q(X_1, \ldots, X_{n-1}, h).$$

We claim that r is absolutely irreducible and that $\mathrm{rank}(r_0) \geqslant 2$. The latter follows on noting that $r_0(X_1, \ldots, X_{n-1}) = Q(X_1\mathbf{m}_1 + \cdots + X_{n-1}\mathbf{m}_{n-1})$, which must have rank at least $n - 2 \geqslant 2$. To check that r is absolutely irreducible

we consider the rank of the quadratic form $X_0^2 r(X_1/X_0, \ldots, X_{n-1}/X_0)$. The latter is equal to

$$Q(X_1\mathbf{m}_1 + \cdots + X_{n-1}\mathbf{m}_{n-1} + hX_0\mathbf{m}_n) + X_0 \left(b_1X_1 + \cdots + b_{n-1}X_{n-1} + b_nhX_0 \right).$$

If $h = 0$ then this has rank at least $n - 2 + 1 = n - 1 \geqslant 3$ since in this scenario $(b_1, \ldots, b_{n-1}) \neq (0, \ldots, 0)$. If $h \neq 0$ then it clearly has rank at least $n - 1 \geqslant 3$. Hence r is indeed absolutely irreducible.

Returning to our estimation of $\#S_\ell(H; \boldsymbol{\xi})$, suppose first that $|\mathbf{m}_n|\ell \gg H$, so that $\lambda_n = 0$ in any solution to be counted. It follows from the condition $b_1\lambda_1 + \cdots + b_n\lambda_n \neq 0$ that we may assume $(b_1, \ldots, b_{n-1}) \neq (0, \ldots, 0)$. Then $q(X_1, \ldots, X_{n-1}, 0)$ satisfies the hypotheses of Theorem 1.11 and we see that there is an overall contribution of $O_\varepsilon((H/\ell)^{n-3+\varepsilon})$ from this case, which is satisfactory.

We proceed under the assumption that $|\mathbf{m}_i|\ell \ll H$ for $1 \leqslant i \leqslant n$. We will fix a value of λ_n and then use Theorem 1.11 to estimate the associated number of $\lambda_1, \ldots, \lambda_{n-1}$. It follows from the condition $b_1\lambda_1 + \cdots + b_n\lambda_n \neq 0$ that when $b_1 = \cdots = b_{n-1} = 0$, any solution with $\lambda_n = 0$ is to be ignored. Let $\lambda_n = h$ be fixed and put $r(X_1, \ldots, X_{n-1}) = q(X_1, \ldots, X_{n-1}, h)$, as before. Then r satisfies the hypotheses of Theorem 1.11 and we deduce that the total number of $\lambda_1, \ldots, \lambda_{n-1}$ associated to h is

$$\ll_\varepsilon \left(\frac{H}{|\mathbf{m}_1|\ell} \right)^{n-3+\varepsilon}.$$

The implied constant in this estimate depends at most on ε and n and, crucially, is independent of h. Summing over h therefore leads to the overall conclusion that

$$\#S_\ell(H; \boldsymbol{\xi}) \ll_\varepsilon \left(\frac{H}{\ell} \right)^{n-3+\varepsilon} + \frac{H^{n-2+\varepsilon}}{\ell^{n-2+\varepsilon}|\mathbf{m}_1|^{n-3}|\mathbf{m}_n|}.$$

It follows from (5.3) that

$$|\mathbf{m}_1|^{n-3}|\mathbf{m}_n| \geqslant |\mathbf{m}_1|^{1/(n-1)}|\mathbf{m}_n| \geqslant (|\mathbf{m}_1| \ldots |\mathbf{m}_n|)^{1/(n-1)} \gg (\det \Lambda_{\boldsymbol{\xi}})^{1/(n-1)}$$
$$\gg \ell^{1/(n-1)}.$$

Taking this lower bound in our estimate for $\#S_\ell(H; \boldsymbol{\xi})$ concludes the proof. $\square$

The case $n = 3$. We henceforth take $n = 3$ and concern ourselves with the proof of Proposition 4.6. For $j \geqslant 2$ we are interested in estimating the quantity $U_{k^j}(H)$, when $k^{4j/3} \leqslant H$. Developing $U_{k^j}(H)$ as in (5.4), our starting point is the inequality

$$\#S_\ell(H; \boldsymbol{\xi}) \leqslant 2\#\{\boldsymbol{\lambda} \in \mathbb{Z}^3 : \lambda_i \ll (|\mathbf{m}_i|\ell)^{-1}H \text{ for } 1 \leqslant i \leqslant 3, \ q(\boldsymbol{\lambda}) = 0\},$$

where q is given by (5.6) and $\ell = k^j$. When $n = 3$ we have been unable to produce a satisfactory estimate for this quantity by first fixing λ_3 and analysing the resulting binary quadratic polynomial. The problem is that for certain choices of λ_3 it may happen that the resulting polynomial is reducible over $\mathbb{Q}$, which would then contribute too much to $\#S_\ell(H; \boldsymbol{\xi})$. Instead we will apply Theorem 1.11 directly to the ternary quadratic polynomial and attempt to show that $|\mathbf{m}_1|$ cannot often be very small.

Since q_0 has rank 3 it follows that q is absolutely irreducible. By (5.3) one has $|\mathbf{m}_1|^3 \ll \det(\Lambda_{\boldsymbol{\xi}}) \leqslant \ell$. Moreover, $\ell^{4/3} = k^{4j/3} \leqslant H$. Hence $H \gg |\mathbf{m}_1|\ell$ and so Theorem 1.11 implies that

$$\#S_\ell(H; \boldsymbol{\xi}) \ll_\varepsilon \left(\frac{H}{|\mathbf{m}_1|\ell} \right)^{1+\varepsilon}. \tag{5.7}$$

This estimate is sharpest when $|\mathbf{m}_1|^3$ has exact order ℓ. Let us put

$$\kappa = \frac{1}{3d},$$

for convenience. Returning to (5.4) we will denote by $U_\ell^{(I)}(H)$ the overall contribution to the right hand side from $\boldsymbol{\xi}$ for which the corresponding basis vector $\mathbf{m}_1 = \mathbf{m}_1(\boldsymbol{\xi})$ satisfies $|\mathbf{m}_1| \geqslant \ell^\kappa$. We denote by $U_\ell^{(II)}(H)$ the contribution from $\boldsymbol{\xi}$ for which the smallest basis vector satisfies $|\mathbf{m}_1| < \ell^\kappa$. The estimation of $U_\ell^{(I)}(H)$ is a straightforward consequence of (5.4), (5.7) and Lemma 4.2. Thus we deduce that

$$U_\ell^{(I)}(H) \ll_\varepsilon \frac{\varrho(\ell)H^{1+\varepsilon}}{\ell^{(1+\kappa)(1+\varepsilon)}} \leqslant \frac{C_j^{\omega(\ell)}H^{1+\varepsilon}}{\ell^{\kappa+\varepsilon}} \ll_{\varepsilon,j} \frac{H^{1+\varepsilon}}{\ell^\kappa},$$

for an appropriate constant $C_j > 0$ depending on j. In view of our choice of κ, this is clearly satisfactory for Proposition 4.6.

In order to estimate $U_\ell^{(II)}(H)$, we introduce averaging over the least non-zero elements of $\Lambda_{\boldsymbol{\xi}}$. Thus, given a non-zero vector $\mathbf{m} \in \mathbb{Z}^3$ satisfying $|\mathbf{m}| < \ell^\kappa$, we shall return to (5.4) and consider the overall contribution from $\boldsymbol{\xi}$ and $\mathbf{x}$ such that $\mathbf{m}.\nabla Q(\boldsymbol{\xi}) \equiv 0 \,(\mathrm{mod}\,\ell)$. Note that $\mathbf{m}.\nabla Q(\boldsymbol{\xi}) = (2\mathbf{Bm}).\boldsymbol{\xi}$, where $\mathbf{B}$ is the symmetric matrix associated to Q. Applying (5.7), we obtain

$$U_\ell^{(II)}(H) \ll_\varepsilon \sum_{\substack{\mathbf{m}\in\mathbb{Z}^3 \\ 0<|\mathbf{m}|<\ell^\kappa}} \varrho(\ell; 2\mathbf{Bm}) \left(\frac{H}{|\mathbf{m}|\ell} \right)^{1+\varepsilon}.$$

Invoking Lemma 4.3, we deduce the existence of a constant $C_j > 0$ such that

$$\varrho(\ell; 2\mathbf{Bm}) \leqslant C_j^{\omega(\ell)}\ell^{1-1/d} \gcd(\ell, 2\mathbf{Bm})^3.$$

Thus

$$U_\ell^{(II)}(H) \ll_\varepsilon \frac{C_j^{\omega(\ell)} H^{1+\varepsilon}}{\ell^{1/d+\varepsilon}} \sum_{\substack{\mathbf{m} \in \mathbb{Z}^3 \\ 0 < |\mathbf{m}| < \ell^\kappa}} \frac{\gcd(\ell, 2\mathbf{Bm})^3}{|\mathbf{m}|}.$$

The contribution to the inner sum from $\mathbf{m}$ such that $M < |\mathbf{m}| \leqslant 2M$ is at most

$$\sum_{h|\ell} \frac{h^3}{M} \#\{\mathbf{m} \in \mathbb{Z}^3 : M < |\mathbf{m}| \leqslant 2M, \ h \mid 2\mathbf{Bm}\} \ll \sum_{h|\ell} \frac{h^3}{M} \left(\frac{M}{h}\right)^3$$
$$\ll \tau(\ell) M^2.$$

Summing over dyadic intervals for $M < \ell^\kappa$, we now find that

$$U_\ell^{(II)}(H) \ll_\varepsilon \frac{\tau(\ell) C_j^{\omega(\ell)} H^{1+\varepsilon}}{\ell^{1/d+\varepsilon}} \sum_{M < \ell^\kappa} M^2 \ll_{\varepsilon,j} \ell^{2\kappa - 1/d} H^{1+\varepsilon}.$$

Recalling our choice of κ, this is also satisfactory for Proposition 4.6 and so completes its proof.

References

[1] R.C. Baker, The values of a quadratic form at square-free points. *Acta Arith.* **124** (2006), 101–137.

[2] Y. Benoist and H. Oh, Effective equidistribution of S-integral points on symmetric varieties. *Ann. Inst. Fourier* **62** (2012), 1889–1942.

[3] M. Berger, Les espaces symétriques noncompacts. *Ann. Sci. École Norm. Sup.* **74** (1957), 85–177.

[4] A. Borel and Harish-Chandra, Arithmetic subgroups of algebraic groups. *Ann. Math.* **75** (1962), 485–535.

[5] M. Borovoi, On representations of integers by indefinite ternary quadratic forms. *J. Number Theory* **90** (2001) 281–293.

[6] M. Borovoi and Z. Rudnick, Hardy–Littlewood varieties and semisimple groups. *Invent. Math.* **119** (1995) 37–66.

[7] T.D. Browning and R. Munshi, Rational points on singular intersections of quadrics. *Compositio Math.* **149** (2013), 1457–1494.

[8] T.D. Browning, D.R. Heath-Brown and P. Salberger, Counting rational points on algebraic varieties. *Duke Math. J.* **132** (2006), 545–578.

[9] M. Burger and P. Sarnak, Ramanujan duals. II. *Invent. Math.* **106** (1991), 1–11.

[10] L. Clozel, Démonstration de la conjecture τ. *Invent. Math.* **151** (2003), 297–328.

[11] J.-L. Colliot-Thélène and F. Xu, Brauer–Manin obstruction for integral points of homogeneous spaces and representation by integral quadratic forms, *Compositio Math.* **145** (2009), 309–363.

[12] H. Davenport, Cubic forms in 16 variables. *Proc. Roy. Soc. A* **272** (1963), 285–303.

[13] W. Duke, Z. Rudnick and P. Sarnak, Density of integer points on affine homogeneous varieties. *Duke Math. J.* **71** (1993), 143–179.

[14] P. Erdős, Arithmetical properties of polynomials. *J. London Math. Soc.* **28** (1953), 416–425.

[15] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.* **71** (1993), 181–209.

[16] A. Eskin, S. Mozes, and N. Shah, Unipotent flows and counting lattice points on homogeneous varieties. *Ann. Math.* **143** (1996), 149–159.

[17] W. Fulton and R. Lazarsfeld, Connectivity and its applications in algebraic geometry. *Algebraic geometry (Chicago, Ill., 1980)*, 26–92, Lecture Notes in Math. **862**, Springer, 1981.

[18] D.R. Heath-Brown, The density of rational points on curves and surfaces. *Ann. Math.* **155** (2002), 553–595.

[19] A. Gorodnik, H. Oh and N. Shah, Integral points on symmetric varieties and Satake compactifications. *Amer. J. Math.* **131** (2009), 1–57.

[20] D. Kleinbock and G. Margulis, Logarithm laws for flows on homogeneous spaces. *Invent. Math.* **138** (1999), 451–494.

[21] A. Nevo and P. Sarnak, Prime and almost prime integral points on principal homogeneous spaces. *Acta Math.* **205** (2010), 361–402.

[22] J. Pila, Density of integral and rational points on varieties. *Astérisque* **228** (1995), 183–187.

[23] B. Poonen, Squarefree values of multivariable polynomials. *Duke Math. J.* **118** (2003), 189–373.

[24] H. Schlichtkrull, *Hyperfunctions and harmonic analysis on symmetric spaces*. Progress in Mathematics **49**, Birkhaüser Boston, Boston, MA, 1984.

[25] R. Steinberg, *Endomorphisms of linear algebraic groups*, Memoirs of the American Mathematical Society **80**, American Mathematical Society, Providence, R.I., 1968.

[26] C.L. Stewart, On the number of solutions of polynomial congruences and Thue equations. *J. Amer. Math. Soc.* **4** (1991), 793–835.

[27] H.P.F. Swinnerton-Dyer, Rational zeros of two quadratic forms. *Acta Arith.* **9** (1964), 261–270.

School of Mathematics, University of Bristol, Bristol, BS8 1TW, UK
E-mail address: t.d.browning@bristol.ac.uk, a.gorodnik@bristol.ac.uk