Counting rational points on a certain exponential-algebraic surface

Jonathan Pila

1. Introduction

This paper is devoted to giving an upper estimate for the number of non-trivial rational points (or algebraic points over a given real numberfield) up to a given height on the surface $X \subset \mathbb{R}^3$ defined by

$$X = \{(x, y, z) \in (0, \infty)^3 : \log x \log y = \log z\}.$$

The half-lines $L_x = \{(x, 1, 1) : x > 0\}$ and $L_y = \{(1, y, 1) : y > 0\}$ contained in $X$ evidently contain rational (or algebraic) points $(r, 1, 1), (1, s, 1) \in X$, where $r, s \in \mathbb{Q}_{>0}$ (or $r, s \in \mathbb{Q} \cap \mathbb{R}_{>0}$), and these algebraic points we call trivial. Schanuel’s conjecture implies (as we elaborate in §4) that there are no non-trivial algebraic points on $X$, and hence that there are no rational points on $X^0 = X - (L_x \cup L_y)$.

Our result is that this conjecturally empty set is fairly sparse.

For a set $Y \subset \mathbb{R}^n$ put $Y(\mathbb{Q}) = Y \cap \mathbb{Q}^n$ and define, for $T \geq 1$ (which we assume throughout),

$$Y(\mathbb{Q}, T) = \{x = (x_1, \ldots, x_n) \in Y \cap \mathbb{Q}^n : H(x_1), \ldots, H(x_n) \leq T\}$$

where $H(a/b) = \max(|a|, |b|)$ for a rational number $a/b$ in lowest terms. The cardinality of a set $A$ will be denoted $\# A$. Note that $\#(L_x \cup L_y)(\mathbb{Q}, T) \geq cT^2$, where $c$ is some positive constant. In the sequel, $c(\alpha, \beta, \ldots), C(\alpha, \beta, \ldots)$ denote positive constants that depend only on $\alpha, \beta, \ldots$, and that may differ at each occurrence.

1.1. Theorem. Let $\epsilon > 0$. Then

$$\#X^0(\mathbb{Q}, T) \leq c(\epsilon) \left(\log T\right)^{44 + \epsilon}.$$

This result may be viewed as a statement about the set of points $(x, y) \in (0, \infty)^2$ at which the three algebraically independent real-analytic functions $x, y, \exp(\log x \log y)$ are simultaneously rational, or alternatively about the points $(u, v) \in \mathbb{R}^2$ at which the functions $e^u, e^v, e^{uv}$ are simultaneously rational. The set of points at which algebraically independent meromorphic functions of several complex variables simultaneously assume values in a number field has been quite extensively studied in connection with transcendental number theory, especially functions generating rings closed under partial differentiation [8, 1]. Without such assumptions, results of Lang [9], systematizing methods going back to Schneider, have been improved and extended by Waldschmidt [17] and others (see e.g. [20, 19]), and are intimately connected to interpolation problems and Schwarz Lemmas in several variables, see e.g. papers of Roy [16]. See also [18]. Note that we do not assume any hypotheses on the points $(u, v)$, such as lying in a Cartesian product, nor is the ring of functions $\mathbb{C}[e^u, e^v, e^{uv}]$ closed under partial differentiation, while the function $\exp(\log x \log y)$ is not meromorphic in $\mathbb{C}^2$. Nevertheless, complex variable methods may well yield results along the lines of 1.1, although I am not aware of any explicit statements in the literature that imply such a result. We will employ real variable methods and draw on the theory of o-minimal structures.

To contextualise our result, we review the background results and conjectures. An o-minimal structure over $\mathbb{R}$ is, informally speaking, a sequence $\mathcal{S} = (\mathcal{S}_n), n = 1, 2, \ldots$ with each $\mathcal{S}_n$ a collection of subsets of $\mathbb{R}^n$ such that $\cup_n \mathcal{S}_n$ contains all semi-algebraic sets and is closed under certain operations (boolean operations, products and projections), but nevertheless has strong finiteness properties (the boundary of every set in $\mathcal{S}_1$ is finite). A formal definition is given in the Appendix (§7), or see [5]. If $\mathcal{S}$ is an o-minimal structure over $\mathbb{R}$, a set $Y \subset \mathbb{R}^n$ belonging to $\mathcal{S}_n$ is said to be definable in $\mathcal{S}$. A set $Y \subset \mathbb{R}^n$ will be called definable if it is definable in some o-minimal structure over $\mathbb{R}$.
The paradigm example of an o-minimal structure is the collection of semi-algebraic sets. Another example is provided by the collection $\mathbb{R}_{an}$ of \textit{globally subanalytic} sets (see [6]), and the crucial example for this paper is the collection $\mathbb{R}_{\exp}$ of sets definable using the exponential function (see §7). The o-minimality of $\mathbb{R}_{\exp}$ is due to Wilkie [21], whose result yields the elegant description of the sets definable in $\mathbb{R}_{\exp}$ given in 7.2. The set $X$ is definable in $\mathbb{R}_{\exp}$ (see 7.3).

Suppose then that $Y \subset \mathbb{R}^n$ is definable, and consider the counting function $\#Y(Q, T)$. If $Y$ contains semialgebraic sets of positive dimension (such as rational curves, as is the case for the set $X$), then one can certainly have

$$\#Y(Q, T) \geq c(Y)T^\delta$$

for some positive $\delta$. If on the other hand $Y$ contains no semialgebraic sets of positive dimension then, according to [15], one has

$$\#Y(Q, T) \leq c(Y, \epsilon)T^\epsilon$$

for every $\epsilon > 0$. Indeed if we define, for any $Y \subset \mathbb{R}^n$, the \textit{algebraic part} $Y_{\text{alg}}$ of $Y$ to be the union of all connected semialgebraic subsets of $Y$ of positive dimension, then an estimate as above holds for the rational points of the \textit{transcendental part} $Y_{\text{trans}} = Y - Y_{\text{alg}}$ of any definable set $Y$.

\begin{thm} ([15]) \ Let $Y$ be definable in an o-minimal structure over $\mathbb{R}$ and $\epsilon > 0$. Then

$$\#Y_{\text{trans}}(Q, T) \leq c(Y, \epsilon)T^\epsilon.$$  

\end{thm}

Examples show (see [10, 7.5 and 7.6], elaborating a remark from [3]) that this estimate cannot be much improved in general. For example one can construct sets definable in $\mathbb{R}_{an}$ such that no estimate of the form

$$\#Y_{\text{trans}}(Q, T) \leq c(Y)(\log T)^C(Y)$$

holds. However, Wilkie conjectured in [15] that such an estimate always holds for a set definable in $\mathbb{R}_{\exp}$.

\begin{conjecture} \ Suppose $Y$ is definable in $\mathbb{R}_{\exp}$. Then

$$\#Y_{\text{trans}}(Q, T) \leq c(Y)(\log T)^C(Y).$$

\end{conjecture}

Thus Theorem 1.1 affirms this conjecture for the particular set $X$. In fact $X_{\text{alg}}$ consists of $L_x$ and $L_y$ together with infinitely many other rational curves defined over $\mathbb{R}$ (see 4.1). However these other rational curves do not contain any algebraic points (see 4.3).

Consider now the question of estimating the number of points of a definable set $Y$ up to a given height defined over a real numberfield. Set $Y(F) = Y \cap F^n$ for a field $F \subset \mathbb{R}$ and put (again for $T \geq e$)

$$Y(F, T) = \{(x_1, \ldots, x_n) \in Y(F) : H(x_1), \ldots, H(x_n) \leq T\}.$$  

where $H(x)$ is the absolute multiplicative height of an algebraic number, as defined in [2], which agrees with the previous definition of $H(x)$ for rational $x$. Theorem 1.2 may be extended quite straightforwardly to an estimate of the same form for $\#Y_{\text{trans}}(F, T)$ when $F$ is a numberfield (i.e. $[F : \mathbb{Q}] < \infty$), in which the implicit constant depends on $Y, \epsilon$, and $[F : \mathbb{Q}]$.

Less straightforwardly, a much stronger result holds. For an integer $k \geq 1$, denote by

$$Y(k) = \{(x_1, \ldots, x_n) \in Y : [\mathbb{Q}(x_1) : \mathbb{Q}], \ldots, [\mathbb{Q}(x_n) : \mathbb{Q}] \leq k\}$$

the set of algebraic points of $Y$ of degree $\leq k$. Observe that the definition permits the coordinates of a point in $Y(k)$ to be defined over different fields. Put (for $T \geq e$)

$$Y(k, T) = \{(x_1, \ldots, x_n) \in Y(k) : H(x_1), \ldots, H(x_n) \leq T\}.$$
Then for a definable set \( Y \subset \mathbb{R}^n, k \geq 1, \) and \( \epsilon > 0 \) we have ([14])
\[
\#Y^{\text{trans}}(k, T) \leq c(Y, k, \epsilon)T^\epsilon.
\]

To obtain this result one studies the rational points of a suitable definable set \( Y_k \) of higher dimension than \( Y \) whose rational points correspond to points of \( Y \) of degree \( \leq k \). However \( Y_k^{\text{trans}} \) is empty, and a closer study of the proof structure of 1.2 is required.

In view of the above results for \( Y(F, T) \) and \( Y(k, T) \), it seems likely that if Conjecture 1.3 is affirmed, then the following stronger versions will also be affirmed. First, a version for varying number field with exponent independent of the number field.

1.4. Conjecture. Let \( Y \subset \mathbb{R}^n \) be definable in \( \mathbb{R}_{\exp} \) and \( F \subset \mathbb{R} \) a numberfield of degree \( f = [F : \mathbb{Q}] < \infty \). Then
\[
\#Y^{\text{trans}}(F, T) \leq c(Y, f)\left(\log T\right)^{C(Y)}.
\]

Second, a version for algebraic points of bounded degree.

1.5. Conjecture. Let \( Y \subset \mathbb{R}^n \) be definable in \( \mathbb{R}_{\exp} \) and \( k \geq 1 \). Then
\[
\#Y^{\text{trans}}(k, T) \leq c(Y, k)\left(\log T\right)^{C(Y, k)}.
\]

The following theorem affirms 1.4 for \( X \). For the time being I cannot establish 1.5 for \( X \). However I frame in Section 3 a conjecture (3.4) that would imply 1.4 and 1.5 in general.

1.6. Theorem. Let \( F \subset \mathbb{R} \) be a numberfield of degree \( f \) over \( \mathbb{Q} \), and let \( \epsilon > 0 \). Then
\[
\#X^{\text{trans}}(F, T) \leq c(f, \epsilon)\left(\log T\right)^{44 + \epsilon}.
\]

That the exponent of \( \log T \) in 1.6 is independent of \( F \) is a feature related to transcendence theory. In [13] I affirmed Wilkie’s conjecture for pfaff curves (see 5.2). (This class of plane curves does not contain all plane curves definable in \( \mathbb{R}_{\exp} \), but on the other hand there are pfaff curves that are not definable in \( \mathbb{R}_{\exp} \).) In [14] I observed that the result held for the points of a pfaff curve defined over a real number field \( F \), and with an exponent of \( \log T \) independent of \( F \). This result applies in particular to the graph \( W_\alpha \) for \( \alpha = x^\alpha, x \in (0, \infty) \), for positive irrational \( \alpha \), though it gives a result weaker than previously known results in that case. According to [13] and (for algebraic points) [14], if \( F \subset \mathbb{R} \) is a numberfield with \([F : \mathbb{Q}] = f\) then
\[
\#W_\alpha(F, T) \leq C(f)\left(\log T\right)^{20}.
\]

This estimate directly implies a weak form of the “Six Exponentials Theorem” as follows. Suppose that \( W_\alpha \) has 21 algebraic points \((x_i, y_i)\) on \( W_\alpha \) with the \( x_i \) multiplicatively independent. Then, considering the \( 21 \times 21 \) matrix \( \begin{vmatrix} x_1^{a_1} & \cdots & x_1^{a_{21}} \\ \vdots & \ddots & \vdots \\ x_n^{a_1} & \cdots & x_n^{a_{21}} \end{vmatrix} \) for 21-tuples of integers \( a_i \), we would have \#\( W_\alpha(F, T) \geq c(W_\alpha, F)\left(\log T\right)^{21} \) for suitable \( F \), giving a contradiction. Therefore, we conclude that if \( w_i \) are 21 real numbers, linearly independent over \( \mathbb{Q} \), then at least one of the 42 exponentials \( \exp w_i, \exp(\alpha w_i) \) must be transcendental.

In fact the same conclusion holds if there are just 3 linearly independent \( w_i \), namely that at least one of the \( 21 \) exponentials \( \exp w_i, \exp(\alpha w_i) \) is transcendental. This is the Six Exponentials Theorem, and our “Forty-Two Exponentials Theorem” is rather weak. However the point I wish to observe is that any estimate \#\( W_\alpha(F, T) \leq c(W_\alpha, F)\left(\log T\right)^{C(W_\alpha)} \) with \( C(W_\alpha) \) independent of \( F \) entails a transcendence result because the curve \( W_\alpha \) is a group (with suitable height growth in the group law), so that finitely many independent points generate an infinite set of a certain log-power density. The surface \( X \) is not a group, and so our \#\( X^{\text{trans}}(F, T) \leq c(f)\left(\log T\right)^{C} \) estimate does not yield a transcendence result, even though it is – qualitatively speaking – of the same quality.
Thus a uniform version of Wilkie’s conjecture i.e. that \( \#Y^{\text{trans}}(F,T) \leq c(Y,F)(\log T)^{C(Y)} \) for a set \( Y \) definable in \( \mathbb{R}^{\exp} \) and a real numberfield \( F \) with an exponent \( C(Y) \) independent of \( F \) (just as we affirm for \( X \) in 1.6) can be viewed as a qualitative transcendence-type statement, and for suitable sets \( Y \) it would indeed imply a transcendence result.

Our strategy combines elements of the approaches of several previous papers. The key to the method of [15] is the possibility of parameterizing a definable subset of \((0,1)^n\) of dimension \( k \) by finitely many functions \((0,1)^k \rightarrow (0,1)^n\) all of whose partial derivatives up to a prescribed order are bounded in absolute value by 1. In [12] I showed that Wilkie’s conjecture holds for pfaff curves that are mild, i.e. admit a parameterization in which derivatives to all orders are suitably controlled (see §2). Later, I established Wilkie’s conjecture in the form 1.3 for all pfaff curves by a different method in [13], and in the form 1.4 in [14]. Here, a mild parameterization of \( X \) is used to show that \( X(F,T) \) is contained in \( \ll (\log T)^C \) intersections of \( X \) with hypersurfaces of degree \( \ll (\log T)^2 \). These intersection curves are treated by adapting the methods of [13]. Here, as in [12, 13], a crucial role is played by results of Gabrielov and Vorob’ev [7] estimating the topological complexity of Pfaffian sets (see §5). As it stands, this combination of methods — mild parameterization for the initial set and Pfaffian bounds for the intersection curves — is applicable only to surfaces. Our surface \( X \) was selected as being related to the threefold \( \log x \log y = \log z \log t \) associated with the Four Exponentials Conjecture (see [17]). The present method is generalized by Butler [4] to further surfaces definable in \( \mathbb{R}^{\exp} \).

Acknowledgements. My thanks to Lee Butler for detailed corrections to a previous version of this paper, to Eric Descheemaeker for assistance, and to the referee for helpful comments and suggestions. I am grateful to Roger Heath-Brown and the Mathematical Institute, Oxford, for affording me hospitality as an Academic Visitor, and to the Leverhulme Trust for supporting my work through a Research Fellowship.

2. Mild functions

We write \( x = (x_1, \ldots, x_k) \) etc as variables in \( \mathbb{R}^k \). For a function \( \phi : U \rightarrow \mathbb{R} \) defined on some domain \( U \subset \mathbb{R}^k \) and \( \mu = (\mu_1, \ldots, \mu_k) \in \mathbb{N}^k \) we set \( |\mu| = \sum \mu_i \) and denote by \( \partial^\mu \phi \) the partial derivative

\[
\partial^\mu \phi = \phi^{(\mu)} = \frac{\partial^{|\mu|} \phi}{\partial x_1^{\mu_1} \cdots \partial x_k^{\mu_k}}
\]

of order \( |\mu| \). We denote by \( x^\mu \) the monomial \( \prod x_i^{\mu_i} \) of degree \( |\mu| \). We set \( \mu! = \prod \mu_i! \) and \( \pi = \max \mu_i \).

2.1. Definition. A function \( \phi : (0,1)^k \rightarrow (0,1) \) is called \((A,C)\)-mild if it is \( C^\infty \) and, for all \( \mu \in \mathbb{N}^k \) and all \( z \in (0,1)^k \),

\[
|\partial^\mu \phi(z)| \leq \mu!(A|\mu|^C)^{|\mu|}.
\]

2.2. Remark. One could define a finer notion \((A,B,C)\)-mild with a term \((|\mu|+1)^B\) to enable finer estimates. However only the parameter \( C \) survives to influence the exponent of \( \log T \) in the density estimate, so the above notion was preferred for simplicity.

2.3. Definition. A function \( \theta : (0,1)^k \rightarrow (0,1)^n, \theta(x) = (\theta_1(x), \ldots, \theta_n(x)) \) is called \((A,C)\)-mild if each of its coordinate functions \( \theta_i \) is \((A,C)\)-mild.

2.4. Definition. A set \( Y \subset (0,1)^n \) of dimension \( k \) is called \((J,A,C)\)-mild if there exists a collection \( \Theta \) of \((A,C)\)-mild maps \( \theta : (0,1)^k \rightarrow (0,1)^n \) such that \( \# \Theta = J \) and

\[
\bigcup_{\theta \in \Theta} \theta((0,1)^k) = Y.
\]

A set \( Y \subset (0,1)^n \) is called mild if it is \((J,A,C)\)-mild for some \( J, A, C \).
2.5. Conjecture. Every set $Y \subset (0,1)^n$ definable in $\mathbb{R}_{exp}$ is mild.

A more precise version of this conjecture is formulated in 3.4. A more optimistic version would require a fixed value of $C$. The following property of mild functions will be used in the sequel.

2.6. Proposition. Suppose $\phi_1, \ldots, \phi_\ell : (0,1)^k \to (0,1)$ are $(A,C)$-mild, $\mu \in \mathbb{N}^k$ and $z \in (0,1)^k$. Then

$$|\partial^\mu \phi_1 \ldots \phi_\ell(z)| \leq \mu!(\bar{p} + 1)^{(\ell-1)k} (A|\mu|_C)^{|\mu|}.$$ 

Proof. We have

$$\partial^\mu \phi_1 \ldots \phi_\ell = \sum_{\mu_1 + \ldots + \mu_\ell = \mu} \text{Ch}(\mu_1, \ldots, \mu_\ell) \prod_{i=1}^\ell \partial^{\mu_i} \phi_i$$

where, for $\alpha = (\alpha_1, \ldots, \alpha_k), \beta = (\beta_1, \ldots, \beta_k)$, etc

$$\text{Ch}(\alpha, \beta, \ldots, \zeta) = \prod_{j=1}^k \frac{(\alpha_j + \beta_j + \ldots + \zeta_j)!}{\alpha_j!\beta_j!\ldots\zeta_j!}.$$

Therefore

$$\frac{|\partial^\mu \phi_1 \ldots \phi_\ell(z)|}{\mu!} \leq \sum_{\mu_1 + \ldots + \mu_\ell = \mu} \prod_{i=1}^k |\partial^{\mu_i} \phi_i| \mu_i! \leq (\bar{p} + 1)^{(\ell-1)k} \prod (A|\mu_i|_C)^{|\mu_i|} \leq (\bar{p} + 1)^{(\ell-1)k} (A|\mu|_C)^{|\mu|}$$

as required. $\square$

We next establish that certain functions that we will use in our parameterizations are mild. First observe that the function

$$\psi(r) = r^e e^{1-r} = \exp(r \log r + 1 - r)$$

is increasing for $r \geq 1$, as the derivative $\log r$ of the exponent is positive for $r > 1$, and has $\psi(1) = 1$. We define $\psi(0) = 1$.

2.7. Lemma. Let $m = (m_1, \ldots, m_k) \in (0,\infty)^k, a = (a_1, \ldots, a_k) \in [0,\infty)^k$ and suppose that, for each $i$, either $a_i = 0$ or $a_i \geq m_i$. Define $E_{m,a} : (0,1)^k \to \mathbb{R}$ by

$$E_{m,a}(z) = \exp\left(1 - \frac{1}{z^m}\right) \frac{1}{z^a}.$$ 

Then

$$\sup_{z \in (0,1)^k} |E_{m,a}(z)| = \psi\left(\max_i (a_i/m_i)\right).$$

Proof. If all $a_j = 0$ then the supremum is clearly 1, which agrees with our definition of $\psi(0) = 1$. So we can assume that some $a_j > 0$, so that $a_j \geq m_j$ by our hypothesis, and then $\max_i (a_i/m_i) \geq a_j/m_j \geq 1$.

We proceed by induction on $k$. If $k = 1$ we have $E_{m,a}(z) = E(t) = \exp(1-t^{-1})t^{-r}$ where $t = z^m$, $t \in (0,1)$, $r \geq 1$. The maximum of the function for $t \in [0,\infty)$ occurs at $t = 1/r \in (0,1]$ and has the value $\psi(r)$.

Suppose the result true for $k - 1$ variables, $k \geq 2$. We have

$$\partial^r E_{m,a}(z) = \frac{E_{m,a}(z)}{z^i} (m_iz^{-m_i} - a_i).$$

If all $a_i/m_i = r$, the function again reduces to a function of one variable, $E_{m,a}(z) = \exp(1-t^{-1})t^{-r}$, where $t = z^m, r \geq 1, t \in (0,1)$. As before the maximum of the function for $t \in [0,\infty)$ occurs at $t = 1/r$ and has the value $\psi(r)$, affirming the conclusion.
If the $a_i/m_i$ are not all equal, then there is no stationary point inside $(0,1)^k$ and the supremum is given by the maximum of the function on $[0,1]^k$, which is attained on a boundary, and moreover on a boundary where some $x_i = 1$, as the function is flat at the $x_i = 0$ boundaries.

By induction, the supremum on a boundary $x_j = 1$ is $\psi(\max(r_j,j \neq i))$. As the function $\psi$ is increasing for arguments $\geq 1$, we get the desired conclusion in this case too, and complete the induction and the proof. □

2.8. Proposition. For $m = (m_1,\ldots,m_k) \in (0,\infty)^k$ define $E_m : (0,1)^k \to \mathbb{R}$ by

$$E_m(z) = \exp\left(1 - \frac{1}{z^m}\right).$$

Then $E_m$ is $(A,C)$-mild with $C = \max((m_i + 1)/m_i)$ and $A = (\bar{m} + 1)C^{1/m_i}$.\n
Proof. Write $E = E_m$. For $\mu \in \mathbb{N}^k$ we have

$$\partial^\mu E = E \sum_{m'} a_{m'}^{(\mu)} z^{-m'},$$

over suitable $m' \in (0,\infty)^k$. The $m'$ that appear all have, for each $i$, $m'_i = 0$ or $m'_i > m_i$. Furthermore, for each $i$, the largest $m'_i$ occurring is $\mu_i(m_i + 1)$.

Set, for $\mu \in \mathbb{N}^k$,

$$\alpha_\mu = \sum_{m'} |a_{m'}^{(\mu)}|$$

and, for $\ell \in \mathbb{N}$,

$$\alpha_\ell = \max_{|\mu| = \ell} \alpha_\mu.$$

Denote by $e_i$ the element of $\mathbb{N}^k$ that has zero entries except for an entry 1 in the $i$-th place, so that $\partial^{e_i} = \partial^{e_i}$. Observe that

$$\partial^{e_i} \partial^\mu E = \partial^{e_i} \left(E \sum_{m'} a_{m'}^{(\mu)} z^{-m'}\right) = E \sum_{m'} a_{m'}^{(\mu)} \left(z^{-m'} \frac{m_i}{z^{m+e_i}} - \frac{m'_i}{z^{m'+e_i}}\right).$$

Therefore

$$\alpha_{\mu + e_i} \leq m_i \alpha_\mu + \max_{m'} (m'_i) \alpha_\mu = m_i \alpha_\mu + \mu_i(m_i + 1) \alpha_\mu \leq (\mu_i + 1)(\bar{m} + 1) \alpha_\mu,$$

and so, by induction on $|\mu|$,\n
$$\alpha_\mu \leq \mu!(\bar{m} + 1)^{|\mu|}.$$\n
The largest “$a/m$” occurring is\n
$$\max_i \frac{\mu_i(m_i + 1)}{m_i} \leq \bar{m} \lambda$$

where\n
$$\lambda = \max_i \frac{m_i + 1}{m_i}.$$

By Lemma 2.7,

$$\frac{|\partial^\mu E(z)|}{\mu!} \leq (\bar{m} + 1)^{|\mu|} \left(\frac{\bar{m} \lambda}{e}\right)^{\lambda/\bar{m}}.$$

This establishes that $E_m$ is $(A,C)$-mild with\n
$$A = (\bar{m} + 1)\left(\frac{\lambda}{e}\right)^\lambda, \quad C = \lambda$$

as required. □
3. Exploring mild sets with algebraic hypersurfaces

3.1. Proposition. For integers \( a > 0, x \geq a(a + 1)/2 \),
\[
\frac{x^a}{a!} \leq \binom{a + x}{a} \leq \frac{x^a}{a!} \left( 1 + \frac{a(a + 1)}{x} \right).
\]

Proof. We have
\[
\binom{a + x}{a} = \frac{(a + x)!}{a!x!} = \frac{x^a}{a!} \left( 1 + \frac{a}{x} \left( \frac{a - 1}{x} \right) \cdots \left( \frac{1}{x} \right) \right).
\]
So the left-hand inequality of the Proposition is immediate provided only \( a, x \) are positive, while
\[
\log \left( \left( 1 + \frac{a}{x} \left( \frac{a - 1}{x} \right) \cdots \left( \frac{1}{x} \right) \right) \leq \frac{a}{x} + \cdots + \frac{1}{x} = \frac{a(a + 1)}{x} \right).
\]
Since \( e^y \leq 1 + 2y \) for \( 0 \leq y \leq 1 \), the assumption \( x \geq a(a + 1)/2 \) implies
\[
\left( 1 + \frac{a}{x} \left( \frac{a - 1}{x} \right) \cdots \left( \frac{1}{x} \right) \right) \leq \exp \left( \frac{a(a + 1)}{2x} \right) \leq 1 + \frac{a(a + 1)}{x}
\]
giving the right-hand inequality provided \( x \geq a(a + 1)/2 \).

We observe the following consequences of this Lemma, in which the expression “1+o(1)” is to apply
for \( d \to \infty \) with \( k, n \) fixed.

Let \( \Lambda_k(d) \) denote the set of monomials of exact degree \( d \) in \( k \) variables, and \( L_k(d) = \#\Lambda_k(d) \). Then
\[
L_k(d) = k - 1 + d = \frac{d^{k-1}}{(k-1)!} \left( 1 + o(1) \right).
\]
Let \( \Delta_k(d) \) denote the set of monomials of degree \( \leq d \) in \( k \) variables, and \( D_k(d) = \#\Delta_k(d) \). Then
\[
D_k(d) = L_{k+1}(d) = \frac{d^k}{k!} \left( 1 + o(1) \right).
\]
Let \( b(k, n, d) \) be the unique positive integer \( b \) with
\[
D_k(b) \leq D_n(d) < D_k(b+1).
\]
Then
\[
b(k, n, d) = \left( \frac{k!d^n}{n!} \right)^{1/k} \left( 1 + o(1) \right).
\]
Let
\[
B(k, n, d) = \sum_{\beta=0}^{b} L_k(\beta) \beta + \left( D_n(d) - \sum_{\beta=0}^{b} L_k(\beta) \right) (b+1).
\]
Then
\[
B(k, n, d) = \frac{1}{(k+1)!} \left( \frac{k!}{n!} \right)^{(k+1)/k} d^{n(k+1)/k} \left( 1 + o(1) \right).
\]
Finally, let
\[
V(n, d) = \sum_{\beta=0}^{d} L_n(\beta) \beta.
\]
Then
\[
V(n, d) = \frac{1}{(n+1)(n-1)!} d^{n+1} \left( 1 + o(1) \right).
\]
The following are the results showing that, for a mild set \( Y \subset (0,1)^n \) of dimension \( k \), \( Y(F,T) \) is contained in “few” algebraic hypersurfaces. It is convenient to establish the result first using a different height function.

For an algebraic number \( \alpha \) we denote by \( \text{den}(\alpha) \) the denominator of \( \alpha \), namely, the least positive integer \( m \) such that \( m \alpha \) is an algebraic integer. If \( \alpha_i \in \mathbb{C} \) are the conjugates of \( \alpha \) we set

\[
H^{\text{size}}(\alpha) = \max\{\text{den}(\alpha), |\alpha_i|\}.
\]

Suppose \( \alpha \), of degree \( f \), with minimal polynomial (over \( \mathbb{Z} \)) \( a_f(t - \alpha_1)\ldots(t - \alpha_f) \). Then \cite[1.6.5, 1.6.6]{1} \( H^{\text{size}}(\alpha) \leq |a_f| \prod_i \max(1, |\alpha_i|) = H(\alpha)^f \).

For \( Y \subset \mathbb{R}^n \) we set

\[
Y^{\text{size}}(F,T) = \{(x_1,\ldots,x_n) \in Y(F) : H^{\text{size}}(x_1),\ldots,H^{\text{size}}(x_n) \leq T\}.
\]

For \( \alpha \in \mathbb{R} \) we let \([\alpha]\) denote the integer part (least integer not exceeding \( \alpha \)).

3.2. Theorem. Suppose \( Y \subset (0,1)^n \) of dimension \( k \) has a \((J,A,C)\)-mild parameterization. Let \( f \) be a positive integer and \( F \subset \mathbb{R} \) a numberfield of degree \( f \) over \( \mathbb{Q} \). Then \( Y^{\text{size}}(F,T) \) is contained in at most

\[
Jc(k,n)^f A^{(k+1)(1+o(1))} \left( \log T \right)^{C \left( \frac{n(k+1)}{\sigma-1} \right)^{(1+o(1))}}
\]

intersections of \( Y \) with hypersurfaces (possibly reducible) of degree

\[
d = \left[ \left( \log T \right)^\frac{k}{n} \right]
\]

where “\(1+o(1)\)” is taken as \( T \to \infty \) with implicit constants depending only on \( k,n \).

Proof. Since \( Y \) is the union of \( J \) images of mild maps, it suffices (given the factor \( J \) in the conclusion) to suppose that \( Y \) is the image of a single \((A,C)\)-mild map \( \theta : (0,1)^k \to (0,1)^n \).

Consider a \( D_n(d) \times D_n(d) \) determinant \( \Delta \) of the form

\[
\Delta = \det \left( (x^{(i)})^j \right)
\]

where \( j \in \mathbb{N}^n \) with \( |j| \leq d \) indexes the columns, \( x^{(i)} \in Y(F,T) \), \( i = 1,\ldots,D_n(d) \), and \( x^j \) denotes as usual the monomial \( \prod x_i^{j_i} \).

Each coordinate of each \( x^{(i)} \) has denominator \( \leq T \). The entries in row \( i \) consist of monomials in which each coordinate is raised to power \( \leq d \). Therefore \( K\Delta \) is an algebraic integer for some positive integer \( K \) with

\[
K \leq T^{ndD_n(d)},
\]

and then

\[
\prod_\sigma (K\Delta)^\sigma \in \mathbb{Z}
\]

where \( \sigma \) runs over the embeddings \( F \to \mathbb{C} \).

Let us estimate \( |\Delta^\sigma| \) (later we will use the mild parameterization to get a better estimate for \( \Delta \) itself, i.e. when \( \sigma = \text{id} \)). Expand \( \Delta^\sigma \) into a sum of \( D_n(d)! \) terms. Since \( \Delta \) has \( L_n(\beta) \) columns of degree \( \beta \), for \( \beta = 0,\ldots,d \), and in each column the entries have absolute value at most \( T^{\beta} \), the largest term in the expansion has complex absolute value

\[
\leq T^{\sum L_n(\beta)\beta} = T^{V(n,d)}
\]
so that, for any $\sigma$,
\[ |\Delta^\sigma| \leq D_n(d)!T^{V(n,d)}. \]

Therefore, if $\Delta \neq 0$ then $\prod_{\sigma}(K\Delta)^\sigma$ is a non-zero integer and
\[
1 \leq |K\Delta| \prod_{\sigma \neq \text{id}} |K\Delta^\sigma| \leq |\Delta|T^{fndD_n(d)+(f-1)V(n,d)}(D_n(d)!)^{f-1}.
\]

To estimate $|\Delta|$, suppose that the points $x^{(i)}$ are the images of some points $z^{(i)} \in (0,1)^k$ under $\theta$ where the $z^{(i)}$ in fact belong to some cube of side $\leq r \leq 1$, and so are at a distance $\leq r$ in each coordinate from the centre $z^{(0)}$ of the cube, which contains also all the lines segments from $z^{(0)}$ to $z^{(i)}$. We have then that
\[
\Delta = \det(\phi_j(z^{(i)}))
\]
where $\phi_j$ is the monomial function indexed by $j$, namely
\[
\phi_j(z^{(i)}) = \left(\theta_1(z^{(i)}), \ldots, \theta_n(z^{(i)})\right)^j = \left(x_1^{(i)}, \ldots, x_n^{(i)}\right)^j.
\]

We expand each entry of $\Delta$ as a Taylor series about $z^{(0)}$ of order $b = b(k,n,d)$ with remainder terms of order $b + 1$:
\[
\phi_j(z^{(i)}) = \sum_{\mu \in \Delta_k(b)} \frac{\partial^\mu \phi_j(z^{(0)})}{\mu!} \left(z^{(i)} - z^{(0)}\right)^\mu + \sum_{\mu \in \Delta_k(b+1)} \frac{\partial^\mu \phi_j(z^{(i)})}{\mu!} \left(z^{(i)} - z^{(0)}\right)^\mu
\]
where $\zeta = \zeta_j$ is a suitable intermediate point on the line segment from $z^{(0)}$ to $z^{(i)}$.

Now we expand out the determinant. In doing so, terms of low degree as products of terms of the form $(z^{(i)} - z^{(0)})$ cancel out, as observed in [10, Proof of 3.1]. Specifically, consider the totality of terms corresponding to a particular specification of the number of multiplicands of each order of derivative. Consider a minor of size $h \times h$ of $\det(\phi_j(z^{(i)}))$ comprising the expansion terms of degree $\beta \leq b$ only. That is, select $h$ points $\zeta^{(i)}$ from among the $z^{(i)}$, and $h$ functions $\psi_j$ from among the $\phi_j$ and consider
\[
\det\left(\sum_{\mu \in \Delta_k(\beta)} \frac{\partial^\mu \psi_j(z^{(0)})}{\mu!} \left(\zeta^{(i)} - z^{(0)}\right)^\mu\right).
\]
If $h > L_k(\beta)$ then the columns are dependent and the minor vanishes. Thus if, for a particular specification of orders, there are more than $L_k(\beta)$ multiplicands of order $\beta$ for some $\beta$, then the totality of terms corresponding to this choice vanishes.

Therefore, all surviving terms are products of $B(k,n,d)$ or more terms of the form $(z^{(i)} - z^{(0)})$. The number of surviving terms is estimated by the number of terms assuming no cancellation, i.e. for each term we consider which row the multiplicand from column $j$ came from, for which there are $D_n(d)!$ possibilities, and given this choice we can then choose, for each column, one of the $D_k(b+1)$ terms in the Taylor expansion, giving an estimate for the number of terms of at most
\[
D_n(d)!D_k(b+1)^{D_n(d)}.
\]

Finally, each term is a product of $D_n(d)$ terms, each one of the summands in the Taylor formula for $\phi_j$ which, neglecting the terms $(z^{(i)} - z^{(0)})$, takes the form
\[
\frac{\partial^\mu \left(\theta^j\right)(\zeta)}{\mu!}
\]
for some suitable $\zeta$, and some $\mu$ with $|\mu| \leq b+1$. By Proposition 2.6, as $\theta$ is $(A,C)$-mild and $|\mu| \leq b+1$,
\[
\left|\frac{\partial^\mu \left(\theta^j\right)(\zeta)}{\mu!}\right| \leq (\pi + 1)^{|j|-1}k \left(A(b+1)^C\right)^{b+1} \leq (b+2)^{|j|k} \left(A(b+1)^C\right)^{b+1}.
\]

9
Now
\[ \sum_{j \in \mathbb{N}^n : |j| \leq d} |j| = \sum_{\beta=0}^d \beta L_n(\beta) = V(n, d) \]
so that
\[ \prod_{j \in \mathbb{N}^n : |j| \leq d} (b+2)^{|j|k} \leq (b+2)^{kV(n,d)}. \]

Therefore, since \(|x^{(1)}_r - z^{(0)}_r| \leq r \leq 1,\)
\[ |\Delta| \leq D_n(d)!D_k(b+1)^{D_{n,d}}(b+2)^{kV(n,d)} \left( (A(b+1)^C)^{k+1} \right)^{D_{n,d}} r^{B(k,n,d)}, \]
and if the points \(x^{(i)}\) do not lie on any hypersurface in \(\mathbb{R}^n\) of degree \(d\) then \(\Delta \neq 0\) and
\[ 1 \leq (D_n(d))^f D_k(b+1)^{D_{n,d}}(b+2)^{kV(n,d)} T^{fndD_{n,d}+fV(n,d)} \left( (A(b+1)^C)^{b+1} \right)^{D_{n,d}} r^{B(k,n,d)}. \]

Now we take the \(B(k,n,d)\)-th root of this inequality. In the following discussion, the expression “\(1 + o(1)\)” is to be taken as \(d \to \infty\) with \(k,n\) fixed, while \(c(k,n)\) is a positive constant that may differ at each occurrence.

First we observe that
\[ \frac{D_n(d)}{B(k,n,d)} = \frac{d^n (k+1)(k-1)!}{n!} \frac{n^k}{d^{n(k+1)/k}} \left( \frac{n!}{k!} \right)^{1/k} (1 + o(1)) = \frac{c(k,n)}{d^{n/k}}, \]
where
\[ c(k,n) = \frac{k+1}{k} \left( \frac{n!}{k!} \right)^{1/k} (1 + o(1)), \]
and that
\[ \frac{V(n,d)}{B(k,n,d)} = c(k,n)(1 + o(1)) \frac{d^{n+1}}{d^{n(k+1)/k}} = \frac{c(k,n)}{d^{n/k-1}}. \]

So
\[ \left( \frac{D_n(d)}{B(k,n,d)} \right)^f \leq \frac{D_n(d)^f}{B(k,n,d)^{fD_{n,d}}} = c(k,n)(1 + o(1)d^n)^f c(k,n)^{d^{n+k}} (1 + o(1))^f, \]
and similarly
\[ D_k(b+1)^{\frac{D_{n,d}}{B(k,n,d)}} = \left( \frac{(b+1)^k}{k!} (1 + o(1)) \right)^{c(k,n)(1 + o(1)d^n)^f} c(k,n)(1 + o(1)d^n)^f c(k,n)^{d^{n+k}} = 1 + o(1). \]

Next,
\[ (b+2)^{\frac{kV(n,d)}{B(k,n,d)}} = c(k,n)(1 + o(1))d^{n/k} = c(k,n)^f \]
provided
\[ d = \left[ \left( \log T \right)^{\frac{k}{n}} \right]. \]

Finally
\[ \frac{(b+1)D_n(d)}{B(k,n,d)} = \frac{k+1}{k} (1 + o(1)), \]
so that
\[ \left( A(b+1)^C \right)^{\frac{D_{n,d}}{B(k,n,d)}} = \left( Ac(k,n)(1 + o(1))d^{n+k} \right)^{\frac{k+1}{k} (1 + o(1))} = c(k,n)^f d^{nC/k}. \]
where

\[ P = \frac{k+1}{k}(1+o(1)). \]

Thus if \( \Delta \neq 0 \) we find that

\[ 1 \leq c(k,n) PA^{P} d^{nCP/k} r \]

where

\[ d = \left[ \left( \log T \right)^{\frac{1}{k-1}} \right] \]

and all the preimages \( z^{(i)} \) of the points \( x^{(i)} \) lie in a cube of side \( r \) in \( (0,1)^{k} \). The points \( x^{(i)} \) whose coordinates have \( H^\text{size}(x^{(i)}) \leq T \) and whose preimages lie in such a cube must therefore all lie on one hypersurface (possibly reducible) of degree \( d \) provided

\[ r < c(k,n)f^{kP} d^{nCP/k}, \]

and since \( (0,1)^{k} \) may be covered by at most

\[ c(k,n)f^{kP} d^{nCP} \]

such cubes, and \( T, d \) go to infinity together, the proof is complete. \( \square \)

3.3. Corollary. Suppose \( Y \subset (0,1)^{n} \) of dimension \( k \) has a \((J,A,C)\)-mild parameterization. Let \( f \) be a positive integer and \( F \subset \mathbb{R} \) a numberfield of degree \( f \) over \( \mathbb{Q} \). Then \( Y(F,T) \) is contained in at most

\[ Jc(k,n)f^{A(k+1)(1+o(1))} \left( f \log T \right)^{C \left( \frac{n(k+1)}{k} \right)(1+o(1))} \]

intersections of \( Y \) with hypersurfaces (possibly reducible) of degree

\[ d = \left[ \left( f \log T \right)^{\frac{1}{k-1}} \right] \]

where “\( 1+o(1) \)” is taken as \( T \to \infty \) with implicit constants depending only on \( k,n \).

Proof. We have \( Y(F,T) \) contained in \( Y^\text{size}(F,T^f) \). \( \square \)

3.4. Conjecture. Let \( Y \subset (0,1)^{n} \) be definable in \( \mathbb{R}_{\exp} \). There exist constants \( C_{1},C_{2},C_{3},C_{4},C_{5} \) depending only on \( Y \) with the following property. Let \( \mathcal{F} \) be an algebraic family of closed algebraic sets in \( \mathbb{R}^{n} \) of degree \( d = d(\mathcal{F}) \), and suppose \( V \in \mathcal{F} \). Then \( Y \cap V \) is \((C_{2}d^{C_{5}},C_{4}d^{C_{5}},C_{1})\)-mild.

3.5. Corollary. Conjecture 3.4 implies Conjectures 1.4 and 1.5.

Proof. It suffices to work with \( H^\text{size} \). By maps \( x \to \pm x^{\pm 1} \) it suffices, as in [15], to consider sets \( Y \subset (0,1)^{n} \). Then one iteratively intersects with hypersurfaces. Assuming 3.4, all the sets involved are \((J,A,C_{1})\)-mild with \( C_{1} \) fixed and \( J,A \) depending polynomially on the degree of the family. For 1.5, imitate the proof of Theorem 5.3 in [14] using 3.3 to estimate the number of intersections required at each stage, rather than the appeal in [14] (via [15]) to [10, Lemma 4.4]. For 1.4, use 3.3 on \( Y \) and then on the intersections given by the conclusion of 3.3 repeatedly. In both cases the degrees of the families are polynomial in \((\log T)\) at each stage. \( \square \)

4. The algebraic part, Schanuel’s conjecture and algebraic points

4.1. Proposition. Let

\[ X = \{(x,y,z) \in (0,\infty)^{3} : \log x \log y = \log z \} \]

then \( X^\text{alg} \) consists of the lines \( L_{x} = \{(x,1,1) : x \in (0,\infty)\} \) and \( L_{y} = \{(1,y,1) : y \in (0,\infty)\} \) and, for \( q \in \mathbb{Q}^{*} \), the curves \( \Gamma_{x,q} = \{(x,e^{q},z) : z = x^{q}, x > 0\} \) and \( \Gamma_{y,q} = \{(e^{q},y,z) : z = y^{q}, y > 0\} \).
Proof. Suppose that $\Gamma$ is an arc of an algebraic curve contained in $X$. Suppose $x$ is constant on $\Gamma$. If $x = 1$ then also $z = 1$ and $\Gamma$ is an arc of the line $L_y$. If $x$ is constant but not equal to 1 then $q = \log x$ must be rational, and $\Gamma$ is contained in the curve $\Gamma_{y,q}$. Similarly, if $y$ is constant we find $\Gamma$ contained in $L_x$ or $\Gamma_{x,q}$. If $z$ is constant, we get no algebraic curves unless $z = 1$ and we find that either $x = 1$ or $y = 1$ identically on $\Gamma$ and revert to the previous cases. Otherwise, $x, y, z$ are non-constant and further $y, z$ are algebraic functions of $x$. We then have

$$x = \exp \left( \frac{\log z(x)}{\log y(x)} \right)$$

on $\Gamma$ and, by analytic continuation, this relation holds also for large (possibly complex) $x$. Then $y(x), z(x)$ are given by some convergent Puiseaux series,

$$z(x) = z_0 x^s + \ldots, \quad y(x) = y_0 x^r + \ldots$$

and we have

$$x = \exp \left( \frac{\zeta \log x + \log z_0 + \log(1+ \ldots)}{\eta \log x + \log y_0 + \log(1+ \ldots)} \right)$$

which is clearly untenable for large $|x|$ as the right hand side tends to a finite limit. \(\square\)

We now elaborate the implications of Schanuel’s conjecture for algebraic points on $X$. Schanuel’s conjecture implies that the logarithms of multiplicatively independent algebraic numbers are algebraically independent over $\mathbb{Q}$ (see e.g. [19]).

4.2. Proposition. Assume Schanuel’s conjecture (or just that the logarithms of multiplicatively independent algebraic numbers are algebraically independent). Then if $x, y, z \in (0, \infty)$ are algebraic with $\log x \log y = \log z$ then either $(x, y, z) = (x, 1, 1)$ for some $x \in \mathbb{Q}$, or $(x, y, z) = (1, y, 1)$ for some $y \in \mathbb{Q}$.

Proof. Suppose $x, y, z$ are algebraic numbers in $(0, \infty)$ with $\log x \log y = \log z$. Then $x, y, z$ are multiplicatively dependent, and we have

$$x^a y^b z^c = 1$$

for certain integers $a, b, c$. If two of $a, b, c$ equal 0 then one of $x, y, z = 1$ and then we have either $x = z = 1$ and $y$ arbitrary or $y = z = 1$ and $x$ arbitrary.

Suppose that just one of $a, b, c$ is zero, assuming $x, y, z \neq 1$. If $c = 0$ we have $y = x^r$ for some rational $r \neq 0$ and $r(\log x)^2 = \log z$. Then $x, z$ must (by Schanuel) be multiplicatively related, say $z = x^s$ for some $s \in \mathbb{Q}^*$ and $r(\log x)^2 = s \log x$ implies $\log x = 0$ (contrary to our assumptions) or $\log x \in \mathbb{Q}^*$, whence $x$ is non-algebraic. If $a = 0$, then $z = y^r$ for some $r \in \mathbb{Q}^*$ and $\log x \log y = r \log y$ implies (as $\log y \neq 0$) that $\log x \in \mathbb{Q}^*$ and is not algebraic.

Suppose then that none of $a, b, c$ is zero. Then $z$ depends multiplicatively on $x$ and $y$ and we get a relation $r \log x + s \log y = \log x \log y$ with $r, s$ non-zero rational numbers. Then $x, y$ must be multiplicatively related, and we find that $\log x$ is algebraic and hence $x = 1$. \(\square\)

4.3. Summary. The set $X^\mathrm{als}$ consists of infinitely many real semi-algebraic curves: the lines $L_x, L_y$ and, for each $q \in \mathbb{Q}^*$, the curves $\Gamma_{x,q}, \Gamma_{y,q}$. By the Hermite-Lindemann theorem the curves $\Gamma_{x,q}, \Gamma_{y,q}$ contain no algebraic points. The lines $L_x, L_y$ evidently contain algebraic points. Under Schanuel’s Conjecture, $X^\alpha(\supset X^{\mathrm{trans}})$ contains no algebraic points.

5. Pfaffian sets and Gabrielov-Vorobjov bounds

Definition 5.1 and the key result Theorem 5.3 are taken from the paper [7] of Gabrielov and Vorobjov.

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

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

12
for $1 \leq j \leq r$, were $g_{ij} \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_j]$ are polynomials of degree not exceeding $\alpha$. A function

$$f = P(x_1, \ldots, x_n, f_1, \ldots, f_r)$$

where $P \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_r]$ is a polynomial of degree not exceeding $\beta \geq 1$ is called a pfaffian function of order $r$ and degree $(\alpha, \beta)$.

5.2. Definition. By a pfaffian set we will mean the set of common zeros of some pfaffian functions. By a pfaff curve we mean the graph of a pfaffian function of one variable on a connected subset of $\mathbb{R}$.

In the above definition no restriction is placed on the domain $G$. To obtain complexity bounds on pfaffian sets, one must impose restrictions on $G$ (as we will do, following [7]), or allow more complicated domains whose complexity contributes to the complexity of the pfaffian sets. By a simple domain $G \subset \mathbb{R}^n$ we mean, as in [7], that $G$ is a domain of the form $\mathbb{R}^n, [-1, 1]^n, (0, \infty)^n$ or $\{x : ||x||^2 < 1\}$. The number of connected components of a set $Y$ is denoted $cc(Y)$.

5.3. Theorem. ([7, Corollary 3.3]) Let $h_1, \ldots, h_\ell$ be pfaffian functions in a simple domain $G \subset \mathbb{R}^n$ having a common pfaffian chain of order $r$ and degrees $(\alpha, \beta_i)$ respectively. Put $\beta = \max_i \beta_i$. Let $Y$ be the pfaffian set

$$Y = \{x \in G : h_1(x) = \ldots = h_\ell(x) = 0\}$$

Then

$$cc(Y) \leq 2^{r(r-1)/2+1}\beta(\alpha + 2\beta - 1)^{n-1}((2n-1)(\alpha + \beta) - 2n + 2)^\ell. \quad \square$$

Observe that the bound on $cc(Y)$ does not depend on $\ell$. When the ambient space $\mathbb{R}^n$ and the pfaffian chain are fixed, as they will be, this fixes $n, r, \alpha$ and then we have

$$cc(Y) \leq c(n, r, \alpha)\beta^{n+r}.$$ 

6. Proof of Theorems 1.1 and 1.6

Theorems 1.1 and 1.6 concern the surface

$$X = \{(x, y, z) \in (0, \infty)^3 : \log x \log y = \log z\}.$$ 

If $\log x = 0$ then $\log z = 0$ also, so $X \cap \{x = 1\} = \{(x, y, z) : x = z = 1\} \subset X^{\text{alg}}$. Likewise $X \cap \{y = 1\} \subset X^{\text{alg}}$, while if $\log z = 0$ we must have $\log x = 0$ or $\log y = 0$, so that $X \cap \{z = 1\} \subset X^{\text{alg}}$ too. In studying $(X - X^{\text{alg}})(F, T)$ we may therefore assume that $x, y, z \neq 1$. Let

$$\mathcal{X} = \{(x, y, z) \in (0, 1)^3 : \log x \log y = -\log z\}.$$ 

The surface $\mathcal{X}$ contains semi-algebraic curves corresponding to fixing a rational negative value for $\log x$ or $\log y$. However, these curves contain no algebraic points (the corresponding $x$ or $y$ is transcendental by the Hermite-Lindemann Theorem). Thus $\mathcal{X}^{\text{alg}}(\mathbb{Q})$ is empty, and we need not restrict our counting to $\mathcal{X}^{\text{trans}}$.

If $(x, y, z) \in X(F, T)$ with $x > 1, y > 1$ then $z > 1$ also. Since $H(\alpha) = H(1/\alpha)$ for any nonzero algebraic number $\alpha$, we see that $(1/x, 1/y, 1/z) \in X(F, T)$. If $(x, y, z) \in X(F, T)$ with $x < 1, y > 1$ then $z < 1$ and now $(x, 1/y, z) \in X(F, T)$. The cases $x > 1, y < 1$ and $x, y < 1$ are similar and we see that, up to a finite multiplicative factor, Theorems 1.1 and 1.6 follow from the following result concerning $\mathcal{X}$.

6.1. Theorem. Let $F \subset \mathbb{R}$ be a numberfield of degree $f$ over $\mathbb{Q}$ and let $\epsilon > 0$. Then

$$|X(F, T)| \leq c(\mathcal{X}, f, \epsilon)(\log T)^{44+\epsilon}.$$ 

13
Proof. It suffices to prove a bound of the stated form for \(|X|\) of \(F, T\). For each integer \(g > 1\) we have a \((J(g), A(g), 1 + 1/g)\)-mild parameterization (with \(J(g) = 1\)) of \(X\) given by

\[
\theta : (0, 1)^2 \to (0, 1)^3,
\]

\[
\theta(s, t) = \left( \exp \left( 1 - \frac{1}{s^g} \right), \exp \left( 1 - \frac{1}{t^g} \right), \exp \left( - \left( 1 - \frac{1}{s^g} \right) \left( 1 - \frac{1}{t^g} \right) \right) \right).
\]

By Theorem 3.2, \(|X|\) is contained in

\[
\leq c(g, f) \left( \log T \right)^{9(1+1/g)(1+o(1))}
\]

intersections of \(X\) with hypersurfaces of degree

\[
\left( \left( \log T \right)^2 \right)
\]

with the \(1 + o(1)\) as \(T \to \infty\) (and implicit constants depending only on \(g, f\)). These intersections all have dimension 1, since \(X\) is not semi-algebraic, and we may ignore any semi-algebraic components, as the semi-algebraic curves in \(X\) contain no algebraic points.

The mild parameterization plays no further role in the study of these hypersurface intersections. In applying the Gabrielov-Vorobjov bounds it is advantageous to define them as pfaffian sets with as low degree as possible. For the remainder of the proof we therefore consider \(X\) to be parameterized by

\[
(0, \infty)^2 \to (0, 1)^3,
\]

\[
(p, q) \mapsto (e^{-p}, e^{-q}, e^{-pq}) = (x, y, z) \in X.
\]

If \(H \in \mathbb{R}[x, y, z]\) defines the hypersurface \(V_H : H(x, y, z) = 0\) then the intersection \(X \cap V_H\) is the image of the exponential-algebraic curve (not necessarily connected) in the \((p, q)\)-plane defined by

\[
K(p, q) = H(e^{-p}, e^{-q}, e^{-pq}) = 0, \quad p, q > 0.
\]

We observe that, for \(H \neq 0\), the equation \(K(p, q) = 0\) defines a curve \(V = V_K\), i.e. a set of dimension 1, again because \(X\) is not semi-algebraic. The set of singular points \(V_s\) of \(V\) is defined by

\[
K = 0, \quad K_p = -H_x e^{-p} - qH_z e^{-pq} = 0, \quad K_q = -H_y e^{-q} - pH_z e^{-pq} = 0.
\]

It is a finite set (definable of dimension zero).

We now follow the procedure of [10, 11], substituting Gabrielov-Vorobjov bounds for the appeals made in [10, 11] to Gabrielov’s Theorem for subanalytic sets.

Let then \(\Pi\) be a coordinate plane in \(\mathbb{R}^3\) whose coordinates we denote \((u, v)\). Projection of \(\mathbb{R}^3\) onto \(\Pi\) takes the curve \(V\) defined by \(K(p, q) = 0\) into some curve in \(\Pi\). At a point \(P = (p, q)\) of \(V - V_s\), \(V\) is locally an analytic curve. If \(K_q \neq 0\) at \(P\) then we may use \(q\) as a local parameter and we find that \(u\) is nonconstant at \(P\) unless

\[
u_p K_q - u_q K_p = 0.
\]

Similarly, \(v\) is nonconstant at \(P\) unless

\[
u_p K_q - v_q K_p = 0.
\]

Let \(V_u\) be the subset of \(V - V_s\) where one or more of these quantities vanish. At points of \(V - V_s - V_u\) the slope \(du/dv\) is well defined, and the image of \(V\) in \(\Pi\) is locally the graph of a function. We proceed to derive an expression for its derivatives. We have, locally,

\[
u = u(p(v), q(v)), \quad v = v(p(v), q(v)), \quad K(p(v), q(v)) = 0.
\]
Differentiating the second and third equations implicitly,

$$1 = v_p p' + v_q q', \quad K_p p' + K_q q' = 0$$

which we may write as a matrix equation

$$\begin{pmatrix} v_p & v_q \\ K_p & K_q \end{pmatrix} \begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

giving

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \frac{1}{v_p K_q - v_q K_p} \begin{pmatrix} K_q \\ -v_p \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{v_p K_q - v_q K_p} \begin{pmatrix} K_q \\ -K_p \end{pmatrix}.$$ 

We have then

$$\frac{du}{dv} = \frac{u_p K_q - u_q K_p}{v_p K_q - v_q K_p}.$$ 

To get expressions for higher derivatives, we differentiate this expression with respect to $v$ and use the expressions we have for $p', q'$. For points $(u, v)$ with $v_p K_q - v_q K_p \neq 0$ and a positive integer $m$ we will have

$$\frac{d^m u}{dv^m} = \frac{R_m(u, v, K)}{(v_p K_q - v_q K_p)^{2m-1}}$$

for a suitable differential polynomial $R_m$.

We want to estimate the number of zeros of $R_m$ which we will do by controlling its order and degree as a pfaffian function. Let us write

$$\Delta = v_p K_q - v_q K_p$$

(no confusion should arise with the previous use of $\Delta$), which we consider as a function of $v$, so that

$$p' = \frac{K_q}{\Delta}, \quad q' = -\frac{K_p}{\Delta}$$

and

$$\Delta' = \frac{v_p K_q^2 - v_q K_p K_q + v_p K_{qp} K_q - v_p K_q K_p - v_{pp} K_q K_p + v_{pq} K_p^2 - v_q K_p K_q + v_q K_{pq} K_p}{\Delta} = \frac{\Gamma}{\Delta}.$$ 

If we now write

$$\frac{d^m u}{dv^m} = \frac{R_m}{\Delta^{2m-1}}, \quad R_m = \frac{S_m}{\Delta}$$

then

$$\frac{d^{m+1} u}{dv^{m+1}} = \frac{\Delta^{2m-2} S_m - (2m-1) \Delta^{2m-2} \frac{\Gamma}{\Delta} R_m}{\Delta^{4m-2}} = \frac{R_{m+1}}{\Delta^{2(m+1)-1}}$$

gives a recurrence for $R_m$ (and validates the asserted form for $d^m u/dv^m$), starting with

$$R_1 = u_p K_q - u_q K_p.$$ 

Consider the pfaffian chain of functions on $(0, \infty)^2$

$$f_1 = e^{-p}, f_2 = e^{-q}, f_3 = e^{-pq}$$

where we have $\partial_p f_3 = -q f_3, \partial_q f_3 = -p f_3$. This is then a pfaffian chain of order $r = 3$ and degree $\alpha = 2$. The function $u, v, K$ and their partial derivatives with respect to $p, q$ are pfaffian with this chain, i.e. they are polynomials in $p, q, f_1, f_2, f_3$, and therefore so are all the functions $R_m$ and $S_m$, and they therefore have order 3 and degree $(2, \beta)$, where $\beta \geq 1$ is their degree as a polynomial in $p, q, f_1, f_2, f_3$.

Claim. $u_\mu$ has degree $(2, |\mu| + 1)$. 

15
Proof of Claim. By induction. It holds for $|\mu| = 1$, the "worst case" being $u = f_3 = e^{-pq}$ for which $u_p = -q f_3$ is a polynomial of degree 2. Suppose the Claim is true for all $\mu$ with $|\mu| \leq m$. Then, with some polynomial $P$ of degree $\leq |\mu| + 1$,
\[
\partial_p \partial_u u = \partial_q P(p, q, f_1, f_2, f_3) = P_p + P_q f_1 + P_{q^2} f_2 + P_{q^3} f_3 =
\]
having degree $\leq |\mu| + 2$. The $\partial_q \partial_u$ calculation is similar. \(\square\)

If $H$ has degree $d$ then $K = H(f_1, f_2, f_3)$ is pfaffian with the chain $f_1, f_2, f_3$ and degree $(2, d)$. Generalizing the previous Claim we find:

Claim. $K_\mu$ has degree $(2, d + |\mu|)$. \(\square\)

Returning to our functions $R_m$ and $S_m$, we have that $R_1 = u_p K_q - u_q K_p$ has degree $(2, d + 3)$. Suppose $R_m$ has degree $(2, \rho_m)$, so $R_m = P(p, q, f_1, f_2, f_3)$ for suitable polynomial $P$ of degree $\leq \rho_m$. Then
\[
R'_m = P_p p' + P_q q' + P_{f_1} f'_1 + P_{f_2} f'_2 + P_{f_3} f'_3 =
\]
and therefore
\[
\rho_{m+1} = \max(d + 3 + \rho_m + d + 2, 2d + 5 + \rho_m) = \rho_m + 2d + 5
\]
and therefore
\[
\rho_m = m(2d + 5) - (d + 2).
\]

With these degrees in hand, we consider the decomposition of the curve $V_K$ defined by $K(p, q) = 0$ into "good" curves, where a "good" curve is a connected subset whose projection into each coordinate plane $\Pi$ is a "good" graph with respect to one or other of the axes, namely, the graph of a function $\phi$ which is smooth (indeed analytic) on an interval, has slope of absolute value at most 1 at each point, and such that the derivative of $\phi^{(m)}$ of each order $m = 1, \ldots, M$ is either non-vanishing in the interior of the interval or identically zero.

In the following, constants in $\ll$ depend on a pfaffian chain on a simple domain $G$. This will always be the chain $f_1, f_2, f_3$ of order 3 and degree 2 in the simple domain $p, q > 0$ in $\mathbb{R}^2$, so that the implicit constant is then absolute and explicit from Theorem 5.3.

The set $V = V_K$ has $\ll d^5$ connected components. Its singular set $V_s$ is defined by $K = 0, K_p = K_q = 0$ where $K_p, K_q$ have degree at most $d + 1$. So
\[
cc(V_s) \ll d^5
\]
and therefore also
\[
cc(V - V_s) \ll d^5.
\]
Let $V_u$ be the subset of $V - V_s$ where $du/dv$ is undefined. Considering the conditions exhibited above for such points, and also for the set $V_s$ where the slope of the graph in $\Pi$ is $\pm 1$ we have again
\[
cc(V - V_s - V_u - V_u) \ll d^5.
\]

Now take one such component, fix a coordinate plane $\Pi$, and consider the points where some $R_m = 0$. Since $\deg(R_m) \leq (2, (2d + 5)m)$, we have at most $m^5(2d + 5)^5$ points where $R_m = 0$, unless it vanishes
identically on the component. In this case the image in \( \Pi \) is the graph of a polynomial with respect to one of the axes. If the graph is not a polynomial than, summing over \( m = 1, 2, \ldots, M \), we have at most \( << M^6d^5 \) further components, whose slope lies in \([-1, 1]\), and for which no derivative up to order \( M \) vanishes. Taking the isolated points where some \( R_m = 0 \) for \( m = 1, 2, \ldots, M \) for each of the 3 coordinate planes \( \Pi \), we find that \( V_K \) decomposes into \( << M^6d^5 \) connected components whose image in each coordinate plane is a graph with respect to one of the axes with slope in \([-1, 1]\) and such that, for each \( m = 1, 2, \ldots, M \), \( R_m \) is nonzero in the interior or identically zero on the component, i.e. “good” components.

If such a connected component of \( V_K \) is semi-algebraic then its projection in each coordinate plane \( \Pi \) will be algebraic, and conversely if all the projections are semi-algebraic then the component is semi-algebraic. Now we need not consider algebraic components, therefore we can assume that every component has a non-algebraic (and hence non-polynomial) projection into one of the planes \( \Pi \).

Let \( W \) be a “good” component of \( V_K \), and \( Y \) its non-semi-algebraic image in some \( \Pi \). If we intersect \( Y \) with a plane algebraic curve (in \( \Pi \)) defined by \( L(u, v) = 0 \) of degree \( b \), then since the function \( L(u(p, q), v(p, q)) \) is pfaffian of degree \((2, b)\), intersecting with \( Y \) gives again at most 

\[ << (b, d)^5 \]

connected components. So \( Y \cap \{ L = 0 \} \) consists of at most this many isolated points.

Since \( Y \) is a “good” graph then, by [13] (for rational points) and [14, 6.7] (for \( F \)-points), \( Y^{\text{size}}(F, T) \) is contained in

\[ c(f)M \log T \]

plane algebraic curves of degree \( b \) where \( M = (b + 1)(b + 2)/2 \). So we get

\[ \#Y^{\text{size}}(F, T) \leq c(f) \max(b, d)^5 M \log T \]

and the same estimate holds for the corresponding component of \( V_H \), where having a point of \( X^{\text{size}}(F, T) \) requires that the other coordinate be also in \( F \) with its \( H^{\text{size}} \) bounded by \( T \).

Putting all the above together, we find

\[ \#X^{\text{size}}(F, T) \leq c(X, f, g) \left( \log T \right)^{9(1+1/\epsilon)(1+o(1))} M^6d^5 M \log T \max(b, d)^5 \]

where \( d = \left[ (\log T)^2 \right] \), \( M = (b + 1)(b + 2)/2 \), and \( b = \left[ \log T \right] \), giving

\[ \#X^{\text{size}}(F, T) \leq c(X, f, g) \left( \log T \right)^{9(1+1/\epsilon)(1+o(1)) + 35}. \]

This completes the proof of Theorem 6.1, and thereby establishes Theorems 1.1 and 1.6 as well. \( \square \)

7. Appendix: O-minimal structures

We give the basic definitions, following [22], referring the reader to [5, 6, 21, 22] for more information.

**7.1. 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).
7.2. Definition. [5, p3] We denote by $\mathbb{R}^{\exp}$ the prestructure consisting of those sets in $\mathbb{R}^n$ arising as the image under projection maps $\mathbb{R}^{n+k} \to \mathbb{R}^n$ of sets of the form \{$(x,y) \in \mathbb{R}^{n+k} : P(x,y,e_x,e_y) = 0$\}, where $P$ is a real polynomial in $2(n+k)$ variables, and where $x = (x_1,\ldots,x_n), y = (y_1,\ldots,y_k), e_x = (e^{x_1},\ldots,e^{x_n}), e_y = (e^{y_1},\ldots,e^{y_k})$.

7.3. Example. The set $X$ is the image under the projection $\mathbb{R}^6 \to \mathbb{R}^3$ of $Y = \{(x,y,z,u,v,w) : (x-e^u)^2 + (y-e^v)^2 + (z-e^w)^2 + (uv-w)^2 = 0\}$.

7.4. Theorem. (Wilkie [21]) $\mathbb{R}^{\exp}$ is an o-minimal structure. □

References
5. L. van den Dries, Tame topology and o-minimal structures, CUP 1998.

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

j.pila@bristol.ac.uk
Counting rational points 
on a certain exponential-algebraic surface

Jonathan Pila

Abstract

We study the distribution of rational points on a certain exponential-algebraic surface and we prove, for this surface, a conjecture of A. J. Wilkie.

2000 Mathematics Subject Classification: 11G99, 03C64
Keywords: O-minimal structure, rational points, transcendental numbers.

Running title:

Counting rational points

Comptage des points rationnels 
sur une surface exponentielle-algébrique particulière

Jonathan Pila

Résumé

Nous étudions la répartition des points rationnels sur une certaine surface exponentielle-algébrique et prouvons, pour cette surface, une conjecture de A. J. Wilkie.

Classification math.: 11G99, 03C64
Mots-clés: Structure o-minimale, points rationnels, nombres transcendants.