Rational points on finite covers of $\mathbb{P}^1$ and $\mathbb{P}^2$

Niklas Broberg

Department of Mathematics,
Chalmers University of Technology and Göteborg University,
SE-412 96 Göteborg, Sweden

1 Introduction

Let $\mathbb{P}^n$ denote the n–dimensional projective space over the rational numbers. For any morphism $f : X \to \mathbb{P}^n$ from a scheme X, one can define a counting function

$$N(f, B) = \#\left\{\mathbf{x} \in f(X(\mathbb{Q})) \ : \ H(\mathbf{x}) \leq B\right\},$$

where $H : \mathbb{P}^n(\mathbb{Q}) \to \mathbb{R}_{>0}$ is the standard multiplicative height on $\mathbb{P}^n(\mathbb{Q})$. If X is integral and $f : X \to \mathbb{P}^n$ is finite dominant of degree at least 2, then

$$N(f, B) = O_{f,\varepsilon}(B^{n+1/2+\varepsilon}),$$

for every $\varepsilon > 0$. The proof is by sieve-methods. See chapter 13 of [9]. In the same book Serre conjectures that

$$N(f, B) = O_{f,\varepsilon}(B^{n+\varepsilon})$$

for such morphisms $f : X \to \mathbb{P}^n$. We intend to prove the following results.

Theorem 1. *If X is integral and $f : X \to \mathbb{P}^1$ is finite dominant of degree d, then*

$$N(f, B) = O_{f,\varepsilon}(B^{2/d+\varepsilon})$$

for every $\varepsilon > 0$.

Theorem 2. *If X is integral and $f : X \to \mathbb{P}^2$ is finite dominant of degree at least 3, then*

$$N(f, B) = O_{f,\varepsilon}(B^{2+\varepsilon}).$$

If $f : X \to \mathbb{P}^2$ is finite dominant of degree 2, then

$$N(f, B) = O_{f,\varepsilon}(B^{9/4+\varepsilon}).$$

The paper is structured as follows.

In section 2 we fix the notation and state some preliminary results.

In section 3 we obtain estimates of $N(f,B)$ for a special kind of covers $f: X \to \mathbb{P}^1$. The result is formulated in terms of polynomials and the estimates depend explicitly on the coefficients of the polynomials. The proof is by means of a method due to Heath-Brown [6]. As a corollary we obtain theorem 1.

In section 4 we prove theorem 2. The idea of the proof is simple. First we reduce to a case where the cover $f: X \to \mathbb{P}^2$ is given by a single equation. We then choose $O(B^{3/2})$ lines $H_1, \ldots, H_k \subset \mathbb{P}^2$ such that every rational point in the plane of height at most B is contained in $H_i(\mathbb{Q})$ for some i. By using the explicit estimates of $N(f|_{f^{-1}H_i}, B)$ from section 3 we are able to estimate the right-hand side of

$$N(f,B) \leq \sum_{i=1}^{k} N(f|_{f^{-1}H_i}, B).$$

Acknowledgement. The author would like to thank Per Salberger for valuable discussions during the course of this work. In particular, the idea of using Hilbert schemes is due to him.

2 Preliminaries

We shall use the following notations throughout the paper.

- $\ll$ is used synonymously with the big Oh notation. If f, g are real-valued functions with g positive, then $f = O_{p_1,\ldots,p_k}(g)$ and $f \ll_{p_1,\ldots,p_k} g$ means that $|f| \leq cg$ for some positive constant c which depends on the parameters $p_1, \ldots, p_k$.
- $|\mathbf{x}|$ is the Euclidean length of $\mathbf{x} \in \mathbb{R}^n$.
- $\langle \mathbf{x}, \mathbf{y} \rangle$ is the standard inner product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
- $\|F\|$ is the maximum modulus of the coefficients of $F \in \mathbb{R}[\mathbf{x}]$.
- P^n is the set of all non-zero $\mathbf{x} \in \mathbb{Z}^{n+1}$ for which $\gcd(x_0, \ldots, x_n) = 1$.
- $P^n(\mathbf{B})$ is the set of all $\mathbf{x} \in P^n$ for which $|x_i| \leq B_i$ for $i = 0, 1, \ldots, n$. The components of $\mathbf{B} \in \mathbb{R}^{n+1}$ are supposed to satisfy $B_i \geq 1$.

$P^n(B)$ is short-hand for $P^n(\mathbf{B})$ in case all the components of $\mathbf{B}$ are equal to B.

$\mathbb{Z}_w[t, \mathbf{x}]$ is the set of polynomials $F \in \mathbb{Z}[t, \mathbf{x}]$ which are monic in the variable t and such that $F(t^w, \mathbf{x})$ are homogeneous. Here w is a positive integer.

$N(F; \mathbf{B})$ is a counting function assigned to $F \in \mathbb{Z}_w[t, \mathbf{x}]$. It counts the number of $\mathbf{x} \in P^n(\mathbf{B})$ for which $F(t, \mathbf{x}) = 0$ has a rational solution.

$N(F, B)$ is short-hand for $N(F; \mathbf{B})$ in case all the components of $\mathbf{B}$ are equal to B.

We will require two auxiliary results.

Lemma 1. *Let $f : X \to \mathbb{P}^n$ be a finite dominant morphism of degree d, where X is integral. Then there exist a proper closed subset $V \subset X$ and an irreducible polynomial $F \in \mathbb{Z}_w[t, \mathbf{x}]$ of degree d in t such that*

$$N(f, B) \leq N(f|_V, B) + N(F, B).$$

Proof. Let $U \subset X$ be the affine variety $f^{-1}\mathbb{A}^n$, where $\mathbb{A}^n \subset \mathbb{P}^n$ is given by $x_0 \neq 0$, and let $y_i = x_i/x_0$ for $i = 1, 2, \ldots, n$. Let t be a generator of the function field $K(U)$ over $K(\mathbb{A}^n) = \mathbb{Q}(y_1, \ldots, y_n)$ such that the coefficients of its minimal polynomial

$$p(t, \mathbf{y}) = t^d + p_1(\mathbf{y})t^{d-1} + \cdots + p_d(\mathbf{y})$$

are polynomials with integer coefficients. Let $Y \subset \mathbb{A}^{n+1}$ be the variety defined by $p(t, \mathbf{y}) = 0$ and $h : Y \to \mathbb{A}^n$ the projection $H(t, \mathbf{y}) = \mathbf{y}$. Finally, let $V_0 \subset U$ be a proper closed subset such that $f : U \setminus V_0 \to \mathbb{A}^n$ factors through $h : Y \to \mathbb{A}^n$. Such a set exists since the isomorphism between the function fields $K(Y)$ and $K(U)$ induces a birational map $U \dashrightarrow Y$.

We now define $V \subset X$ by $V = V_0 \cup (X \setminus U)$ and $F \in \mathbb{Z}_w[t, \mathbf{x}]$ by

$$F(t, x_0, \ldots, x_n) = x_0^{dw} p(x_0^{-w} t, x_0^{-1} x_1, \ldots, x_0^{-1} x_n),$$

where w is the smallest positive integer for which $\deg p_i \leq iw$ for all $i = 1, 2, \ldots, d$. Then $f : X \setminus V \to \mathbb{P}^n$ factors through $h : Y \to \mathbb{A}^n \subset \mathbb{P}^n$ and $F(x_0^w t, \mathbf{x}) = 0$ whenever $p(t, x_0^{-1} x_1, \ldots, x_0^{-1} x_n) = 0$. Hence,

$$N(f, B) \leq N(f|_V, B) + N(h, B) \leq N(f|_V, B) + N(F, B).$$

Note that F is irreducible since p is. $\square$

Lemma 2. *Let $f : X \to \mathbb{P}^n$ be a finite morphism over a field K of characteristic 0 and let C be a closed subscheme of X. Assume that the invertible sheaf $f^*\mathcal{O}(m)$ on X is very ample for some $m > 0$. Then $f : C \to \mathbb{P}^n$ is a closed immersion onto a k–dimensional linear subvariety of $\mathbb{P}^n$ if and only if the Hilbert polynomial of C with respect to $f^*\mathcal{O}(m)$ is equal to*

$$P(q) = \binom{mq+k}{mq}.$$

Proof. Recall that the Hilbert polynomial of C with respect to $f^*\mathcal{O}(m)$ is defined by

$$P_C(q) = \chi(i^*f^*\mathcal{O}(mq)) = \sum (-1)^i \dim_K H^i(C, i^*f^*\mathcal{O}(mq)),$$

where $i : C \to X$ is the embedding. Analogously, the Hilbert polynomial $P_D(q)$ of the scheme-theoretic image $D = f(C)$ with respect to $\mathcal{O}(1)$ is given by $P_D(q) = \chi(\mathcal{O}_D(q))$. The restriction $g : C \to D$ of f to C is affine, so

$$H^i(C, i^*f^*\mathcal{O}(q)) = H^i(C, g^*\mathcal{O}_D(q)) = H^i(D, g_*\mathcal{O}_C \otimes \mathcal{O}_D(q))$$

for every $q \geq 0$. Consider the exact sequence

$$0 \to \mathcal{O}_D \xrightarrow{g^\#} g_*\mathcal{O}_C \to \mathcal{F} \to 0 \tag{1}$$

of coherent sheaves on D. The map $g^\#$ is injective by the definition of scheme-theoretic image, and $\mathcal{F}$ is coherent since $g_*\mathcal{O}_C$ is. If we twist (1) by q and compare the Euler characteristics we get

$$\chi(g_*\mathcal{O}_C \otimes \mathcal{O}_D(q)) = \chi(\mathcal{O}_D(q)) + \chi(\mathcal{F}(q)).$$

Now $\mathcal{O}_D(1)$ is ample and $\mathcal{F}$ coherent so $\chi(\mathcal{F}(mq)) = 0$ for every $q \geq 0$ only if $\mathcal{F} = 0$. Consequently, $g : C \to D$ is an isomorphism if and only if $P_C(q) = P_D(mq)$ for every $q \geq 0$. It is well-known that

$$P_D(q) = \binom{q+k}{q}$$

exactly when $D \subset \mathbb{P}^n$ is a k–dimensional linear variety. ☐

3 Covers of $\mathbb{P}^1$

The main result of this section is the following:

Theorem 3. *Let $F \in \mathbb{Z}_w[t, x, y]$ be irreducible of degree d in t. Then*

$$N(F; A, B) \ll_{d,w,\varepsilon} (AB)^{1/d+\varepsilon} \|F\|^{\varepsilon},$$

for every $\varepsilon > 0$.

We get theorem 1 as a corollary by noting that a proper closed subset of a curve is finite and then refer to lemma 1. The rest of this section is devoted to the proof of theorem 3. It is similar to the proof of theorem 3 of Heath-Brown [6].

Let $\Delta_F \in \mathbb{Z}[x, y]$ be the discriminant of F and let $N_0(F; A, B)$ be the number of $(x, y) \in P^1(A, B)$ for which $F(t, x, y) = 0$ has a rational solution and $y\Delta_F(x, y) \neq 0$. Since F is irreducible, Δ_F is a non-trivial form of degree $d(d-1)w$. There are thus $O_{d,w}(1)$ points $(x, y) \in P^1$ for which $y\Delta_F(x, y) = 0$. Hence, it is sufficient to prove theorem 3 with $N(F; A, B)$ replaced by $N_0(F; A, B)$. Following [6], we define $N_p(F; A, B)$ to be the number of $(x, y) \in P^1(A, B)$ for which $F(t, x, y) = 0$ has a rational solution and $p \nmid y\Delta_F(x, y)$.

Lemma 3. *Suppose that P satisfies*

$$P \geq \log^2 \max_{\substack{(x,y)\in P^1(A,B) \\ y\Delta_F(x,y)\neq 0}} |y\Delta_F(x, y)| .$$

Then there exist distinct primes $p_1, \ldots, p_r$ such that

- $P \ll p_i \ll P$,
- $r \ll_{d,w} \log(AB \|F\|)$,
- $N_0(F; A, B) \leq \sum_{i=1}^{r} N_p(F; A, B)$.

Proof. See lemma 4 of [6]. □

Theorem 3 is proved by combining lemma 3 and the following result.

Lemma 4. *Assume that*

$$F(t, x, y) = t^d + F_1(x, y)t^{d-1} + \cdots + F_d(x, y),$$

where $F_i \in \mathbb{Z}[x, y]$ and $\deg F_i = iw$. If

$$p \geq 4(2dw \|F_d\|)^{\varepsilon}(AB)^{1/d+\varepsilon}, \tag{2}$$

then $N_p(F; A, B) = O_{d,w,\varepsilon}(p)$.

Proof. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be the points counted by $N_p(F; A, B)$ for some p satisfying (2). If $t \in \mathbb{Z}$ satisfies $F(t, x, y) = 0$ for some $(x, y) \in P^1$, then $t \mid F_d(x, y)$. Moreover, if $F_d(x, y) = 0$, then $F(0, x, y) = 0$. Hence, we can choose integers $t_1, \ldots, t_n$ such that $F(t_i, x_i, y_i) = 0$ and

$$|t_i| \leq |F_d(x_i, y_i)| \leq 2dw(AB)^{dw} \|F_d\| . \tag{3}$$

Let $f(u,v) = F(u,v,1)$ and $(u_i,v_i) = (y_i^{-w}t_i, y_i^{-1}x_i)$ for $i = 1,2,\ldots,n$. Then $f(u_i,v_i) = 0$, and $(u_i,v_i) \in \mathbb{Z}_p^2$ by the assumption $p \nmid y_i$.

There are $O_{dw}(p)$ solutions of $f(u,v) = 0$ modulo p. In order to establish the estimate $n = O_{d,w,\varepsilon}(p)$ it is thus sufficient to show that there are $O_{d,w,\varepsilon}(1)$ points among $(u_1,v_1),\ldots,(u_n,v_n)$ which belong to a general class modulo p. Assume that

$$\begin{cases} (u_1,v_1) \equiv (u_i,v_i) \pmod{p}, & 1 \le i \le k \\ (u_1,v_1) \not\equiv (u_i,v_i) \pmod{p}, & k < i \le n. \end{cases}$$

We claim that there exists a polynomial g of degree $D = O_{d,w,\varepsilon}(1)$ such that $f \nmid g$ but $g(u_i,v_i) = 0$ for $i = 1,2,\ldots,k$. According to Bezout's theorem, we then have $k = O_{d,w,\varepsilon}(1)$.

Let $(a_1,b_1),\ldots,(a_e,b_e)$ be an enumeration of the set

$$\{0,1,\ldots,d-1\} \times \{0,1,\ldots,D-1\},$$

where D is the smallest integer satisfying

$$D \ge 2\max(1, dw/\varepsilon).$$

A non-trivial polynomial of the shape

$$g(u,v) = \sum_{j=1}^{e} g_j u^{a_j} v^{b_j}$$

is not divisible by f (g has degree at most $d-1$ in the variable u while f has degree d). Moreover, if the matrix

$$M = [u_i^{a_j} v_i^{b_j}]_{\substack{1 \le i \le k \\ 1 \le j \le e}}$$

has rank less than e, then there exist $g_1,\ldots,g_e \in \mathbb{Z}$, which are not all equal to 0, such that $g(u_i,v_i) = 0$ for $i = 1,2,\ldots,k$. This is obviously the case if $k < e$ so assume that $k \ge e$. Then $\operatorname{rank} M < e$ if and only if all the $e \times e$–minors of M vanishes. Without loss of generality we may consider

$$\Delta = \det[u_i^{a_j} v_i^{b_j}]_{\substack{1 \le i \le e \\ 1 \le j \le e}}.$$

The assumption $p \nmid \Delta_F(x_1,y_1)$ implies that $\frac{\partial f}{\partial u}(u_1,v_1) \not\equiv 0 \pmod{p}$. We can thus apply the lifting argument from the proof of Hensel's lemma to construct a polynomial $h \in \mathbb{Z}_p[v]$ such that $u_i \equiv h(v_i) \pmod{p^{e^2}}$ for $i = 1,2,\ldots,k$ (see lemma 5 of [6]). Let $z_i = v_i - v_1 \in p\mathbb{Z}_p$ and

$$r_j(z) = h(v_1+z)^{a_j}(v_1+z)^{b_j} \in \mathbb{Z}_p[z].$$

Then

$$\Delta \equiv \det[r_j(z_i)] \pmod{p^{e^2}},$$

and we can act on $[r_j(z_i)]$ by elementary column operations over $\mathbb{Z}_p$ to obtain a new matrix $[z_i^{j-1} s_j(z_i)]$ for some $s_j \in \mathbb{Z}_p[z]$. Hence, if $z_i = pw_i$, then

$$\Delta \equiv \det[p^{j-1} w_i^{j-1} s_j(z_j)] \equiv p^{e(e-1)/2} \det[w_i^{j-1} s_j(z_j)] \pmod{p^{e^2}}.$$

This shows that $p^{e(e-1)/2}$ divides the integer

$$\Delta' = \left(\prod_{i=1}^{e} y_i^{(D-1)+w(d-1)} \right) \Delta = \det[t_i^{a_j} x_i^{b_j} y_i^{c_j}],$$

where

$$c_j = D - 1 - b_j + w(d - 1 - a_j).$$

By expanding Δ' we find that

$$|\Delta'| \le e^e A^{e_x} B^{e_y} C^{e_t},$$

where C is the right-hand side of (3), and

$$\begin{aligned} e_x &= \sum_{j=1}^{e} b_j = \frac{dD(D-1)}{2} \\ e_y &= \sum_{j=1}^{e} c_j = \frac{dD(D + w(d-1) - 1)}{2} \\ e_t &= \sum_{j=1}^{e} a_j = \frac{d(d-1)D}{2}. \end{aligned}$$

Thus, Δ vanishes if

$$p^{e(e-1)/2} > e^e A^{e_x} B^{e_y} C^{e_t} = e^e (2dw \, \|F_d\|)^{e_t} A^{e_x + dwe_t} B^{e_y + dwe_t}.$$

One can check that $e^{2/(e-1)} \le 4$ and

$$\begin{aligned} \frac{2(e_x + dwe_t)}{e(e-1)} &= \frac{1}{d} + \frac{wd^3 - wd^2 - d + 1}{d(dD-1)} \le \frac{1}{d} + \varepsilon, \\ \frac{2(e_y + dwe_t)}{e(e-1)} &= \frac{1}{d} + \frac{wd^3 - wd - d + 1}{d(dD-1)} \le \frac{1}{d} + \varepsilon, \\ \frac{2e_t}{e(e-1)} &= \frac{d-1}{dD-1} \le \varepsilon. \end{aligned}$$

Hence, Δ vanishes due to (2). □

4 Covers of $\mathbb{P}^2$

Suppose that X is integral and $f : X \to \mathbb{P}^2$ is finite dominant of degree d. By lemma 1 there exist a proper closed subset $V \subset X$ and an irreducible polynomial $F \in \mathbb{Z}_w[t, x_0, x_1, x_2]$ of degree d in t such that

$$N(f, B) \leq N(f|_V, B) + N(F, B).$$

Obviously, $N(f|_V, B)$ is dominated by the number of rational points of height at most B on $f(V) \subset \mathbb{P}^2$. There are $O_{f(V)}(B^2)$ such points (see theorem 1 in [6]). Theorem 2 thus follows from:

Theorem 4. *Let $F \in \mathbb{Z}_w[t, x_0, x_1, x_2]$ be irreducible of degree d in the variable t. Then $N(F, B) = O_{F,\varepsilon}(B^{9/4+\varepsilon})$ if $d = 2$, and $N(F, B) = O_{F,\varepsilon}(B^{2+\varepsilon})$ if $d \geq 3$.*

The rest of this section is concerned with the proof of this result.

Siegel's lemma states that there exists a constant c such that the equation $\langle \mathbf{x}, \mathbf{u} \rangle = 0$ has a solution $\mathbf{u} \in P^2(cB^{1/2})$ for every $\mathbf{x} \in P^2(B)$ (see lemma 1 in [6]). Hence,

$$N(F, B) \leq \sum_{\mathbf{u} \in P^2(cB^{1/2})} N_{\mathbf{u}}(F, B), \tag{4}$$

where $N_{\mathbf{u}}(F, B)$ is the number of $\mathbf{x} \in P^2(B)$ for which $F(t, \mathbf{x}) = 0$ has a rational solution and $\langle \mathbf{x}, \mathbf{u} \rangle = 0$. Let $\mathbf{x}_0$, $\mathbf{x}_1 \in P^2$ be a basis of the lattice $\langle \mathbf{x}, \mathbf{u} \rangle = 0$ for a given $\mathbf{u} \in P^2$. Then

$$F(t, y_0\mathbf{x}_0 + y_1\mathbf{x}_1) = G_1(t, \mathbf{y}) \cdots G_k(t, \mathbf{y}) \tag{5}$$

for some irreducible $G_i \in \mathbb{Z}_w[t, \mathbf{y}]$ of corresponding degrees d_i in the variable t. Note that $G_i(t^w, \mathbf{y})$ is homogeneous since a divisor of a homogeneous polynomial is homogeneous. The numbers $d_1, \ldots, d_k$ are independent of the choice of the basis so it makes sense to use the notation

$$\delta_{\mathbf{u}} F = \min_{1 \leq i \leq k} d_i.$$

If we define

$$N^{(\delta)}(F, B) = \sum_{\substack{\mathbf{u} \in P^2(cB^{1/2}) \\ \delta_{\mathbf{u}} F = \delta}} N_{\mathbf{u}}(F, B)$$

for $\delta \leq d$, we have

$$N(F, B) \leq \sum_{\delta=1}^{d} N^{(\delta)}(F, B)$$

from (4). We will prove theorem 4 by finding estimates of $N^{(\delta)}(F,B)$ for various δ. The following lemma will be essential in every case.

Lemma 5. *If* $|\mathbf{u}| \ll B$, *then*

$$N_{\mathbf{u}}(F,B) \ll_{d,w,\varepsilon} B^{2/\delta_{\mathbf{u}}F+\varepsilon} |\mathbf{u}|^{-1/\delta_{\mathbf{u}}F} \|F\|^{\varepsilon}.$$

Proof. Lemma 1 in [6] states that we can find a basis $\mathbf{x}_0, \mathbf{x}_1 \in P^2$ of the lattice $\langle \mathbf{x}, \mathbf{u}\rangle = 0$ such that

(i) $|y_i| \ll |\mathbf{x}| / |\mathbf{x}_i|$ whenever $\mathbf{x} = y_0\mathbf{x}_0 + y_1\mathbf{x}_1$,

(ii) $|\mathbf{u}| \ll |\mathbf{x}_0|\,|\mathbf{x}_1| \ll |\mathbf{u}|$.

Let $G_1, \dots G_k \in \mathbb{Z}_w[t, \mathbf{y}]$ be as in (5). By (i) we have

$$N_{\mathbf{u}}(F,B) \le \sum_{i=1}^{k} N\left(G_i, \frac{c'B}{|\mathbf{x}_0|}, \frac{c'B}{|\mathbf{x}_1|}\right),$$

for some constant c'. Since (ii) and $|\mathbf{u}| \ll B$ implies that $|\mathbf{x}_i| \ll B$, we may assume that $c'B/|\mathbf{x}_i| \ge 1$. Theorem 3 then gives

$$N\left(G_i, \frac{c'B}{|\mathbf{x}_0|}, \frac{c'B}{|\mathbf{x}_1|}\right) \ll_{d,w,\varepsilon} \left(\frac{B^2}{|\mathbf{x}_0|\,|\mathbf{x}_1|}\right)^{1/\delta+\varepsilon} \|G_i\|^{\varepsilon} \ll B^{2/\delta+\varepsilon} |\mathbf{u}|^{-1/\delta} \|F\|^{\varepsilon},$$

where $\delta = \delta_{\mathbf{u}}F$. The second inequality follows from (ii) and

$$\|G_1\| \cdots \|G_k\| = \|G_1 \cdots G_k\| \ll_{dw} |\mathbf{u}|^{dw} \|F\| \ll B^{dw} \|F\|. \tag{6}$$

The equality on the left of (6) is Gauss' lemma (see chapter I, proposition 2.1 in [8]). □

The following lemma gives $N^{(\delta)}(F,B) = O_{F,\varepsilon}(B^{2+\varepsilon})$, provided that $\delta \ge 3$.

Lemma 6. *For every* $\delta \le d$, *we have*

$$N^{(\delta)}(F,B) \ll_{d,w,\varepsilon} B^{3/2(1+1/\delta)+\varepsilon} \|F\|^{\varepsilon}.$$

Proof. Suppose that $T \le |\mathbf{u}| < 2T$, where $1 \le T \ll B^{1/2}$. Then

$$N_{\mathbf{u}}(F,B) \ll_{d,w,\varepsilon} B^{2/\delta+\varepsilon} T^{-1/\delta} \|F\|^{\varepsilon},$$

by lemma 5. There are $O(T^3)$ elements of $P^2(T)$, so

$$\sum_{\substack{T \le |\mathbf{u}| < 2T \\ \delta_{\mathbf{u}}F = \delta}} N_{\mathbf{u}}(F,B) \ll_{d,w,e} B^{2/\delta+\varepsilon} T^{3-1/\delta} \|F\|^{\varepsilon} \ll B^{3/2(1+1/\delta)} \|F\|^{\varepsilon}. \tag{7}$$

We finish the proof by dividing $1 \le T \le cB^{1/2}$ into dyadic intervals and summing the corresponding bounds (7). □

We will need some arguments from algebraic geometry in order to find sufficient estimates of $N^{(1)}(F,B)$ and $N^{(2)}(F,B)$. We can associate a geometric object to F as follows:

Let $R_i = \mathbb{Q}[t_i, \mathbf{x}_i]/\langle F(t_i, \mathbf{x}_i)\rangle$ for $i = 0,1,2$, where $\mathbf{x}_i = (x_{i0}, x_{i1}, x_{i2})$ and $x_{ii} = 1$ at all times. Then R_i are domains so the affine schemes $X_i = \operatorname{Spec} R_i$ are integral. For each pair (i,j), the homomorphisms

$$\phi_{ij} : (R_i)_{x_{ij}} \to (R_j)_{x_{ji}}, \quad (t_i, \mathbf{x}_i) \mapsto (x_{ji}^{-w} t_j, x_{ji}^{-1}\mathbf{x}_j),$$

are well-defined because of the relations

$$F(\phi_{ij}(t_i, \mathbf{x}_i)) = x_{ji}^{-w} F(t_j, \mathbf{x}_j).$$

One can check that $\phi_{jk} \circ \phi_{ij} = \phi_{ik}$ whenever the equality makes sense. Hence, we can construct an integral scheme X by gluing X_0, X_1, X_2 along the open sets $\operatorname{Spec}(R_i)_{x_{ij}}$. If we treat $\mathbb{P}^2$ similarly as $\operatorname{Spec}\mathbb{Q}[\mathbf{x}_i]$ glued along $\operatorname{Spec}\mathbb{Q}[\mathbf{x}_i, x_{ij}^{-1}]$, we see that the natural homomorphisms $\mathbb{Q}[\mathbf{x}_i] \to R_i$ are well matched and define a finite morphism $f : X \to \mathbb{P}^2$. The invertible sheaf $f^*\mathcal{O}(1)$ on X is ample (see proposition I.4.4 in [4]). Hence, X is a projective variety over $\mathbb{Q}$.

Lemma 7. *If F is absolutely irreducible and $d > 2$, then*

$$N^{(2)}(F,B) \ll_{F,\varepsilon} B^{2+\varepsilon}.$$

Proof. Let $H_{\mathbf{u}} \subset \mathbb{P}^2$ be the line defined by $\langle \mathbf{x}, \mathbf{u}\rangle = 0$. It is clear from the construction of X that there is a one-to-one correspondence between the different irreducible factors G_i in (5) and the reduced irreducible components of $f^{-1}H_{\mathbf{u}} \subset X$. If F is absolutely irreducible, then $\overline{X} = X \times \operatorname{Spec}\overline{\mathbb{Q}}$ is a projective variety of dimension 2 over an algebraically closed field. In that case it is well-known that $f^{-1}H \subset \overline{X}$ is irreducible for a generic line $H \subset \overline{\mathbb{P}^2}$ (see the proof of proposition 18.10 in [3], or corollary 10.9 and the following remarks in [5], chapter III). Hence, there exists a non-trivial form $G \in \overline{\mathbb{Q}}[\mathbf{u}]$ such that $G(\mathbf{u}) = 0$ whenever $\delta_{\mathbf{u}}F < d$. By theorem 1 in [6], there are $O_G(B)$ points $\mathbf{u} \in P^2(cB^{1/2})$ for which $G(\mathbf{u}) = 0$. By lemma 5, $N_{\mathbf{u}}(F,B) = O_{F,\varepsilon}(B^{1+\varepsilon})$ when $\delta_{\mathbf{u}}F = 2$ and $|\mathbf{u}| \ll B$. Hence, $N^{(2)}(F,B) = O_{F,\varepsilon}(B^{2+\varepsilon})$. □

Lemma 8. *If F is absolutely irreducible, then*

$$N^{(1)}(F,B) \ll_{F,\varepsilon} B^{2+\varepsilon}.$$

Proof. We claim that there are $O_{F,\varepsilon}(T^{1+\varepsilon})$ points $\mathbf{u} \in P^2(T)$ for which $\delta_{\mathbf{u}}F = 1$. Assuming this and referring to lemma 5, we have

$$\sum_{\substack{T \le |\mathbf{u}| < 2T \\ \delta_{\mathbf{u}}F = 1}} N_{\mathbf{u}}(F,B) \ll_{F,\varepsilon} B^{2+\varepsilon}T^{\varepsilon} \ll B^{2+\varepsilon},$$

for $1 \leq T \ll B^{1/2}$. By summing over dyadic intervals we get the promised result. Now lemma 2 states that if $\delta_{\mathbf{u}}F = 1$, then $f^{-1}H_{\mathbf{u}} \subset X$ contains a curve with Hilbert polynomial

$$P(q) = \binom{mq+1}{mq}$$

with respect to some very ample sheaf $f^*\mathcal{O}(m)$ on X. Let $\mathcal{H}$ be the Hilbert scheme parametrising such curves and let $\mathcal{X} \subset X \times \mathcal{H}$ be the corresponding universal family. Consider the composed map

$$\mathcal{X} \to X \times \mathcal{H} \xrightarrow{f \times \mathrm{id}} \mathbb{P}^2 \times \mathcal{H}.$$

over $\mathcal{H}$. By lemma 2, $\mathcal{X}_P \to \mathbb{P}^2_{k(P)}$ is a closed immersion onto a line for every $P \in \mathcal{H}$. Hence, $\mathcal{X}$ is isomorphic to its image $\mathcal{Y} \subset \mathbb{P}^2 \times \mathcal{H}$ (see corollary 18.12.6 in [2]). Moreover, $\mathcal{Y} \to \mathcal{H}$ has the Hilbert polynomial of a line, so $\mathcal{Y} = \mathcal{H} \times_{\mathbb{P}^{2*}} \mathcal{L}$ for a unique morphism $g : \mathcal{H} \to \mathbb{P}^{2^*}$, where $\mathcal{L} \subset \mathbb{P}^2 \times \mathbb{P}^{2^*}$ is the universal line. There are at most d curves on X which map to a given line in $\mathbb{P}^2$. In other words, g is quasi-finite. It is also proper so it is finite (see corollary 18.12.4 in [2]). The claim above is implied by the statement $N(g, T) = O_{g,\varepsilon}(T^{1+\varepsilon})$.

Assume that F is absolutely irreducible. If $\mathcal{H}$ has dimension 2, then $g(\overline{\mathcal{H}}) = \overline{\mathbb{P}^{2*}}$. This means that $f^{-1}H \subset \overline{X}$ is reducible for every $H \subset \overline{\mathbb{P}^2}$. We know that this is not the case so $\mathcal{H}$ has dimension at most 1. Let C be a reduced and irreducible curve on $\mathcal{H}$. If $g(C) \subset \mathbb{P}^{2*}$ has degree at least 2, then $N(g|_C, T) = O_{g,\varepsilon}(T^{1+\varepsilon})$ according to theorem 3 in [6]. Suppose that $g : C \to g(C)$ is an isomorphism onto a line and let $\pi^{-1}C$ and $\pi^{-1}g(C)$ be the preimages under the projections $\pi : \mathcal{X} \to \mathcal{H}$ and $\pi : \mathcal{L} \to \mathbb{P}^{2^*}$, respectively. The restriction $\pi^{-1}C \to \pi^{-1}g(C)$ of $\mathcal{X} \to \mathcal{L}$ is then an isomorphism since $(\pi^{-1}C)_P \to (\pi^{-1}g(C))_{g(P)}$ is an isomorphism for every $P \in C$. One can check that $\pi^{-1}g(C) \to \mathbb{P}^2$ is the blow up at the point of $\mathbb{P}^2$ corresponding to the line $g(C) \subset \mathbb{P}^{2*}$. That is, the composition $\pi^{-1}C \to X \to \mathbb{P}^2$ is a birational equivalence. This contradicts the assumption that f is not an isomorphism. Hence, if $g(C) \subset \mathbb{P}^{2^*}$ is a line, then the finite dominant morphism $g : C \to g(C)$ has degree at least 2. By theorem 1, $N(g|_C, T) = O_{g,\varepsilon}(T^{1+\varepsilon})$. □

To sum up, if F is absolutely irreducible and $d \geq 3$, then

$$N(F,B) \leq \underbrace{N^{(1)}(F,B)}_{\substack{O_{F,\varepsilon}(B^{2+\varepsilon}) \\ \text{by lemma 8}}} + \underbrace{N^{(2)}(F,B)}_{\substack{O_{F,\varepsilon}(B^{2+\varepsilon}) \\ \text{by lemma 7}}} + \underbrace{\sum_{\delta=3}^{d} N^{(\delta)}(F,B)}_{\substack{O_{F,\varepsilon}(B^{2+\varepsilon}) \\ \text{by lemma 6}}} \ll_{F,\varepsilon} B^{2+\varepsilon}.$$

If $d = 2$, then

$$N(F,B) = \underbrace{N^{(1)}(F,B)}_{\substack{O_{F,\varepsilon}(B^{2+\varepsilon}) \\ \text{by lemma 8}}} + \underbrace{N^{(2)}(F,B)}_{\substack{O_{F,\varepsilon}(B^{9/4+\varepsilon}) \\ \text{by lemma 6}}} \ll_{F,\varepsilon} B^{9/4+\varepsilon}.$$

The following observation completes the proof of theorem 4.

Lemma 9. *If F is not absolutely irreducible, then*

$$N(F,B) \ll_F B^2.$$

Proof. Let

$$F(t,\mathbf{x}) = F_1(t,\mathbf{x}) \cdots F_k(t,\mathbf{x}),$$

for some irreducible polynomials $F_i \in \overline{\mathbb{Q}}[t,\mathbf{x}]$ which are monic in the variable t. Such a factorisation is unique except for the arrangement of the factors, so the Galois group of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$ acts transitively on $F_1, \ldots, F_k$. Consequently, $\frac{\partial F}{\partial t}(t,\mathbf{x})$ has the same rational roots as $F(t,\mathbf{x})$ for any given $\mathbf{x} \in P^2$. The counting function $N(f,B)$ is thus dominated by the number of rational points of height at most B on the discriminant locus $\Delta_F(\mathbf{x}) = 0$. There are $O_F(B^2)$ such points (see theorem 1 in [6]). □

References

[1] D. Eisenbud and J. Harris, *The geometry of schemes*, Springer-Verlag, 2000.

[2] A. Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas IV, *Inst. Hautes Études Sci. Publ. Math.*, No. 32, 1967.

[3] J. Harris, *Algebraic Geometry*, Springer-Verlag, 1995.

[4] R. Hartshorne, *Ample subvarieties of algebraic varieties*, Springer-Verlag, 1970.

[5] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, 1977.

[6] R. Heath-Brown, The density of rational points on curves and surfaces, Preprint.

[7] J. Kollár, *Rational curves on algebraic varieties*, Springer-Verlag, 1996.

[8] S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, 1983.

[9] J.-P. Serre, *Lectures on the Mordell-Weil theorem*, Friedr. Vieweg & Sohn, 1989.