

# COMPONENTWISE HEIGHT BOUNDS FOR POLYNOMIAL VALUE-SET LIFTING

HENRY SHIN

ABSTRACT. Let  $f, g \in k[x]$  be nonconstant polynomials over a number field. We count  $S$ -integer inputs  $a$  for which  $f(a)$  has a  $k$ -rational preimage under  $g$ , after removing the polynomial graph components  $Y = h(X)$  with  $f = g \circ h$ . The main theorem gives componentwise height bounds. For a rational component of  $f(X) - g(Y) = 0$  with one geometric point at infinity and projection degree  $d_X(C)$  to the  $X$ -line, the corresponding contribution has the sharp power-log order

$$B^{[k:\mathbb{Q}]/d_X(C)}(\log B)^{q_{k,S}}, \quad q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1,$$

precisely when its  $X$ -parametrization is  $S$ -active. Rational components with two geometric points at infinity contribute only polylogarithmically, and all other components contribute finitely many inputs. Over  $\mathbb{Q}$ , square-root growth after graph removal occurs exactly from active rational one-infinity components with  $d_X(C) = 2$ . We give an explicit thin exceptional set for the generic multiplicity theorem and prove that every square-root source forces  $g$  to have an involutive affine symmetry.

## 1. INTRODUCTION

Let  $K$  be a Hilbertian field of characteristic zero. A basic application of Hilbert irreducibility says that if  $f, g \in K[x]$  are nonconstant and

$$\{x \in K : f(x) \in g(K)\}$$

is non-thin, then  $f = g \circ h$  for some  $h \in K[x]$ . Indeed, rational lifts of  $f(x)$  through  $g$  are rational points on the finite cover

$$C_{f,g} := (f(X) - g(Y) = 0)_{\text{red}} \longrightarrow \mathbb{A}_X^1.$$

A finite cover whose rational image contains a non-thin set must have a component of generic degree one over the base. In the polynomial situation such components are exactly the graph components  $Y = h(X)$ , with  $f = g \circ h$ .

This observation is useful, but by itself it lies very close to the definition of thinness. The purpose of this note is to record the next layer: what remains *inside* the sparse exceptional set after the graph components have been removed. For integer and  $S$ -integer inputs, the answer is geometric and componentwise. Power-size exceptional sets come from rational components with one geometric point at infinity; the exponent is the reciprocal of the projection degree to the  $X$ -line. Rational components with two geometric

---

2020 *Mathematics Subject Classification.* 14G05, 11D45, 11G50, 12E05, 11R58.

*Key words and phrases.* polynomial value sets, integral points, thin sets, Hilbert irreducibility,  $S$ -integers, separated variables.

points at infinity give  $S$ -unit, hence polylogarithmic, growth. All other components give only finitely many  $S$ -integral inputs.

**What is and is not new.** This paper does not classify the separated-variable components of  $f(X) - g(Y)$ . That problem belongs to a substantial literature. Bilu–Tichy classify the pairs  $(f, g)$  for which  $f(x) = g(y)$  has infinitely many rational solutions with bounded denominator [5]; Davenport, Lewis and Schinzel initiated related questions about separated polynomials [8, 9]; and Dvornicich–Zannier study closely related value questions for algebraic functions [10, 11]. König–Neftin study reducible fibers of polynomial maps from a quantitative Hilbert-irreducibility viewpoint [14]. The recent preprint of Behajaina–König–Neftin announces a solution to the Davenport–Lewis–Schinzel reducibility problem for  $f(X) - g(Y)$  over  $\mathbb{C}$ , with applications to the Hilbert–Siegel problem [6].

Our contribution is different and deliberately narrower. We assume the component geometry of  $f(X) - g(Y)$  has been found, or can be found by existing factorization and normalization methods, and then extract the integer and  $S$ -integer lifting exponents after polynomial graph components have been removed. In this form the theorem records three pieces of arithmetic data which are not separated out by the classification statements alone: the projection degree  $d_X(C)$ , the distinction between arithmetically active and inactive one-infinity parametrizations, and the fact that two-infinity components are only unit-theoretic and hence polylogarithmic.

This is why the main theorem should be read as a componentwise height-counting refinement of Hilbertian value-set lifting, not as a replacement for the Bilu–Tichy or Davenport–Lewis–Schinzel classification theorems. We also use Bilu’s paper on quadratic factors [4] only as nearby low-degree separated-variable literature. Its condition concerns factors of small total degree, whereas the square-root boundary in this paper is the projection-degree condition  $d_X(C) = 2$ . These are related but not identical. Work of Avanzi–Zannier on separated-variable genus-one curves and polynomial Pell equations is also close to the two-infinity examples below [2].

The final section adds one useful structural restriction on the square-root boundary. A quadratic source forces the polynomial  $g$  to have a nontrivial involutive affine symmetry. Equivalently, after a translation of the  $Y$ -line,  $g$  factors through the squaring map. Thus, for example, no square-root source can occur when  $\deg g$  is odd or when  $g$  has no such symmetry.

**Standard inputs and counting convention.** The proof uses three standard inputs: Hilbert irreducibility through thin sets, Siegel–Mahler finiteness for  $S$ -integral points on affine curves, and height counting for  $S$ -integers. On the integral-point side, Alvanos–Bilu–Poulakis give necessary and sufficient conditions for an affine curve to have infinitely many  $S$ -integral points [1]; here we use only the curve-wise finiteness dichotomy and then refine the count by the  $X$ -projection. The last input is used in the form

$$\#\{a \in \mathcal{O}_{k,S} : H(a) \leq B\} \asymp B^{[k:\mathbb{Q}]}(\log B)^{q_{k,S}}, \quad q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1,$$

following Schanuel and Barroero. The logarithmic exponent includes the Archimedean unit rank; it is not merely the number of finite places of  $S$ .

The main upper bounds count  $X$ -values, not solution pairs. Summing component counts gives an upper bound because the same input may occur on several components. Lower bounds are proved componentwise: on an active component with one geometric point at infinity a polynomial parametrization gives many  $S$ -integer  $X$ -values, and the bounded degree of the parametrization controls collisions. This is why the paper obtains sharp power-log orders without asserting a global leading constant.

**Proof roadmap and component table.** The proof has the following schematic form:

$$\begin{aligned} \text{new rational lift} &\implies \text{non-graph component of } C_{f,g}, \\ \text{component type} &\implies \text{Siegel–Mahler/height estimate} \\ &\implies \text{input exponent.} \end{aligned}$$

More explicitly, after denominator control places all relevant lifts on a fixed integral model, each non-graph component contributes according to the following trichotomy.

component type	geometric description	input contribution
Siegel-finite	not geometrically integral; or positive genus; or at least three boundary points; or no $k$ -rational genus-zero model	$O(1)$
one-infinity	$\tilde{C} \cong \mathbb{P}_k^1$ , $\#D_C = 1$ , $X = A(t)$	$B^{[k:\mathbb{Q}]/\deg A} (\log B)^{q_{k,S}}$ if active; otherwise $O(1)$
two-infinity	$\tilde{C} \cong \mathbb{P}_k^1$ , with $D_C$ consisting of two geometric points	$(\log B)^\rho$

Here a “one-infinity” component means a rational component with one geometric point at infinity, and similarly for two-infinity components. The power terms are sharp componentwise when the one-infinity component is active. No leading constant is asserted globally, because different components may produce the same input values; the theorem is a sharp exponent statement, not a disjointness or main-term theorem.

**Generic multiplicity and the explicit thin set.** Let  $p : W \rightarrow X$  be a finite morphism with  $X$  normal and integral. Let  $s(p)$  denote the number of irreducible components of  $W_{\text{red}}$  which dominate  $X$  with generic degree one.

**Theorem 1.1** (Generic rational fiber count). *Let  $K$  be a field, let  $X/K$  be a normal integral variety, and let  $p : W \rightarrow X$  be finite with  $W$  reduced. Assume that every irreducible component of  $W$  which dominates  $X$  is generically separable over  $X$ . Then there is a thin subset  $T \subseteq X(K)$  such that*

$$\#p^{-1}(x)(K) = s(p) \quad (x \in X(K) \setminus T),$$

where fibers are counted as reduced sets of rational points.

For polynomials, put

$$\mathcal{H}_{f,g}(K) = \{h \in K[x] : f = g \circ h\}.$$

Then [Theorem 1.1](#) gives

$$\#\{y \in K : g(y) = f(x)\} = \#\mathcal{H}_{f,g}(K)$$

outside a thin subset of  $K$ . We also make this exceptional set explicit. If

$$(f(X) - g(Y))_{\text{red}} = c \prod_{i=1}^s (Y - h_i(X)) \prod_{j=1}^m F_j(X, Y)$$

is the squarefree factorization in  $K[X, Y]$ , where  $h_i \in K[X]$ ,  $f = g \circ h_i$ , and  $\deg_Y F_j \geq 2$ , then outside an explicitly computable finite collision set the exceptional inputs are exactly

$$\bigcup_{j=1}^m \{x \in K : F_j(x, Y) \text{ has a root in } K\}.$$

For Hilbertian  $K$  of characteristic zero this is a thin subset of  $\mathbb{A}^1(K)$ . See [Proposition 3.3](#).

**The sparse number-field theorem.** Let  $k$  be a number field, let  $S$  be a finite set of places of  $k$  containing the Archimedean places, and let  $\mathcal{O}_{k,S}$  be the ring of  $S$ -integers. We write  $H$  for the absolute multiplicative Weil height and  $n_k = [k : \mathbb{Q}]$ . Define

$$N_{f,g,k,S}^{\text{new}}(B) = \#\left\{a \in \mathcal{O}_{k,S} : H(a) \leq B, \exists y \in k \text{ with } g(y) = f(a), \right. \\ \left. y \neq h(a) \text{ for every } h \in \mathcal{H}_{f,g}(k)\right\}.$$

Thus we count only the inputs admitting a *new* rational lift, not explained by polynomial graph sections. Throughout, the quantities  $N^{\text{new}}$  count *inputs*  $a$ , not points on the curve. Consequently the upper bounds are obtained by summing componentwise contributions, while the lower bounds use one active component, bounded degree of its  $X$ -parametrization, and deletion of only finitely many graph intersections.

Let  $C$  be a non-graph irreducible component of  $C_{f,g}$ . Let  $\tilde{C}$  be the smooth projective normalization and let  $D_C$  be the boundary of the affine normalization. Let  $\mathcal{R}_1(f, g; k)$  denote the set of non-graph components such that  $C$  is geometrically integral,  $\tilde{C} \cong \mathbb{P}_k^1$ , and  $\#D_C = 1$ . For such a component, choose a coordinate  $t$  on the affine normalization and write

$$X = A_C(t), \quad Y = B_C(t), \quad A_C, B_C \in k[t].$$

Then

$$d_X(C) := [k(C) : k(X)] = \deg A_C.$$

We call  $C$   $S$ -active if

$$A_C(k) \cap \mathcal{O}_{k,S} \neq \emptyset.$$

This condition is independent of the chosen coordinate  $t$ . Let  $\mathcal{R}_1^{\text{act}}(f, g; k, S) \subseteq \mathcal{R}_1(f, g; k)$  be the active subset. Let  $\mathcal{R}_2(f, g; k)$  be the set of non-graph irreducible components  $C$  such that  $C$  is geometrically

integral,  $\tilde{C} \cong \mathbb{P}_k^1$ , and  $D_C$  consists of two geometric boundary points. The two boundary points need not be individually defined over  $k$ ; in the proof they are split after a finite extension.

**Theorem 1.2** (Sparse exceptional lifting over number fields). *Let  $k$  be a number field, let  $S$  be a finite set of places containing the Archimedean places, put  $q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1$ , and let  $f, g \in k[x]$  be nonconstant. Then there are nonnegative integers  $\rho_C$ , for  $C \in \mathcal{R}_2(f, g; k)$ , depending only on  $k, S, f, g$ , such that*

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} 1 + \sum_{C \in \mathcal{R}_1^{\text{act}}(f,g;k,S)} B^{n_k/d_X(C)} (\log(2B))^{q_{k,S}} \\ + \sum_{C \in \mathcal{R}_2(f,g;k)} (\log(2B))^{\rho_C}.$$

In particular, if  $\mathcal{R}_1^{\text{act}}(f, g; k, S) \neq \emptyset$  and

$$\theta_{f,g,k,S} = \max_{C \in \mathcal{R}_1^{\text{act}}(f,g;k,S)} \frac{1}{d_X(C)},$$

then

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} B^{n_k \theta_{f,g,k,S}} (\log(2B))^{q_{k,S}} + (\log(2B))^c$$

for some  $c \geq 0$ . If  $\mathcal{R}_1^{\text{act}}(f, g; k, S) = \emptyset$ , then

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} (\log(2B))^c.$$

The power-log orders in [Theorem 1.2](#) are sharp when an active component is present.

**Theorem 1.3** (Active one-infinity components give lower bounds). *Let  $C \in \mathcal{R}_1^{\text{act}}(f, g; k, S)$ , and put  $d = d_X(C)$ . Then*

$$N_{f,g,k,S}^{\text{new}}(B) \gg_{k,S,f,g,C} B^{n_k/d} (\log(2B))^{q_{k,S}},$$

where  $q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1$ . Consequently, the power-log exponent  $n_k/d$  attached to an active one-infinity component cannot be improved in general.

For ordinary integer inputs over  $\mathbb{Q}$ , no logarithm appears in the power terms.

**Corollary 1.4** (Integer inputs over  $\mathbb{Q}$ ). *Let  $f, g \in \mathbb{Q}[x]$  be nonconstant and put*

$$N_{f,g}^{\text{new}}(B) = \#\{n \in \mathbb{Z} : |n| \leq B, \exists y \in \mathbb{Q} \text{ with } g(y) = f(n), \\ y \neq h(n) \text{ for every } h \in \mathcal{H}_{f,g}(\mathbb{Q})\}.$$

Then

$$N_{f,g}^{\text{new}}(B) \ll_{f,g} 1 + \sum_{C \in \mathcal{R}_1^{\text{act}}(f,g;\mathbb{Q},\{\infty\})} B^{1/d_X(C)} + \sum_{C \in \mathcal{R}_2(f,g;\mathbb{Q})} (\log(2B))^{\rho_C}.$$

In particular,

$$N_{f,g}^{\text{new}}(B) = O_{f,g}(B^{1/2}).$$

Moreover,

$$N_{f,g}^{\text{new}}(B) \asymp_{f,g} B^{1/2}$$

if there is an active non-graph rational one-infinity component with  $d_X(C) = 2$ ; if no such component exists, then

$$N_{f,g}^{\text{new}}(B) = O_{f,g}(B^{1/3}) + O_{f,g}((\log(2B))^e).$$

Thus square-root growth is not an undifferentiated failure of Hilbert irreducibility. It is caused exactly by active rational one-infinity components with quadratic projection to the  $X$ -line.

**Quadratic sources.** The square-root case can be isolated without importing the full Bilu–Tichy standard-pair classification.

**Proposition 1.5** (Quadratic-source reformulation). *Let  $k$  be a field of characteristic zero, and let  $f, g \in k[x]$  be nonconstant. The curve  $C_{f,g}$  has a non-graph rational one-infinity component  $C$  with  $d_X(C) = 2$  if and only if there exist polynomials  $A, B \in k[t]$  such that*

$$\deg A = 2, \quad k(A(t), B(t)) = k(t), \quad f(A(t)) = g(B(t)),$$

and

$$B(t) \notin k[A(t)].$$

Equivalently, after a linear change of the parameter and a linear change of the  $X$ -coordinate, every such source has the form

$$A(t) = \alpha t^2 + \beta, \quad \alpha \in k^*, \beta \in k.$$

We prove a sharper structural form in [Section 8](#): every quadratic source comes from a nontrivial affine involution of  $g$ . Equivalently, after a linear change of the  $Y$ -coordinate,  $g$  is an even polynomial. This gives a simple obstruction: if  $g$  has no such involution, then no square-root source exists. The source-even normal form classifies the quadratic source mechanism; a further sorting of those identities into Bilu–Tichy standard-pair families would require additional separated-variable classification machinery. In particular, one must not confuse the condition  $d_X(C) = 2$  with the small-total-degree condition in Bilu’s quadratic-factor theorem; that theorem is adjacent background, not a direct black-box classification of the square-root sources considered here.

## 2. THIN SETS, MULTIPLICITY, AND EXPLICIT EXCEPTIONAL SETS

Throughout, a variety over a field is an integral separated scheme of finite type over that field. If  $\pi : Y \rightarrow X$  is a dominant generically finite morphism of integral varieties, its generic degree is

$$\deg_{\text{gen}}(\pi) = [K(Y) : K(X)].$$

**Definition 2.1** (Thin sets). Let  $K$  be a field and let  $X$  be an integral  $K$ -variety. A subset  $A \subseteq X(K)$  is *thin* if it is contained in a finite union of sets of the following two types:

- (i)  $Z(K)$ , where  $Z \subsetneq X$  is a proper closed subvariety;
- (ii)  $\varphi(V(K))$ , where  $V$  is an integral  $K$ -variety and  $\varphi : V \rightarrow X$  is dominant, generically finite, generically separable, and  $\deg_{\text{gen}}(\varphi) \geq 2$ .

The variety  $X$  has the *Hilbert property* over  $K$  if  $X(K)$  is not thin. We use the standard language of thin sets and Hilbert irreducibility as in Serre and Fried–Jarden [17, 13].

**Lemma 2.2** (Thin image of a cover without degree-one components). *Let  $K$  be a field, let  $X$  be an integral  $K$ -variety, and let  $\pi : Y \rightarrow X$  be finite. Assume that every irreducible component of  $Y_{\text{red}}$  which dominates  $X$  is generically separable over  $X$  and has generic degree at least 2. Then  $\pi(Y(K))$  is thin in  $X(K)$ .*

*Proof.* Write  $Y_1, \dots, Y_m$  for the irreducible components of  $Y_{\text{red}}$ . If  $Y_i$  is not dominant over  $X$ , then  $\pi(Y_i)$  is contained in a proper closed subset of  $X$ , since  $\pi$  is finite.

Suppose  $Y_i$  is dominant. Let  $\nu_i : Y_i^\nu \rightarrow Y_i$  be the normalization. There is a dense open subset  $U_i \subseteq Y_i$  over which  $\nu_i$  is an isomorphism. Put  $E_i = Y_i \setminus U_i$ . Then  $\pi(E_i)$  is contained in a proper closed subset of  $X$ , and every  $K$ -point of  $Y_i \setminus E_i$  lifts uniquely to  $Y_i^\nu(K)$ . Hence

$$\pi(Y_i(K)) \subseteq (\pi \circ \nu_i)(Y_i^\nu(K)) \cup \pi(E_i)(K).$$

The morphism  $\pi \circ \nu_i : Y_i^\nu \rightarrow X$  is dominant, generically finite, generically separable, and of generic degree at least 2. Its image on  $K$ -points is thin of type (ii), up to the proper closed exceptional image  $\pi(E_i)$ . Taking the finite union over all  $i$  proves the lemma.  $\square$

**Definition 2.3** (Degree-one component count). Let  $p : W \rightarrow X$  be a finite morphism with  $X$  integral. Define  $s(p)$  to be the number of irreducible components  $W_i$  of  $W_{\text{red}}$  such that  $W_i$  dominates  $X$  and  $\deg_{\text{gen}}(W_i/X) = 1$ .

**Theorem 2.4** (Generic rational fiber count). *Let  $K$  be a field, let  $X/K$  be a normal integral variety, and let  $p : W \rightarrow X$  be finite with  $W$  reduced. Assume that every irreducible component of  $W$  which dominates  $X$  is generically separable over  $X$ . Then there is a thin subset  $T \subseteq X(K)$  such that, for every  $x \in X(K) \setminus T$ ,*

$$\#p^{-1}(x)(K) = s(p),$$

where the fiber is counted as a reduced set of  $K$ -points.

*Proof.* Let  $W_1, \dots, W_m$  be the irreducible components of  $W$ . We build the exceptional set  $T$ . Include the rational points of the images of all non-dominant components. These images are proper closed subsets of  $X$ . Also include the rational images of all dominant components of generic degree at least 2. By Lemma 2.2, their union is thin.

If  $W_i$  has generic degree one over  $X$ , then  $W_i^\nu \rightarrow X$  is finite birational. Since  $X$  is normal, this morphism is an isomorphism. Include in  $T$  the image in  $X(K)$  of the proper closed locus of  $W_i$  over which  $W_i^\nu \rightarrow W_i$  is not an isomorphism, and include the images in  $X(K)$  of all pairwise intersections of distinct degree-one components. These images lie in proper closed subsets of  $X$ .

For  $x \notin T$ , no non-dominant or degree-at-least-two component contributes a rational point over  $x$ . Each degree-one component contributes

exactly one rational point, and the points contributed by distinct degree-one components are distinct. Thus the fiber cardinality is  $s(p)$ .  $\square$

**Corollary 2.5** (Multi-section criterion). *With hypotheses as in Theorem 2.4, suppose*

$$\Omega_r = \{x \in X(K) : \#p^{-1}(x)(K) \geq r\}$$

*is non-thin. Then  $s(p) \geq r$ .*

*Proof.* If  $s(p) < r$ , Theorem 2.4 gives a thin set outside which all rational fibers have cardinality less than  $r$ . Hence  $\Omega_r$  would be thin, a contradiction.  $\square$

### 3. POLYNOMIAL MULTIPLICITY AND AN EXPLICIT EXCEPTIONAL THIN SET

Let  $K$  be a field and let  $f, g \in K[x]$  be nonconstant. Put

$$\mathcal{H}_{f,g}(K) = \{h \in K[x] : f = g \circ h\},$$

and

$$C_{f,g} = (f(X) - g(Y) = 0)_{\text{red}} \subset \mathbb{A}_{X,Y}^2.$$

The set  $\mathcal{H}_{f,g}(K)$  is finite. Indeed, its elements are in bijection with those irreducible components of  $C_{f,g}$  which have generic degree one over  $\mathbb{A}_X^1$ , and the curve  $C_{f,g}$  has only finitely many irreducible components.

**Theorem 3.1** (Generic rational preimage formula). *Let  $K$  be a field of characteristic zero, and let  $f, g \in K[x]$  be nonconstant. Then there is a thin subset  $T \subseteq K$  such that*

$$\#\{y \in K : g(y) = f(x)\} = \#\mathcal{H}_{f,g}(K) \quad (x \in K \setminus T).$$

*Proof.* Projection  $p : C_{f,g} \rightarrow \mathbb{A}_X^1$  is finite because  $g$  is nonconstant. By Theorem 2.4, outside a thin subset of  $K$ , the rational fiber cardinality equals the number of components of  $C_{f,g}$  of generic degree one over  $\mathbb{A}_X^1$ .

Such components are exactly the polynomial graph components. Indeed, if  $D$  is a component of generic degree one over the  $X$ -line, then in  $K(D) = K(X)$  we have  $Y = h(X)$  for some  $h \in K(X)$ , and

$$f(X) = g(h(X)).$$

If  $h$  had a finite pole, then  $g(h)$  would have a finite pole: at that place the leading term of  $g(h)$  has strictly larger pole order than every lower term. This contradicts  $f \in K[X]$ . Hence  $h \in K[X]$ . Conversely, every  $h \in \mathcal{H}_{f,g}(K)$  gives the graph component  $Y = h(X)$ . This identifies the degree-one components with  $\mathcal{H}_{f,g}(K)$ .  $\square$

**Corollary 3.2** (Multiplicity lifting for polynomials). *Let  $K$  be a Hilbertian field of characteristic zero, let  $f, g \in K[x]$  be nonconstant, and let  $r \geq 1$ . Suppose there is a non-thin subset  $\Omega \subseteq K$  such that, for every  $x \in \Omega$ , the equation*

$$g(y) = f(x)$$

has at least  $r$  distinct solutions  $y \in K$ . Then there exist distinct polynomials  $h_1, \dots, h_r \in K[x]$  satisfying

$$f = g \circ h_i \quad (1 \leq i \leq r).$$

*Proof.* If fewer than  $r$  such polynomial sections existed, then [Theorem 3.1](#) would give a thin set outside which the rational fiber cardinality is less than  $r$ , contradicting the non-thinness of  $\Omega$ .  $\square$

**Proposition 3.3** (Explicit exceptional set for polynomial multiplicity). *Let  $K$  be a field of characteristic zero, and let  $f, g \in K[x]$  be nonconstant. Write the squarefree factorization in  $K[X, Y]$  as*

$$(f(X) - g(Y))_{\text{red}} = c \prod_{i=1}^s (Y - h_i(X)) \prod_{j=1}^m F_j(X, Y),$$

where  $c \in K^*$ ,  $h_i \in K[X]$ ,  $f = g \circ h_i$ , and each  $F_j$  is irreducible over  $K$  with  $\deg_Y F_j \geq 2$ . Let

$$P_i = Y - h_i(X) \quad (1 \leq i \leq s), \quad P_{s+j} = F_j \quad (1 \leq j \leq m).$$

The pairwise resultants are nonzero because the factors  $P_a$  are distinct irreducible factors, hence coprime in  $K(X)[Y]$ . Define

$$R(X) = \prod_{a < b} \text{Res}_Y(P_a(X, Y), P_b(X, Y)),$$

after omitting constant nonzero factors, with the convention  $R = 1$  if no nonconstant factor remains, and put

$$Z_R = \{x \in K : R(x) = 0\}.$$

For each  $j$ , put

$$T_j = \{x \in K : F_j(x, Y) \text{ has a root in } K\}.$$

Then

$$\{x \in K : \#\{y \in K : g(y) = f(x)\} \neq s\} \subseteq Z_R \cup \bigcup_{j=1}^m T_j.$$

Moreover, for every  $x \in K \setminus Z_R$ ,

$$\#\{y \in K : g(y) = f(x)\} = s + \sum_{j=1}^m \#\{y \in K : F_j(x, y) = 0\}.$$

In particular,

$$\{x \in K \setminus Z_R : \#\{y \in K : g(y) = f(x)\} \neq s\} = \left( \bigcup_{j=1}^m T_j \right) \setminus Z_R.$$

Each  $T_j$  is thin in  $\mathbb{A}^1(K)$ . If  $K$  is Hilbertian, this is a genuine Hilbertian exceptional set:  $\mathbb{A}^1(K)$  itself is not thin.

*Proof.* The graph factors contribute the points  $y = h_i(x)$ . Away from  $Z_R$ , no two distinct component factors have a common  $Y$ -root after specialization at  $X = x$ . Hence the graph contributions are distinct, and no root coming from a non-graph component collides with a graph root or with a root

from another component. Therefore the reduced rational fiber cardinality is exactly

$$s + \sum_{j=1}^m \#\{y \in K : F_j(x, y) = 0\}.$$

This proves the equality outside  $Z_R$ , and the containment follows.

Because  $f(X) - g(Y)$  has constant nonzero leading coefficient in  $Y$ , namely  $-\text{lc}(g)$ , each positive- $Y$ -degree factor  $F_j$  also has constant nonzero leading coefficient in  $Y$ . Hence the projection  $F_j(X, Y) = 0 \rightarrow \mathbb{A}_X^1$  is finite of degree  $\deg_Y F_j$ . Since  $\deg_Y F_j \geq 2$  and characteristic is zero, this is a generically separable degree-at-least-two cover. Thus  $T_j$  is the rational image of a degree-at-least-two cover of  $\mathbb{A}^1$ , up to the finite exceptional set already absorbed into  $Z_R$ , so it is thin.  $\square$

*Remark 3.4.* One may multiply  $R$  by the discriminants  $\text{disc}_Y(F_j)$  if one wants a good specialization locus on which all specialized fibers are also separable. For the reduced fiber cardinality formula above, pairwise resultants are the essential finite collision terms.

#### 4. INTEGRAL POINTS ON COMPONENTS OVER NUMBER FIELDS

Let  $k$  be a number field,  $S$  a finite set of places containing the Archimedean places, and  $\mathcal{O}_{k,S}$  the ring of  $S$ -integers. We write  $H$  for the absolute multiplicative height on  $\mathbb{P}^1(k)$ , put  $n_k = [k : \mathbb{Q}]$ , and put  $q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1$ .

**Lemma 4.1** (Counting  $S$ -integers of bounded height). *Let  $k$  be a number field, let  $S$  be a finite set of places containing the Archimedean places, and put*

$$q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1.$$

*There are constants  $c_1, c_2 > 0$  such that, for  $R \geq 2$ ,*

$$c_1 R^{n_k} (\log R)^{q_{k,S}} \leq \#\{a \in \mathcal{O}_{k,S} : H(a) \leq R\} \leq c_2 R^{n_k} (\log R)^{q_{k,S}}.$$

*The same upper and lower bounds hold for any nonempty affine coset  $a_0 + \lambda \mathcal{O}_{k,S}$ ,  $\lambda \neq 0$ , and for any nonzero fractional  $\mathcal{O}_{k,S}$ -ideal, with constants depending on the coset or ideal.*

*Proof.* This is the degree-one case of the standard height asymptotics for  $S$ -integers; see Schanuel's theorem for bounded-height points over number fields [16] and Barroero's  $S$ -integer counting theorem [3]. The logarithmic exponent is the full  $S$ -unit rank, not only the number of finite places in  $S$ :

$$q_{k,S} = r_1 + r_2 - 1 + |S_f| = |S| - 1,$$

where  $S_f$  is the set of finite places in  $S$ . For fractional ideals and affine cosets, we recall the standard reduction. If  $I$  is a nonzero fractional  $\mathcal{O}_{k,S}$ -ideal, choose nonzero  $\alpha, \beta \in k$  with

$$\alpha \mathcal{O}_{k,S} \subseteq I \subseteq \beta \mathcal{O}_{k,S}.$$

The existence of  $\beta$  is the definition of a fractional ideal, and any nonzero  $\alpha \in I$  gives the left inclusion. Multiplication by a fixed nonzero element

distorts absolute height by a bounded factor, so the same upper and lower estimates hold for  $I$ . If  $a_0 + I$  is a nonempty affine coset, the inclusions

$$a_0 + \alpha\mathcal{O}_{k,S} \subseteq a_0 + I \subseteq a_0 + \beta\mathcal{O}_{k,S}$$

reduce the matter to cosets of principal fractional ideals. For fixed  $a_0 \in k$  and  $\gamma \in k^*$ , the elementary height inequalities

$$H(a_0 + \gamma u) \leq C_{a_0, \gamma} H(u), \quad H(u) \leq C'_{a_0, \gamma} H(a_0 + \gamma u)$$

hold for all  $u \in k$ . The first inequality follows from the standard height bounds for addition and multiplication, and the second follows by writing  $u = \gamma^{-1}((a_0 + \gamma u) - a_0)$ . Hence replacing  $u$  by  $a_0 + \gamma u$  distorts height balls only by fixed multiplicative constants, and the same upper and lower estimates hold for the coset.  $\square$

**Lemma 4.2** (Height growth under rational functions). *Let  $A \in k(t)$  be a nonconstant rational function of degree  $d$ . There are positive constants  $c_1, c_2$ , depending on  $A$ , such that*

$$c_1 H(t)^d \leq H(A(t)) \leq c_2 H(t)^d$$

for all  $t \in \mathbb{P}^1(k)$  for which  $A(t)$  is defined.

*Proof.* This is the standard functoriality of Weil heights for morphisms  $\mathbb{P}^1 \rightarrow \mathbb{P}^1$ ; see, for example, [15, Ch. 3].  $\square$

**Lemma 4.3** (Geometrically reducible curves have finitely many rational points). *Let  $C/k$  be an irreducible affine curve over a number field. If  $C$  is not geometrically integral, then  $C(k)$  is finite.*

*Proof.* Let  $C^\nu \rightarrow C$  be the normalization. The nonnormal locus of  $C$  is finite, and the normalization is an isomorphism over its complement. Let  $L$  be the algebraic closure of  $k$  in  $k(C)$ . Since  $k$  is perfect, non-geometric integrality is equivalent to  $L \neq k$ . A  $k$ -point of the normal open part would lift uniquely to  $C^\nu(k)$ , and evaluation at that point would embed  $L$  into  $k$  over  $k$ , impossible. Thus all  $k$ -points lie in the finite nonnormal locus.  $\square$

Let  $C/k$  be an irreducible affine curve, let  $C^\nu$  be its affine normalization, and let  $\tilde{C}$  be the smooth projective normalization. When  $C$  is geometrically integral, write

$$D_C = \tilde{C} \setminus C^\nu$$

for the geometric boundary.

**Proposition 4.4** (Componentwise  $S$ -integral value count). *Let  $S$  and  $T$  be finite sets of places of  $k$  containing the Archimedean places. Let  $C/k$  be an irreducible affine curve, let  $x \in k[C]$  be nonconstant, and let  $\mathcal{A} \subseteq C(k)$  be contained in the set of  $T$ -integral points for a fixed affine model of  $C$ . Define*

$$M_C(B; S, T) = \#\{a \in \mathcal{O}_{k,S} : H(a) \leq B, a = x(P) \text{ for some } P \in \mathcal{A}\}.$$

Then:

- (a) *If  $C$  is not geometrically integral, then  $M_C(B; S, T) = O_C(1)$ .*

- (b) If  $C$  is geometrically integral and either  $\text{genus}(\tilde{C}) > 0$  or  $\#D_C \geq 3$ , then  $M_C(B; S, T) = O_{C,k,T}(1)$ .
- (c) If  $\text{genus}(\tilde{C}) = 0$  but  $\tilde{C} \not\cong \mathbb{P}_k^1$ , then  $M_C(B; S, T) = O_C(1)$ .
- (d) If  $\tilde{C} \cong \mathbb{P}_k^1$  and  $\#D_C = 1$ , then there is a coordinate  $t$  on  $\tilde{C} \setminus D_C \cong \mathbb{A}^1$  such that

$$x = A(t), \quad A \in k[t].$$

Let  $d = \deg A$ . If  $A(k) \cap \mathcal{O}_{k,S} = \emptyset$ , then  $M_C(B; S, T) = O_C(1)$ . Otherwise

$$M_C(B; S, T) \ll_{C,k,S,T} B^{n_k/d} (\log(2B))^{q_{k,S}}.$$

- (e) If  $\tilde{C} \cong \mathbb{P}_k^1$  and  $\#D_C = 2$ , then there is a finite extension  $L/k$  splitting the two boundary points and a finite set of places  $U$  of  $L$ , depending only on  $C, S, T$  and the affine model, such that

$$M_C(B; S, T) \ll_{C,k,S,T} (\log(2B))^{\rho_C}, \quad \rho_C = \text{rk } \mathcal{O}_{L,U}^*.$$

*Proof.* Part (a) is [Lemma 4.3](#). Passing from  $C$  to its affine normalization changes the count by at most a constant, since the normalization is an isomorphism away from finitely many affine points.

Part (b) is Siegel–Mahler: an affine curve over a number field has finitely many  $T$ -integral points unless its normalization has genus zero and at most two points at infinity; see Lang [15, Ch. 7] or the subspace-theorem proof of Corvaja–Zannier [7].

For (c), a smooth projective genus-zero curve over  $k$  which is not isomorphic to  $\mathbb{P}_k^1$  has no  $k$ -rational point. Thus rational points of the affine curve can occur only in the finite exceptional locus where the normalization map is not an isomorphism.

For (d), the unique geometric point of  $D_C$  is Galois invariant, hence defined over  $k$ . Its complement is  $\mathbb{A}_k^1$ ; choose a coordinate  $t$ . Since  $x$  is regular away from the unique point at infinity,  $x = A(t)$  with  $A \in k[t]$ . The normalized part contributes only values in  $A(k) \cap \mathcal{O}_{k,S}$ , so if this set is empty the only possible contribution is from the finite normalization-exceptional locus.

Assume now that  $A(k) \cap \mathcal{O}_{k,S} \neq \emptyset$ . There exists a nonzero  $\delta \in k$ , depending only on  $A$  and  $S$ , such that

$$A(t) \in \mathcal{O}_{k,S} \implies \delta t \in \mathcal{O}_{k,S}.$$

Indeed, outside a fixed finite set of places containing  $S$ , the coefficients of  $A$  are integral and the leading coefficient is a unit; at such a place,  $v(t) < 0$  would force  $v(A(t)) < 0$ . At the remaining finitely many places outside  $S$ , the possible negative values of  $v(t)$  are uniformly bounded by the same leading-term comparison. A single  $\delta$  clears these bounded negative valuations. Thus the relevant parameters lie in a fixed fractional  $\mathcal{O}_{k,S}$ -ideal. If  $H(A(t)) \leq B$ , then [Lemma 4.2](#) gives  $H(t) \ll_A B^{1/d}$ . Applying [Lemma 4.1](#) to that fractional ideal gives the asserted upper bound.

For (e), pass to a finite extension  $L/k$  over which the two boundary points are rational. Choose a coordinate  $u$  sending them to 0 and  $\infty$ . The

affine normalization becomes  $\mathbb{G}_m$ , and  $x = A(u)$  for a Laurent polynomial  $A \in L[u, u^{-1}]$ . The normalization ring  $L[u, u^{-1}]$  is finite over the coordinate ring of the chosen affine model after base change to  $L$ . After enlarging  $T$  to a fixed finite set  $U$  of places of  $L$ , all structure constants for this finite integral extension are  $U$ -integral. If  $P \in \mathcal{A}$ , the values of the original affine coordinate ring at  $P$  are  $U$ -integral; since  $u$  and  $u^{-1}$  are integral over that ring, their values are integral over  $\mathcal{O}_{L,U}$ . The ring  $\mathcal{O}_{L,U}$  is integrally closed, so

$$u(P), u(P)^{-1} \in \mathcal{O}_{L,U}.$$

Thus  $u(P) \in \mathcal{O}_{L,U}^*$ . The absolute multiplicative Weil height is unchanged by passing from  $k$  to the finite extension  $L$ , since it is normalized with respect to the product formula. Therefore the condition  $H(x(P)) \leq B$ , with  $x(P) \in k$ , is the same height condition after viewing  $x(P)$  as an element of  $L$ . The Laurent polynomial  $A(u)$  defines a nonconstant rational map  $\mathbb{P}_L^1 \rightarrow \mathbb{P}_L^1$ . By the height-growth estimate for rational maps, [Lemma 4.2](#), the condition  $H(x(P)) = H(A(u(P))) \leq B$  gives  $H(u(P)) \ll_A B^{C_A}$  for some constant  $C_A > 0$ . Dirichlet's unit theorem then gives

$$\#\{u \in \mathcal{O}_{L,U}^* : H(u) \leq B^{C_A}\} = O((\log(2B))^{\rho_C}),$$

where  $\rho_C = \text{rk } \mathcal{O}_{L,U}^*$ . This proves (e).  $\square$

## 5. SPARSE EXCEPTIONAL LIFTING OVER NUMBER FIELDS

Let  $k$  be a number field, let  $S$  be a finite set of places containing the Archimedean places, and let  $f, g \in k[x]$  be nonconstant. Define

$$C_{f,g} = (f(X) - g(Y) = 0)_{\text{red}} \subset \mathbb{A}_k^2.$$

A component  $C$  of  $C_{f,g}$  is a *graph component* if

$$C = (Y = h(X))$$

for some  $h \in \mathcal{H}_{f,g}(k)$ . By [Theorem 3.1](#), these are exactly the components of generic degree one over  $\mathbb{A}_X^1$ . For a non-graph component  $C$ , put

$$d_X(C) = [k(C) : k(X)].$$

Equivalently,  $d_X(C)$  is the degree of the projection  $C \rightarrow \mathbb{A}_X^1$ . Since graph components have been removed,  $d_X(C) \geq 2$ .

Let  $\mathcal{R}_1(f, g; k)$  be the set of non-graph irreducible components  $C$  of  $C_{f,g}$  such that  $C$  is geometrically integral,  $\tilde{C} \cong \mathbb{P}_k^1$ , and  $\#D_C = 1$ . For  $C \in \mathcal{R}_1(f, g; k)$ , choose

$$X = A_C(t), \quad Y = B_C(t), \quad A_C, B_C \in k[t].$$

We call  $C$  *S-active* if  $A_C(k) \cap \mathcal{O}_{k,S} \neq \emptyset$ , and write

$$\mathcal{R}_1^{\text{act}}(f, g; k, S) = \{C \in \mathcal{R}_1(f, g; k) : A_C(k) \cap \mathcal{O}_{k,S} \neq \emptyset\}.$$

This definition is independent of the chosen affine coordinate  $t$ , because another coordinate has the form  $at+b$  with  $a \in k^*$ ,  $b \in k$ . Let  $\mathcal{R}_2(f, g; k)$  be the set of non-graph irreducible components  $C$  such that  $C$  is geometrically

integral,  $\tilde{C} \cong \mathbb{P}_k^1$ , and  $D_C$  consists of two geometric points. These two points may be conjugate over  $k$ .

**Lemma 5.1** (Uniform denominator control). *There is a finite set of places  $S' \supseteq S$ , depending only on  $f, g, k, S$ , such that for every  $a \in \mathcal{O}_{k,S}$  and every  $y \in k$ ,*

$$g(y) = f(a) \implies y \in \mathcal{O}_{k,S'}.$$

*Proof.* Let  $S'$  be obtained from  $S$  by adjoining the finite places at which some coefficient of  $f$  or  $g$  is not integral, or at which the leading coefficient of  $g$  is not a unit. Let  $v \notin S'$  be non-Archimedean and suppose  $a \in \mathcal{O}_{k,S}$ . Since  $S \subseteq S'$ , the element  $a$  is  $v$ -integral, and hence  $f(a)$  is  $v$ -integral. If  $v(y) < 0$ , the leading term of  $g(y)$  has strictly smaller  $v$ -adic valuation than all lower terms, because the leading coefficient is a  $v$ -adic unit and the other coefficients are  $v$ -integral. Thus  $v(g(y)) < 0$ , contradicting  $g(y) = f(a)$ . Hence  $v(y) \geq 0$  for all  $v \notin S'$ , which is precisely  $y \in \mathcal{O}_{k,S'}$ .  $\square$

Define the new lifting count

$$N_{f,g,k,S}^{\text{new}}(B) = \#\left\{a \in \mathcal{O}_{k,S} : H(a) \leq B, \exists y \in k \text{ with } g(y) = f(a), \right. \\ \left. y \neq h(a) \text{ for every } h \in \mathcal{H}_{f,g}(k)\right\}.$$

**Lemma 5.2** (New lifts lie on non-graph components). *Let  $a \in k$  and  $y \in k$  satisfy  $g(y) = f(a)$ . If  $y \neq h(a)$  for every  $h \in \mathcal{H}_{f,g}(k)$ , then  $(a, y)$  lies on at least one non-graph component of  $C_{f,g}$ , and on no graph component. Consequently, the inputs counted by  $N_{f,g,k,S}^{\text{new}}(B)$  are accounted for by non-graph components; component intersections affect only finitely many affine points and are harmless for the stated upper bounds.*

*Proof.* If  $(a, y)$  lies on a graph component  $Y = h(X)$ , then  $y = h(a)$ , contrary to the defining condition for a new lift. Since  $(a, y) \in C_{f,g}$ , it lies on some irreducible component; if no graph component can contain it, at least one component containing it is non-graph. Intersections of distinct components form a finite subset of the affine curve and therefore affect the input count by at most a constant.  $\square$

**Theorem 5.3** (Sparse exceptional lifting over number fields). *Let  $k$  be a number field, let  $S$  be a finite set of places containing the Archimedean places, and let  $f, g \in k[x]$  be nonconstant. Put*

$$q_{k,S} = \text{rk } \mathcal{O}_{k,S}^* = |S| - 1.$$

*Then there are nonnegative integers  $\rho_C$ , for  $C \in \mathcal{R}_2(f, g; k)$ , depending only on  $k, S, f, g$ , such that*

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} 1 + \sum_{C \in \mathcal{R}_1^{\text{act}}(f,g;k,S)} B^{n_k/d_X(C)} (\log(2B))^{q_{k,S}} \\ + \sum_{C \in \mathcal{R}_2(f,g;k)} (\log(2B))^{\rho_C}.$$

The integer  $\rho_C$  may be taken to be the rank of the  $U_C$ -unit group of a finite extension splitting the two boundary points of  $C$ , for a finite set  $U_C$  depending only on  $C$ ,  $S$ , and the affine model.

*Proof.* By [Lemma 5.1](#), choose a finite set of places  $S' \supseteq S$  such that every relevant lift satisfies  $y \in \mathcal{O}_{k,S'}$ . Thus the relevant pairs  $(a, y)$ , with  $a \in \mathcal{O}_{k,S}$ , are  $S'$ -integral points on the affine curve  $C_{f,g}$ . The original set  $S$  controls the  $X$ -coordinate being counted; the enlarged set  $S'$  is used only to place the points on a fixed integral model and to invoke Siegel–Mahler and unit estimates.

By [Lemma 5.2](#), the counted lifts are accounted for by non-graph components; graph components are exactly the excluded polynomial sections. Thus it suffices to sum over the non-graph components of  $C_{f,g}$ . For a non-graph component  $C$ , apply [Proposition 4.4](#) to the coordinate function  $x = X|_C$ , with the original set  $S$  in the counted  $X$ -coordinate and with  $T = S'$  for integrality of points on the curve. Non-geometrically integral components, components of positive genus, components with at least three points at infinity, and genus-zero components whose normalization is not  $\mathbb{P}_k^1$ , all contribute only  $O(1)$ .

If  $C \in \mathcal{R}_1(f, g; k)$ , the parametrization  $X = A_C(t)$  shows that the normalized one-infinity part contributes a power term only when  $A_C(k) \cap \mathcal{O}_{k,S} \neq \emptyset$ , i.e. only when  $C$  is  $S$ -active; an inactive component contributes only  $O(1)$ . In the active case [Proposition 4.4](#) gives the stated bound with logarithmic exponent  $q_{k,S} = |S| - 1$ . Components in  $\mathcal{R}_2(f, g; k)$  contribute the stated polylogarithmic bound. Since there are finitely many components, summing proves the theorem.  $\square$

**Corollary 5.4** (Power exponent over number fields). *If  $\mathcal{R}_1^{\text{act}}(f, g; k, S) \neq \emptyset$ , define*

$$\theta_{f,g,k,S} = \max_{C \in \mathcal{R}_1^{\text{act}}(f,g;k,S)} \frac{1}{d_X(C)}.$$

Then

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} B^{n_k \theta_{f,g,k,S}} (\log(2B))^{q_{k,S}} + (\log(2B))^c.$$

for some  $c \geq 0$ , coming only from two-infinity components. If  $\mathcal{R}_1^{\text{act}}(f, g; k, S) = \emptyset$ , then

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} (\log(2B))^c.$$

In all cases,

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} B^{n_k/2} (\log(2B))^{q_{k,S}} + (\log(2B))^c,$$

and if no active one-infinity component has  $d_X(C) = 2$ , this improves to

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} B^{n_k/3} (\log(2B))^{q_{k,S}} + (\log(2B))^c.$$

*Proof.* Every non-graph one-infinity component has  $d_X(C) \geq 2$ , because the  $d_X = 1$  components are exactly graph components by [Theorem 3.1](#). If no active component has  $d_X = 2$ , then all active one-infinity components have  $d_X \geq 3$ . The claims follow from [Theorem 5.3](#).  $\square$

## 6. ACTIVE COMPONENTS AND LOWER BOUNDS

The upper bound in [Theorem 5.3](#) distinguishes geometric possibility from arithmetic activity. We now show that activity gives the expected lower power.

**Lemma 6.1** (Integer-valued cosets). *Let  $A \in k[t]$ , and suppose  $A(t_0) \in \mathcal{O}_{k,S}$  for some  $t_0 \in k$ . Then there exists  $\lambda \in k^*$  such that*

$$A(t_0 + \lambda u) \in \mathcal{O}_{k,S} \quad (u \in \mathcal{O}_{k,S}).$$

*Proof.* Write the Taylor expansion

$$A(t_0 + Z) = A(t_0) + c_1 Z + \cdots + c_d Z^d, \quad c_i \in k.$$

Choose  $\lambda \in k^*$  so that  $c_i \lambda^i \in \mathcal{O}_{k,S}$  for all  $i$ . Then  $A(t_0 + \lambda u) \in \mathcal{O}_{k,S}$  for all  $u \in \mathcal{O}_{k,S}$ .  $\square$

**Theorem 6.2** (Lower bound from an active component). *Let  $k$  be a number field, let  $S$  contain the Archimedean places, and let  $f, g \in k[x]$  be nonconstant. Suppose  $C \in \mathcal{R}_1^{\text{act}}(f, g; k, S)$  is a non-graph component with  $d = d_X(C)$ . Then*

$$N_{f,g,k,S}^{\text{new}}(B) \gg_{k,S,f,g,C} B^{n_k/d} (\log(2B))^{q_{k,S}}.$$

*In particular, the power-log order attached to  $C$  is sharp.*

*Proof.* Choose a parametrization

$$X = A(t), \quad Y = B(t), \quad A, B \in k[t],$$

with  $\deg A = d$ . Since  $C$  is active, there exists  $t_0 \in k$  with  $A(t_0) \in \mathcal{O}_{k,S}$ . By [Lemma 6.1](#), choose  $\lambda \in k^*$  such that  $A(t_0 + \lambda u) \in \mathcal{O}_{k,S}$  for all  $u \in \mathcal{O}_{k,S}$ .

For  $H(u) \leq cB^{1/d}$ , with  $c > 0$  small enough, the height estimate gives  $H(A(t_0 + \lambda u)) \leq B$ . By [Lemma 4.1](#), there are  $\gg B^{n_k/d} (\log(2B))^{q_{k,S}}$  such parameters  $u$ . The polynomial  $A$  has degree  $d$ , so each  $x$ -value has at most  $d$  preimages in the parameter line.

It remains only to remove graph intersections. For a fixed  $h \in \mathcal{H}_{f,g}(k)$ , the identity

$$B(t) = h(A(t))$$

can hold for infinitely many  $t$  only if the component  $C$  is contained in the graph  $Y = h(X)$ , contrary to the hypothesis that  $C$  is non-graph. Since there are finitely many graph components, excluding all graph intersections removes only finitely many parameters. The remaining parameters give distinct new lifts counted by  $N_{f,g,k,S}^{\text{new}}(B)$ , up to the bounded multiplicity of  $A$ . This proves the lower bound.  $\square$

**Corollary 6.3** (Order of magnitude when an active component exists). *If  $\mathcal{R}_1^{\text{act}}(f, g; k, S) \neq \emptyset$  and*

$$\theta_{f,g,k,S} = \max_{C \in \mathcal{R}_1^{\text{act}}(f,g;k,S)} \frac{1}{d_X(C)},$$

*then*

$$N_{f,g,k,S}^{\text{new}}(B) \asymp_{k,S,f,g} B^{n_k \theta_{f,g,k,S}} (\log(2B))^{q_{k,S}}.$$

If there is no active one-infinity component, then the only possible infinite contribution is the polylogarithmic contribution from two-infinity components.

*Proof.* The upper bound is [Corollary 5.4](#). For the lower bound, choose an active component  $C$  with  $1/d_X(C) = \theta_{f,g,k,S}$  and apply [Theorem 6.2](#). Since the number of components is finite, overlaps among components can only decrease the componentwise upper sum and do not affect the lower bound obtained from this single component.  $\square$

**Corollary 6.4** (Sharp integer exponents over  $\mathbb{Q}$ ). *Let  $f, g \in \mathbb{Q}[x]$  be non-constant. Then*

$$N_{f,g}^{\text{new}}(B) \ll_{f,g} 1 + \sum_{C \in \mathcal{R}_1^{\text{act}}(f,g;\mathbb{Q},\{\infty\})} B^{1/d_X(C)} + \sum_{C \in \mathcal{R}_2(f,g;\mathbb{Q})} (\log(2B))^{\rho_C}.$$

If  $\mathcal{R}_1^{\text{act}}(f,g;\mathbb{Q},\{\infty\}) \neq \emptyset$  and

$$\theta_{f,g} = \max_{C \in \mathcal{R}_1^{\text{act}}(f,g;\mathbb{Q},\{\infty\})} \frac{1}{d_X(C)},$$

then

$$B^{\theta_{f,g}} \ll_{f,g} N_{f,g}^{\text{new}}(B) \ll_{f,g} B^{\theta_{f,g}} + (\log(2B))^c.$$

If there is no active one-infinity component, then

$$N_{f,g}^{\text{new}}(B) \ll_{f,g} (\log(2B))^c.$$

In particular, square-root growth

$$N_{f,g}^{\text{new}}(B) \asymp_{f,g} B^{1/2}$$

occurs if and only if there is an active non-graph rational one-infinity component with  $d_X(C) = 2$ . If no such component exists, then

$$N_{f,g}^{\text{new}}(B) = O_{f,g}(B^{1/3}) + O_{f,g}((\log(2B))^c).$$

*Proof.* This is [Theorem 5.3](#) and [Theorem 6.2](#) with  $k = \mathbb{Q}$  and  $S = \{\infty\}$ . For this  $S$ , there is no logarithmic factor in the one-infinity power terms. Since non-graph components have  $d_X(C) \geq 2$ , the maximal possible exponent is  $1/2$ , and it is attained exactly by an active component with  $d_X(C) = 2$ .  $\square$

**Corollary 6.5** (No-composition form). *If  $f, g \in \mathbb{Q}[x]$  are nonconstant and there is no  $h \in \mathbb{Q}[x]$  with  $f = g \circ h$ , then*

$$\#\{n \in \mathbb{Z} : |n| \leq B, f(n) \in g(\mathbb{Q})\} = O_{f,g}(B^{1/2}).$$

The refined componentwise bound of [Corollary 6.4](#) also holds.

*Proof.* Under the no-composition hypothesis  $\mathcal{H}_{f,g}(\mathbb{Q}) = \emptyset$ , so  $N_{f,g}^{\text{new}}(B)$  is exactly the displayed count.  $\square$

**Example 6.6** (Sharpness of every exponent). For  $d \geq 2$ , take

$$f(X) = X, \quad g(Y) = Y^d.$$

There is no  $h \in \mathbb{Q}[X]$  with  $X = h(X)^d$ . The curve  $X = Y^d$  is rational, has one geometric point at infinity, and is parametrized by

$$X = t^d, \quad Y = t.$$

Thus  $d_X = d$ , and the integer inputs which lift are the integer  $d$ -th powers, up to the usual sign convention when  $d$  is odd. The count is  $\asymp B^{1/d}$ . Over a number field, the same example gives the sharp order

$$B^{[k:\mathbb{Q}]/d}(\log B)^{q_{k,S}}.$$

For instance, over a real quadratic field with  $S = S_\infty$ , the logarithmic factor is  $\log B$ , because  $q_{k,S} = 1$ .

**Example 6.7** (A square-root singular component). For

$$f(X) = X^3, \quad g(Y) = Y^2,$$

the curve  $X^3 = Y^2$  is parametrized by

$$X = t^2, \quad Y = t^3.$$

It is rational with one geometric point at infinity and  $d_X = 2$ , so it gives square-root growth. The square-root source need not be a smooth conic; singular rational components also occur.

**Example 6.8** (A local obstruction). For

$$f(X) = X, \quad g(Y) = Y^2 + \frac{1}{2},$$

the curve is rational and has the parametrization

$$X = t^2 + \frac{1}{2}, \quad Y = t.$$

However it produces no integer inputs. If  $n = y^2 + 1/2$  with  $n \in \mathbb{Z}$  and  $y = a/b \in \mathbb{Q}$  in lowest terms, then

$$2nb^2 = 2a^2 + b^2.$$

If  $b$  is odd, the right side is odd and the left side is even. If  $b$  is even, then  $a$  is odd, and after writing  $b = 2c$  the equation becomes

$$4nc^2 = a^2 + 2c^2,$$

whose two sides have opposite parity. Thus no such rational  $y$  exists. This explains why the upper-bound theorem must distinguish component geometry from arithmetic activity.

**Example 6.9** (A reducible example with graph removal). Let

$$f(X) = X^4, \quad g(Y) = Y^4.$$

Then

$$f(X) - g(Y) = (X - Y)(X + Y)(X^2 + Y^2).$$

The first two factors are graph components, corresponding to  $Y = X$  and  $Y = -X$ . The remaining factor is irreducible over  $\mathbb{Q}$  but not geometrically integral; over  $\mathbb{Q}(i)$  it splits into the two graph lines  $Y = \pm iX$ . Hence it contributes only finitely many rational points, in fact only  $(0, 0)$  over  $\mathbb{Q}$ . Thus after graph removal one has  $N_{f,g}^{\text{new}}(B) = O(1)$ . This example illustrates why graph components and geometrically reducible components are separated before the sparse exponent is read off.

**Example 6.10** (A Chebyshev standard-pair component). Let  $T_2(z) = 2z^2 - 1$  and  $T_3(z) = 4z^3 - 3z$ . Take

$$f(X) = T_2(X), \quad g(Y) = T_3(Y).$$

The identity  $T_2(T_3(t)) = T_3(T_2(t))$  gives a rational component of  $T_2(X) - T_3(Y) = 0$  parametrized by

$$X = T_3(t), \quad Y = T_2(t).$$

It is birational because  $X = t(2Y - 1)$ , so  $t = X/(2Y - 1)$  in the function field. The component is not a graph component:  $T_2(t) \notin \mathbb{Q}[T_3(t)]$ , since a nonconstant polynomial in  $T_3(t)$  has degree divisible by 3. Thus it is an active one-infinity component with  $d_X = 3$ , and it gives

$$N_{f,g}^{\text{new}}(B) \gg B^{1/3}.$$

This illustrates how the theorem reads a Bilu–Tichy-type component as a sparse integer exponent without attempting to reclassify the component.

**Example 6.11** (A two-infinity Pell component). Let  $D > 1$  be a squarefree integer, and take

$$f(X) = X^2, \quad g(Y) = DY^2 + 1.$$

The curve  $f(X) = g(Y)$  is

$$X^2 - DY^2 = 1.$$

It is rational over  $\mathbb{Q}$ , has two geometric points at infinity, and is not a graph component. Its integer points are Pell solutions. For every nonsquare  $D > 1$ , the norm-one unit group in  $\mathbb{Q}(\sqrt{D})$  has rank one; even when a chosen fundamental unit has norm  $-1$ , its square has norm 1. Hence the integer solutions give  $\asymp_D \log B$  inputs  $|X| \leq B$ , in agreement with the two-infinity polylogarithmic term in [Corollary 6.4](#). This example also shows why the one-infinity power terms and the two-infinity unit terms must be separated.

## 7. COMPUTING THE EXPONENT

For explicit polynomials over  $\mathbb{Q}$ , the exponent in [Corollary 6.4](#) can be extracted by a finite procedure. The procedure is not a new factorization algorithm; it is an explicit translation of the theorem into standard operations on plane curves: factorization, normalization, genus and boundary computation, parametrization of rational components, and a finite integrality test for the  $X$ -parametrization. General algorithmic criteria for infinitude of integral points on affine curves are developed by Alvanos–Bilu–Poulakis [\[1\]](#);

the test below is the special one-variable denominator reduction needed for the activity condition in the present counting theorem.

**Lemma 7.1** (Finite denominator test over  $\mathbb{Q}$ ). *Let  $A \in \mathbb{Q}[t]$  be nonconstant. There is an effectively computable integer  $M \geq 1$  such that, whenever  $t = a/b \in \mathbb{Q}$  is in lowest terms and  $A(t) \in \mathbb{Z}$ , one has  $b \mid M$ . Consequently, the condition  $A(\mathbb{Q}) \cap \mathbb{Z} \neq \emptyset$  can be checked by finitely many explicit polynomial congruence conditions.*

*Proof.* Write  $A(t) = q^{-1}P(t)$ , with  $P \in \mathbb{Z}[t]$ ,  $q \in \mathbb{Z}_{>0}$ , and let  $m = \deg P$ . Let  $p$  be a prime. If the coefficients of  $P$  are  $p$ -integral and the leading coefficient is a  $p$ -adic unit, then  $v_p(t) < 0$  forces  $v_p(P(t)) = mv_p(t) < 0$ , because the leading term has strictly smaller valuation than every lower term. Hence, at primes not dividing  $q$  or the leading coefficient of  $P$ , integrality of  $A(t)$  forces  $v_p(t) \geq 0$ .

It remains only to bound denominators at the finite set of bad primes. This is effective. Fix such a prime  $p$ . For each lower term  $p_i t^i$  with  $i < m$ , the inequality

$$v_p(p_m) + mv_p(t) < v_p(p_i) + iv_p(t)$$

holds for all sufficiently negative  $v_p(t)$ ; the explicit threshold is obtained by taking the maximum of the finitely many ratios  $(v_p(p_m) - v_p(p_i))/(m - i)$ , together with the valuation of  $q$ . Below this threshold the leading term strictly dominates, so  $v_p(A(t)) < 0$ . Therefore  $v_p(t)$  is bounded below by a computable integer for every bad prime. Taking  $M$  to contain the corresponding prime powers clears all possible bounded negative valuations and gives the first claim.

For each divisor  $b \mid M$ , the condition  $A(a/b) \in \mathbb{Z}$ , with  $(a, b) = 1$ , is completely explicit. Namely, write

$$P_b(T) = b^m P(T/b) = \sum_i p_i T^i b^{m-i} \in \mathbb{Z}[T],$$

where  $P(T) = \sum_i p_i T^i$ . Then

$$A(a/b) \in \mathbb{Z} \iff qb^m \mid P_b(a).$$

For fixed  $b$ , this divisibility depends only on the residue class of  $a$  modulo  $qb^m$ , together with the condition  $(a, b) = 1$ . Thus activity is reduced to finitely many polynomial congruence checks.  $\square$

- Step 1. Factor  $P(X, Y) = f(X) - g(Y)$  in  $\mathbb{Q}[X, Y]$ , and take the squarefree part.
- Step 2. Remove graph factors. Equivalently, remove the factors  $Y - h(X)$  with  $h \in \mathbb{Q}[X]$  and  $f = g \circ h$ . These are exactly the components of degree one over the  $X$ -line.
- Step 3. For each remaining factor, test geometric irreducibility. A factor irreducible over  $\mathbb{Q}$  but not geometrically integral contributes only finitely many rational points.
- Step 4. For each geometrically integral component, compute the genus of the smooth projective normalization and the number of boundary points above the line at infinity of the affine model.

Step 5. Keep the rational one-infinity components. For each such component, choose a polynomial parametrization

$$X = A(t), \quad Y = B(t), \quad A, B \in \mathbb{Q}[t],$$

and record  $d_X(C) = \deg A$ .

Step 6. Test activity by checking whether  $A(\mathbb{Q}) \cap \mathbb{Z}$  is nonempty. If  $A(t_0) \in \mathbb{Z}$ , [Lemma 6.1](#) constructs a rational arithmetic progression on which  $A$  is integer-valued. Conversely, if no such  $t_0$  exists, the normalized one-infinity part contributes no integer inputs, apart from the finite normalization-exceptional locus.

Step 7. The power exponent is

$$\theta_{f,g} = \max_{C \text{ active}} \frac{1}{d_X(C)}.$$

If no active one-infinity component survives, only the polylogarithmic contribution from two-infinity rational components remains.

The finite nature of Step 6 is justified by [Lemma 7.1](#). The procedure is effective for the exponent once the standard curve operations above have been carried out. The finite exceptional sets whose finiteness comes from Siegel–Mahler are not claimed to be effectively enumerated here; the paper uses them only through their  $O(1)$  contribution to the height count.

## 8. QUADRATIC SOURCES AND THE SQUARE-ROOT BOUNDARY

We now prove the reduction stated in the introduction.

**Proposition 8.1** (Quadratic-source reformulation). *Let  $k$  be a field of characteristic zero, and let  $f, g \in k[x]$  be nonconstant. The curve  $C_{f,g}$  has a non-graph rational one-infinity component  $C$  with  $d_X(C) = 2$  if and only if there exist polynomials  $A, B \in k[t]$  such that*

$$\deg A = 2, \quad k(A(t), B(t)) = k(t),$$

$$f(A(t)) = g(B(t)),$$

and

$$B(t) \notin k[A(t)].$$

*Equivalently, after a linear change of the parameter and a linear change of the  $X$ -coordinate, one may write*

$$A(t) = \alpha t^2 + \beta, \quad \alpha \in k^*, \beta \in k.$$

*Proof.* Suppose first that  $C$  is such a component. Its smooth projective normalization is  $\mathbb{P}_k^1$ , and the affine normalization has one point at infinity. Hence there is a coordinate  $t$  on the affine normalization for which

$$X = A(t), \quad Y = B(t), \quad A, B \in k[t].$$

The projection degree to the  $X$ -line is  $\deg A$ , so  $d_X(C) = 2$  gives  $\deg A = 2$ . Since  $t$  is a normalization parameter,

$$k(C) = k(t) = k(A(t), B(t)).$$

The identity  $f(X) = g(Y)$  on  $C$  gives  $f(A(t)) = g(B(t))$ . If  $B(t) \in k[A(t)]$ , then  $Y$  is a polynomial function of  $X$ , so  $C$  is a graph component. Thus  $B(t) \notin k[A(t)]$ .

Conversely, such  $A, B$  define a rational curve contained in  $C_{f,g}$ . Its Zariski closure is irreducible, and since  $C_{f,g}$  is a curve this closure is an irreducible component of  $C_{f,g}$ . The condition  $k(A, B) = k(t)$  says that the parametrization is birational onto its image. Since  $\deg A = 2$ , the projection degree to the  $X$ -line is 2. The condition  $B \notin k[A]$  excludes graph components. Finally, because  $A, B \in k[t]$ , the affine normalization has one geometric point at infinity: indeed  $k[t]$  is integral over  $k[A, B]$ , has the same fraction field, and is integrally closed, so it is the normalization of the affine coordinate ring.

In characteristic zero, every quadratic polynomial can be completed to a square. After a linear change of the parameter and a linear change of the  $X$ -coordinate we may therefore write  $A(t) = \alpha t^2 + \beta$ , with  $\alpha \neq 0$ .  $\square$

The following classical cancellation result is used only in the equal-outer-degree case needed for quadratic sources.

**Lemma 8.2** (Engstrom cancellation in equal outer degree). *Let  $k$  be a field of characteristic zero. Suppose  $P, Q, R, S \in k[x]$  are nonconstant polynomials satisfying*

$$P \circ Q = R \circ S, \quad \deg P = \deg R.$$

*Then there is a linear polynomial  $\ell \in k[x]$  such that*

$$P = R \circ \ell, \quad S = \ell \circ Q.$$

*In particular, if  $P \circ Q = P \circ S$ , then*

$$S = \ell \circ Q, \quad P \circ \ell = P$$

*for some linear  $\ell$ .*

*Proof.* This is the equal-outer-degree case of Engstrom's cancellation theorem for polynomial substitutions [12]. The theorem applied over  $\bar{k}$  gives a linear polynomial  $\ell(Y) = aY + b$ . Since  $Q, S \in k[x]$  and  $S = \ell \circ Q$ , comparison of leading coefficients gives  $a \in k$ , and then comparison of constant terms gives  $b \in k$ . Hence  $\ell$  is defined over  $k$ .  $\square$

**Theorem 8.3** (Quadratic sources and affine involutions). *Let  $k$  be a field of characteristic zero, and let  $f, g \in k[x]$  be nonconstant. The following are equivalent.*

- (i) *The curve  $C_{f,g}$  has a non-graph rational one-infinity component  $C$  with  $d_X(C) = 2$ .*
- (ii) *There exist  $\alpha \in k^*$ ,  $\beta, c \in k$ , a nonzero polynomial  $E \in k[U]$ , and a nonconstant polynomial  $G \in k[Z]$  such that*

$$g(Y) = G((Y - c)^2)$$

and

$$f(\alpha U + \beta) = G(UE(U)^2).$$

In this case a corresponding component is parametrized by

$$X = \alpha t^2 + \beta, \quad Y = c + tE(t^2).$$

Consequently, if  $g$  has no nontrivial affine involution  $Y \mapsto 2c - Y$  with  $g(2c - Y) = g(Y)$ , then  $C_{f,g}$  has no non-graph rational one-infinity component with  $d_X = 2$ . In particular, if  $\deg g$  is odd, no such component exists.

*Proof.* Assume first that  $C$  is a non-graph rational one-infinity component with  $d_X(C) = 2$ . By [Proposition 8.1](#), after a linear change of the parameter and of the  $X$ -coordinate we may write

$$X = A(t) = \alpha t^2 + \beta, \quad Y = B(t),$$

with  $\alpha \neq 0$ ,  $k(A, B) = k(t)$ ,  $f(A(t)) = g(B(t))$ , and  $B(t) \notin k[A(t)] = k[t^2]$ . Since  $A(t) = A(-t)$ , we have

$$g(B(t)) = g(B(-t)).$$

The two polynomials  $B(t)$  and  $B(-t)$  have the same degree. Applying [Lemma 8.2](#) to  $g \circ B = g \circ (B \circ (-1))$ , there is a linear polynomial  $\ell \in k[Y]$  such that

$$B(-t) = \ell(B(t)), \quad g \circ \ell = g.$$

Applying  $t \mapsto -t$  again gives  $B(t) = \ell^2(B(t))$ , hence  $\ell^2 = \text{id}$ . The map  $\ell$  is not the identity, because otherwise  $B(t) = B(-t)$  and  $B \in k[t^2] = k[A]$ , contradicting the non-graph condition. Thus  $\ell$  is a nontrivial affine involution. In characteristic zero,  $\ell(Y) = 2c - Y$  for some  $c \in k$ . Its invariant ring is

$$k[Y]^\ell = k[(Y - c)^2],$$

so  $g \circ \ell = g$  is equivalent to

$$g(Y) = G((Y - c)^2)$$

for some  $G \in k[Z]$ . The identity  $B(-t) = 2c - B(t)$  says that  $B(t) - c$  is odd, hence

$$B(t) = c + tE(t^2)$$

with  $E \in k[U]$ ,  $E \neq 0$ . Substituting  $U = t^2$  into  $f(A(t)) = g(B(t))$  gives

$$f(\alpha U + \beta) = G(UE(U)^2).$$

This proves (ii).

Conversely, assume (ii). Then

$$X = \alpha t^2 + \beta, \quad Y = c + tE(t^2)$$

defines a rational curve contained in  $C_{f,g}$ , since

$$g(Y) = G((tE(t^2))^2) = G(t^2E(t^2)^2) = f(\alpha t^2 + \beta).$$

Because  $E \neq 0$ , the function field generated by  $t^2$  and  $tE(t^2)$  is  $k(t)$ : explicitly,  $t = (tE(t^2))/E(t^2)$ . Therefore the parametrization is birational onto its image. The Zariski closure of the image is an irreducible component of  $C_{f,g}$ ,

its projection degree to the  $X$ -line is two, and its affine normalization has one geometric point at infinity. Finally,  $c + tE(t^2) \notin k[t^2]$ , so the component is not a graph component. This proves (i).

The final assertions follow immediately. A nontrivial affine involution has the form  $Y \mapsto 2c - Y$ , and its invariants are polynomials in  $(Y - c)^2$ . If  $\deg g$  is odd,  $g(Y) = G((Y - c)^2)$  is impossible for nonconstant  $G$ .  $\square$

**Corollary 8.4** (Explicit fixed- $g$  square-root sources). *Let  $k$  be a field of characteristic zero and let  $g \in k[Y]$  be nonconstant. The polynomials  $f \in k[X]$  for which  $C_{f,g}$  has a non-graph rational one-infinity component with  $d_X = 2$  are precisely the following. There exist  $c, \beta \in k$ ,  $\alpha \in k^*$ , a nonzero polynomial  $E \in k[U]$ , and a nonconstant polynomial  $G \in k[Z]$  such that*

$$g(Y) = G((Y - c)^2)$$

and

$$f(X) = G\left(\frac{X - \beta}{\alpha} E\left(\frac{X - \beta}{\alpha}\right)^2\right).$$

The corresponding component is parametrized by

$$X = \alpha t^2 + \beta, \quad Y = c + tE(t^2).$$

In particular, once the involutive quotient  $g(Y) = G((Y - c)^2)$  is fixed, quadratic sources are completely parametrized by the choice of  $\alpha, \beta$  and  $E$ .

*Proof.* This is [Theorem 8.3](#) with  $U = (X - \beta)/\alpha$ , rewritten as a formula for  $f$ . Conversely, the displayed formula gives

$$f(\alpha t^2 + \beta) = G(t^2 E(t^2)^2) = g(c + tE(t^2)),$$

and  $k(t^2, tE(t^2)) = k(t)$ , so the parametrized curve is a non-graph rational one-infinity component of projection degree two.  $\square$

**Corollary 8.5** (Monomial square-root sources). *Let  $k$  be a field of characteristic zero and let  $g(Y) = Y^m$  with  $m \geq 2$ . If  $m$  is odd, then no non-graph rational one-infinity component with  $d_X = 2$  can occur for any  $f \in k[X]$ . If  $m$  is even, then such components occur exactly for polynomials of the form*

$$f(X) = \left(\frac{X - \beta}{\alpha}\right)^{m/2} E\left(\frac{X - \beta}{\alpha}\right)^m, \quad \alpha \in k^*, \beta \in k,$$

with  $E \in k[U] \setminus \{0\}$ . The corresponding component is

$$X = \alpha t^2 + \beta, \quad Y = tE(t^2).$$

*Proof.* For  $g(Y) = Y^m$ , the condition  $g(Y) = G((Y - c)^2)$  is impossible when  $m$  is odd. When  $m$  is even, the only affine involution fixing  $Y^m$  in characteristic zero is  $Y \mapsto -Y$ , so  $c = 0$  and  $G(Z) = Z^{m/2}$ . Apply [Corollary 8.4](#).  $\square$

**Corollary 8.6** (Generic absence of square-root sources). *Fix an integer  $m \geq 3$ . For a Zariski-generic polynomial  $g \in k[Y]$  of degree  $m$ , there is no non-graph rational one-infinity component with  $d_X = 2$  in  $C_{f,g}$ , for any polynomial  $f \in k[X]$ . More precisely, the exceptional set of degree- $m$*

polynomials  $g$  for which such a source is not excluded is contained in the proper algebraic locus of polynomials satisfying  $g(Y) = G((Y - c)^2)$  for some  $c \in k$  and  $G \in k[Z]$ .

*Proof.* By [Theorem 8.3](#), a square-root source forces  $g(Y) = G((Y - c)^2)$ . We work in the coefficient space  $\mathcal{P}_m \subset \mathbb{A}_k^{m+1}$  of degree-exactly- $m$  polynomials, where the leading coefficient is nonzero. If  $m$  is odd, no polynomial of the form  $G((Y - c)^2)$  has degree  $m$ , so the exceptional locus is empty.

Assume now that  $m = 2r \geq 4$ . Consider the incidence parametrizing source-even degree- $m$  polynomials:

$$(c, G) \mapsto G((Y - c)^2), \quad c \in \mathbb{A}^1, \quad G \in k[Z], \quad \deg G = r.$$

The parameter space has dimension at most  $1 + (r + 1) = r + 2$ . Its image in  $\mathcal{P}_m$  is constructible, and the Zariski closure of that image has dimension at most  $r + 2$ . But

$$\dim \mathcal{P}_m = m + 1 = 2r + 1,$$

and  $r + 2 < 2r + 1$  for  $r \geq 2$ . Hence the Zariski closure of the source-even locus is a proper algebraic subset of the coefficient space. On the complement of this proper closed subset, [Theorem 8.3](#) excludes square-root sources. Here genericity is taken in the coefficient space of degree- $m$  polynomials.  $\square$

**Definition 8.7** (Source-even polynomials). Let  $k$  be a field of characteristic zero. We say that  $g \in k[Y]$  is *source-even over  $k$*  if there are  $c \in k$  and  $G \in k[Z]$  such that

$$g(Y) = G((Y - c)^2).$$

Equivalently,  $g$  is invariant under the nontrivial affine involution  $Y \mapsto 2c - Y$ . This terminology avoids the phrase “linearly conjugate,” which can mean a different two-sided operation in dynamics.

**Proposition 8.8** (Detecting source-even polynomials). *Let  $k$  be a field of characteristic zero, and let*

$$g(Y) = a_n Y^n + a_{n-1} Y^{n-1} + \cdots + a_0 \in k[Y]$$

*be nonconstant. If  $g$  is source-even over  $k$ , then  $n$  is even and the center  $c$  is unique. For even  $n$ , the only possible center is*

$$c = -\frac{a_{n-1}}{na_n}.$$

*Thus source-evenness is checked by one finite coefficient test: translate by this candidate  $c$  and verify that all odd powers of  $Z$  in  $g(c + Z)$  vanish.*

*Proof.* Source-evenness is the existence of an affine reflection  $Y \mapsto 2c - Y$  fixing  $g$ , equivalently  $g(c + Z) = g(c - Z)$ . If two different centers  $c_1, c_2$  worked, then the two corresponding reflections would generate the nontrivial translation  $Y \mapsto Y + 2(c_2 - c_1)$ . A nonconstant polynomial in characteristic zero cannot be invariant under a nontrivial translation, so the center is unique.

If  $g(c + Z)$  is even, its degree must be even. Also the coefficient of  $Z^{n-1}$  in  $g(c + Z)$  must vanish. That coefficient is  $na_n c + a_{n-1}$ , giving the displayed

formula for  $c$ . Once this unique candidate is known, the condition is exactly the vanishing of the finitely many odd coefficients in  $g(c + Z)$ .  $\square$

**Corollary 8.9** (A practical square-root obstruction). *Let  $k$  be a number field, let  $S$  contain the Archimedean places, and let  $f, g \in k[x]$  be nonconstant. If  $g$  is not source-even over  $k$ , then*

$$N_{f,g,k,S}^{\text{new}}(B) \ll_{k,S,f,g} B^{[k:\mathbb{Q}]/3} (\log(2B))^{q_{k,S}} + (\log(2B))^c$$

for some  $c \geq 0$ . In particular, no square-root-order new lifting set can occur over  $\mathbb{Q}$ . If  $\deg g$  is odd, then  $g$  is not source-even.

*Proof.* By [Theorem 8.3](#), a quadratic one-infinity source forces  $g(Y) = G((Y - c)^2)$  for some  $c \in k$ , i.e.  $g$  is source-even over  $k$ . If this does not happen, there is no active component with  $d_X = 2$ , and [Corollary 5.4](#) gives the asserted bound. The degree of a source-even polynomial is even, so odd degree excludes source-evenness.  $\square$

*Remark 8.10.* [Proposition 8.1](#) reduces square-root sources to nontrivial quadratic polynomial pullback identities

$$f(\alpha t^2 + \beta) = g(B(t)), \quad k(\alpha t^2 + \beta, B(t)) = k(t), \quad B(t) \notin k[\alpha t^2 + \beta].$$

The source-even theorem above classifies the quadratic sources at the level needed for the height-counting problem: they are exactly the identities coming from an involutive quotient  $g(Y) = G((Y - c)^2)$ . A different, more taxonomic problem is to place these identities inside the Bilu–Tichy standard-pair families after imposing  $k$ -rationality and arithmetic activity. Bilu’s theorem on quadratic factors is relevant low-degree background, but it concerns total degree of a factor rather than the projection degree  $d_X$ , so it does not directly settle this projection-degree-two taxonomy.

## 9. FINAL REMARKS

The main theorem should be read as a componentwise sparse-counting refinement of the Hilbertian value-set lifting principle. It does not decide which separated-variable components occur; rather, once those components are known, it determines which ones can contribute integer or  $S$ -integer lifting inputs and with which exponent. The additional symmetry theorem in [Theorem 8.3](#), together with the source-even test in [Proposition 8.8](#), gives the square-root source classification needed for this paper: every square-root source is obtained from an involutive quotient  $g(Y) = G((Y - c)^2)$  and a polynomial  $E$  through

$$X = \alpha t^2 + \beta, \quad Y = c + tE(t^2).$$

A finer classification of which such identities fall into particular Bilu–Tichy standard-pair families is a separate classification problem, not a needed input for the componentwise height theorem. The natural remaining task is therefore narrower than the square-root classification itself: the present paper already gives the source-even normal form, and what remains is to compare that normal form with the standard-pair taxonomy and with the arithmetic activity condition.

*Question 9.1.* How does the fixed- $g$  source-even family of [Corollary 8.4](#) decompose inside the Bilu–Tichy standard-pair taxonomy after imposing  $k$ -rationality, removing graph components, and imposing  $S$ -activity? In particular, over  $\mathbb{Q}$ , can one give a short list of the source-even standard-pair subfamilies for which the associated quadratic  $X$ -parametrization is arithmetically active?

## REFERENCES

- [1] P. Alvanos, Y. F. Bilu, and D. Poulakis, *Characterizing algebraic curves with infinitely many integral points*, Int. J. Number Theory **5** (2009), no. 4, 585–590, DOI 10.1142/S1793042109002274.
- [2] R. M. Avanzi and U. M. Zannier, *Genus one curves defined by separated variable polynomials and a polynomial Pell equation*, Acta Arith. **99** (2001), no. 3, 227–256, DOI 10.4064/aa99-3-2.
- [3] F. Barroero, *Algebraic  $S$ -integers of fixed degree and bounded height*, Acta Arith. **167** (2015), no. 1, 67–90, DOI 10.4064/aa167-1-4.
- [4] Y. F. Bilu, *Quadratic factors of  $f(x) - g(y)$* , Acta Arith. **90** (1999), no. 4, 341–355, DOI 10.4064/aa-90-4-341-355.
- [5] Y. F. Bilu and R. F. Tichy, *The Diophantine equation  $f(x) = g(y)$* , Acta Arith. **95** (2000), no. 3, 261–288, DOI 10.4064/aa-95-3-261-288.
- [6] A. Behajaina, J. König, and D. Neftin, *The Davenport–Lewis–Schinzel problem on the reducibility of  $f(X) - g(Y)$* , preprint, arXiv:2603.27728v1 (2026).
- [7] P. Corvaja and U. Zannier, *A subspace theorem approach to integral points on curves*, C. R. Acad. Sci. Paris Sér. I Math. **334** (2002), no. 4, 267–271.
- [8] H. Davenport, D. J. Lewis, and A. Schinzel, *Equations of the form  $f(x) = g(y)$* , Quart. J. Math. Oxford Ser. (2) **12** (1961), 304–312, DOI 10.1093/qmath/12.1.304.
- [9] H. Davenport, D. J. Lewis, and A. Schinzel, *Polynomials of certain special types*, Acta Arith. **9** (1964), no. 1, 107–116, DOI 10.4064/aa-9-1-107-116.
- [10] R. Dvornicich and U. Zannier, *Fields containing values of algebraic functions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **21** (1994), no. 3, 421–443.
- [11] R. Dvornicich and U. Zannier, *Fields containing values of algebraic functions. II. On a conjecture of Schinzel*, Acta Arith. **72** (1995), no. 3, 201–210.
- [12] H. T. Engstrom, *Polynomial substitutions*, Amer. J. Math. **63** (1941), 249–255, DOI 10.2307/2371520.
- [13] M. D. Fried and M. Jarden, *Field Arithmetic*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 11, Springer, 2008.
- [14] J. König and D. Neftin, *Reducible fibers of polynomial maps*, Int. Math. Res. Not. IMRN **2024**, no. 6, 5373–5402, DOI 10.1093/imrn/rnad251.
- [15] S. Lang, *Fundamentals of Diophantine Geometry*, Springer, 1983.
- [16] S. H. Schanuel, *Heights in number fields*, Bull. Soc. Math. France **107** (1979), 433–449, DOI 10.24033/bsmf.1905.
- [17] J.-P. Serre, *Topics in Galois Theory*, 2nd ed., Research Notes in Mathematics, vol. 1, A K Peters, 2008.

INDEPENDENT RESEARCHER

*Email address:* hkshin@gmail.com