On the algebraic points of a definable set

Jonathan Pila

Abstract

This paper studies diophantine properties of sets definable in an o-minimal structure over the real field. It observes a refinement of the main theorem of the author’s recent paper with A J Wilkie, and uses it to deduce a strong result on the density of algebraic points of such sets.

2000 Mathematics Subject Classification: 11G99, 03C64

Keywords: o-minimal structure, definable set, rational point, algebraic point

1. Introduction

This paper builds on [15] in considering diophantine properties of certain non-algebraic subsets \(X \subset \mathbb{R}^n\). We use geometric means to control the rational or algebraic points of \(X\), so it is natural to restrict to sets \(X\) with suitably nice geometric properties. The sets definable in an \(o\)-minimal structure (over \(\mathbb{R}\)) provide a natural setting for the methods employed in [15] and pursued here. A definition is given below in Section 2.

Examples of \(o\)-minimal structures (over \(\mathbb{R}\)) include: the structure comprising all semialgebraic sets (the paradigm example, whose \(o\)-minimality follows from the Tarski-Seidenberg Theorem); the structure comprising globally subanalytic sets \(\mathbb{R}_{an}\) (for which \(o\)-minimality is afforded by Gabrielov’s Theorem [6]); the structure \(\mathbb{R}_{exp}\) of sets definable using the exponential function (Wilkie [20]); the structure \(\mathbb{R}_{pfaff}\) generated by pfaffian sets (Wilkie [21]). Thus \(o\)-minimal structures provide rich and flexible categories to work in while maintaining strong geometric finiteness properties. For a description of these and further examples see [4, 5, 17, 18, 19] in addition to the references above, for a development of the theory of \(o\)-minimal structures see [3, 5], and for the context of the diophantine results pursued here see [15].

1.1. Definition. Let \(S\) be an \(o\)-minimal structure over \(\mathbb{R}\), which will remain fixed. We call \(X \subset \mathbb{R}^n\) definable if it definable in \(S\).

In studying the diophantine properties of such sets, our guiding idea is that a transcendental set should contain only few rational points, in a suitable sense. However, a non-algebraic set \(X\) may contain semi-algebraic subsets of positive dimension. Such semi-algebraic subsets may contain many rational points (where “many” means \(\gg T^3\) points of height \(\leq T\)). To isolate the truly transcendental part of a set \(X\), we make the following definition.

1.2. Definition. Let \(X \subset \mathbb{R}^n\). The algebraic part of \(X\), denoted \(X^{alg}\), is the union of all connected semi-algebraic subsets of \(X\) of positive dimension.

Our primary interest is in diophantine properties of the transcendental part \(X - X^{alg}\) of \(X\). The transcendental part of a definable set may contain infinitely many rational points (e.g. \(y = 2^x\) definable in \(\mathbb{R}_{exp}\)). So we define a height function on the points of interest and a function counting points up to given height. We then seek upper estimates for this counting function. Excluding \(X^{alg}\) is thus crudely analogous to excluding the special set when counting rational points up to height \(T\) on an algebraic variety.

By \(H(\alpha)\) we denote the (multiplicative) height of an algebraic number \(\alpha\), i.e. \(H(\alpha) = \exp h(\alpha)\), where \(h(\alpha)\) is the absolute logarithmic height, as defined in [1, p16] (see the characterization of \(h(\alpha)\) in Section 5 below). As \(h(\alpha) \geq 0\), \(H(\alpha) \geq 1\). For a rational number \(a/b\) in lowest terms, \(H(a/b) = \max(|a|, |b|)\).

For an \(n\)-tuple \((\alpha_1, \ldots, \alpha_n)\) of algebraic numbers we define \(H(\alpha_1, \ldots, \alpha_n) = \max_i H(\alpha_i)\). Note that this is not the usual projective height \(H^{proj}\) of an \(n\)-tuple, but simply coordinate-wise (affine) height, which seems more natural when considering non-algebraic sets. As \(H \leq H^{proj}\) ([1, p16 cf p15], e.g. \(H^{proj}\) of a tuple of rational numbers involves a least common denominator for all the coordinates), our results are marginally stronger than if stated using \(H^{proj}\), though this would be seen only in the implicit constants.
1.3. Definition. For a set \( X \subset \mathbb{R}^n \) let \( X(\mathbb{Q}) = X \cap \mathbb{Q}^n \). For a real number \( T \geq 1 \), set
\[
X(\mathbb{Q}, T) = \{ x \in X(\mathbb{Q}) : H(x) \leq T \},
\]
and define the counting function
\[
N(X, T) = \#X(K, T).
\]

With these definitions, the basic result of [15] is as follows.

1.4. Theorem. Let \( X \subset \mathbb{R}^n \) be definable and let \( \epsilon > 0 \). Then, for \( T \geq 1 \),
\[
N(X - X^{\text{alg}}, T) = O_{X, \epsilon}(T^\epsilon).
\]

This result can be adapted to estimate points of \( X \) with coordinates in a number field up to height \( T \) in a straightforward way (for the graph of a transcendental real analytic function on a compact interval this is already in [10]). For a set \( X \subset \mathbb{R}^n \) and a number field \( K \subset \mathbb{R} \) let \( X(K) = X \cap K^n \) and define, for a real number \( T \geq 1 \),
\[
X(K, T) = \{ x \in X(K) : H(x) \leq T \}, \quad N_K(X, T) = \#X(K, T).
\]

1.5. Theorem. Let \( X \subset \mathbb{R}^n \) be definable, let \( K \subset \mathbb{R} \) be a number field of degree \( k \), and let \( \epsilon > 0 \). Then
\[
N_K(X - X^{\text{alg}}, T) = O_{X, k, \epsilon}(T^\epsilon).
\]

Theorem 1.4 is established in a refined form in [15]. The main point of this paper is that the proof in [15] actually yields a further strengthening. We defer stating this result to Section 3, but it enables us to obtain a result about algebraic points that is much stronger than 1.5 in which we consider, given a positive integer \( k \), points \( x \) of \( X \) that are defined over any algebraic number field of degree \( k \) — indeed different coordinates of \( x \) may be defined over different fields. We define, for a set \( X \subset \mathbb{R}^n \), a positive integer \( k \), and a real number \( T \geq 1 \),
\[
X(k, T) = \{ x = (x_1, \ldots, x_n) \in X : \max_i \lfloor x_i / q_i \rfloor \leq k, \max_i H(x_i) \leq T \}, \quad N_k(X, T) = \#X(k, T).
\]

1.6. Theorem. Let \( X \subset \mathbb{R}^n \) be definable, let \( k \) be a positive integer, and let \( \epsilon > 0 \). Then
\[
N_{k}(X - X^{\text{alg}}, T) = O_{X, k, \epsilon}(T^\epsilon).
\]

The results of [15] have been used to prove diophantine results. A new proof of the Manin-Mumford conjecture by the present author and Zannier [16], and affirmation of a special case of the Zilber- Pink conjecture by Masser and Zannier [9] both make essential use of [15] or its predecessors. It is anticipated that Theorem 1.6 will also have applications in this circle of diophantine problems.

Our strengthening of 1.4 also yields a result for points of \( X \) whose coordinates lie in some finite-dimensional \( \mathbb{Q} \) vector subspace of \( \mathbb{R} \). For \( \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k \) let \( \mathbb{Q}(\lambda) = \{ \sum_i q_i \lambda_i : q_i \in \mathbb{Q} \} \). For \( x \in \mathbb{Q}(\lambda) \) set \( H_\lambda(x) = \min \{ \max_i H(q_i) : q_1, \ldots, q_k \in \mathbb{Q}, \sum_i q_i \lambda_i = x \} \), with \( H_\lambda(x_1, \ldots, x_n) = \max_i H_\lambda(x_i) \) for \( x = (x_1, \ldots, x_n) \in \mathbb{Q}(\lambda)^n \). We let \( X(\mathbb{Q}(\lambda)) = X \cap \mathbb{Q}(\lambda)^n \), and define
\[
X(\lambda, T) = \{ x \in X(\mathbb{Q}(\lambda)) : H_\lambda(x) \leq T \}, \quad N_\lambda(X, T) = \#X(\lambda, T).
\]

1.7. Theorem. Let \( X \subset \mathbb{R}^n \) be definable, \( k \) a positive integer, \( \lambda \in \mathbb{R}^d \) and \( \epsilon > 0 \). Then
\[
N_\lambda(X - X^{\text{alg}}, T) = O_{X, k, \epsilon}(T^\epsilon).
\]

While in all the statements above we exclude the algebraic part \( X^{\text{alg}} \) in its entirety, it is the exploration of precisely which parts of \( X^{\text{alg}} \) need to be removed from \( X \) to achieve the \( O(T^\epsilon) \) estimate that is the focus of our strengthening of 1.4, and the key to proving 1.6 and 1.7.

Theorem 1.7 is a somewhat easier consequence of our main result than 1.6 and we prove it first, in Section 4, before establishing 1.6 in Section 5.
As remarked in [15], the $O(T^s)$ estimate in 1.4 cannot be much improved for a general o-minimal structure, in view of one-dimensional examples in $\mathbb{R}_{an}$ constructed in [12, 7.5]. However, Wilkie has conjectured ([15, 1.11]) that a much better estimate holds for sets $X$ definable in $\mathbb{R}_{exp}$, namely, for $T \geq e$,

$$N(X - X^{alg}, T) = O_X \left( (\log T)^{O_X(1)} \right)$$

One would expect such estimates to hold likewise for the strengthenings of 1.1 obtained in [15] and here, and also for algebraic points.

In [14] I established such an estimate for the rational points on a pfaff curve. A pfaff curve $X$ is the graph of a pfaffian function $f$ of one variable on some connected subset of its domain. A definition of pfaffian function is given in Section 6. Pfaffian functions and the sets defined by them arose in the work of Khovanskii [8] on “fewnomials”. The reader is referred there and to [7] for further information, and to [21] for o-minimal aspects.

This paper concludes with the observation that an adaptation of the method of [14] obtains a result for points on a pfaff curve defined over a number field in which the exponent of $\log T$ is independent of the field.

1.8. Theorem. Let $X \subset \mathbb{R}^2$ be a non-algebraic pfaff curve and $K \subset \mathbb{R}$ a number field of degree $k$. Then, for $T \geq e$,

$$N_K(X, T) = O_{X,k} \left( (\log T)^{O_X(1)} \right).$$

The proof of 1.8 is carried out in Section 6. It is straightforward but unfortunately seems to require revisiting several results of [14], and some earlier papers, in order to make the requisite minor modifications. I have tried to make the ideas of the proof and the modifications required clear without too much repetition.

In this paper, $\#A$ denotes the cardinality of a set $A$, the inclusion $A \subset B$ need not be strict, and $\mathbb{N}$ denotes the set of non-negative integers. By $f(T) = O(g(T))$, for real functions $f, g$ defined on some specified domain, we mean that there is an absolute constant $C$ such that $|f(T)| \leq Cg(T)$ for all $T$ in the domain. If $f(T) = f_{a,b,c,...,\alpha,\beta,\gamma,...}(T)$ depends on some parameters then by $f(T) = O_{a,b,c,...}(g(T))$ we indicate that the implied constant $C = C_{a,b,c,...}$ is permitted to depend on those exhibited parameters $a, b, c, \ldots$ but not on $\alpha, \beta, \gamma, \ldots$. By $O_{a,b,c,...}(1)$ we denote a constant that depends on $a, b, c, \ldots$.

2. Definable sets and families

We give a definition of o-minimal structure over $\mathbb{R}$, set up some notation and recall some well-known properties of sets definable in o-minimal structures, and explicate a consequence of the Trivialization Theorem.

2.1. Definition. A pre-structure is a sequence $S = (S_n : n \geq 1)$ where each $S_n$ is a collection of subsets of $\mathbb{R}^n$. A pre-structure $S$ is called a structure (over the real field) if, for all $n, m \geq 1$, the following conditions are satisfied:

1. $S_n$ is a boolean algebra (under the usual set-theoretic operations)
2. $S_n$ contains every semi-algebraic subset of $\mathbb{R}^n$
3. if $A \in S_n$ and $B \in S_m$ then $A \times B \in S_{n+m}$
4. if $m \geq n$ and $A \in S_m$ then $\pi[A] \in S_n$, where $\pi : \mathbb{R}^m \to \mathbb{R}^n$ is projection onto the first $n$ coordinates.

If $S$ is a structure and $X \subset \mathbb{R}^n$, we say $X$ is definable in $S$ if $X \in S_n$. If $S$ is a structure and, in addition, (5) the boundary of every set in $S_1$ is finite then $S$ is called an o-minimal structure (over the real field).

Let $X \subset \mathbb{R}^n$ be definable. Then $X$ has finitely many connected components and each of them is definable ([5, 4.3]).

For each pair $\kappa, p \in \mathbb{N}$ we define the $p$-regular points of dimension $\kappa$ of $X$, denoted $\text{reg}_p^\kappa(X)$, to be the set of $x \in X$ such that there is an open neighbourhood $U$ of $x$ with $U \cap X$ a $C^p$ (embedded) submanifold of $\mathbb{R}^n$ of dimension $\kappa$. Each $\text{reg}_p^\kappa(X)$ is definable ([5, B.9]). A regular point of dimension $\kappa$ will mean a 1-regular point of dimension $\kappa$. 3
The dimension of $X$ is the maximum $\kappa \in \mathbb{N}$ such that $X$ has a regular point of dimension $\kappa$. Therefore, if $X$ has dimension $\kappa$, then $X - \text{reg}^0(X)$ has dimension $\leq \kappa - 1$.

We need to consider families of definable sets. In considering subsets of $\mathbb{R}^n \times \mathbb{R}^m$, let $\pi_1$ denote projection on the first factor, and $\pi_2$ projection on the second factor. For a set $Z \subset \mathbb{R}^n \times \mathbb{R}^m$ put $Y = Y_Z = \pi_2(Z)$ and, for $y \in Y$, put $Z_y = \{ z \in \mathbb{R}^n \times \mathbb{R}^m : \pi_2(z) = y \}$ and $X_{Z,y} = \{ \pi_1(z) : z \in Z_y \}$. Thus, for any $y \in Y$, the restriction of $\pi_1$ to $Z_y$ identifies $Z_y$ with $X_{Z,y}$.

2.2. Definition. A definable family of sets is a definable set $Z \subset \mathbb{R}^n \times \mathbb{R}^m$ considered as the family of fibres $X_{Z,y} \subset \mathbb{R}^n$ for $y \in Y$.

2.3. Proposition. ([5, B.10]) Let $Z$ be a definable family, and $\kappa, p \in \mathbb{N}$. Then

$$\{ z = (x, y) \in Z : x \in \text{reg}^p(X_{Z,y}) \}$$

is definable. □

2.4. Proposition. Let $Z \subset \mathbb{R}^n \times \mathbb{R}^m$ be a definable family. There is a positive integer $J$ and definable families $Z_j \subset \mathbb{R}^n \times \mathbb{R}^m$, $j = 1, \ldots, J$, with the following property.

For each $j$, every fibre of $Z_j$ is connected. The set $Y_j = Y_{Z_j}$ is the subset of $Y_Z$ such that the fibre $X_{Z,y}$ has at least $j$ connected components, and for each $y \in Y_Z$, the fibres $(Z_j)_y$ are disjoint and

$$Z_y = \bigcup_{j=1}^J (Z_j)_y.$$

Proof. We apply the Trivialization Theorem [3, Ch. 9, 1.2] to the restriction of $\pi_2$ to $Z$ to obtain: a finite partition $A_1 \cup \ldots A_t$ of $Y_Z$ (i.e. the $A_i$ are disjoint and their union is $Y_Z$) into definable sets $A_i$, and definable families $F_i$, and, for each $i$, a definable map $\theta_i : A_i \times F_i \to \mathbb{R}^n \times \mathbb{R}^m$ that, for each $y \in A_i$, the map $\theta_i(y, \eta)$ gives a definable homeomorphism of $F_i$ with the fibre $Z_y$, and hence with $X_{Z,y}$.

Now each $F_i$ is a union of finitely many connected components $F_{ij}$, each of which is definable. By restricting to these components, and taking the family of images of $\theta_i$ restricted to $F_{ij}$ on $A_i$ we obtain a finite collection $Z'_i$ of families whose fibres are all connected, and, at each $y \in Y$, the union of the fibres is $Z_y$. As $\theta_i(y, \eta)$ restricted to $y$ is a homeomorphism, the images of the $F_{ij}$ are the connected components of $Z_y$, and the projections of these under $\pi_1$ are the connected components of $X_{Z,y}$.

We now need to re-organize these families to achieve the properties required. Let $J$ be the maximum number of connected components in a fibre $Z_y$ (the finiteness of which follows from the trivialization). For $j = 1, \ldots, J$ let $Y_j$ be the union of the sets $A_i$ such that $F_i$ has at least $i$ connected components. Over $Y_1$, choosing arbitrarily one of the components of $F_i$ for each $A_i$, we form the family $Z_1$. Over $Y_2$, choosing from each of the constituent $A_i$ a component different to the one chosen for $Y_1$, we construct $Z_2$. Continue to define $Y_3, Y_4, \ldots$ until all components of all $F_i$ are exhausted. □

3. A further strengthening of Theorem 1.4

Theorem 1.4 is strengthened in [15] in two ways. First, it is made uniform for sets in definable families. Indeed this is an artifact of the proof. Such uniformity will also be obtained for 1.6 and 1.7. Second, there is some refinement concerning the subset $X_\epsilon$ of $X^{\text{alg}}$ that needs to be excluded in order to get a $O_{X_\epsilon}(T^*)$ estimate for $N(X - X_\epsilon, T)$. The following is the final version of 1.4 established in [15].

3.1. Theorem. ([15, 1.10]) Let $Z$ be a definable family, and $\epsilon > 0$. There is a definable family $W = W(Z, \epsilon)$ with the following property. Let $y \in Y_Z$, put $X = X_{Z,y}$ and $X_\epsilon = X_{W,y}$. Then $X_\epsilon \subset X^{\text{alg}}$ and

$$N(X - X_\epsilon, T) = O_{Z,\epsilon}(T^*).$$

As remarked in [15, p597], Theorem 3.1 makes a non-empty assertion in certain situations where Theorem 1.4 does not. This and the further refinement that we will obtain are best illustrated by some simple examples. Each of the following sets are definable in $\mathbb{R}_{\text{an}}$ (or alternatively in $\mathbb{R}_{\text{exp}}$).
Let $X_1 = \{(x, y, z) \in \mathbb{R}^3 : 2 < x, y < 3, z = x^y\}$. Then $X_1^{\text{alg}}$ consists of a union of segments of algebraic curves: for each $q \in \mathbb{Q}, 2 < q < 3$, the curve $y = q, z = x^q$. Thus $(X_1 - X_1^{\text{alg}})(\mathbb{Q})$ is empty, and the conclusion of 1.4 is trivial. However, $X_1^{\text{alg}}$ is not definable in $\mathbb{R}_{\text{an}}$, or indeed in any o-minimal structure, as it has infinitely many connected components – a definable set has only finitely many connected components. So 3.1 makes the non-trivial (though not difficult) assertion that a $O_\epsilon(T^\ell)$ estimate for $N(X_1 - X_{1, \epsilon}, T)$ can be obtained by removing a definable subset $X_{1, \epsilon}$ of $X_1^{\text{alg}}$, depending on $\epsilon$, which will consist of finitely many curve segments.

Let $X_2 = \{(x, y, z) \in \mathbb{R}^3 : 2 < x, y < 3, z = 2x^y\}$. Here every point lies on some line segment $z = 2^c, x + y = c$ contained in $X_2$. So $X_2^{\text{alg}} = X_2$. Now both 1.4 and 3.1 make trivial assertions, because $X_2^{\text{alg}}$ is definable. However, to achieve a $O_\epsilon(T^\ell)$ estimate for $N(X_2 - X_2, T)$ one need remove only those $\log T = O_\epsilon(T^\ell)$ segments with $c \in \mathbb{Z}, 0 \leq c \leq \log T/\log 2$. Our strengthening of 3.1 will make a non-trivial (though not difficult) assertion in this example.

Let $X_3 = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < e^1\}$. Here every $(x, y) \in X_3$ is contained in some small open disk also contained in $X_3$. The disk is semi-algebraic, also it contains $T^4$ rational points up to height $T$. Therefore, $X_3^{\text{alg}} = X_3$ and definable. The assertions of 1.4 and 3.1 are trivial, but now we cannot hope to recover a $O_\epsilon(T^\ell)$ estimate for $N(X_3 - X_3, T)$ without removing essentially all of $X_3$.

The set $X_3$ is not semi-algebraic (if it were then its bounding curves would also be semi-algebraic), but it is a “definable piece of a semi-algebraic set” — which is made precise below. Examination of the proof of 3.1 in [15] shows that all the sets excluded in the course of the proof have this form, and come from a finite number of definable families.

3.2. Definition.

1. A basic definable semi-algebraic block or basic block of dimension $\kappa$ in $\mathbb{R}^n$ is a connected definable set $U \subseteq \mathbb{R}^n$ of dimension $\kappa$ contained in some semi-algebraic set $A$ of dimension $\kappa$ such that every point $x$ of $U$ is regular of dimension $\kappa$ in $U$ and in $A$ and in $U \cap A$. Dimension zero is allowed: a point is a basic block.

2. A definable semi-algebraic block, or block, is the image of a basic block $U$ under a semi-algebraic map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ defined and continuous on a semi-algebraic set containing $U$.

3. A basic definable semi-algebraic block family, or basic block family, is a definable family $Z$ whose fibres are all basic blocks.

4. A definable semi-algebraic block family, or block family, is the family $W \subseteq \mathbb{R}^\ell \times \mathbb{R}^m$ of images of a basic block family $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ under a semi-algebraic map $\phi : \mathbb{R}^n \to \mathbb{R}^\ell$ defined and continuous on a semi-algebraic set containing all the fibres.

3.3. Remark. We have framed the above definitions to meet our requirements. Presumably a block is always a union of finitely many basic blocks (even if the semi-algebraic map is not continuous), and then we could dispense with this distinction and restriction to continuous semi-algebraic maps. This is presumably entirely a question about semi-algebraic sets and maps. In defining a block family, one could allow the semi-algebraic map to vary with the parameters.

We observe the following facts.

3.4. Proposition. Let $X \subseteq \mathbb{R}^n$ be definable.

1. Let $U \subseteq X$ be a basic block of dimension $\kappa > 0$. Then $U \subseteq X^{\text{alg}}$.

2. Let $U \subseteq X$ be a block of dimension $\kappa > 0$. Then $U \subseteq X^{\text{alg}}$.

3. Let $W$ be a basic block family such that each fibre of $W$ is a subset of $X$, and all the fibres have dimension $\kappa > 0$. Then the union of all the fibres is definable and a subset of $X^{\text{alg}}$.

4. Let $W$ be a block family such that every fibre of $W$ is a subset of $X$. Then the union of fibres of positive dimension is definable and is a subset of $X^{\text{alg}}$.

Proof. 1. Let $A$ be a semi-algebraic set with the properties set out in 3.2.1. Each point $x \in U$ has a neighbourhood $B_x$ such that $B_x \cap U$ is regular of dimension $\kappa$ in $U$, $B_x \cap A$ is regular of dimension $\kappa$ in $X$, $B_x \cap U \cap A$ is regular of dimension $\kappa$ in $U \cap A$. These three intersection sets therefore coincide. Now $B_x \cap A$ is a connected semi-algebraic set of positive dimension contained in $X$. So $x \in B_x \cap A \subseteq X^{\text{alg}}$. 


2. The block $U$ is the image of some basic block $V$ under a continuous semialgebraic map $\phi$. Let $x, y \in U$ be distinct points, and let $P, Q$ be preimages of them in $U$. As $U$ is connected, it is path-connected [5, 4.21], and their is a definable compact map of $g : [0, 1] \to U$ with $G(0) = P, g(1) = Q$. Then $g([0, 1])$ is compact and covered by finitely many balls $B_x$ as in part 1. On each such $B_x$ we may replace the corresponding portion of $g([0, 1])$ by a semialgebraic path. Thus there is a semialgebraic path joining $P, Q$. Then the image of this path is a connected semialgebraic set of positive dimension containing $x$.

3. Suppose $W \subset \mathbb{R}^n \times \mathbb{R}^m$. The union of the fibres is the projection on $\mathbb{R}^n$ of $W$, and so definable. Call this set $A$. If $x \in X$ is a point in $A$, then for some $y$, the fibre $W_y$ contains $x$, and so contains a semialgebraic set $B_x \cap W_y$ that is connected of positive dimension contained in $X$. So $x \in X^{alg}$, and so $A \subset X^{alg}$.

4. The union of fibres of positive dimension is definable (it is the projection of $W$ minus the set of fibres that are single points). Each fibre of positive dimension is contained in $X^{alg}$ by part 2. \( \square \)

We can now state our further refinement of 3.1.

3.5. Theorem. Let $Z \subset \mathbb{R}^n \times \mathbb{R}^m$ be a definable family, and $\epsilon > 0$. There exist $J = J(Z, \epsilon) \in \mathbb{N}$ and a collection of basic block families $W_j \subset \mathbb{R}^n \times (\mathbb{R}^m \times \mathbb{R}^{m_j})$, $j = 1, \ldots, J$, such that each point in each fibre of $W_j$ is regular of dimension $w_j$, and with the following properties.

1. For each $J$ and $(y, \eta) \in \mathbb{R}^m \times \mathbb{R}^{m_j}$, $X_{W_j}(y, \eta) \subset X_{Z, \epsilon}$.

2. If $X = X_{Z, \epsilon}$ is a fibre $Z$ and $T \geq 1$ then $X(\mathbb{Q}, T)$ is contained in $O_{Z, \epsilon}(T^*)$ basic blocks, each of which is a fibre of one of the $W_j$ at some $(y, \eta) \in \mathbb{R}^m \times \mathbb{R}^{m_j}$.

3. Let $W = W(Z, \epsilon)$ be the family whose fibre at $y$ is the union of all fibres of the $W_j$ over $y$ over all $j$ with $w_j > 0$. Then $W$ is definable. If $X_{\epsilon}$ is the fibre of $W$ at $y$ then $X_{\epsilon} \subset X^{alg}$ and $N(X - X_{\epsilon}, T) = O_{Z, \epsilon}(T^*)$.

Proof. Examining the proof of [15, Theorem 1.10] shows that, given $Z$ and $\epsilon$, one considers various families obtained by taking regular points of highest dimension in the fibres previously defined families and intersecting with suitable families of semialgebraic sets of the same dimension. Only finitely many such families arise. Among them are a finite number of basic block families, $W_j$, whose fibres are all contained in the corresponding fibres of $Z$, and the proof shows that, given $T$ and $y \in Y_Z$, $X_{Z, \epsilon}(\mathbb{Q}, T)$ is contained in $O_{Z, \epsilon}(T^*)$ fibres of the $W_j$. This gives the first two statements of 3.5. If $w_j > 0$ then the union of fibres over $y$ is a definable subset of $Z_{y}^{alg}$, by 3.4.3. Let $W$ be the family whose fibre is the union of these definable sets over $j$ with $w_j > 0$. Since $X_{Z, \epsilon}(\mathbb{Q}, T)$ is contained in $O_{Z, \epsilon}(T^*)$ fibres of the $W_j$, the same estimate applies to the number of fibres coming from $W_j$ with $w_j = 0$. This gives the final statement. \( \square \)

4. Proof of Theorem 1.3

We state a refined version of 1.7 for a definable family $Z$ and then prove it by applying 3.5 to a suitable family obtained from $Z$.

Note that the family $Z_k$ constructed in the proof has the property that every set $X_{Z_k, \epsilon}$ in $Z_k$ is fibred by $(k - 1)$-planes, so that $X_{Z_{k, \epsilon}} = X_{Z_k, \epsilon}$ for all $y$ and Theorems 1.4 and 3.1 make a trivial assertion.

4.1. Theorem. Let $Z \subset \mathbb{R}^n \times \mathbb{R}^m$ be a definable family, $k$ a positive integer, and $\epsilon > 0$. There exists $J = J(Z, k, \epsilon) \in \mathbb{N}$ and a collection $W_j \subset \mathbb{R}^n \times (\mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^{m_j})$, $j = 1, \ldots, J$, of block families with the following properties.

1. For each $j$ and $(y, \lambda, \eta) \in \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^{m_j}$, $X_{W_j}(y, \lambda, \eta) \subset X_{Z, \epsilon}$.

2. If $X = X_{Z, \epsilon}$ is a fibre of $Z$ and $\lambda \in \mathbb{R}^k$ then $X(\lambda, T)$ is contained in $O_{Z, \epsilon}(T^*)$ blocks contained in $X$, each a fibre of one of the families $W_j$ at some $(y, \lambda, \eta) \in \mathbb{R}^{m_j}$. Hence

$$N_{\lambda}(X - X^{alg}, T) = O_{Z, \epsilon}(T^*).$$

3. Further, let $W \subset \mathbb{R}^n \times \mathbb{R}^m$ be the family whose fibre at $y$ is the union over all $j$ of all fibres of $W_j$ at $(y, \lambda, \eta)$ of positive dimension. Then $W$ is definable, and, for any $y \in \mathbb{R}^m$, if $X = X_{Z, \epsilon}$ and $X_{\epsilon} = X_{W_{j, \epsilon}}$ then $X_{\epsilon} \subset X^{alg}$ and, for any $\lambda \in \mathbb{R}^k$,

$$N_{\lambda}(X - X_{\epsilon}, T) = O_{Z, \epsilon}(T^*).$$
Proof. Define $\pi : (\mathbb{R}^k)^n \times \mathbb{R}^k \to \mathbb{R}^n$ by

$$
\pi(x_1, \ldots, x_n, \lambda) = \xi, \text{ where } x_i = (x_{i1}, \ldots, x_{ik}) \in \mathbb{R}^k, \text{ and } \xi_i = \sum_{j=1}^{k} x_{ij} \lambda_j, i = 1, \ldots, n.
$$

Put

$$
Z_k = \{(x_1, \ldots, x_n, \lambda, y) \in (\mathbb{R}^k)^n \times \mathbb{R}^k \times \mathbb{R}^m : (\pi(x_1, \ldots, x_n, \lambda), y) \in Z\}.
$$

We view $Z_k$, which is evidently definable, as a family of sets in $(\mathbb{R}^k)^n$ parameterized by $\mathbb{R}^k \times Y_Z$, and apply 3.5. Thus, given $\epsilon > 0$, we have $J = J(Z_k, \epsilon) \in \mathbb{N}$ and a collection of basic block families $V_j \subset (\mathbb{R}^k)^n \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^\mu$, with the properties asserted in 3.5.

If now $y \in Y$ and $\lambda \in \mathbb{R}^k$, the set of points $X_{Z_k,y}(\lambda, T)$ is precisely the image of the set of points $X_{Z_k,y}(\mathbb{Q}, T)$ under the projection $\pi$.

The map $\pi$ is evidently semialgebraic and continuous on all its domain. The image of a basic block family is therefore a block family. We put

$$
W_j = \{(\pi(x_1, \ldots, x_n, \lambda), y, \eta) : (x_1, \ldots, x_n, \lambda, y, \eta) \in V_j\}.
$$

The asserted properties of $W_j$ follow immediately from the properties of $V_j$. □

5. Algebraic points and proof of Theorem 1.2

The construction required to prove 1.6 is an elaboration of the one used to prove 1.7. We will parameterize algebraic coordinates $\alpha \in \mathbb{R}$ with $[\mathbb{Q}(\alpha) : \mathbb{Q}] \leq k$ by means of their minimal polynomials. 5.1. Definition. Let $k$ be a positive integer. For a real number $\alpha$ we define $H^\text{poly}_k(\alpha)$ as follows. If $[\mathbb{Q}(\alpha) : \mathbb{Q}] > k$ we set $H^\text{poly}_k(\alpha) = \infty$. Otherwise we set

$$
H^\text{poly}_k(\alpha) = \min\{H(\xi) : \xi = (\xi_1, \ldots, \xi_k) \in \mathbb{Q}^k \setminus \{(0, \ldots, 0), \sum_{j=1}^{k} \xi_j \alpha^j = 0\},
$$

(Recall that $H(\xi)$ is the maximum height of the coordinates.) Thus $H^\text{poly}_k(\alpha) < \infty$ when $[\mathbb{Q}(\alpha) : \mathbb{Q}] \leq k$. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we set $H^\text{poly}_k(x) = \max_i H^\text{poly}_k(x_i)$.

Now suppose $\alpha \in \overline{\mathbb{Q}}$ of degree $k$ with minimal polynomial (over $\mathbb{Z}$)

$$
f(t) = a_k(t - \alpha_1) \ldots (t - \alpha_k) \in \mathbb{Z}[t].
$$

Then ($[1, 1.6.5 \text{ and } 1.6.6]$)

$$
\log |a_k| + \sum_{i=1}^{k} \max(0, \log |a_i|) = kh(\alpha)
$$

($= \log M(f)$, the Mahler measure of $f(t)$). Hence

$$
|a_k| \prod_{i=1}^{k} \max(1, |a_i|) = H(\alpha)^k.
$$

Thus we have $f(\alpha) = 0$ where $f$ has integer coefficients of maximum absolute value at most

$$
\max_i \left(\begin{array}{c} k \\ i \end{array}\right) \prod_{j=1}^{i} |\alpha_j| |a_k| \leq 2^k H(\alpha)^k
$$

the product being the largest absolute value of the product of $i$ distinct roots of $f$. Therefore

$$
H^\text{poly}_k(\alpha) \leq 2^k H(\alpha)^k,
$$

and it suffices to prove 1.6, which we establish in a strengthened form for families below, using $H^\text{poly}_k$ rather than $H$. 7
5.2. Lemma. Let \( n \) and \( k \) be positive integers. Define

\[
A_{n,k} = \{(\xi, x) = (\xi_1, \ldots, \xi_n, x) \in (\mathbb{R}^k - \{0\})^n \times \mathbb{R}^n : \xi_i = (\xi_{i1}, \ldots, \xi_{ik}), \sum_{j=1}^k \xi_{ij}x_j^{j-1} = 0, i = 1, \ldots, n\}.
\]

\[
A_{n,k}^\pi = \pi_1(A_{n,k}) \subset (\mathbb{R}^k - \{0\})^n,
\]

where \( \pi_1 : (\mathbb{R}^k)^n \times \mathbb{R}^n \rightarrow (\mathbb{R}^k)^n \) is projection on the first factor.

There are a finite number of semi-algebraic maps \( \phi_i : (\mathbb{R}^k)^n \rightarrow (\mathbb{R}^k)^n \times \mathbb{R}^n \), sections of \( \pi_1 \), defined and continuous on semi-algebraic sets \( U_i \subset (\mathbb{R}^k)^n \), with the following property.

If \((x_1, \ldots, x_n) \in \mathbb{R}^n \) with \( H_k^{\text{poly}}(x) \leq T \) then there is an index \( i \) such that \( U_i \cap A_{n,k}^\pi \) contains a preimage \( \xi = (\xi_1, \ldots, \xi_n) \) of \( x \) under the composition \( \pi \phi_i \) such that \( \xi \in (\mathbb{Q}^k - \{0\})^n \) with \( H(\xi) \leq T \).

Proof. Given \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) with \( H_k^{\text{poly}}(x) \leq T \) we can, by definition, find \( \xi \in (\mathbb{Q}^k - \{0\})^n \) with \( H(\xi) \leq T \) and \( \sum \xi_{ij}x_j^i = 0 \) for each \( i \). Therefore \( \xi \) is in \( \pi_1 \) of the preimage of \( x \) under \( \pi_2 \) in \( (\mathbb{R}^d - \{0\})^n \times \mathbb{R}^n \).

Now \( \pi_1 \) has finite fibres, as a non-zero polynomial of degree \( \leq k \) has at most \( k \) roots. We can therefore construct semi-algebraic sections. Say \( \psi_1 \) is the largest root, where a root exists, \( \psi_2 \) is the second largest root, where there are 2 or more roots, etc up to \( \psi_k \). These maps are evidently definable in the structure of semi-algebraic sets, and therefore they are defined on some semi-algebraic sets \( V_i \). Then each \( V_i \) may be decomposed into a finite number of semi-algebraic sets on which the restriction of \( \psi_i \) is continuous. We take the collection of maps so constructed as our \( \phi_j \). \( \square \)

5.3. Theorem. Let \( Z \subset \mathbb{R}^n \times \mathbb{R}^m \) be a definable family, \( k \) a positive integer, and \( \epsilon > 0 \). There exists \( J = J(Z, k, \epsilon) \in \mathbb{N} \) and block families \( W_j \subset \mathbb{R}^n \times (\mathbb{R}^m \times \mathbb{R}^{\mu_j}), j = 1, \ldots, J, \) with the following properties.

1. For each \( j \) and \( (y, \eta) \in \mathbb{R}^m \times \mathbb{R}^{\mu_j}, X_{W_j,(y,\eta)} \subset X_{Z,y}. \)
2. If \( X = X_{Z,y} \) is a fibre of \( Z \) then \( X(k,T) \) is contained in \( O_{Z,k,\epsilon}(T^\epsilon) \) blocks contained in \( X \), each a fibre of one of the families \( W_j \) at some \( (y,\eta), \eta \in \mathbb{R}^{\mu_j} \). Hence

\[
N_k(X - X^{\text{alg}}, T) = O_{Z,k,\epsilon}(T^\epsilon).
\]

3. Further, let \( W \subset \mathbb{R}^n \times \mathbb{R}^m \) be the family whose fibre at \( y \) is the union over all \( j \) of all the fibres of \( W_j \) at \( (y,\eta) \) of positive dimension. Then \( W \) is definable, and, for any \( y \in Y_Z \), if \( X = X_{Z,y} \) and \( X_{\epsilon} = X_{W,y} \), then \( X_{\epsilon} \subset X^{\text{alg}} \) and

\[
N(X - X_{\epsilon}, T) = O_{Z,k,\epsilon}(T^\epsilon).
\]

Proof. Define

\[
Z_{n,k} = \{(\xi, x, y) \in (\mathbb{R}^k - \{0\})^n \times \mathbb{R}^n \times \mathbb{R}^m : (\xi, x) \in A_{n,k}, (x, y) \in Z\},
\]

a definable family in \((\mathbb{R}^k)^n \times \mathbb{R}^n) \times \mathbb{R}^m\). Let \( \pi \) be the projection of \((\mathbb{R}^k)^n \times \mathbb{R}^n \) onto \((\mathbb{R}^k)^n\), and

\[
Z_{n,k}^\pi = \{(\xi, y) : \xi \in \pi(Z_{n,k,y})\},
\]

a definable family in \((\mathbb{R}^k)^n \times \mathbb{R}^m\).

We apply Theorem 3.5 to \( Z_{n,k}^\pi \) with \( \epsilon \), to obtain a finite collection of basic block families having the requisite properties.

Applying the semi-algebraic maps \( \phi_i \) of Lemma 5.2 we get a finite number of block families whose fibres are contained in the fibres of \( Z \), and which satisfy the required properties by 5.2. \( \square \)
5.4. Remarks.
1. In defining $H^\text{poly}_k$ in 5.1 we did not stipulate that the polynomials involved be irreducible, though this will be true when the degree of $\alpha$ is equal to $k$. It may be possible that an algebraic $\alpha$ of degree $\ell < k$ has $H^\text{poly}_k(\alpha) < H^\text{poly}_\ell(\alpha)$.

2. Let us define $H^\text{module}_k(\alpha)$, for $\alpha \in \mathbb{R}$ of degree $\leq k$ as the minimum height of a $k \times k$ matrix $(q_{ij})$ of rational numbers (considered as a $k^2$-tuple) such that there exist complex numbers $z_1, \ldots, z_k$, not all zero, such that $\alpha z_i = \sum q_{ij} z_j$. This height, constructed directly from the definition of an algebraic number as one which preserves a finite (positive) dimensional $\mathbb{Q}$-subspace of $\mathbb{C}$, seems quite natural. If $\alpha \neq 0$ then multiplication by $\alpha$ is an invertible $\mathbb{Q}$-linear map on the $\mathbb{Q}$-subspace of $\mathbb{C}$ generated by $z_i$, and so $1/\alpha$ also preserves this space, so that $H^\text{module}_k(1/\alpha) = H^\text{module}_k(\alpha)$.

Of course Theorem 5.3 restricts to algebraic numbers of degree $\leq k$, and one can ask whether it is possible to get some comparable result about all algebraic points of height $\leq T$ on a definable set $X$.

5.5. Question. Let $X : y = f(x), x \in [0,1]$ be definable and non-algebraic – or even more specifically real analytic and non-algebraic. Let $X(\overline{\mathbb{Q}}, T)$ denote the set of algebraic points of $X$ of (coordinate-wise) height $\leq T$, and $N_{\overline{\mathbb{Q}}}(X, T) = \# X(\overline{\mathbb{Q}}, T)$. Is it always true that $N_{\overline{\mathbb{Q}}}(X, T) = O_{X, \varepsilon}(T^\varepsilon)$?

6. Algebraic points on a pfaff curve

6.1. Definition. ([7, 2.1]) Let $U \subset \mathbb{R}^n$ be an open domain. A pfaffian chain of order $r \geq 0$ and degree $\alpha \geq 1$ in $U$ is a sequence of real analytic functions $f_1, \ldots, f_r$ in $U$ satisfying differential equations

$$df_j = \sum_{i=1}^n g_{ij}(x, f_1(x), f_2(x), \ldots, f_r(x))dx_i$$

for $j = 1, \ldots, r$, where $x = (x_1, \ldots, x_n)$ and $g_{ij} \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_j]$ of degree $\leq \alpha$. A function $f$ on $U$ is called a pfaffian function of order $r$ and degree $(\alpha, \beta)$ if $f(x) = P(x, f_1(x), \ldots, f_r(x))$, where $P$ is a polynomial of degree at most $\beta \geq 1$.

The order and degree of a pfaff curve $X$ will be taken to be the order and degree of the pfaffian function $f$ whose graph is $X$.

In this section we prove Theorem 1.8 by establishing the following somewhat more precise version.

6.2. Theorem. Let $X \subset \mathbb{R}^2$ be a non-algebraic pfaff curve of order $r$ and degree $(\alpha, \beta)$. Let $K \subset \mathbb{R}$ be a number field of degree $k$. Then

$$N_K(X, T) = O_{r, \alpha, \beta, k}((\log T)^{5(r+2)}).$$

It is convenient to work with yet another height on algebraic points, one that is often used in transcendence theory.

6.3. Definition. For an algebraic number $\alpha$, we let $\text{den}(\alpha)$ denote the denominator of $\alpha$, namely the smallest positive integer $\delta$ such that $\delta \alpha$ is an algebraic integer. If $\alpha_1, \ldots, \alpha_k$ are the conjugates of $\alpha$ we set $H^\text{size}(\alpha) = \max(\text{den}(\alpha), |\alpha_1|, \ldots, |\alpha_k|)$.

Now suppose $\alpha \in \overline{\mathbb{Q}}$ of degree $k$ with minimal polynomial (over $\mathbb{Z}$) $f(t) = a_d(t - \alpha_1) \ldots (t - \alpha_k) \in \mathbb{Z}[t]$ then

$$H^\text{size}(\alpha) \leq |a_k| \prod \max(1, |\alpha_i|) = H(\alpha)^k.$$ 

So it suffices to prove 6.2 using $H^\text{size}$ rather than $H$, and in this section we will work throughout with $H^\text{size}$. Accordingly we define

$$X^\text{size}(K, T) = \{ x \in X(K) : H^\text{size}(x) \leq T \}, \quad N^\text{size}_K(X, T) = \# X^\text{size}(K, T).$$

6.4. Definition. Let $I$ be an interval (which may be closed, open or half-open; bounded or unbounded), $k \in \mathbb{N} = \{0, 1, 2, \ldots\}$, $L > 0$ and $f : I \rightarrow \mathbb{R}$ a function with $k$ continuous derivatives on $I$. Set $A_{L,0}(f) = 1$ and, for positive $k$,

$$A_{L,k}(f) = \max_{1 \leq i \leq k} \left( 1, \sup_{x \in I} \left( \frac{|f^{(i)}(x)|L^{i-1}}{i!} \right)^{1/i} \right).$$

(so possibly $A_{L,k}(f) = \infty$ if a derivative of order $i$, $1 \leq i \leq k$, is unbounded).
6.5. **Lemma.** Let $K \subseteq \mathbb{R}$ be a number field with $|K : \mathbb{Q}| = k$. Let $d \geq 1, T \geq 1, L \geq 1/T^4k$. Let $I$ be an interval of length $\leq L$. Suppose that the function $f$ possess $D - 1$ continuous derivatives on $I$, with $|f'| \leq 1$, and let $X$ be the graph of $f$ on $I$. Put $D = (d + 1)(d + 2)/2$. Then $X^{\text{size}}(K, T)$ is contained in the union of at most
\[ 6^k (LT^{4k})^{4/(3(d + 3))} A_{L,D-1} \]
plane algebraic curves of degree $\leq d$ (possibly reducible).

**Proof.** This is a variant of [14, 2.2]. Suppose that $x_1 < x_2 < \ldots < x_D$ and that $(x_1, y_1), \ldots, (x_D, y_D)$ are points of $X^{\text{size}}(K, T)$ that do not lie on any plane algebraic curve of degree $\leq d$. Then
\[ 0 \neq \Delta = \det (\phi_j(x_i)) \]
where $\phi_j, j = 1, \ldots, D$ ranges over all the functions $x^a f(x)^b$ with $0 \leq a, b \leq a + b \leq d$.

Let $s_i, t_i \leq T$ be the denominators of $x_i, y_i$, and $\sigma_1, \ldots, \sigma_k$ the embeddings of $K$ into $\mathbb{C}$. Since $\prod_s \prod t^d \Delta$ is an algebraic integer, and non-zero, the product of its conjugates is a non-zero rational integer. Hence
\[ 1 \leq \left| \prod_{i=1}^{k} \left( \prod_j s_j \prod_j t^d \Delta \right)^{\sigma_i} \right| . \]

We pursue an upper bound for this quantity. For any $\sigma$, using $H(x_i), H(y_i) \leq T$, a straightforward estimation gives
\[ |\Delta^\sigma| \leq D! T^{2dD}/3 . \]

Now $\Delta$ is an alternant formed by evaluating the $D$ functions $\phi_j$, which have $D - 1$ continuous derivatives, at the $D$ points $x_i$, so by the “Mean Value Theorem” for alternants (see e.g. [2, Prop. 2])
\[ \Delta = V(x_1, \ldots, x_D) \det \left( \frac{\phi_j^{(i-1)}(\xi_{ij})}{(i-1)!} \right) \]
where $V(x_1, \ldots, x_D)$ is the Vandermonde determinant and $\xi_{ij}$ are suitable intermediate points. Since the functions $\phi_j$ are the monomial functions $x^a f(x)^b, a + b \leq d$, we can use column operations to replace the points $x_i$ by translations $x_i - b$, so that $x_i \in [0, L]$, and $f$ by a translation $f(x) - c$ for suitable $c$ so that, since $|f'| \leq 1$, we have
\[ |f^{(i)}(x)| \leq i! L^{1-i} (A_{L,k}(f))^i \]
for $i = 0$ as well as $i = 1, \ldots, k$, and all $x \in I$. Then ([11, Lemma 1]) for $0 \leq a, b$ and $2, i + 1 \leq D$
\[ \left| \frac{(x^a f(x)^b)^{(i)}}{i!} \right| \leq L^{a+b-i} 2^{a(i + 1)} (A_{L,k}(f))^i \leq L^{a+b-i} D^{a+b} (A_{L,k}(f))^i . \]

We use this to estimate the entries in $\Delta$. Expand $\Delta$ into $D!$ products of entries. In each term the multiplicands run over $0 \leq a, b \leq a + b \leq d$ and over $0 \leq i \leq D - 1$. Since $\sum (a + b) = 2dD/3$ and $\sum i = D(D - 1)/2$ over these ranges we find, as in [13, 2.4] and the surrounding discussion, that
\[ |\Delta| \leq |V(x_1, \ldots, x_D)| D! L^{2dD/3 - D(D-1)/2} D^{2dD/3} (A_{L,D-1}(f))^{D(D-1)/2} . \]

We therefore have
\[ 1 \leq |x_D - x_1|^{D(D-1)/2} (D!)^k L^{2dD/3} L^{2dD/3 - D(D-1)/2} T^{8kdD/3} (A_{L,D-1}(f))^{D(D-1)/2} . \]

Now computation shows $(D!D^{2dD/3})^{2/(D(D-1))} \leq 5$ for all $d$, so that the above implies
\[ x_D - x_1 \geq 5^{-k} L \left( L^{2dD/3} T^{8kdD/3} (A_{L,D-1}(f))^{2/(D(D-1))} \right) \leq 5^{k} (LT^{4k})^{4/(3(d + 3))} A_{L,D-1} . \]

Since the interval $I$ has length $\leq L$, it is thus covered by at most
\[ 5^{k} (LT^{4k})^{4/(3(d + 3))} A_{L,D-1} + 1 \leq 6^{k} (LT^{4k})^{4/(3(d + 3))} A_{L,D-1} \]
(here we use $LT^{4k} \geq 1$) subintervals on each of which the points of $X^{\text{size}}(K, T)$ must lie on a single algebraic curve of degree $\leq d$ (possibly reducible). \[ \square \]
6.6. Proposition. Let \( k \in \mathbb{N} \), \( L > 0 \), \( A \geq 1 \) and let \( I \) be an interval of length \( \leq L \). Suppose \( g : I \to \mathbb{R} \) has \( k \) continuous derivatives on \( I \). Suppose that \( |g'| \leq 1 \) throughout \( I \) and that
\begin{enumerate}[(a)]  \item |\( g^{(i)}(x) \)| \( \leq i! A^i L^{-i} \), all \( 1 \leq i \leq k-1 \), \( x \in I \), and  \item |\( g^{(k)}(x) \)| \( \geq k! A^k L^{1-k} \), all \( x \in I \).\end{enumerate}
Then \( I \) has length \( \leq 2L/A \).

Proof. Let \( a, b \in I \). By Taylor’s formula, for a suitable intermediate point \( \xi \),
\[
g(b) - g(a) = \sum_{i=1}^{k-1} \frac{g^{(i)}(a)}{i!} (b-a)^i + \frac{g^{(k)}(\xi)}{k!} (b-a)^k.
\]
Therefore
\[
L \left( \frac{(b-a)A}{L} \right)^k \leq (b-a)^k A^k L^{1-k} \leq \sum_{i=1}^{k-1} (b-a)^i A^i L^{-i} + L \sum_{i=0}^{k-1} \left( \frac{(b-a)A}{L} \right)^i.
\]
Thus, if \( q = (b-a)A/L \), then \( q^k \leq \sum_{i=0}^{k-1} q^i \), whence \( q \leq 2 \), completing the proof. \( \square \)

6.7. Proposition. Let \( d \geq 1 \), \( k \geq 1 \), \( T \geq e \), \( L \geq 1/T^{4k} \) and \( I \subset \mathbb{R} \) an interval of length \( \leq L \). Let \( K \subset \mathbb{R} \) be a numberfield of degree \( k \). Let \( f : I \to \mathbb{R} \) have \( D \) continuous derivatives, with \( |f'| \leq 1 \) and \( f^{(j)} \) either non-vanishing in the interior of \( I \) or identically vanishing, for \( j = 1, \ldots, D \). Let \( X \) be the graph of \( f \). Then \( X^{size}(K, T) \) is contained in the union of at most
\[
9 \cdot 6^k D LT^{4k} 4/(3(d+3)) \log(eLT^{4k})
\]
real algebraic curves of degree \( \leq d \) (possibly reducible).

Proof. This follows closely the scheme of proof of [14], and we just sketch the key points of the argument.

The interval \( I \) may or may not include either of its endpoints. Let us write \( I = (a, b) \) allowing \( \langle \) to be either \( ( \) or \( [ \) and similarly for \( \rangle \), and for the other intervals that arise.

For a subinterval \( J \) of \( I \) write \( f|_J \) for the restriction of \( f \) to \( J \), and denote its graph by \( X|_J \). Let \( G(f, J) \) be the minimal number of algebraic curves of degree \( \leq d \) required to contain \( X|_J^{size}(K, T) \). We pursue an upper bound on \( G(f, I) \) by a recursion argument, using the fact that \( f \) and its derivatives up to order \( D \) are monotonic to subdivide the given interval into subintervals on which \( A_{L,K}(f) \) is controlled, the residual intervals being short, as afforded by 6.6, and few, due to the hypothesis on the derivatives of \( f \) up to order \( D \).

Let \( A \geq 2D \).

Subdivide \( I \) at suitable points (i.e. where \( f \) and its derivatives assume suitable values) to obtain a “good” subinterval \( I_0 = [s, t] \) of \( I \) (possibly empty!) on which \( A_{L,D-1}(f) \leq A \), and such that the residual “bad” intervals \( J^L_t = (a, s) \) and \( J^R_t = (t, b) \) each comprise at most \( D \) intervals of the form \( (a, s') \) or \( (t', b) \) (by the monotonicity) of length \( \leq 2L/A \) (by 6.6), so that \( J^L_t, J^R_t \) each have length \( \leq 2DL/A \).

Thus, setting \( \lambda = 2D/A, \sigma = 4/(3(d+3)), \tau = 16k/(3(d+3)), \) we have
\[
G(f, I) \leq G(f, I_0) + G(f, J^L_t) + G(f, J^R_t) \leq 6^k AT^\sigma L^\sigma + G(f, J^L_t) + G(f, J^R_t)
\]
where \( J^L_t, J^R_t \) have length \( \leq 2DL/A \).

We now repeat the subdivision using the same \( A \) but with \( \lambda L \) instead of \( L \). The monotonicity of each derivative up to order \( D - 1 \) ensures that the “bad” part of \( J^L_t \) is a single subinterval of \( J^L_t \) of the form \( (a, u) \), and likewise the “bad” subinterval of \( J^R_t \) is of the form \( (v, b) \).

In particular, we have only 2 “bad” intervals, and 2 further “good” intervals of length \( \leq \lambda L \) on which \( A_{L,D-1}(f) \leq A \). After \( n \) iterations of this process we have (provided \( \lambda^{n-1} LT^{4k} \geq 1 \) to retain the assumptions of 6.5)
\[
G(f, I) \leq 6^k AT^\sigma L^\sigma (1 + 2\lambda + \ldots + 2\lambda^{n-1}) + G(f, J^L_n) + G(f, J^R_n)
\]
\[
\leq 6^k AT^\sigma L^\sigma (2n - 1) + G(f, J^L_n) + G(f, J^R_n)
\]
(since \( \lambda \leq 1 \)) where \( J^L_n, J^R_n \) have (each) length \( \leq \lambda^n L \).
Now consider $\alpha, \beta \in K$ with $\alpha \neq \beta$ and $H^{\text{size}}(\alpha), H^{\text{size}}(\beta) \leq T$. Then

$$\prod_{i=1}^{k} \left( \frac{\text{den}(\alpha) \text{den}(\beta) (\alpha - \beta)}{\sigma_{i}} \right)^{\sigma_{i}},$$

is a non-zero rational integer. Since $|\alpha - \beta|^{\sigma_{i}} \leq 2T$ for each $\sigma_{i}$ and $\text{den}(\alpha) \text{den}(\beta) \leq t^2$ we see that

$$|\alpha - \beta| \geq \frac{1}{(2T)^{k-1}T^{2k}} \geq \frac{1}{T^{4k}}.$$

(since $T \geq e$).

Take $n$ such that $\lambda/(LT^{4k}) \leq \lambda/n < 1/(LT^{4k})$. Then an interval of length $\leq \lambda/n L$ contains at most one point $\alpha \in K$ with $H^{\text{size}}(\alpha) \leq T$, while the assumption $\lambda^{n-1}LT^{4k} \geq 1$ allows the subdivision process to be repeated $n$ times. We take $A = 2eD$, i.e. $\lambda = \frac{1}{e}$. We then have $n \leq \log(eLT^{4k})$ and

$$G(f, I) \leq 6^k 2eD(LT^{4k})^{4/(3(d+3))}(2\log(eLT^{4k}) - 1) + 2 \leq 9 \cdot 6^k D(LT^{4k})^{4/(3(d+3))} \log(eLT^{4k})$$

as required. $\square$

**6.8 Corollary.** With the hypotheses of 6.7, if also $L \leq 2T$ and $T \geq e$ then $X^{\text{size}}(K, T)$ is contained in at most

$$12 \cdot (4k + 3) \cdot 6^k DT^{4(4k+1)/(3(d+3))} \log T$$

real algebraic curves of degree $\leq d$ (possibly reducible).

The key point behind the uniform exponent in 6.2 is here in the Corollary. Eventually we will take $d = \lfloor \log T \rfloor$. Then $T^{1/(d+3)} \leq e$, and the effect of the number field $K$ is reduced to a constant factor $c(k)$ in the number of curves required.

**Proof of 6.2.** As already observed it suffices to prove the result with the height $H^{\text{size}}$. The proof follows exactly the proof of [14], save that we appeal to 6.8 rather than the corresponding result (Corollary 3.3) in [14].

Since $f$ is a pfaffian function, its derivatives are also pfaffian. Indeed, $f^{(i)}$ is pfaffian of order $r$ and degree $(\alpha, \beta + i(\alpha - 1))$ ([7, 3.3 or 14, 4.1]). Using the bounds from [7], as quoted in [14, 4.1], we can subdivide $I$ into at most $c_1(\alpha, \beta, r)D^{2r+3}$ subintervals on which either $f$ or its inverse (is defined and) satisfies the hypotheses of 6.6.

We intersect these intervals with the interval $[-T, T]$ of the relevant axis, in which all points $\alpha \in K$ with $H^{\text{size}}(\alpha) \leq T$ lie. In each such interval $J$, $X^{\text{size}}(K, T)$ is, by Corollary 6.8, contained in at most

$$12 \cdot (4k + 3) \cdot 6^k DT^{4(4k+1)/(3(d+3))} \log T$$

algebraic curves of degree $\leq d$, while, using again the bounds from [7, or 14, 4.1], the intersection of $X$ with such a curve comprises at most $c_2(r, \alpha, \beta)d^{r+1}$ points.

Combining, we have

$$N_K^{\text{size}}(X, T) \leq c_3(r, \alpha, \beta, k)d^{5r+7} DT^{4(4k+1)/(3(d+3))} \log T$$

and taking $d = \lfloor \log T \rfloor$ completes the proof. $\square$
References


School of Mathematics
University of Bristol
Bristol, BS8 1TW
UK