A TOPOLOGICAL APPROACH TO UNDEFINABILITY IN ALGEBRAIC EXTENSIONS OF $\mathbb{Q}$

KIRSTEN EISENTRÄGER, RUSSELL MILLER, CALEB SPRINGER, AND LINDA WESTRICK

ABSTRACT. For any subset $Z \subseteq \mathbb{Q}$, consider the set S_Z of subfields $L \subseteq \overline{\mathbb{Q}}$ which contain a co-infinite subset $C \subseteq L$ that is universally definable in L such that $C \cap \mathbb{Q} = Z$. Placing a natural topology on the set $\mathrm{Sub}(\overline{\mathbb{Q}})$ of subfields of $\overline{\mathbb{Q}}$, we show that if Z is not thin in $\mathbb{Q}$, then S_Z is meager in $\mathrm{Sub}(\overline{\mathbb{Q}})$. Here, *thin* and *meager* both mean "small", in terms of arithmetic geometry and topology, respectively. For example, this implies that only a meager set of fields L have the property that the ring of algebraic integers $\mathcal{O}_L$ is universally definable in L. The main tools are Hilbert's Irreducibility Theorem and a new normal form theorem for existential definitions. The normal form theorem, which may be of independent interest, says roughly that every $\exists$-definable subset of an algebraic extension of $\mathbb{Q}$ is a finite union of single points and projections of hypersurfaces defined by absolutely irreducible polynomials.

1. INTRODUCTION

Let $\mathrm{Sub}(\overline{\mathbb{Q}})$ denote the set of subfields of $\overline{\mathbb{Q}}$. Given a field $L \in \mathrm{Sub}(\overline{\mathbb{Q}})$ and a set $C \subseteq L$, it is a question of general interest whether C is first-order definable in L using the language of rings. If so, one also wants to know how simple a defining formula can be. For example, results of Koenigsmann [10], extended by Park [15], have shown that in every number field K, the ring $\mathcal{O}_K$ of algebraic integers is defined by a universal formula. Here we show that the usual situation is the opposite, not only for rings of integers but for any subset $A \subseteq \overline{\mathbb{Q}}$ satisfying a rather general condition on $A \cap \mathbb{Q}$. Just as $\mathcal{O}_L = \mathcal{O}_{\overline{\mathbb{Q}}} \cap L$, we write $A_L = A \cap L$. Placing a natural topology on $\mathrm{Sub}(\overline{\mathbb{Q}})$, we will show that in most cases there is a comeager set of fields $L \in \mathrm{Sub}(\overline{\mathbb{Q}})$ such that A_L cannot be defined in L by any universal formula.

Theorem 1.1. *If $A \subseteq \overline{\mathbb{Q}}$ is a subset for which $A_{\mathbb{Q}}$ is coinfinite and not thin (as a subset of the Hilbertian field $\mathbb{Q}$), then the following class is meager in* $\mathrm{Sub}(\overline{\mathbb{Q}})$*:*

$$\mathcal{U}_A = \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : A_L \text{ is universally definable in } L\}.$$

Indeed, a stronger statement holds, and depends only on the subset of $\mathbb{Q}$ in question.

Theorem 1.2 (Theorem 5.6)**.** *If $Z \subset \mathbb{Q}$ is not thin, then the following is meager in* $\mathrm{Sub}(\overline{\mathbb{Q}})$*:*

$$S_Z = \bigcup_{A \subseteq \overline{\mathbb{Q}}:\ A \cap \mathbb{Q} = Z} \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : A_L \text{ is coinfinite and universally definable in } L\}$$

The second theorem implies the first by setting $Z = A_{\mathbb{Q}}$. Thus the irrational portion of A is irrelevant: in all subfields L outside the meager class $S_{A_{\mathbb{Q}}}$, neither A_L nor any other set that intersects $\mathbb{Q}$ in $A_{\mathbb{Q}}$ can be universally defined. Clearly this is much stronger than the first statement.

MSC codes: 03C57 (primary); 12L05, 11U05, 03C40, 03D45 (secondary). Key words: algebraic fields, algebraic integers, definability, Hilbert Irreducibility Theorem, Hilbert's Tenth Problem.

Dually (with $B = \overline{\mathbb{Q}} \setminus A$), if $B_{\mathbb{Q}}$ is infinite and not co-thin in $\mathbb{Q}$, then the class

$$\mathcal{E}_B = \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : B_L \text{ is existentially definable in } L\}$$

equals $\mathcal{U}_A$, hence is meager. The second statement can also be applied in a dual form to existentially definable sets.

The notion of thinness which appears in the theorem is due to Serre. Intuitively, a set is thin if it is "small" in the sense of arithmetic geometry; see Section 2.3. Initially we did not expect definability of a set to be intertwined with any notion of its size apart from finiteness, but this condition arose naturally in our investigations.

The topology on $\mathrm{Sub}(\overline{\mathbb{Q}})$ is defined by considering it as a subset of the power set $2^{\overline{\mathbb{Q}}}$, from which it inherits the product topology. In this topology, every nonempty open set is non-meager. The topology also coincides (via the Galois correspondence) with the Vietoris topology on the space of closed subgroups of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. We thank Florian Pop for pointing out this connection to us, and for alerting us that the same topology appears in [17], where it is called the strict topology. The topology has also been used by other authors: for examples, see [9, 4, 5, 8].

The theorem also remains true when replacing $\mathrm{Sub}(\overline{\mathbb{Q}})$ with the quotient space $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ considered in [14], which only considers fields up to isomorphism; see Corollary 5.18.

Using the fact that neither $\mathbb{Z}$ nor $\mathbb{Q} \setminus \mathbb{Z}$ is thin in $\mathbb{Q}$, we obtain the following corollary.

Theorem 1.3 (Theorem 5.7)**.** *The set of algebraic extensions K of $\mathbb{Q}$ for which $\mathcal{O}_K$ is existentially or universally definable is a meager subset of* $\mathrm{Sub}(\overline{\mathbb{Q}})$.

After seeing one of the authors speak on these results, Philip Dittmann and Arno Fehm extended Theorem 1.3 in a different way, in [2], improving "existentially or universally definable" to "definable" by explicitly using the fact that $\mathcal{O}_K$ forms a ring. Their proof uses techniques from model theory, entirely different from those employed here.

1.1. **Outline of the paper.** To prove Theorem 1.2, we study the existential definability of sets $Y \subseteq \mathbb{Q}$ whose complement is not thin, in the sense of Serre. These are the complements of the sets Z described above. The necessary background of algebraic number theory, arithmetic geometry and thin sets is recalled in Section 2. In order to prove the main theorem, we introduce a new notion of rank in Section 3 that applies to existential formulas. This notion generalizes the multidegree of a polynomial in a way that we found to be both natural and quite useful, providing a pre-well-ordering of existential formulas. Thus, if Y is existentially definable within $\mathbb{Q}$ over some field $L \subseteq \overline{\mathbb{Q}}$, then there is a formula of least rank which does the job. By studying such minimal-rank formulas in Section 4, we obtain the following normal form for existential definitions, which may be of independent interest.

Theorem 1.4 ((Theorem 4.8))**.** *For any field $L \subseteq \overline{\mathbb{Q}}$, if $A \subseteq L$ is existentially definable in L, then A is definable in L by a formula of the form*

$$\alpha(X) = \vee_{i=1}^{r} \beta_i(X),$$

where each $\beta_i(X)$ has one of the following forms:

(i) The quantifier-free formula $X = z_0$ for a fixed $z_0 \in L$.

(ii) A formula of the form

$$\exists Y_1 \ldots \exists Y_e \; [f(X, Y_1, \ldots, Y_e) = 0 \neq g(X, Y_1, \ldots, Y_e)]$$

for polynomials $f, g \in L[X, Y_1, \ldots, Y_e]$, where f is absolutely irreducible and does not divide g.

Finally, we introduce the topological spaces of $\mathrm{Sub}(\overline{\mathbb{Q}})$ and $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ in Section 5, and use the normal form to deduce the main result via Hilbert's Irreducibility Theorem. In fact, the proof also leads to an algorithm which, given a basic open subset $U \subseteq \mathrm{Sub}(\overline{\mathbb{Q}})$, produces a computable field $L \in U$ in which the ring of integers $\mathcal{O}_L$ is neither existentially or universally definable; see Theorem 5.12.

1.2. **Previous work on definability of rings of integers.** Much of the previous work on the definability of subsets $A \subseteq K \in \mathrm{Sub}(\overline{\mathbb{Q}})$ has focused on the case where $A = \mathcal{O}_K$. We conclude the introduction with a overview of the literature on this case.

The existential definability of $\mathcal{O}_K$ in K is an ingredient that would assist a standard reduction argument for proving undecidability results for generalizations of Hilbert's Tenth Problem. In its original form, this problem asked for an algorithm that decides, given a polynomial equation $f(x_1, \ldots, x_n) = 0$ with coefficients in the ring $\mathbb{Z}$ of integers, whether there is a solution with $x_1, \ldots, x_n \in \mathbb{Z}$. Matiyasevich [13], building on earlier work by Davis, Putnam, and Robinson [1], proved that no such algorithm exists, i.e., Hilbert's Tenth Problem is undecidable. Since then, analogues of this problem have been studied by asking the same question for polynomial equations with coefficients and solutions in other recursive commutative rings. One of the most important unsolved questions in this area is Hilbert's Tenth Problem over the field of rational numbers $\mathbb{Q}$, and more generally over number fields. If $\mathbb{Z}$ is existentially definable in $\mathbb{Q}$, then a reduction argument shows that Hilbert's Tenth Problem for $\mathbb{Q}$ must be undecidable.

However, if Mazur's Conjecture holds, then $\mathbb{Z}$ is not existentially definable in $\mathbb{Q}$. Proving this unconditionally currently appears to be out of reach. In fact, it seems generally very difficult to prove undefinability results for individual fields. One example of success is the field of all totally real algebraic numbers $\mathbb{Q}^{\mathrm{tr}}$. Fried, Haran and Völklein showed that its first-order theory is decidable [6], while J. Robinson showed that the first-order theory of the ring of all totally real integers $\mathbb{Z}^{\mathrm{tr}}$ is undecidable [22]. This difference in decidability implies that $\mathbb{Z}^{\mathrm{tr}}$ cannot be first-order definable in the field $\mathbb{Q}^{\mathrm{tr}}$. Another example is the ring $\overline{\mathbb{Z}}$ of all algebraic integers inside $\overline{\mathbb{Q}}$, which is undefinable by the strong minimality of $\overline{\mathbb{Q}}$. In both examples, the facts used for proving undefinability are not remotely close to necessary conditions for undefinability. Instead, they simply reflect the available pathways for unconditionally proving undefinability in a limited number of cases.

While it is still an open question whether $\mathbb{Z}$ is existentially definable in $\mathbb{Q}$, it is possible to give a first-order definition of $\mathbb{Z}$ in $\mathbb{Q}$, i.e. a definition that uses both existential and universal quantifiers. This was first done by J. Robinson [20], who generalized this result to define the ring of integers $\mathcal{O}_K$ inside any number field K [21]. Later, Rumely [23] was able to make the definition of the ring of integers uniform across number fields. Robinson's definition was improved by Poonen [16] who gave a $\forall\exists$-definition that in every number field K defines its ring of integers. Following this, Koenigsmann [10] proved that it is possible to give a universal definition of $\mathbb{Z}$ in $\mathbb{Q}$, i.e. a definition that only involves universal ($\forall$) quantifiers, and Park extended his result to show that $\mathcal{O}_K$ is universally definable in K for every number field K [15]. This raises the question of whether we can expect universal and first-order definability to continue to hold for many infinite algebraic extensions of $\mathbb{Q}$.

Currently, first-order definability results are only known for certain classes of infinite extensions of the rationals. These are usually proved in order to establish the first-order undecidability of certain infinite extensions via reductions. For example, Videla proved the definability of the ring of integers over certain infinite algebraic pro-p extensions of $\mathbb{Q}$ [26], while Fukuzaki was able to define the ring of integers in infinite extensions in which every finite subextension has odd degree and that satisfy certain ramification conditions [7]. These results were further generalized by Shlapentokh in [25], to which we refer readers for more extensive background on known results for the first-order definability and decidability of infinite algebraic extensions of $\mathbb{Q}$. In Shlapentokh's framework, all known examples of algebraic extensions of $\mathbb{Q}$ with first-order definable rings of integers can be viewed as relatively small extensions which are somehow "close" to $\mathbb{Q}$. On the other hand, although first-order definability seems less likely for extensions which are similarly "far from" $\mathbb{Q}$, very few negative examples are known, as mentioned above.

Acknowledgements. This project began during a workshop at the American Institute of Mathematics in May 2019. It is based upon work supported by the National Science Foundation under Grant # DMS-1928930 while the authors participated in a program hosted by the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2020 semester. The authors wish to acknowledge useful conversations with Tom Tucker. Eisenträger was partially supported by National Science Foundation awards CNS-1617802, CNS-2001470, and a Vannevar Bush Faculty Fellowship from the US Department of Defense. Miller was partially supported by Grant # 581896 from the Simons Foundation and by the City University of New York PSC-CUNY Research Award Program. Springer was partially supported by National Science Foundation award CNS-1617802. Westrick was partially supported by the Cada R. and Susan Wynn Grove Early Career Professorship in Mathematics.

2. Background from number theory and algebraic geometry

In this section, we will recall some of the basic facts that we will require for fields, thin sets, and affine varieties. Readers can find additional background in the books of Lang [11], Serre [24] and Liu [12], respectively.

2.1. Field extensions and the irreducibility of polynomials.

In the material that follows, we will be presented with the following question: Given number fields $F \subseteq K$, which field extensions of F contain elements of the complement $K \setminus F$? This question is intimately related to the irreducibility of polynomials. First, we recall a basic result on the irreducibility of multivariable polynomials.

Lemma 2.1. *If K/F is an extension of fields within a larger field L, and $z \in L$ is algebraic over F with $F(z) \cap K \neq F$, then the minimal polynomial $h(Z)$ of f over F must be reducible over K.*

Proof. By hypothesis $1 < [F(z) \cap K : F]$, so

$$[F(z) : F(z) \cap K] < [F(z) : F(z) \cap K] \cdot [F(z) \cap K : F] = [F(z) : F].$$

From this it follows that $h(Z)$ must factor over $F(z) \cap K$, so it certainly also factors over the larger field K. □

The next proposition forms a kind of converse to Lemma 2.1 when K/F is a finite Galois extension. Given an algebraic function field $E = \mathrm{Frac}(F[Y_0, Y_1, \ldots, Y_m]/(f))$ where $f \in F[Y_0, Y_1, \ldots, Y_m]$ is an irreducible polynomial, the *constant field* of E is the set of elements which are algebraic over F.

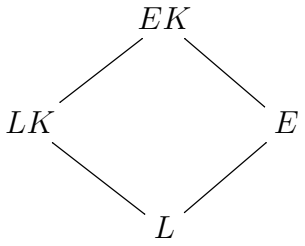
Proposition 2.2. *Let F be a number field, and K a finite Galois extension of F. If $m \geq 0$ and $f \in F[Y_0, Y_1, \ldots, Y_m]$ is an irreducible polynomial that becomes reducible in $K[Y_0, Y_1, \ldots, Y_m]$, then the constant field of $E = \mathrm{Frac}(F[Y_0, Y_1, \ldots, Y_m]/(f))$ is larger than F. In particular, there is an element $z \in E \setminus F$ such that there is an F-linear field embedding of $F(z)$ into K with the image of z lying in $K \setminus F$.*

Proof. Assume without loss of generality that Y_m appears nontrivially in f, and write $L = F(Y_0, Y_1, \ldots, Y_{m-1})$. We will view $E = L(\theta)$ for an element θ in the algebraic closure $\overline{L}$ with minimal polynomial f. Similarly consider K to be an extension of F inside $\overline{L}$.

Suppose that E contains no elements of $K \setminus F$. Then $E \cap K = L \cap K = F$, and a basic theorem of Galois theory [11, Theorem 1.12] implies the following because K is a Galois extension of F:

$$[EK : E] = [K : E \cap K] = [K : F] = [K : L \cap K] = [LK : L].$$

Using the diamond written below, we deduce that $[E : L] = [EK : LK]$. Importantly, these field extension degrees are also the degrees of the minimal polynomial of θ over L and LK, respectively.



This shows that f remains irreducible over the field $L = K(Y_0, Y_1, \ldots, Y_{m-1})$ as a polynomial in Y_m. We claim that f is actually irreducible as an element of the ring $K[Y_0, Y_1, \ldots, Y_m]$, which contradicts the hypothesis. To prove this, it only remains to show that the coefficients of f lying in $K[Y_0, \ldots, Y_{m-1}]$ have no common factor; see [11, IV.2.3]. Clearly, as a polynomial in Y_m, the coefficients of f lying in $F[Y_0, \ldots, Y_m]$ have no common factor over F because f is irreducible over F. In fact, this implies that the coefficients also have no common factor over any algebraic extension of F by the following lemma, which completes the proof. $\square$

Lemma 2.3. *Let F be a field and let F' be a separable extension. If $f_0, f_1, \ldots, f_k$ are a collection of polynomials in $F[Y_0, \ldots, Y_m]$ with no common factor, then $f_0, \ldots, f_k$ also have no common factor over the extension F'.*

Proof. By writing $f_0, \ldots, f_k$ in terms of their irreducible factors, we can reduce without loss of generality to the case of two irreducible polynomials $f_0, f_1 \in F[Y_0, \ldots, Y_m]$. Indeed, for every irreducible factor p of f_0, there is a polynomial f_j for $1 \leq j \leq m$ which is not divisible by p, and it suffices to show that the irreducible factors of f_j remain relatively prime to p over the larger field F'.

Notice that irreducible polynomials f_0 and f_1 are relatively prime over F if and only if $f_0 f_1$ generates a radical ideal in $F[Y_0, \ldots, Y_m]$, i.e. if and only if $F[Y_0, \ldots, Y_m]/(f_0 f_1)$ is a reduced

ring. The latter condition is stable under separable field extensions, i.e. $F'[Y_0, \ldots, Y_m]/(f_0 f_1)$ is also reduced; see [12, Proposition 3.2.7.(b)]. Therefore f_0 and f_1 have no common factors over F'. □

2.2. **Dimensions of rings and affine varieties.** We will require a usable notion of dimension, which can equivalently be viewed as a geometric or algebraic phenomenon. In particular, there are related notions of the dimension of a commutative ring A, and the dimension of the associated topological space $\operatorname{Spec} A$ consisting of all prime ideals of A with the Zariski topology. In this section, we will review some basic facts of commutative algebra and algebraic geometry, limiting the discussion to only what is necessary for our purposes.

First, let us recall this topology and some basic notation. Given a commutative ring A, the set $\operatorname{Spec} A$ is endowed with the *Zariski topology* by defining the following as basic closed and open sets, respectively. For any ideal $I \subseteq A$, we define $V(I)$ to be the subset of $\operatorname{Spec} A$ consisting of all prime ideals that contain I, and $D(f) = \operatorname{Spec} A \setminus V(f)$. Notice that it is natural via the isomorphism theorems for rings to identify $V(I)$ with $\operatorname{Spec} A/I$. With this notation, the closed subsets of $\operatorname{Spec} A$ in the Zariski topology are precisely the sets of the form $V(I)$ where $I \subseteq A$ is an ideal, and sets of the form $D(f)$ for $f \in A$ form a base for the open subsets of $\operatorname{Spec} A$. In fact, $\operatorname{Spec} A$ is an *affine scheme*, meaning that it has even more structure than just a topology, although we will not require this full structure; see [12, Chapter 2] for more background.

In this paper, we consider the ring $A = F[Y_0, \ldots, Y_m]$ and its quotients, where F is a subfield of $\overline{\mathbb{Q}}$. An *affine variety over* F is an object of the form $V(I) = \operatorname{Spec} F[Y_0, \ldots, Y_m]/I$ for some $m \geq 0$ and some ideal $I \subseteq F[Y_0, Y_1, \ldots, Y_m]$. Furthermore, if the quotient $F[Y_0, \ldots, Y_m]/I$ is an integral domain, then the corresponding affine variety is called *integral*. We will write $V(I) = V(f_1, \ldots, f_k)$ when the ideal $I \subseteq F[X, Y_1, \ldots, Y_m]$ is generated by $\{f_1, \ldots, f_k\}$. If there is ambiguity about the base field, then we will write V_F instead of V for clarity.

Given an affine variety $V = \operatorname{Spec} F[Y_0, \ldots, Y_m]/I$, the *rational points* of V (over F) are the tuples $(y_0, \ldots, y_m) \in F^m$ such that $f(y_0, \ldots, y_m) = 0$ for all $f \in I$. The set of rational points can be identified with the set of all F-algebra homomorphisms $\varphi : F[Y_0, \ldots, Y_m]/I \to F$. We refer the reader to [12, Section 2.3.2] for more details. As we are frequently working over non-algebraically closed fields, it is possible for nontrivial affine varieties to have no rational points, such as the affine variety $\operatorname{Spec} \mathbb{Q}[Y_0, \ldots, Y_m]/(Y_0^2 + \cdots + Y_m^2 + 1)$ for any $m \geq 0$. We can view the varieties as geometric objects which help us find and describe the rational points.

The *Krull dimension* of a ring A, written $\dim(A)$, is the supremal length r of a chain of prime ideals $\mathfrak{p}_0 \subsetneq \cdots \subsetneq \mathfrak{p}_r$ in A. Similarly, given a topological space X, we define $\dim(X)$ to be the supremal length r of a chain of irreducible closed subsets $Z_0 \subsetneq \cdots \subsetneq Z_r$ in X. The following proposition equates these two notions of dimension. Recall that the nilradical of a commutative ring is the set of all nilpotent elements, or equivalently the intersection of all prime ideals.

Proposition 2.4 (Proposition 2.5.8, [12])**.** *Let A be a (commutative) ring and let N be the nilradical of A. Then* $\dim(\operatorname{Spec} A) = \dim(A) = \dim(A/N)$.

In our applications, we need to understand the dimension of subsets of affine varieties. Recall that if X is any topological space and Y is any subset of X endowed with the subset

topology, then $\dim(Y) \leq \dim(X)$ [12, Proposition 2.5.5]. In the context of affine varieties and open subsets, this inequality is often an equality due to the fact that open subsets in the Zariski topology are "large". This idea is formulated precisely in the following proposition. Given a field extension L/F, we write $\operatorname{trdeg}_F L$ for the *transcendence degree* of L over F. If $X = \operatorname{Spec} A$ is an integral affine variety, we call $\operatorname{Frac}(A)$ the *function field* of X.

Proposition 2.5 (Proposition 2.5.19, [12])**.** *If $X = \operatorname{Spec} A$ is an integral affine variety over a field F, then*

$$\dim(U) = \dim(X) = \operatorname{trdeg}_F \operatorname{Frac}(A)$$

for each nonempty open subset $U \subseteq X$.

Similarly, it is helpful to know when a subset of a topological space X has strictly smaller dimension than X. In contrast to the result immediately above, this often happens for proper closed subsets of an affine variety.

Proposition 2.6 (Corollary 2.5.26, [12])**.** *Let $X = \operatorname{Spec} A$ be an integral affine variety. If $f \in A$ is nonzero, then every irreducible component of $V(f)$ has dimension $\dim(X) - 1$. In particular, every proper closed subset of X has strictly smaller dimension than X.*

So far in this section, the definition of dimension depends on the base field $F \subseteq \overline{\mathbb{Q}}$, *a priori*. However, the result below clarifies that dimension stays the same under base extension. This allows us to ignore the field of definition to some extent, especially when defining the rank of a formula below, although the notion of integrality truly does depend on the base field, so care is still required when applying the previous two propositions.

Proposition 2.7 (Proposition 3.2.7, [12])**.** *Let $F \subseteq L \subseteq \overline{\mathbb{Q}}$ be fields. Given an affine variety $V_F(f_1, \ldots, f_k) = \operatorname{Spec} F[Y_0, \ldots, Y_m]/(f_1, \ldots, f_k)$, the affine variety*

$$V_L(f_1, \ldots, f_k) = \operatorname{Spec} L[Y_0, \ldots, Y_m]/(f_1, \ldots, f_k)$$

is the base extension *of the variety $V_F(f_1, \ldots, f_k)$ to L, and these affine varieties have the same dimension.*

To apply this proposition to open sets, we remark that open sets can be equivalently viewed as affine varieties themselves, albeit in a different ambient space with an extra variable.

Corollary 2.8. *Let $F \subseteq \overline{\mathbb{Q}}$ be a field. For polynomials $g, f_1, \ldots, f_k \in F[Y_0, \ldots, Y_m]$, define $A = F[Y_0, \ldots, Y_m]/(f_1, \ldots, f_k)$ and let A_g be the localization of A be the element g. Then there are isomorphisms of ringed topological spaces*

$$V_F(f_1, \ldots, f_k) \cap D(g) \cong \operatorname{Spec}(A_g) \cong V_F(f_1, \ldots, f_k, Y_{m+1}g - 1).$$

In particular, $\dim(V_F(f_1, \ldots, f_k) \cap D(g)) = \dim(V_L(f_1, \ldots, f_k) \cap D(g))$ for any algebraic extension of fields $L \supseteq K$.

Proof. The first isomorphism is [12, Lemma 2.3.7]. The second isomorphism actually follows from a well-known isomorphism of underlying rings

$$A_g \cong F[Y_0, \ldots, Y_{m+1}]/(f_1, \ldots, f_k, Y_{m+1}g - 1);$$

see [19, Lemma §6.2]. Therefore, the statement on dimension follows immediately from Proposition 2.7. □

2.3. **Thin sets.** Hilbert's Irreducibility Theorem can take many different forms, but we put a simple version here that suffices for the purposes of this article. For brevity, we present *thin sets* as a black box, and refer the reader to [24, Prop. 3.3.5] for more details. Essentially, a thin subset $T \subseteq K$ of a number field is small, in the view of arithmetic geometry. For example, any set of points that is contained in a closed subvariety of affine n-space K^n, and which is different from the entire space, is thin with respect to K. All necessary details can be deduced from the results we recall below .

Theorem 2.9 (Hilbert's Irreducibility Theorem)**.** *Let $f(Y_0, Y_1, \ldots, Y_m)$ be a polynomial with coefficients in a number field K which is irreducible as an $(m+1)$-variable polynomial. There exists a thin set $T \subseteq K^m$ such that if $(y_1, \ldots, y_m) \in K^m \setminus T$, then $f(Y_0, y_1, \ldots, y_m)$ is an irreducible single-variable polynomial of degree $\deg_{Y_0}(f)$.*

In order for the theorem above to be non-trivial, we need to know that K^m is not a thin subset of itself, and this is indeed true for all number fields [24, Prop 3.4.1]. Moreover, the propositions below show that thin sets cannot contain arithmetically important subsets, which will allow us to use Hilbert's Irreducibility Theorem in the cases we care about.

Proposition 2.10 (Proposition 3.2.1, [24])**.** *If L/K is a finite extension of fields and $T \subseteq L^m$ is thin with respect to L, then $T \cap K^m$ is thin with respect to K.*

Proposition 2.11. *If K is a number field, then no thin subset of K contains either $\mathbb{Z}$ or $\mathbb{Q} \setminus \mathbb{Z}$.*

Proof. Thin sets of $\mathbb{Q}$ cannot contain $\mathbb{Z}$ or $\mathbb{Q} \setminus \mathbb{Z}$ by [24, Theorem 3.4.4] and [24, Prop. 3.4.2], respectively. Thus, the result for arbitrary number fields follows from Proposition 2.10. □

Moreover, we can understand thin sets in products. This lemma will be used to show that if a set $Z \subseteq \mathbb{Q}$ is not thin, then the product $Z \times \mathbb{Q}^n$ cannot be thin, either.

Lemma 2.12. *If $n \geq 0$ and $S \subseteq \mathbb{Q}$ is a set such that $S \times \mathbb{Q}^n \subseteq \mathbb{Q}^{n+1}$ is thin, then $S \subseteq \mathbb{Q}$ is thin.*

Proof. There is a line $\mathcal{L} \subseteq \mathbb{Q}^{n+1}$ such that $\mathcal{L} \cap (S \times \mathbb{Q}^n)$ is thin in $\mathcal{L}$ and the projection of $\mathcal{L}$ to the first coordinate is all of $\mathbb{Q}$ [24, Proposition 3.2.3]. As $\mathcal{L}$ is a line, this projection is an isomorphism and $\mathcal{L} \cap (S \times \mathbb{Q}^n)$ maps onto to the set S. Therefore, S is thin in $\mathbb{Q}$. □

Finally, we prove a proposition that lets us stitch this material together. This is ultimately the result that is required in the proof of our main theorem.

Proposition 2.13. *Let K be a number field and let $f(X, Y_1, \ldots, Y_m), g(X, Y_1, \ldots, Y_m) \in K[X, Y_1, \ldots, Y_m]$ be relatively prime irreducible polynomials. Then there is a thin set $T \subseteq K^m$ such that $f(x, y_1, \ldots, y_{m-1}, Y)$ and $g(x, y_1, \ldots, y_{m-1}, Y)$ are relatively prime irreducible single-variable polynomials for every $(x, y_1, \ldots, y_{m-1}) \in K^m \setminus T$, of degrees $\deg_{Y_m}(f)$ and $\deg_{Y_m}(g)$, respectively.*

Proof. Take T_0 to be the union of the two thin sets given by applying Hilbert's Irreducibility Theorem to f and g separately. By construction, $f(x, y_1, \ldots, y_{m-1}, Y)$ and $g(x, y_1, \ldots, y_{m-1}, Y)$ are irreducible polynomials in Y for every $(x, y_1, \ldots, y_{m-1}) \in K^m \setminus T_0$, and it only remains to check the claim of relative primality.

If $\deg_{Y_m}(f) \neq \deg_{Y_m}(g)$, then this claim is trivial. Therefore, write $d = \deg_{Y_m}(f) = \deg_{Y_m}(g)$, and consider $(x, \ldots, y_{m-1}) \in K^m \setminus T$. Since the polynomials $f(x, y_1, \ldots, y_{m-1}, Y)$

and $g(x, y_1, \ldots, y_{m-1}, Y)$ are irreducible, the failure of relative primality implies that they are unit multiples of each other, i.e., $f(x, y_1, \ldots, y_{m-1}, Y) = zg(x, y_1, \ldots, y_{m-1}, Y)$ for some nonzero $z \in K$. In particular, if we write

$$f(X, Y_1, \ldots, Y_m) = \sum_{i=0}^{d} f_i(X, Y_1, \ldots, Y_{m-1}) Y_m^i,$$

$$g(X, Y_1, \ldots, Y_m) = \sum_{i=0}^{d} g_i(X, Y_1, \ldots, Y_{m-1}) Y_m^i,$$

where $f_i, g_i \in K[X, Y_1, \ldots, Y_{m-1}]$ are polynomials, then this condition is the same as

$$f_i(x, y_1, \ldots, y_{m-1}) = zg_i(x, y_1, \ldots, y_{m-1})$$

for all $0 \leq i \leq d$. Multiplying these conditions together, we get the equations

$$f_i g_j = z g_i g_j = g_i f_j$$

for $0 \leq i, j \leq d$. We will show that this system of equations holds only inside a thin set, which completes the proof.

We claim that the polynomial

$$f_i(X, Y_1, \ldots, Y_{m-1}) g_j(X, Y_1, \ldots, Y_{m-1}) - g_i(X, Y_1, \ldots, Y_{m-1}) f_j(X, Y_1, \ldots, Y_{m-1})$$

is nonzero for some choice of i and j. Indeed, if this were not the case, then we would find that

$$\begin{aligned} f_i(X, Y_1, \ldots, Y_{m-1}) g(X, Y_1, \ldots, Y_m) &= \sum_{j=0}^{d} f_i(X, Y_1, \ldots, Y_{m-1}) g_j(X, Y_1, \ldots, Y_{m-1}) Y_m^j \\ &= \sum_{j=0}^{d} g_i(X, Y_1, \ldots, Y_{m-1}) f_j(X, Y_1, \ldots, Y_{m-1}) Y_m^j \\ &= g_i(X, Y_1, \ldots, Y_{m-1}) f(X, Y_1, \ldots, Y_m) \end{aligned}$$

for all i. As g and f are irreducible and the only polynomials on the left and right sides of the equation containing the variable Y_m, we conclude that they are unit multiples of each other, which contradicts the hypothesis of relative primality.

Therefore, let T_1 be the set of all K-rational points on the affine variety

$$V_K(\{f_i g_j - g_i f_j : 0 \leq i < j \leq \deg_{Y_m}(f)\}).$$

Since one of the polynomials in the defining set is nonzero, the affine variety is a proper closed variety, which implies that T_1 is a thin set by definition. By construction, the set $T = T_0 \cup T_1$ is the desired thin set. □

3. Rank of a Formula

The goal of this section is to define a notion of rank for existential formulas in the language of fields, using degrees of polynomials and dimensions of varieties, as well as the number of $\exists$-quantifiers used. Certain formulas will have the same rank, just as certain polynomials have the same degree. Crucially, the ranks are well-ordered.

3.1. **A useful well-ordering.**

Definition 3.1. Let $(\mathcal{L}, <)$ be a linear order. For a finite tuple $(a_0, \ldots, a_n) \in \mathcal{L}^{<\omega}$, write $\vec{a}^*$ for the tuple of the same $(n+1)$ elements (including repetitions) arranged in $<$-descending order: $\vec{a}^* = (a_{\alpha(0)}, \ldots, a_{\alpha(n)})$ where α is a permutation and $a_{\alpha(i+1)} \leq a_{\alpha(i)}$ for all $i < n$. Write $\vec{a} =^* \vec{b}$ just if $\vec{a}^* = \vec{b}^*$.

Then the $*$-order $(\mathcal{L}^*, <^*)$ is the lexicographic order $<^*$ (defined using $<$ on individual coordinates) on the set $\mathcal{L}^*$ of $=^*$-equivalence classes in $\mathcal{L}^{<\omega}$. To be clear: if $\vec{a}^*$ is a proper initial segment of $\vec{b}^*$, then $\vec{a}^* <^* \vec{b}^*$.

Equivalently, one can view the elements of $\mathcal{L}^*$ as finite multisets of elements of $\mathcal{L}$, with the elements of each multiset listed in $<$-nonincreasing order.

Lemma 3.2. *If $(\mathcal{L}, <)$ is a well order, then so is $(\mathcal{L}^*, <^*)$.*

Proof. Clearly $<^*$ is a linear order. If it were not a well order, there would be a least $a \in \mathcal{L}$ such that some infinite $<^*$-descending sequence begins with an $\vec{a}^*$ whose greatest element is a. Choose such an $\vec{a}^* = (a^k, a_1, \ldots, a_n)$, in nonincreasing order with $a_1 < a$ after a appears k times, with k as small as possible (and allowing $n = 0$). Then the infinite descending sequence beginning with this $\vec{a}^*$ can only have finitely many terms that begin with a^k, for if there were infinitely many, then by "chopping off" the a^k from each term, we would get an infinite sequence contradicting the choice of a. But then, immediately after the last term beginning with a^k comes a term beginning with a^j for $j < k$, and this term also begins an infinite descending sequence in $\mathcal{L}^*$, contradicting either the minimality of k (if $j > 0$) or the minimality of a (if $j = 0$). □

3.2. **Definition of rank.** We present an explicit way to put a well-ordering on the set of existential formulas with parameters in any given field. This is done by associating a *rank* to every existential formula.

Every existential formula $\alpha(X)$ can be written in disjunctive normal form

$$\alpha(X) = \exists \vec{Y}(\alpha_1 \vee \alpha_2 \vee \cdots \vee \alpha_n),$$

where each $\alpha_i(X, \vec{Y})$ is a conjunction of equations and inequations. Bringing the existential quantifiers inside the disjunctions and discarding any unused quantifiers, every existential formula can be rewritten as

$$((\exists Y_1 \cdots \exists Y_{m_1})\alpha_1) \vee \cdots \vee ((\exists Y_1 \cdots \exists Y_{m_n})\alpha_n),$$

where all variables $Y_1, \ldots, Y_{m_i}$ appear in α_i. One can also easily rearrange any $\alpha_i(X, \vec{Y})$ into a conjunction of the form

$$f_1(X, \vec{Y}) = \cdots = f_k(X, \vec{Y}) = 0 \ \& \ g(X, \vec{Y}) \neq 0.$$

Only one inequation $g \neq 0$ is needed, as several $g_i(X, \vec{Y})$ could be multiplied together. It is allowed for g to be the constant 1. We call an existential formula *rankable* if it is given in the above format. It is trivial to rearrange any existential formula into rankable format, so in this paper every existential formula which appears is assumed to be rankable.

Before defining rank, we present a way to order tuples of polynomials. Notice that this notion depends on a specific order for the variables.

Definition 3.3. For the variables $X, Y_1, \ldots, Y_m$, the *multidegree* of a monomial $X^c Y_1^{d_1} \cdots Y_m^{d_m}$ is $(c, d_1, \ldots, d_m)$, and these $(m+1)$-tuples are ordered by the reverse lexicographic order. The *multidegree* $\mathrm{mdeg}(f)$ *of a polynomial* f is the maximum of the multidegrees of each monomial appearing (with nonzero coefficient) in it.

Observe that the linear order defined above on multidegrees is a well-ordering.

Definition 3.4. A *basic rankable formula* is an existential formula of the form

$$\exists Y_1 \cdots \exists Y_m \; [f_1(X, Y_1, \ldots, Y_m) = \cdots = f_k(X, Y_1, \ldots, Y_m) = 0 \;\&\; g(X, \vec{Y}) \neq 0],$$

and the *rank* of such a formula is the triple

$$\mathrm{rk}(\beta) = (m, e, (\mathrm{mdeg}(f_1), \ldots, \mathrm{mdeg}(f_k))^*),$$

where the second component is the dimension e of $V_{\overline{\mathbb{Q}}}(\vec{f}) \cap D(g)$, as defined in Section 2.2, and the third component uses the $=^*$-classes of tuples of multidegrees, as in Definition 3.1.

In this definition, we see that $V_{\overline{\mathbb{Q}}}(\vec{f}) \cap D(g)$ is a subset of an ambient space of dimension $m+1$. Therefore, the first coordinate of the definition of rank can be equivalently viewed as a measure of the dimension of this ambient space. Additionally, by Corollary 2.8, the base field does not matter in the definition of the dimension e, so we will usually drop the $\overline{\mathbb{Q}}$ from this notation.

We define an order $\prec$ on ranks of basic rankable formulas in forwards lexicographic order, meaning that

$$(m, e, (d_1, \ldots, d_k)^*) \prec (m', e', (d'_1, \ldots, d'_{k'})^*)$$

if and only if one of the following holds:

- $m < m'$, i.e., the first formula uses fewer $\exists$-quantifiers; or
- $m = m'$ and $e < e'$, so the first formula defines an open variety of lesser dimension than the second; or
- $m = m'$ and $e = e'$ and $(d_1, \ldots, d_k)^* <^* (d'_1, \ldots, d'_{k'})^*$, so the first formula uses polynomials of lower multidegree.

The least possible rank of a (satisfiable) basic rankable formula is $(0, 0, (1)^*)$, which is the rank of the quantifier-free formula $X = x$ for any specific value x: here $m = 0$, $k = 1$ and the variety, which has a single component whose dimension is 0, is defined by $f_1 = X - x = 0$ whose multidegree (in the single variable X, since $m = 0$) is simply 1. (The variety defined by $0 = 0$ has dimension 1, so the formula $0 = 0$ has higher rank.)

Let $\mathcal{R}$ denote the set of all possible ranks of basic rankable formulas. Then $(\mathcal{R}, \prec)$ is a well-ordering. (The third component of $\prec$ is well-ordered by Lemma 3.2.) Let $(R^*, \prec^*)$ be the result of applying Definition 3.1 to $(\mathcal{R}, \prec)$.

Observe that an existential formula is rankable if and only if it is the finite disjunction of basic rankable formulas.

Definition 3.5. If $\alpha = \vee_{i=1}^r \beta_i$ is a rankable formula, the *rank* of α is defined to be

$$\mathrm{rk}(\alpha) = (\mathrm{rk}(\beta_1), \ldots, \mathrm{rk}(\beta_n))^* \in \mathcal{R}^*$$

The rankable formulas can then be compared using the ordering $\prec^*$. By Lemma 3.2, $(\mathcal{R}^*, \prec^*)$ is a well-order.

4. Minimal formulas and hypersurfaces

The well-ordering of ranks means that every nonempty set of existential formulas has an element of least rank. For example, if there exists an existential formula that defines $\mathcal{O}_L$ in L, then there is an existential formula α that accomplishes this which has least rank among all such formulas. Such a formula can be considered a minimal successful formula. This motivates the following general definition.

Definition 4.1. For a field $L \subseteq \overline{\mathbb{Q}}$ and an existential formula $\alpha(X)$ with coefficients from L, we say α is *L-minimal* if α has least rank among all existential formulas α' for which

$$\forall x(\alpha(x) \iff \alpha'(x))$$

holds in L.

In order for the above to make sense, α' ranges only over those existential formulas which have parameters from L. We will show that every L-minimal formula must take the form of a disjunction of formulas with two very simple formats: quantifier-free formulas, and formulas with only one equation and one inequation.

We will start by considering a general rankable formula, then minimize it as much as possible. First, we want to minimize the number of quantifiers, which is the first component of rank. Clearly, we can eliminate the quantifier for any variable that does not appear in any polynomial of the formula. The following simple lemma allows us also to remove any variables that appear in the inequation, but none of the equations.

Lemma 4.2. *Let $1 \leq e < m$ and let $\delta(X)$ be the basic rankable existential formula*

$$\exists Y_1 \cdots \exists Y_m \ [f_1(X, Y_1, \ldots, Y_e) = \cdots = f_k(X, Y_1, \ldots, Y_e) = 0 \neq g(X, Y_1, \ldots, Y_m)],$$

where $f_i \in F[X, Y_1, \ldots, Y_e]$ and $g \in F[X, Y_1, \ldots, Y_m]$ for some field F.

Then there are polynomials $g_1, \ldots, g_r \in F[X, Y_1, \ldots, Y_e]$ such that $\delta(X)$ is equivalent over F to the disjunction of formulas

$$\vee_{i=1}^{r} \exists Y_1 \cdots \exists Y_e \ [f_1(X, \vec{Y}) = \cdots = f_k(X, \vec{Y}) = 0 \neq g_i(X, \vec{Y})].$$

Proof. Write out $g = \sum_{i=0}^{d_m} g_i(X, Y_1, \ldots, Y_{m-1}) Y_m^i$ as a polynomial in Y_m. Notice that if $(x, y_1, \ldots, y_{m-1}) \in \overline{\mathbb{Q}}^m$ is any tuple, then there is a $y_m \in \mathbb{Q}$ such that $g(x, y_1, \ldots, y_m) \neq 0$ if and only if $g_i(x, y_1, \ldots, y_{m-1}) \neq 0$ for some $0 \leq i \leq d_m$. Therefore, we can remove the quantifier for Y_m and instead use a disjunction where g is replaced by g_j for $0 \leq j \leq d_m$ in each formula. By induction, this completes the proof. □

To continue minimizing the number of quantifiers, we can take a more geometric perspective. A basic rankable formula $\beta(X)$ with m quantifiers

$$\exists Y_1 \cdots \exists Y_m \ [f_1(X, \vec{Y}) = \cdots = f_k(X, \vec{Y}) = 0 \neq g(X, \vec{Y})]$$

corresponds to the projection to the X-coordinate of the points on the variety $D(g) \cap V(f_1, \ldots, f_k)$. Minimizing the number of quantifiers m is equivalent to minimizing the dimension $m + 1$ of the ambient space where the variety lives. If k is large, then we expect the dimension e of the variety to be much smaller than $m + 1$, and we can consider this "wasteful," as it uses more variables than necessary. The following proposition uses a basic result of algebraic geometry to show that, in a special case with integral affine varieties, we only need $m = e$ quantifiers and a single equation to describe all but a lower-dimensional

closed subset. To complete the section, we will the show that this is enough to deduce the result in general.

Proposition 4.3. *Let $F \subseteq \overline{\mathbb{Q}}$ be a field and $\mathfrak{p} = (f_1, \ldots, f_k) \subseteq F[X, Y_1, \ldots, Y_m]$ a prime ideal. Define $\beta(X)$ to be the formula*

$$\beta(X) = \exists Y_1, \ldots, Y_m [f_1(X, Y_1, \ldots, Y_m) = \cdots = f_k(X, Y_1, \ldots, Y_m) = 0]$$

and set $e = \dim(V_F(\mathfrak{p}))$. If $\beta(X)$ is satisfied by infinitely many values of X in $\overline{\mathbb{Q}}$ and $e \leq m-1$, then after possibly reordering indices, there are polynomials $h \in F[X, Y_1, \ldots, Y_e]$ and $s \in F[X, Y_1, \ldots, Y_{e-1}]$ with h irreducible and $s \notin \mathfrak{p}$ such that $\beta(X)$ is equivalent to $\gamma_1(X) \vee \gamma_2(X)$ over F, using the formulas

$$\gamma_1(X): \ \exists Y_1 \cdots \exists Y_e \ [h(X, \ldots, Y_e) = 0 \neq s(X, \vec{Y})],$$
$$\gamma_2(X): \ \exists Y_1 \cdots \exists Y_m \ [s(X, \ldots, Y_m) = f_1(X, \ldots, Y_m) = \cdots = f_k(X, \ldots, Y_m) = 0].$$

Proof. Write $L = \mathrm{Frac}(F[X, Y_1, \ldots, Y_m]/\mathfrak{p})$. By Proposition 2.5, we know that e is equal to the transcendence degree of L over F. Since the images of $\{X, Y_1, \ldots, Y_m\}$ generate L over F, there is a transcendence basis consisting of a subset of these elements, and we can force $\bar{X}$ to be in this basis because $\bar{X}$ is not algebraic over F [11, Theorem VIII.1.1]. Indeed, if $\bar{X}$ were algebraic over F, then it would be the root of a single-variable polynomial over F, and therefore $\beta(X)$ would only be solvable over $\overline{\mathbb{Q}}$ by finitely many X, which is not the case by hypothesis.

Reorder the variables so that $\{\bar{X}, \bar{Y}_1, \ldots, \bar{Y}_{e-1}\}$ is a transcendence basis of L over F. Write $L_0 = F(X, Y_1, \ldots, Y_{e-1})$. Although a particular ordering of the variables is used when defining the multidegree component of rank in Definition 3.4, we will produce lower-rank formulas purely in terms of quantifiers and dimension, and therefore the multidegree will not matter here. As L is a finite separable extension of L_0, the primitive element theorem states that $L = L_0(\theta)$ for a single element θ. Write $h \in L_0[Y]$ for the minimal polynomial of θ. By clearing denominators if necessary, we can assume without loss of generality that $h \in F[X, Y_1, \ldots, Y_{e-1}, Y]$ is an irreducible multivariable polynomial. Therefore, writing $\mathfrak{p} = (f_1, \ldots, f_k)$, we have an isomorphism of fields:

$$L_0[Y_e, \ldots, Y_m]/(f_1, \ldots, f_k) \cong L \cong L_0[Y]/(h) \cong \mathrm{Frac}(F[X, Y_1, \ldots, Y_{e-1}, Y]/(h)).$$

Geometrically, this says that the integral affine variety $V_F(\mathfrak{p})$ is birational to the hypersurface $V_F(h)$. In fact, we can see that the two varieties contain isomorphic open sets, as follows.

Using the isomorphism of fields we can write $Y_j = \sum_{\ell=0}^{N_j} c_{j,\ell} Y^\ell$ for each $j = e, \ldots, m$, and $Y = \sum_{\vec{a}} d_{\vec{a}} Y_e^{a_0} \ldots Y_m^{a_{m-e}}$, where $c_{j,\ell}$ and $d_{\vec{a}}$ are elements of L_0, and in particular not contained in $\mathfrak{p}$ because L_0 is a subfield of the function field of $V_F(\mathfrak{p})$. Let s be the products of all denominators appearing in these terms. Then these equations give an isomorphism of the open sets $V_F(\mathfrak{p}) \cap D(s)$ and $V_F(h) \cap D(s)$; see [12, Lemma 3.7]. Moreover, the X-coordinate of rational points is unchanged by the isomorphism because we included X in the transcendence basis. As $V_F(\mathfrak{p}) = (V_F(\mathfrak{p}) \cap D(s)) \cup V_F(\mathfrak{p} + (s))$, this proves the claim that the formula β is equivalent over F to the disjunction stated above. □

Next we show that minimal formulas all have a very convenient structure.

Proposition 4.4. *If $\alpha(X) = \bigvee_{i=1}^r \beta_i(X)$ is a disjunction of basic rankable formulas and is L-minimal for some field $L \subseteq \overline{\mathbb{Q}}$, then each $\beta_i(X)$ has one of the following forms:*

(i) The quantifier-free formula $X = z_0$ for a fixed $z_0 \in L$.
(ii) The "hypersurface formula" $\exists Y_1 \dots \exists Y_e\ [f(X, Y_1, \dots, Y_e) = 0 \neq g(X, Y_1, \dots, Y_e)]$ for an irreducible $f \in L[X, Y_1, \dots, Y_e]$ and a polynomial $g \in L[X, Y_1, \dots, Y_e]$.

Proof. Let $\beta(X)$ be a fixed $\beta_i(X)$ which does not have the desired form. Write β in the form

$$\exists Y_1 \cdots \exists Y_m\ [f_1(X, \dots, Y_m) = \cdots = f_k(X, \dots, Y_m) = 0 \neq g(X, \vec{Y})]$$

and consider the ideal $I = (f_1, \dots, f_k)$. Define $e = \dim(V(I) \cap D(g))$. Without loss of generality, we can assume that each f_i is irreducible. Otherwise, if $f_1 = h_1 h_2$ is a nontrivial factorization, then we could write β as the disjunction of two formulas with f_1 replaced by h_1 and h_2, respectively, which have smaller multidegree.

Since f_1 is irreducible, $V_L(I)$ is a closed subset of the integral affine variety $V_L(f_1)$ which has dimension $\dim(V_L(f_1)) = m$ by Proposition 2.6. In fact, we see that either $V(f_1) = V(I)$, in which case we are done, or we have

$$e = \dim(V(I) \cap D(g)) \leq \dim(V(I)) < \dim(V(f_1)) = m.$$

By assumption, we are in the latter case, and we will produce a set of formulas with parameters in L which explicitly contradicts the minimality of α.

The ideal I has a primary decomposition $I = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_r$ where each $\mathfrak{q}_i$ is a primary ideal associated to a prime ideal $\mathfrak{p}_i$. Indeed, the rational points on $V(I) \cap D(g)$ are the same as the rational points on $\cup_{i=1}^r V(\mathfrak{p}_i) \cap D(g)$. Notice that the open set $V(\mathfrak{p}_i) \cap D(g)$ might be empty for some i, but whenever it is nonempty, $V(\mathfrak{p}_i) \cap D(g)$ has the same dimension as $V(\mathfrak{p}_i)$ by Proposition 2.5.

To summarize, we have shown that the formula $\beta(X)$ is equivalent to the disjunction $\vee_{i=1}^r \delta_{\mathfrak{p}_i}(X)$ where each $\delta_{\mathfrak{p}_i}(X)$ is defined as a formula

$$\delta_{\mathfrak{p}_i}(X) = \exists Y_1, \dots, Y_m [p_1^i(X, \vec{Y}) = \cdots = p_{n(i)}^i(X, \vec{Y}) = 0 \neq g(X, \vec{Y})],$$

where $\mathfrak{p}_i = (p_1^i, \dots, p_{n(i)}^i)$. For each i, we will replace $\delta_{\mathfrak{p}_i}(X)$ itself with an equivalent disjunction of basic rankable formulas, each of which has rank strictly smaller than β. By definition, this contradicts the minimality of α, and the proof will be done.

To this end, we analyze the primes $S = \{\mathfrak{p}_1, \dots, \mathfrak{p}_r\}$ and divide them accordingly. Let S_{finite} be the set of primes $\mathfrak{p} \in S$ such that only finitely many elements of L satisfy $\delta_{\mathfrak{p}}(X)$ in F, and let S_∞ be all other primes of S. Partition $S_\infty = S_{\text{big}} \cup S_{\text{small}}$ where

$$S_{\text{big}} = \{\mathfrak{p} \in S_\infty \mid \dim(V(\mathfrak{p})) = e\},$$
$$S_{\text{small}} = \{\mathfrak{p} \in S_\infty \mid \dim(V(\mathfrak{p})) < e\}.$$

For any prime $\mathfrak{p} \in S_{\text{finite}}$, let $\{z_1, \dots, z_n\}$ be the finite set of elements of L which satisfy $\delta_{\mathfrak{p}}(X)$ in L. We may therefore replace $\delta_{\mathfrak{p}}(X)$ with the disjunction of quantifier-free formulas $\vee_{i=1}^n (X - z_i)$. Each of these quantifier-free formulas consisting of a single-variable polynomial of degree 1 has the smallest rank possible for a nontrivial basic rankable formula and $\beta(X)$ has strictly larger rank.

For any $\mathfrak{p} \in S_{\text{small}}$, the formula $\delta_{\mathfrak{p}}(X)$ is already of smaller rank than β. Indeed, the ambient space is the same, and the dimension is strictly smaller by definition.

For any $\mathfrak{p} \in S_{\text{big}}$, letting $\mathfrak{p} = (p_1, \ldots, p_n)$, we apply Proposition 4.3 to see that $\exists \vec{Y}[p_1(X, \vec{Y}) = \cdots = p_n(X, \vec{Y}) = 0]$ is equivalent to the disjunction of two formulas

$$\exists Y_1 \cdots \exists Y_e \ [f(X, \ldots, Y_e) = 0 \neq s(X, \vec{Y})],$$
$$\exists Y_1 \cdots \exists Y_m \ [s(X, \ldots, Y_m) = p_1(X, \ldots, Y_m) = \cdots = p_n(X, \ldots, Y_m) = 0],$$

where $f \in L[X, Y_1, \ldots, Y_e]$ is irreducible and $s \in L[X, Y_1, \ldots, Y_{e-1}]$ is not contained in $\mathfrak{p}$. Thus, $\delta_{\mathfrak{p}}(X)$ is equivalent to the disjunction of the following two formulas

(1) $\exists Y_1 \cdots \exists Y_m \ [f(X, \ldots, Y_e) = 0 \neq g(X, Y_1, \ldots, Y_m)s(X, Y_1, \ldots, Y_{e-1})]$,

(2) $\exists Y_1 \cdots \exists Y_m \ [s(X, \ldots, Y_{e-1}) = p_1(X, \ldots, Y_m) = \cdots = p_n(X, \ldots, Y_m) = 0 \neq g(X, \vec{Y})]$.

By Lemma 4.2, we can replace the formula (1) with a disjunction of basic rankable formulas, each of which uses only e quantifiers. Since $e < m$, all these formulas have strictly smaller rank than β.

On the other hand, formula (2) has m quantifiers just like β, but we claim the associated variety has smaller dimension. Indeed, we see that

$$\dim(V(\mathfrak{p} + (s)) \cap D(g)) \leq \dim(V(\mathfrak{p} + (s))) < \dim(V(\mathfrak{p})) = \dim(V(\mathfrak{p}) \cap D(g)) = e,$$

where the strict inequality follows by Proposition 2.6. Therefore this formula also has strictly smaller rank than β. This completes the proof. □

We can say more about the hypersurface formula appearing in the previous result. First, we present a simple result on elements of the function field of an irreducible hypersurface.

Lemma 4.5. *Let $F \subseteq \overline{\mathbb{Q}}$ be a field and $f \in F[X, Y_1, \ldots, Y_m]$ an irreducible polynomial whose degree in Y_m is positive. If $\bar{p}/\bar{q} \in \text{Frac}(F[X, Y_1, \ldots, Y_m]/(f))$, then there are lifts of p and q to $F[X, Y_1, \ldots, Y_m]$ such that $\deg_{Y_m}(q) < \deg_{Y_m}(f)$.*

Proof. Write $f = \sum_{i=0}^{d} b_i Y_m^i$ where $b_i \in F[X, Y_1, \ldots, Y_{m-1}]$. Choose arbitrary lifts $p_0, q_0 \in F[Y_0, \ldots, Y_m]$ of $\bar{p}$ and $\bar{q}$. If $\deg_{Y_m} q_0 < \deg_{Y_m} f$, then we are already done. Otherwise, define $p_1 = b_d p_0$ and $q_1 = b_d q_0$, which define the same fraction in the function field because $b_d, q_0 \notin (f)$. Then the leading coefficient of q_1 is divisible by b_d, so we write it as $h_1 b_d$ for $h_1 \in F[Y_0, \ldots, Y_{m-1}]$. Define $q_2 = q_1 - h_1 Y_m^{(\deg_{Y_m} q_1) - d} f_1$, and notice that $\deg_{Y_m} q_2 < \deg_{Y_m} q_1$. Continuing in this way, the claim follows. □

Proposition 4.6. *Suppose $L \subseteq \overline{\mathbb{Q}}$ is a field and $\beta(X)$ is a formula with parameters from L of the following form*

$$\beta(X) = \exists Y_1 \ldots \exists Y_e [f(X, Y_1, \ldots, Y_e) = 0 \neq g(X, Y_1, \ldots, Y_e)]$$

Suppose $\beta(X)$ is L-minimal. Then f is absolutely irreducible.

Proof. First, it is clear that f is irreducible in L; if it were reducible then β could be equivalently expressed as the disjunction of two hypersurface formulas of strictly smaller rank.

Suppose for contradiction that f is not absolutely irreducible. We will use this fact to define $\{x \in L : \beta(x) \text{ holds in } L\}$ by a smaller rank formula using coefficients from L.

Let $F \subseteq L$ be a number field containing all the coefficients which appear anywhere in β. Let K be a finite Galois extension of F containing the coefficients of the absolutely irreducible factors of f over $\overline{\mathbb{Q}}$, and let $F' = K \cap L$. Then $F \subseteq F' \subseteq K$, and K is Galois

over F' because it was Galois over F. We remark that F' is a subfield of L, and therefore f is irreducible over F'.

For each of the finitely many number fields E with $F' \subset E \subseteq K$, let $p_E \in F'[Z]$ be a minimal polynomial for a primitive generator of E over F'. Since K is Galois over F', none of these finitely many p_E have a root in L. Let $h = \prod_{E:F' \subset E \subseteq K} p_E$.

We claim that L has a lower-ranked formula φ with coefficients from F' and with the property that for all $x \in L$, $\varphi(x)$ holds over L if and only if $\beta(x)$ does.

Let M be the function field of f over F'. By Proposition 2.2, M therefore contains some element $z_0 \in K \setminus F'$. Moreover, $F'(z_0)$ is a subfield of K which strictly contains F'. So $F'(z_0)$ contains a root z of h.

As an element of M, the root z will be of the form $\frac{p(X,\vec{Y})+(f)}{q(X,\vec{Y})+(f)}$, with $p, q \in F'[X, \vec{Y}]$. We may view p and q as polynomials $p, q \in F'[X, Y_1, \ldots, Y_e]$, modulo the ideal (f). These polynomials will satisfy

$$h\left(\frac{p(x,\vec{y})}{q(x,\vec{y})}\right) = 0$$

whenever $(x, \vec{y})$ is a solution to $f = 0$ and $q(x, \vec{y}) \neq 0$. Therefore, every solution $(x, \vec{y}) \in L^{m+1}$ to $f = 0$ has $q(x, \vec{y}) = 0$.

By Lemma 4.5, we may choose our specific $q \in F'[X, \vec{Y}]$ so that $\deg_{Y_e}(q) < \deg_{Y_e}(f)$. Notice that $q \notin (f)$ because $q + (f)$ is the denominator of an element of the function field, hence nonzero. Below we will consider q as a polynomial of degree d in Y_e, writing $q = \sum_{i \leq d} c_i Y_e^i$ with all $c_i \in F[X, Y_1, \ldots, Y_{e-1}]$. Without loss of generality, the leading nonzero coefficient c_d does not lie in (f). If it happens that Y_e does not appear in q, then $d = 0$ and $c_0 = q$.

But now we can use these facts to give a lower-ranked disjunction $\varphi(X) = \gamma_0(X) \vee \gamma_1(X)$ which is equivalent to $\beta(X)$ in L. Since Y_e has lower degree in q than in f, the trick is to use the Euclidean algorithm here, using the leading term in the expansion $f = \sum_{i=0}^{d_1} Y_e^i \cdot b_i(X, Y_1, \ldots, Y_{e-1})$ and writing

$$r(X, \vec{Y}) = c_d(X, \ldots, Y_{e-1}) \cdot f(X, \vec{Y}) - b_{d_1}(X, Y_1, \ldots, Y_{e-1}) \cdot Y_e^{d_1 - d} \cdot q(X, \vec{Y})$$

as a remainder with $\deg_{Y_e}(r) < \deg_{Y_e}(q)$. Recall that the polynomial c_d is the coefficient of Y_e^d in q, hence does not involve Y_e. Observe also that all coefficients of r are in F'.

We claim that in this situation, a tuple $(x, \vec{y}) \in L^{m+1}$ is a point on $V(f) \cap D(g)$ if and only if one of the following conditions holds:

$$q(x, \vec{y}) = r(x, \vec{y}) = 0 \neq g(x, \vec{y}) \cdot c_d(x, y_1, \ldots, y_{e-1}) \tag{3}$$

or

$$f(x, \vec{y}) = c_d(x, y_1, \ldots, y_{e-1}) = 0 \neq g(x, \vec{y}). \tag{4}$$

To see the claim, first let $(x, \vec{y})$ be a point on $V(f) \cap D(g)$. As shown above, we must have $q(x, \vec{y}) = 0$. But the Euclidean equation shows that $r(x, \vec{y}) = 0$ as well, so the tuple satisfies one of the conditions, according to whether $c_d(x, y_1, \ldots, y_{m-1}) = 0$ or not. The converse of the claim follows by applying the Euclidean equation to the first condition, and the latter condition directly defines a subset of $V(f) \cap D(g)$.

The formulas $\gamma_0(X)$ and $\gamma_1(X)$ that we promised above are simply the conditions in (3) and (4), each prefixed by $\exists Y_1 \cdots \exists Y_e$. Clearly these formulas have the same number of quantifiers

as β. The first formula corresponds to a subset $V(r,q) \cap D(gc_d)$ of $V(f) \cap D(g)$ because $r + b_{d_1} Y_e^{d_1 - d} q = c_d f$. Hence the dimension of the subset cannot exceed the dimension of $V(f) \cap D(g)$. However, r and q were constructed to have lower multidegree than f, so γ_0 has strictly smaller rank than β.

On the other hand, the affine variety over F' defined by the latter formula is a proper closed subset of $V(f)$, hence

$$\dim(V(f, c_d)) \cap D(g)) \leq \dim(V(f, c_d))) < \dim(V(f)) = \dim(V(f)) \cap D(g))$$

showing that γ_1 has strictly smaller rank than β. □

Putting these results together yields the following normal form theorem for existential formulas in algebraic extensions of $\mathbb{Q}$.

Definition 4.7. An *absolutely irreducible hypersurface formula* is a formula of the form

$$\exists Y_1 \ldots \exists Y_e \; [f(X, Y_1, \ldots, Y_e) = 0 \neq g(X, Y_1, \ldots, Y_e)]$$

for polynomials $f, g \in \overline{\mathbb{Q}}[X, Y_1, \ldots, Y_e]$, where f is absolutely irreducible and does not divide g.

Theorem 4.8 (Normal Form for Existential Definitions)**.** *For any field $L \subseteq \overline{\mathbb{Q}}$, if $A \subseteq L$ is existentially definable in L, then A is definable in L by a formula of the form*

$$\alpha(X) = \vee_{i=1}^{r} \beta_i(X),$$

where each $\beta_i(X)$ has one of the following forms:

(i) The quantifier-free formula $X = z_0$ for a fixed $z_0 \in L$.

(ii) An absolutely irreducible hypersurface formula with coefficients from L which is satisfied by infinitely many $x \in L$.

Proof. Apply Propositions 4.4 and 4.6, plus the following two observations. If f divides g in any of the hypersurface formulas, then that formula is unsatisfiable. If a hypersurface formula is satisfied by at most finitely many $x \in L$ (including if it is unsatisfiable), then it could be replaced by a (possibly empty) disjunction of formulas of the form $X = z_0$, lowering the rank. □

5. The meagerness of definability

Recall that by identifying a subset of $\overline{\mathbb{Q}}$ with its characteristic function, we can consider the set $\mathrm{Sub}(\overline{\mathbb{Q}}) = \{L \subseteq \overline{\mathbb{Q}} : L \text{ is a field}\}$ as a subset of $2^{\overline{\mathbb{Q}}}$, from which it inherits the product topology. A basis for the topology is given by the sets

$$U_{\vec{a},\vec{b}} = \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : a_1, \ldots a_n \in L \text{ and } b_1, \ldots, b_k \notin L\}$$

for any finite sequences of elements $\vec{a}, \vec{b}$ from $\overline{\mathbb{Q}}$. If $\vec{b}$ is empty, we write simply $U_{\vec{a}}$.

Recall that *Cantor space*, denoted 2^ω, is the set of infinite binary sequences with the product topology.

Proposition 5.1. *The space* $\mathrm{Sub}(\overline{\mathbb{Q}})$ *is homeomorphic to Cantor space.*

Proof. Since $\mathrm{Sub}(\overline{\mathbb{Q}})$ is a closed subset of the Cantor-homeomorphic space $2^{\overline{\mathbb{Q}}}$, it suffices to show that $\mathrm{Sub}(\overline{\mathbb{Q}})$ has no isolated points. But it is clear that whenever $U_{\vec{a},\vec{b}}$ is non-empty, there is $c \in \overline{\mathbb{Q}}$ such that both $U_{(\vec{a},c),\vec{b}}$ and $U_{\vec{a},(\vec{b},c)}$ are nonempty. □

The upshot of Proposition 5.1 is a structure on the set $\mathrm{Sub}(\overline{\mathbb{Q}})$ which allows us to describe when a set is "large" or "small" in terms of topology. In particular, we enlist the notions of meager sets and the property of Baire.

Definition 5.2. A subset of a topological space is called *nowhere dense* if its closure has empty interior, and *meager* if it is the countable union of nowhere dense sets. A topological space is *Baire*[1] if every non-empty open subset is non-meager.

Cantor space 2^ω is Baire, and by Proposition 5.1 the same is true for $\mathrm{Sub}(\overline{\mathbb{Q}})$, which allows us to consider meager sets to be small.

Definition 5.3. For any $Z \subseteq \mathbb{Q}$, and formula $\beta(X)$ with coefficients $\vec{a}$ from $\overline{\mathbb{Q}}$, we define $S_\beta(Z)$ to be the set of algebraic fields in which β defines a set disjoint from Z:

$$S_\beta(Z) = \{L \in U_{\vec{a}} : \{x \in L : \beta(x) \text{ holds over } L\} \cap Z = \emptyset\}.$$

Definition 5.4. For any $Z \subseteq \mathbb{Q}$, let $S(Z)$ denote the set

$$S(Z) = \bigcup_\beta S_\beta(Z)$$

where β ranges over absolutely irreducible hypersurface formulas with coefficients from $\overline{\mathbb{Q}}$.

Proposition 5.5. *Let Z be a subset of $\mathbb{Q}$ that is not thin in $\mathbb{Q}$. Then $S(Z)$ is meager. In particular, for every absolutely irreducible hypersurface formula*

$$\beta(X) : \ \exists \vec{Y}[f(X, \vec{Y}) = 0 \neq g(X, \vec{Y})]$$

with coefficients $\vec{a}$ from $\overline{\mathbb{Q}}$, the set $S_\beta(Z)$ is nowhere dense.

Proof. Since $S_\beta(Z) \subseteq U_{\vec{a}}$, it suffices to show that $S_\beta(Z)$ is nowhere dense in $U_{\vec{a}}$. Let $\vec{b}$ and $\vec{c}$ be any sequences of elements of $\overline{\mathbb{Q}}$ such that $U_{(\vec{a},\vec{c}),\vec{b}} \neq \emptyset$. Let $F = \mathbb{Q}(\vec{a}, \vec{c})$ and let $K = F(\vec{b})$. By the application of Hilbert's Irreducibility Theorem in Proposition 2.13, there is a thin set $T \subseteq K^e$ such that for any $(x, y_1, \ldots, y_{e-1}) \in K^e \setminus T$, the polynomial $f(x, y_1, \ldots, y_{e-1}, Y_e)$ is irreducible of degree $\deg_{Y_e}(f)$, and $g(x, y_1, \ldots, y_{e-1}, Y_e)$ is not divisible by f. Because K is a number field, $T_\mathbb{Q} = T \cap \mathbb{Q}^e$ is also a thin set in $\mathbb{Q}^e$ by Proposition 2.10. Further, since Z is not thin in $\mathbb{Q}$, the thin set $T_\mathbb{Q}$ does not contain all of $Z \times \mathbb{Q}^{e-1}$ by Lemma 2.12. For any such tuple $(x, y_1, \ldots, y_{e-1}) \in Z \times \mathbb{Q}^{e-1}$ outside this thin set, the irreducibility of $f(x, y_1, \ldots, y_{e-1}, Y)$ over K implies that adjoining to F any root y of $f(x, y_1, \ldots, y_{e-1}, Y)$ will not generate any element of K: we will have $F(y) \cap K = F$, by Lemma 2.1. Thus $U_{(\vec{a},\vec{c},y),\vec{b}}$ is nonempty. Additionally, the divisibility condition implies that $g(x, y_1, \ldots, y_{m-1}, y) \neq 0$. So for any $L \in U_{(\vec{a},\vec{c},y),b}$, $\beta(x)$ holds in L. Therefore, $U_{(\vec{a},\vec{c},y),b} \cap S_\beta(Z) = \emptyset$.

There are only countably many β, so $S(Z)$ is a countable union of meager sets, and is thus meager. □

[1]Some authors use the terminology *Baire space* to refer to topological spaces with this property. However, we reserve the name Baire space for the particular topological space ω^ω, which is discussed in related papers, such as [14], although we will not use it in this paper.

Theorem 5.6. *If $Z \subset \mathbb{Q}$ is not thin, then the following is meager in* $\mathrm{Sub}(\overline{\mathbb{Q}})$*:*

$$S_Z = \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : \textit{some coinfinite } C \subseteq L, \textit{ universally definable in } L, \textit{ has } C \cap \mathbb{Q} = Z\}$$
$$= \bigcup_{A \subseteq \overline{\mathbb{Q}}:\ A \cap \mathbb{Q} = Z} \{L \in \mathrm{Sub}(\overline{\mathbb{Q}}) : A_L \textit{ is coinfinite and universally definable in } L\}$$

Proof. Let $S = S(Z)$. By Proposition 5.5, S is meager. Let $A \subseteq L$ be coinfinite with $A \cap \mathbb{Q} = Z$. Suppose that A is universally definable in L. Then $L \setminus A$ is existentially definable in L. So by Theorem 4.8, $L \setminus A$ is definable in L by a formula $\alpha = \vee_{i<r}\beta_i$ in normal form. Because $L \setminus A$ is infinite and r is finite, some β_i must be an absolutely irreducible hypersurface formula, and

$$\{x \in L : \beta_i(x) \text{ holds over } L\} \cap A = \emptyset.$$

So $\{x \in L : \beta_i(x) \text{ holds over } L\} \cap Z = \emptyset$. By definition, $L \in S_{\beta_i}(Z)$, as needed. □

As a corollaries we have the following.

Theorem 5.7. *The set of all fields $L \in \mathrm{Sub}(\overline{\mathbb{Q}})$ such that $\mathcal{O}_L$ is either existentially or universally definable in L is meager.*

Proof. Recall that $\mathbb{Z} = \mathbb{Q} \cap \mathcal{O}_L$ for all subfields L. By Proposition 2.11, neither $\mathbb{Z}$ nor $\mathbb{Q} \setminus \mathbb{Z}$ is thin in $\mathbb{Q}$. Applying Theorem 5.6 to $Z = \mathbb{Z}$ and $Z = \mathbb{Q} \setminus \mathbb{Z}$, we see that there is a meager set S such that if either $\mathcal{O}_L$ or $L \setminus \mathcal{O}_L$ is universally definable in L, then $L \in S$. □

Corollary 5.8. *The set of fields $L \in \mathrm{Sub}(\overline{\mathbb{Q}})$ such that $\mathbb{Z}$ itself is either existentially or universally definable in L is meager.*

We can also use the same approach when considering the definability of number fields. Of course $\emptyset$ is a thin subset of $\mathbb{Q}$, so Theorem 5.6 does not directly rule out an existential definition of $\mathbb{Q}$ in a generic algebraic extension L. Nevertheless, we have the following.

Corollary 5.9. *If F is a number field, then the set of fields $L \in \mathrm{Sub}(\overline{\mathbb{Q}})$ containing F such that F has an existential definition in L is a meager set.*

Proof. Let S be the meager set guaranteed by Theorem 5.6 for $Z = \mathbb{Z}$. By Park's generalization [15] of a theorem of Koenigsmann [10], there is a quantifier-free formula $\varphi(X, Y_1, \ldots, Y_n)$, in the language of fields, such that $\exists \vec{Y} \varphi(X, \vec{Y})$ defines the algebraic non-integers $F \setminus \mathcal{O}_F$ in the field F. In particular, it defines $\mathbb{Q} \setminus \mathbb{Z}$ in $\mathbb{Q}$ over F. Now if $\gamma(Y)$ is existential and defines F in L, then the following formula with free variable X,

$$\gamma(X) \ \&\ \exists \vec{Y}[\ \gamma(Y_1)\ \&\cdots\&\ \gamma(Y_n)\ \&\ \varphi(X, \vec{Y})\ ],$$

is an existential definition of $F \setminus \mathcal{O}_F$ in L. So $(L \setminus F) \cup \mathcal{O}_F$ is universally definable in L. Since $((L \setminus F) \cup \mathcal{O}_F) \cap \mathbb{Q} = \mathbb{Z}$, we have $L \in S$. □

5.1. **Computable fields whose algebraic integers are not one-quantifier definable.** Next we effectivize Theorem 5.7 to obtain many computable algebraic extensions of $\mathbb{Q}$ whose algebraic integers are not existentially or universally definable.

Our arguments below will require the decidability of absolute irreducibility. Recall some standard terminology: a computable field E has a *splitting algorithm* if the *splitting set* $S_E = \{f \in E[T] : f \text{ is reducible in } E[T]\}$ is decidable, and has a *root algorithm* if the *root set* $R_E = \{f \in E[T] : f \text{ has a root in } E\}$ is decidable. Notice that these are both stated

for single-variable polynomials. The next lemma is a specific case of the fact that splitting algorithms can be extended to more variables.

Lemma 5.10. *Fix any computable presentation of $\overline{\mathbb{Q}}$. Then it is decidable which polynomials in $\overline{\mathbb{Q}}[X_1, X_2, \ldots]$ are absolutely irreducible.*

Proof. $\overline{\mathbb{Q}}$ has a splitting algorithm, of course: all polynomials in $\overline{\mathbb{Q}}[T]$ of degree > 1 are reducible. The lemma now follows from another theorem of Kronecker (found in [3, §§ 58-59]), stating that whenever a computable field F has a splitting algorithm and t is transcendental over F (within a larger computable field), the field $F(t)$ also has a splitting algorithm. The irreducible polynomials of $\overline{\mathbb{Q}}[X_1, X_2]$ are precisely the irreducible polynomials of $\overline{\mathbb{Q}}[X_1]$ along with the polynomials which are irreducible in $\overline{\mathbb{Q}}(X_1)[X_2]$ and have no common factor among the coefficients lying in $\overline{\mathbb{Q}}[X_1]$; see [11, Theorem IV.2.3]. Therefore, reducibility is clearly decidable using Kronecker's result. Thus we can decide reducibility in $\overline{\mathbb{Q}}[X_1, X_2]$, and one continues by induction on the number n of variables, noting that the resulting decision procedures are uniform in n. □

Therefore, there is a computable listing $\beta_1, \beta_2, \ldots$ of all absolutely irreducible hypersurface formulas. Furthermore, we have the following effective version of Proposition 5.5.

Proposition 5.11. *Let Z be a computable subset of $\mathbb{Q}$ that is not thin in $\mathbb{Q}$. Then there is an algorithm which, given any absolutely irreducible hypersurface formula β with coefficients $\vec{a}$, and any $\vec{c}, \vec{b}$ such that $U_{(\vec{a},\vec{c}),\vec{b}} \neq \emptyset$, returns y such that $U_{(\vec{a},\vec{c},y),\vec{b}}$ is non-empty and has empty intersection with $S_\beta(Z)$.*

Proof. The proof of Proposition 5.5 shows that there is a tuple $(x, y_1, \ldots, y_{e-1}, y) \in Z \times \mathbb{Q}^{e-1} \times \overline{\mathbb{Q}}$ which witnesses that $\beta(x)$ holds in each field extending $\mathbb{Q}(y)$ while keeping $U_{(\vec{a},\vec{c},y),\vec{b}}$ non-empty. So an algorithm can search all such $x, y_1, \ldots, y_{e-1}, y$ until it finds one. This works because Z is computable, and it is computable to check whether a given tuple from $\overline{\mathbb{Q}}$ satisfies the polynomials appearing in β, and computable to check whether $U_{(\vec{a},\vec{c},y),\vec{b}}$ is empty. □

Theorem 5.12. *Let $Z \subseteq \mathbb{Q}$ be a computable subset which is neither thin nor co-thin. For every pair of $\overline{\mathbb{Q}}$-tuples $(\vec{a}, \vec{b})$, if $U_{\vec{a},\vec{b}}$ is nonempty, then there is a computable $L \in U_{\vec{a},\vec{b}}$ which enjoys the following property: If $A \subseteq L$ is any subset such that $A \cap \mathbb{Q} = Z$, then A is neither existentially nor universally definable in L. Moreover, every computable presentation of L has a splitting algorithm.*

Proof. We recursively define sequences $\vec{a} = \vec{a}_0, \vec{a}_1, \ldots$ and $\vec{b} = \vec{b}_0, \vec{b}_1, \ldots$ in stages as follows. Recall that $\beta_1, \beta_2, \ldots$ is a computable listing of all absolutely irreducible hypersurface formulas. Let $c_1, c_2, \ldots$ be a computable listing of all elements of $\overline{\mathbb{Q}}$.

At stages of the form $s = 3t + 1$, given $U_{\vec{a}_{s-1},\vec{b}_{s-1}}$ nonempty, use Proposition 5.11 to find a y such that $U_{(\vec{a}_{s-1},y),\vec{b}_{s-1}}$ is non-empty and disjoint from $S_{\beta_t}(Z)$. Let $\vec{a}_s = (\vec{a}_{s-1}, y)$ and $\vec{b}_s = \vec{b}_{s-1}$.

At stages of the form $s = 3t + 2$, use an analogous process to avoid $S_{\beta_t}(\mathbb{Q} \setminus Z)$.

At stages of the form $s = 3t + 3$, consider $U_{(\vec{a}_{s-1},c_t),\vec{b}_{s-1}}$ and if it is nonempty, set $\vec{a}_s = (\vec{a}_{s-1}, c_t), \vec{b}_s = \vec{b}_{s-1}$. Otherwise, set $\vec{a}_s = \vec{a}_{s-1}$ and $\vec{b}_s = (\vec{b}_{s-1}, c_t)$.

Because $\mathrm{Sub}(\overline{\mathbb{Q}})$ is a compact topological space by Proposition 5.1 and the family of closed subsets $\{U_{\vec{a}_s,\vec{b}_s} : s \geq 0\}$ has the finite intersection property, it follows that the intersection $\cap_s U_{\vec{a}_s,\vec{b}_s}$ is nonempty. Moreover, the intersection is a singleton because after stage $3t+3$, the element c_t is included in either all or none of the fields in $U_{\vec{a_s},\vec{b_s}}$ by construction. In particular, this unique field is $L = \{a \in \overline{\mathbb{Q}} : a \text{ appears in some } \vec{a}_s\}$.

The field L is computable because by stage $3t+3$ it has been decided whether c_t is included. If $A \subset L$ is any subset with $A \cap \mathbb{Q} = Z$, then Theorem 5.6 implies that A is neither existentially nor universally definable in L because the construction of L explicitly avoids the sets $S(Z)$ and $S(\mathbb{Q} \setminus Z)$ by definition.

The splitting algorithm for L follows from Rabin's Theorem (see [18]), since L is given as a decidable subfield of (our computable presentation of) $\overline{\mathbb{Q}}$. Finally, whenever $L \cong \widetilde{L}$ are computable algebraic fields, their splitting sets are Turing-equivalent, so all computable presentations of L have splitting algorithms. □

As an application, because $\mathbb{Z}$ is neither thin nor co-thin in $\mathbb{Q}$, the following is immediate.

Corollary 5.13. *For every pair of $\overline{\mathbb{Q}}$-tuples $(\vec{a}, \vec{b})$, if $U_{\vec{a},\vec{b}}$ is nonempty, then there is a computable $L \in U_{\vec{a},\vec{b}}$ such that $\mathcal{O}_L$ is neither existentially nor universally definable in L. Moreover, every computable presentation of L has a splitting algorithm.*

5.2. **The topological space of algebraic extensions of $\mathbb{Q}$ up to isomorphism.** The questions of definability we have considered have the same answer over isomorphic fields. Although $\mathrm{Sub}(\overline{\mathbb{Q}})$ contains at least one isomorphic copy of every possible algebraic extension of $\mathbb{Q}$, it contains exactly one copy of an algebraic extension L of $\mathbb{Q}$ if and only if L is Galois over $\mathbb{Q}$. A number field F of degree n is isomorphic to at most n fields in $\mathrm{Sub}(\overline{\mathbb{Q}})$, but there are some infinite non-Galois extensions of $\mathbb{Q}$ which are isomorphic to uncountably many elements in $\mathrm{Sub}(\overline{\mathbb{Q}})$. Therefore, given the isomorphism invariance of the property under consideration, one might wonder if the results of the previous section have been skewed by the fact that some isomorphism classes are more represented in $\mathrm{Sub}(\overline{\mathbb{Q}})$ than others.

Thus it is also of interest to consider the collection of algebraic extensions of $\mathbb{Q}$ up to isomorphism as a topological space, as was done in [14]. We denote this set by $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$. From the perspective of number theory, the set $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ can be identified as a quotient of $\mathrm{Sub}(\overline{\mathbb{Q}})$ by the absolute Galois group $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, which equates isomorphic fields. The topology on $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ is the quotient topology which it inherits from $\mathrm{Sub}(\overline{\mathbb{Q}})$.

Alternatively, from the perspective of computability theory, one could begin with the space $\mathcal{ALG}_0^*$ of all possible presentations of algebraic extensions of $\mathbb{Q}$ in a certain language. This is done in [14] and the relevant language in this case is the language of rings enlarged to include additional predicates for the existence of roots of monic one-variable polynomials. Equating isomorphic fields and taking the quotient topology leads to the space $\mathcal{ALG}_0^*/\cong$, which coincides with $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ despite various differences between $\mathcal{ALG}_0^*$ and $\mathrm{Sub}(\overline{\mathbb{Q}})$. For example, in $\mathcal{ALG}_0^*$, every isomorphism class is represented with uncountably many copies. For details about $\mathcal{ALG}_0^*$, we refer the reader to [14].

Returning now to $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$, observe that for any $U_{\vec{a},\vec{b}}$, the following set is the smallest G-invariant subset of $\mathrm{Sub}(\overline{\mathbb{Q}})$ containing $U_{\vec{a},\vec{b}}$. It is also clopen, as there are only finitely

many images $\varphi(\vec{a}), \varphi(\vec{b})$.

$$GU_{\vec{a},\vec{b}} := \{\varphi(L) : L \in U_{\vec{a},\vec{b}}, \varphi \in G\} = \bigcup_{\varphi \in G} U_{\varphi(\vec{a}),\varphi(\vec{b})}$$

It follows that the quotient map $q : \mathrm{Sub}(\overline{\mathbb{Q}}) \to \mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ is open and the images of the sets $GU_{\vec{a},\vec{b}}$ form a clopen basis for $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$.

Proposition 5.14 (Theorem 3.3, [14])**.** $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ *is homeomorphic to Cantor space.*

Proof. The follows because $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ is compact, has a countable clopen basis, and has no isolated points. The last condition follows because every non-empty $GU_{\vec{a},\vec{b}}$ contains at least two non-isomorphic fields. □

Therefore, notions of meager and co-meager make sense in $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$. We can easily transfer the all our results about $\mathrm{Sub}(\overline{\mathbb{Q}})$ to results about $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ by replacing the sets $S_\beta(Z)$ with the following G-invariant sets.

Definition 5.15. For any $Z \subseteq \mathbb{Q}$, and any absolutely irreducible hypersurface formula β with coefficients from $\overline{\mathbb{Q}}$, let

$$GS_\beta(Z) = \bigcup_{\varphi \in G} S_{\varphi(\beta)}(Z),$$

where $\varphi(\beta)$ denotes the result of applying φ to all coefficients appearing in β.

Since every $\varphi \in G$ fixes every element of $\mathbb{Q}$, each $GS_\beta(Z)$ is G-invariant.

Proposition 5.16. *Let β be an absolutely irreducible hypersurface formula with coefficients from $\overline{\mathbb{Q}}$. If $Z \subseteq \mathbb{Q}$ is not thin in $\mathbb{Q}$, then $q(GS_\beta(Z))$ is nowhere dense in $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$, where $q : \mathrm{Sub}(\overline{\mathbb{Q}}) \to \mathrm{Sub}(\overline{\mathbb{Q}})/\cong$ is the quotient map.*

Proof. There are only finitely many coefficients in β, so only finitely many possible outcomes for $\varphi(\beta)$. So $GS_\beta(Z)$ is a finite union of nowhere dense sets, and thus is nowhere dense. Additionally, since each $S_{\varphi(\beta)}(Z)$ is closed, so is $GS_\beta(Z)$. Since $\mathrm{Sub}(\overline{\mathbb{Q}}) \setminus GS_\beta(Z)$ is dense open and q is an open map, its image $q(\mathrm{Sub}(\overline{\mathbb{Q}}) \setminus GS_\beta(Z))$ is dense open. Therefore, by G-invariance of $GS_\beta(Z)$, $q(GS_\beta(Z))$ is nowhere dense. □

Theorem 5.17. *If any set $Z \subset \mathbb{Q}$ is not thin, then the following is meager in $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$:*

$S_Z = \{[L]_\cong \in \mathrm{Sub}(\overline{\mathbb{Q}})/\cong : \textit{some coinfinite } C \subseteq L, \textit{ universally definable in } L, \textit{ has } C \cap \mathbb{Q} = Z\}$.

Proof. Let $S = \bigcup_\beta q(GS_\beta(Z))$, where β ranges over all absolutely irreducible hypersurface formulas. By Proposition 5.16, S is meager. Let $A \subseteq L$ be coinfinite with $A \cap \mathbb{Q} = Z$. Suppose that A is universally definable in L. Then by the same argument as in Theorem 5.6, $L \in S_\beta(Z)$ for some β. So $L \in GS_\beta(Z)$, as needed. □

Therefore, we have the following analogues of the results of the previous section.

Corollary 5.18. *The following sets are meager in $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$:*

(1) The set of isomorphism types of fields L in which $\mathcal{O}_L$ is existentially or universally definable.
(2) The set of isomorphism types of fields in which $\mathbb{Z}$ is existentially or universally definable.

(3) *The set of isomorphism types of fields L in which some number field $F \subset L$ is existentially definable.*

Proof. These sets are all contained in the set S guaranteed by Theorem 5.17 when $Z = \mathbb{Z}$. □

It may seem equally natural to consider the Lebesgue measure on Cantor space and transfer it to $\mathrm{Sub}(\overline{\mathbb{Q}})/\cong$, using some computable homeomorphism such as that obtained in [14, Theorem 3.3]. This is attempted to some extent in [14], but the resulting measure is not canonical: it depends to a great extent on arbitrary choices that are made during the construction of the homeomorphism. Indeed, the notion of *Haar-compatible measure*, put forth in [14], has had to be abandoned, as the reality is more complicated than the analysis in that article recognized. We hope to investigate this situation, and measure-theoretic perspectives in general, more fully in the near future.

References

[1] Davis, M., Putnam, H., and Robinson, J. The decision problem for exponential diophantine equations. *Ann. of Math. (2) 74* (1961), 425–436.

[2] Dittmann, P., and Fehm, A. Non-definability of rings of integers in most algebraic fields. *Notre Dame J. Form. Log. 62*, 3 (2021), 589–592.

[3] Edwards, H. M. *Galois theory*, vol. 101 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984.

[4] Ershov, Y. L. Fields with continuous local elementary properties. II. *Algebra i Logika 34*, 3 (1995), 262–273, 363.

[5] Fisher, S., and Gartside, P. On the space of subgroups of a compact group. I. *Topology Appl. 156*, 5 (2009), 862–871.

[6] Fried, M. D., Haran, D., and Völklein, H. Real Hilbertianity and the field of totally real numbers. In *Arithmetic geometry (Tempe, AZ, 1993)*, vol. 174 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 1994, pp. 1–34.

[7] Fukuzaki, K. Definability of the ring of integers in some infinite algebraic extensions of the rationals. *MLQ Math. Log. Q. 58*, 4-5 (2012), 317–332.

[8] Gartside, P., and Smith, M. Counting the closed subgroups of profinite groups. *J. Group Theory 13*, 1 (2010), 41–61.

[9] Haran, D., and Jarden, M. The absolute Galois group of a pseudo p-adically closed field. *J. Reine Angew. Math. 383* (1988), 147–206.

[10] Koenigsmann, J. Defining $\mathbb{Z}$ in $\mathbb{Q}$. *Ann. of Math. (2) 183*, 1 (2016), 73–93.

[11] Lang, S. *Algebra*, third ed., vol. 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[12] Liu, Q. *Algebraic geometry and arithmetic curves*, vol. 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. Translated from the French by Reinie Erné, Oxford Science Publications.

[13] Matiyasevich, J. V. The Diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR 191* (1970), 279–282.

[14] Miller, R. Isomorphism and classification for countable structures. *Computability 8*, 2 (2019), 99–117.

[15] Park, J. A universal first-order formula defining the ring of integers in a number field. *Math. Res. Lett. 20*, 5 (2013), 961–980.

[16] Poonen, B. Characterizing integers among rational numbers with a universal-existential formula. *Amer. J. Math. 131*, 3 (2009), 675–682.

[17] Pop, F. Classically projective groups and pseudo classically closed fields. In *Valuation theory and its applications, Vol. II (Saskatoon, SK, 1999)*, vol. 33 of *Fields Inst. Commun.* Amer. Math. Soc., Providence, RI, 2003, pp. 251–283.

[18] Rabin, M. Computable algebra, general theory, and theory of computable fields. *Trans. AMS 95* (1960), 341–360.

[19] Reid, M. *Undergraduate commutative algebra*, vol. 29 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1995.
[20] Robinson, J. Definability and decision problems in arithmetic. *J. Symbolic Logic 14* (1949), 98–114.
[21] Robinson, J. The undecidability of algebraic rings and fields. *Proc. Amer. Math. Soc. 10* (1959), 950–957.
[22] Robinson, J. On the decision problem for algebraic rings. In *Studies in mathematical analysis and related topics*. Stanford Univ. Press, Stanford, Calif, 1962, pp. 297–304.
[23] Rumely, R. S. Undecidability and definability for the theory of global fields. *Trans. Amer. Math. Soc. 262*, 1 (1980), 195–217.
[24] Serre, J.-P. *Topics in Galois theory*, second ed., vol. 1 of *Research Notes in Mathematics*. A K Peters, Ltd., Wellesley, MA, 2008. With notes by Henri Darmon.
[25] Shlapentokh, A. First-order decidability and definability of integers in infinite algebraic extensions of the rational numbers. *Israel J. Math. 226*, 2 (2018), 579–633.
[26] Videla, C. R. Definability of the ring of integers in pro-p Galois extensions of number fields. *Israel J. Math. 118* (2000), 1–14.

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA
E-mail address: kxe8@psu.edu

Department of Mathematics, Queens College – City University of New York, 65-30 Kissena Blvd., Flushing, NY 11367 USA
Ph.D. Programs in Mathematics and Computer Science, Graduate Center - City University of New York, 365 Fifth Avenue, New York, NY 10016 USA
E-mail address: Russell.Miller@qc.cuny.edu

Department of Mathematics, University College London, London WC1H 0AY, UK
E-mail address: c.springer@ucl.ac.uk

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA
E-mail address: westrick@psu.edu