STRONGLY MINIMAL EXPANSIONS OF $(\mathbb{C},+)$ DEFINABLE IN O-MINIMAL FIELDS

ASSAF HASSON* AND PIOTR KOWALSKI♣,†

Abstract. We characterize those functions $f:\mathbb{C}\to\mathbb{C}$ definable in o-minimal expansions of the reals for which the structure $(\mathbb{C},+,f)$ is strongly minimal – such functions must be complex constructible, possibly after conjugating by a real matrix. In particular we prove a special case of the Zilber Dichotomy – an algebraically closed field is definable in certain strongly minimal structures which are definable in an o-minimal field.

1. Introduction

This paper is concerned with recognizing algebraicity in model theoretic terms. This line of research is motivated by a conjecture of Zilber in [28] suggesting that certain combinatorial geometries arising naturally in model theory must be either of linear type (locally modular) or coincide with the Zariski geometry coming from an (irreducible) algebraic curve over an algebraically closed field. Although in general false ([7]) Zilber's conjecture has been proved under additional assumptions in [10] and played a crucial role in several applications outside model theory, the first and arguably the best known of which is [8].

Aside from [10] there are several other instances under which Zilber's conjecture has been proved (usually, but not always, by reducing the problem to one close enough to the context treated in [10]). In [9] Hrushovski and Sokolovic prove the conjecture for differentially closed fields of characteristic 0; in [8] the same is done for separably closed fields, and in [4] for difference fields. In [24] Zilber's conjecture has been proved for those combinatorial geometries *interpretable* in algebraically closed fields (under the additional assumption that the interpretation is rank preserving).

Loosely speaking, these results suggest that Zilber's conjecture tends to be true in contexts of geometrical flavor. It seems therefore natural to look for a proof of Zilber's conjecture for *o-minimal* structures. This question has been formulated by Y. Peterzil as:

Problem 1.1. *Let D be a strongly minimal, non-locally modular structure interpretable in an o-minimal structure. Does D interpret an algebraically closed field?*

The combinatorial geometries referred to in Zilber's conjecture are precisely those associated with *strongly minimal structures*. In [19] Peterzil and Starchenko prove a complementary result. They show that if $\mathcal{R}$ is an o-minimal expansion of a real closed field and $\mathcal{K}=\mathcal{R}(\sqrt{-1})$ then $\mathcal{K}$ has no proper expansions which are strongly minimal (or stable). In the present paper we make a first step toward proving

*Supported by the EPSRC grant no. EP C52800X 1.
♣Supported by a MODNET (European Commission Research Training Network) grant.
†Supported by the Polish KBN grant no. 2P03A 018 24.

a generalization of Peterzil and Starchenko's result and give a positive answer to problem 1.1 in a special case. Our main result is:

Theorem 1.2. *Let $f : \mathbb{C} \to \mathbb{C}$ be definable in an o-minimal expansion of the reals and such that $\mathbb{C}_f := (\mathbb{C}; +, f)$ is strongly minimal and non-locally modular. Then there is $M \in \mathrm{GL}_2(\mathbb{R})$ such that $\mathbb{C}_{MfM^{-1}}$ is interdefinable with the complex field. In particular, $\mathbb{C}_f$ is biinterpretable with the complex field.*

Note that the assumptions of the theorem are a special case of the assumptions in 1.1, but our conclusion is sharper than the one in 1.1. Not only do we find a $\mathbb{C}_f$-definable field, but also show that the entire structure of $\mathbb{C}_f$ comes from that field. One can not expect such a strengthening of the conclusion of 1.1 in general, see the remark on page 9 of [29].

Proving Zilber's conjecture for o-minimal expansions of the additive group of the complex field reduces by standard model theoretic arguments to proving it for $(\mathbb{C}; +, X)$ for arbitrary strongly minimal $X \subseteq \mathbb{C}^2$ definable in an o-minimal expansion of the reals such that $(\mathbb{C}; +, X)$ is strongly minimal. Hence, our theorem deals with the special case where X is the graph of a function. Some additional comments regarding the assumptions of the theorem are in place. First, note that if f is conjugate to a complex rational function by a real matrix then $\mathbb{C}_f$ is a strongly minimal structure and by [14] is interdefinable with the complex field. However, examples due to Hrushovski (essentially along the lines of [7]) show that in general if $f : \mathbb{C} \to \mathbb{C}$ is a set theoretic function such that $\mathbb{C}_f$ is strongly minimal no field need be interpretable in the structure even if it is non-locally modular. In view of the above discussion the o-minimality assumption is quite natural. It is also clear that o-minimality is a natural assumption in the (less natural) case of an arbitrary real closed field, since the natural topology in such fields is usually not good enough for any arguments.

Generally speaking, our proof of Theorem 1.2 uses mild consequences of o-minimality, which suggests that considerable parts of the proof may be generalized to weaker contexts. One possible such context is weak o-minimality. Another one (perhaps more natural and challenging) would be to assume only that f is continuous (or C^1). For example, if we assume f to be holomorphic, we get an immediate proof, since strong minimality forbids any essential singularities (infinity included), hence f must be rational.

The idea for our proof is based on a theorem of Marker and Pillay in [13] where they prove - as a special case - the same result as that of our theorem under the assumption that f is definable in the complex field. A key feature in their proof is that f is meromorphic, hence

$$(*) \qquad f'(c) = 0 \iff f \text{ is not a local homeomorphism at } c.$$

The main part of the paper is aimed at proving $(*)$ in our setting. Once this is achieved we can, more or less, follow the steps of [13].

The paper is structured as follows. In Section 2 we give the (rather soft) model theoretic background we need and gather a few easy, though useful, facts concerning our function f. Section 4 is dedicated to the study of the topological properties of our function. We show that (up to finitely many corrections) f is a continuous, open ramified covering of the Riemann sphere. In order to prove these properties of f we use ideas of Peterzil and Starchenko (§3 in [19]) to show that strongly minimal subsets of $\mathbb{C}^2$ are closed (up to possibly finitely many corrections), which is the main result of Section 3. This readily gives us the continuity of f, its properness and its openness. In the second part of Section 4 we use ideas from topological analysis to conclude that f is a ramified covering. We then show that f admits a topological degree and the remainder of the section is dedicated to the investigation of the local behavior of this degree in definable families of definable functions. In Section 5 we extend yet further our use of topological analysis to obtain a weak version of the Cauchy-Riemann equation for f, namely the right-to-left implication in $(*)$: if f is not a local homeomorphism at c and c is a smooth point of f then $f'(c) = 0$. Finding such a point c turned out, unexpectedly, to be a major difficulty.

In Section 6 we prove a local version of the left-to-right implication in $(*)$. The proof makes extensive use of the results of Section 5, and goes through local Lie group theory and the classification of maximal solvable Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$. This allows us to return to the line of proof of [13] in order to produce in Section 7 a certain combinatorial configuration, which by a well known variant of the Hrushovski-Weil group configuration theorem assures the existence of an interpretable field in $\mathbb{C}_f$. Sections 8 and 9 are dedicated to generalizing the proof to o-minimal expansions of arbitrary real closed fields.

Acknowledgment: We would like to thank Y. Peterzil for suggesting the problem to us, and for his helpful remarks and ideas; G. Cherlin for directing us to some useful references; C. Mora-Corral for saving us from trying to re-invent degree theory and directing us to the right books; K. Dymara and J. Dymara for supplying us with examples which kept us off wrong tracks; and the Oxford Logic Group for many tea time discussions.

2. Model theoretic background and preliminaries

In this section we give the formal definitions required for an understanding of Problem 1.1 and the statement of Theorem 1.2. We also gather some basic facts concerning strongly minimal and o-minimal structures which we will use throughout the paper, and prove some immediate implications to our function f, which will be useful later on.

The key to all the definitions that follow are the notions of *structure* and *definability*. Roughly, a structure is a non empty set M with a collection $\mathcal{M}$ of distinguished functions and relations (possibly on powers of M) and a set $S \subseteq M^k$ is definable if it can be obtained from $\mathcal{M}$ by applying logical operations. More formally:

Definition 2.1. Let M be an non-empty set. A structure $\mathcal{S}$ on M is a sequence $\{S_n : n \geqslant 0\}$ of subsets of $\mathcal{P}(M^n)$ satisfying:

(1) Each S_n is a boolean algebra of sets (i.e. closed under finite unions intersections and under taking complements).
(2) $\mathcal{S}$ is closed under cartesian powers (i.e. If $A \in S_n$ and $B \in S_m$ then $A \times B \in S_{n+m}$).
(3) For all n the diagonal $\Delta^n_{i,j} \subseteq M^n$ given by $x_i = x_j$ is a set in S_n.
(4) $\mathcal{S}$ is closed under projections (i.e. if $A \in S_{n+1}$ and $\pi : S_{n+1} \to S_n$ is the projection map on the first n coordinates then $\pi(A) \in S_n$).

A set $A \subseteq M^n$ will be called $\emptyset-$*definable* if $A \in S_n$. A is *definable* if it is a fiber of some $\emptyset-$definable set. A function $f : M^n \to M^k$ is definable if its graph is.

For basic properties of structures we refer the reader to [25]. By the Tarski-Seidenberg theorem the field of real numbers equipped with the semi-algebraic sets is a structure which we will refer to as *the reals*. Similarly, by a theorem of Chevalley's the projection of an algebraically constructible set is constructible, whence any algebraically closed field $\mathcal{K}$ equipped with the $\mathcal{K}-$constructible sets is a structure. If $\mathcal{K} = \mathbb{C}$ we will call the resulting structure *the complex field*. To simplify the notation we will usually refer to "a structure $\mathcal{M}$" without denoting explicitly the pair $(M;\mathcal{S})$. With this notation, for a structure $\mathcal{M}$ the underlying set will be denoted M. We will have two exceptions, $\mathbb{R}$ and $\mathbb{C}$ will denote the (structures of) the real and complex fields, as well as their underlying sets. To get a better feeling of (non-)definability, it is worth remarking that the S_n appearing in the definition, are usually not σ-algebras. For example, the integers are *not* definable in neither the reals nor the complex field but any finite subset thereof is.

Given a non-empty set M and some $\mathcal{S} \subseteq \bigcup_{n\geqslant 0}\mathcal{P}(M^n)$ we denote $(M;\mathcal{S})$ the smallest structure in which every $S \in \mathcal{S}$ is definable. Note that if $\mathcal{S}$ is countable then the class of definable subsets in $(M;\mathcal{S})$ is countable. Identifying functions with their graphs, it is clear by what we have just said that, e.g., $\mathbb{R} = (\mathbb{R};+,\cdot)$.

To simplify the exposition in this section, it will be very useful for us to assume that we consider only such structures $(M;\mathcal{S})$ for which $\mathcal{S}$ is countable and M uncountable. This assumption is not reduce the generality of our discussion – the structure $\mathbb{C}_f$ clearly satisfies it and we can always replace the ambient structure on the reals by $(\mathbb{R},+,\cdot,f)$.

Whereas the notion of definability is essential to the understanding of the proofs in this papers, the remaining model theoretic concepts are - on the technical level - of less importance. Readers who feel uncomfortable with this sort of concepts can skip the next set of definitions and be content with Facts 2.8 and 2.9 which sum up all the consequences of the model theoretic assumptions which will be used in this paper (at least up to Section 8).

The following definition will not be of great importance in this paper, but since it appears in the statement of Problem 1.1 we give it nonetheless:

Definition 2.2. The structure $(N,\mathcal{R})$ is definable in the structure $(M,\mathcal{S})$ if there exist an injection $f : \mathcal{R} \to \mathcal{S}$ which preserves projections and boolean operations. $(M,\mathcal{S})$ is a an expansion of $(N,\mathcal{R})$ if $N = M$ and $\mathcal{S} \supseteq \mathcal{R}$.

We will be mostly interested in structures $(M,\mathcal{S})$ for which a structure is *interpretable* in $(M,\mathcal{S})$ if and only if it is definable there. To avoid technicalities we omit the definition of the former. Section 7 is the only place in the text where a short excursion into the interpretable, rather than the definable, realm seems unavoidable. Roughly, a structure $(N,\mathcal{R})$ is *interpretable* in the structure $(M,\mathcal{S})$ if it is definable in a (canonical) expansion of $(M,\mathcal{S})$ in which every class of every $\emptyset$-definable equivalence relation is given a name. We refer the reader to 1.1 of [21] or 4.3 of [6] for more details.

Note that if $(\mathbb{R},\mathcal{S})$ is any structure expanding the reals then the complex numbers are naturally definable in $(\mathbb{R},\mathcal{S})$. In fact, as in this example, in the present paper a structure $(N,\mathcal{R})$ will be definable in $(M,\mathcal{S})$ if and only if $N \in \mathcal{S}$ and $\mathcal{R} \subseteq \mathcal{S}$.

We are now able to introduce the two main model theoretic notions which will be used in this paper:

Definition 2.3. Let $\mathcal{M}$ be a structure:

(1) $\mathcal{M}$ is *minimal* if every definable subset of M is either finite or co-finite.
(2) $\mathcal{M}$ is *strongly minimal* if it is minimal and every definable family of finite subsets of M is uniformly bounded, i.e. for all definable $X \subset M^{n+1}$ such that for all $\bar{a} \in M^n$
$$X_{\bar{a}} := \{b \in M : (b, \bar{a}) \in X\}.$$
is finite, there exists $N \in \mathbb{N}$ such that $|X_{\bar{a}}| < N$ for all $\bar{a} \in M^n$.
(3) If $<$ is a dense linear definable ordering on M, we say that $(M, <)$ is *o-minimal* if the definable subsets of M are precisely those definable using the ordering alone (i.e. S_1 is precisely the boolean algebra generated by intervals and points in M).

Let $\mathcal{M}$ be any structure and $A \subseteq M$ any set. Define *the algebraic closure of* A to be the union of all finite sets definable with parameters from A. Formally,
$$\operatorname{acl}(A) := \{a \in M : (\exists \bar{b} \subseteq A, S \in S_{1+|\bar{b}|})((a, \bar{b}) \in S \wedge |S(x, \bar{b})| < \infty\}$$

Definition 2.4. Let $\mathcal{M}$ be a structure and $A \subseteq M$. $a \in M$ is *algebraic* over A if $a \in \operatorname{acl}(A)$.

Note that if, e.g. $\mathcal{M}$ is the reals or the complex field then a is algebraic over A in the above sense if and only if it is algebraic over A in the usual sense. In general if $\mathcal{M}$ expands the reals, say, then the model theoretic algebraic closure contains the standard algebraic closure. In this paper we will only be interested in the former notion, so no confusion can arise.

Since, when $\mathcal{M}$ is either a strongly minimal or an o-minimal structure, acl is a closure operator (in the sense of Steinitz) we can define for any $A \subseteq \mathcal{M}$:
$$\operatorname{rk} A := \min\{|a| : a \subseteq A \wedge \operatorname{acl}(a) \cap A = A\}$$
which gives a notion of independence in M in the following way:

Definition 2.5. Let (A, Cl) be a combinatorial pre-geometry and rk the associated rank.

(1) If $B, C \subseteq A$ are finite we write $\operatorname{rk}(C/B) = \operatorname{rk}(BC) - \operatorname{rk}(B)$. For arbitrary B we define $\operatorname{rk}(C/B) = \min\{\operatorname{rk}(C/B') \mid B' \subseteq B \text{ finite}\}$.
(2) $B, C \subseteq A$ are *independent* over $D \subseteq A$ if $\operatorname{rk}(C/BD) = \operatorname{rk}(C/D)$.
(3) (A, Cl) is *locally modular*, if there is a finite set $F \subset A$ such that for each pair of finite sets $B, C \subset A$, $\mathrm{Cl}(B \cup F)$ is independent from $\mathrm{Cl}(C \cup F)$ over $\mathrm{Cl}(B \cup F) \cap \mathrm{Cl}(C \cup F)$.

Definition 2.6. We assume now that a combinatorial geometry comes from a strongly minimal structure or an o-minimal structure.

(1) A strongly minimal structure is *locally modular* if its associated pre-geometry is.
(2) If X is a D-definable set we define
$$\operatorname{rk}(X) = \max\{\operatorname{rk}(d/D) \mid d \in X\}.$$
(3) For a D-definable set X and $D \subseteq D'$ an element $\bar{d} \in X$ is *D'-generic* if $\operatorname{rk}(\bar{d}/D) = \operatorname{rk}(X)$.
(4) For definable sets X, Y we denote $X \sim Y$ if $\operatorname{rk}(X \triangle Y) < \operatorname{rk}(X \cup Y)$.
(5) If $\mathcal{P}$ is any property (not necessarily definable) a definable set X *almost satisfies* property $\mathcal{P}$ if there exists a definable Y such that $X \sim Y$ and Y satisfies property $\mathcal{P}$.

(6) If $\mathcal{M}$ is strongly minimal rk is known as *Morley rank* and is denoted RM.
(7) If $\mathcal{M}$ is o-minimal rk is known as *o-minimal dimension* and is denoted dim.
(8) A *strongly minimal set* in a strongly minimal structure $\mathcal{M}$ is a definable set X with $\mathrm{RM}(X) = 1$ such that for every definable $Y \subseteq X$ either Y is finite or $X \sim Y$.

Remark 2.7. We collect some observations regarding the above definitions.

(1) Throughout this paper our usage of the term "almost" will be reserved exclusively to the one given in the above definition. Thus, an almost continuous (definable) function is one all of whose discontinuities are removable (and not, e.g., one which is continuous outside a set of measure 0).
(2) For a finite set X, a possible confusion may arise – when we write e.g. $\dim(X)$ it could mean either its dimension as a definable set (which is always 0) or its rank as a tuple (which is usually not 0). To prevent this confusion, from now on we distinguish between finite tuples, denoted by lowercase letters $a, b, c, \ldots$ and definable (possibly finite) sets denoted by uppercase letters $A, B, \ldots, X, Y, \ldots$.
(3) Usually, the Morley rank and o-minimal dimension for definable sets are defined otherwise, using properties of definable sets rather then the combinatorial geometry on the universe (see, e.g. [3] and [12], [25] respectively). Then it is shown (see same references) that in particular under our assumptions (uncountable structure, countably many $\emptyset$-definable sets) our definition coincides with the standard one.
Our countability assumption is also used for 2.8(5), 2.8(6) and 2.9(2) below.
(4) It may be worth mentioning that locally modular strongly minimal structures (in any context) are well characterized as, essentially, covers of linear spaces. Hence the assumption of non-local modularity is necessary in Problem 1.1.

We will now collect those facts concerning o-minimal structures which we will use. For a more detailed discussion (and for the proofs) we refer to [12] and [25].

Fact 2.8. *Let $\mathcal{M}$ be an o-minimal expansion of a real closed field, $f : M \to M$ a (partial) definable function then:*

(1) *M can be partitioned into finitely many open intervals $I_0, \ldots, I_k$ and finitely many points $a_1, \ldots, a_k$ such that for each $0 \leq i \leq k$ one of the following holds:*
  (a) *f is not defined on I_i.*
  (b) *f is constant on I_i.*
  (c) *f is monotone on I_i.*
(2) *The limit $\lim_{t \to 1^-} f(t)$ exists in $M \cup \{\pm\infty\}$.*
(3) *For every $n, m, k \in \mathbb{N}$, any definable $g : M^m \to M^k$, there is $X \subseteq M^m$ such that g is C^n on X and* $\dim(M^m \setminus X) \leqslant m - 1$.
(4) *If $X \subseteq \mathbb{R}^n \times \mathbb{R}^m$ is a definable family of definable sets parameterized by $Y \subseteq \mathbb{R}^m$, there exists a definable function $f : Y \to X$ (called* Skolem function*) such that $(f(d), d) \in X$ for all $d \in Y$.*
(5) *If X is a definable set of dimension n, then for any definable $Y \subseteq X$,* $\dim(Y) = n$ *if and only if Y contains a relatively open subset of X if and only if Y contains a generic point of X. In particular if M is the field of real numbers* dim *and the topological dimension coincide.*

(6) *If* $\dim(a/A) = n$*, then there is an* A*-definable,* n*-dimensional set* X *such that* $a \in X$*. If* $a \in Y$*, where* Y *is* A*-definable, then we can find such* X *with* $X \subseteq Y$*.*

The properties of strongly minimal sets we will be using are considerably simpler and will consist mainly in:

Fact 2.9. *Let* X *be definable in a strongly minimal structure* $\mathcal{M}$ *and assume that* X *(with induced structure) is strongly minimal.*

(1) *Let* Y *be a strongly minimal set in* $\mathcal{M}$ *such that* $X \cap Y$ *is infinite, then* $X \sim Y$*.*
(2) *If* $a_1, a_2 \in X$ *are such that* $\mathrm{RM}(a_1/A) = \mathrm{RM}(a_2/A) = 1$ *for some parameter set* A *then* $a_1 \in Y \iff a_2 \in Y$ *for every* $Y \subseteq X$ *definable over* A*.*
(3) *For any definable* $X \subseteq M^{n+m}$*,* $a \in M^n$*, let* $X_a := \{b \in \mathbb{M}^m \mid (a,b) \in X\}$*. Then the set*
$$\{a \in M^n \mid X_a \text{ is finite (resp. infinite)}\}$$
is definable.

From now till Section 8 we fix $\mathcal{R}$, an o-minimal expansion of the reals. We identify $\mathbb{C}$ with $\mathbb{R}^2$ and fix a function $f : \mathbb{C} \to \mathbb{C}$ definable in $\mathcal{R}$ such that the reduct $\mathbb{C}_f := (\mathbb{C}; +, f)$ is strongly minimal and non-locally modular. For simplicity of notation we will assume that f is $\emptyset-$definable in $\mathcal{R}$. Note that $\dim(c) \leq 2\,\mathrm{RM}(c)$ for every $c \in \mathbb{C}^n$ and that if X is a $\mathbb{C}_f-$definable set then $\dim(X) = 2\,\mathrm{RM}(X)$.

We will now fix some (pretty standard) notation.

Notation 2.10. Assume X is a topological space, $Y \subseteq X$ and $g : \mathbb{C} \to \mathbb{C}$.

(1) We will write $g'(c)$ for the Jacobian matrix of g at c (if it exists).
(2) $\mathrm{cl}(Y)$ is the topological closure of Y in X, $\mathrm{int}(Y)$ is its interior.
(3) $\partial(Y) := \mathrm{cl}(Y) \setminus \mathrm{int}(Y)$ is the border of Y, $\mathrm{fr}(Y) := \mathrm{cl}(Y) \setminus Y$ its frontier.
(4) For $x \in \mathbb{C}^n$, $\|x\|$ denotes the norm of x and for $r > 0$, $B_r(x)$ denotes the open ball centered at x and of radius r.
(5) For $a \in \mathbb{R}^n$ and $\mathcal{R}$-definable $X, Y \subset \mathbb{R}^n$ we say that X is transversal to Y at a, denoted $X \pitchfork_a Y$, if $a \in X \cap Y$ and there is an open $U \ni a$ such that $X \cap U, Y \cap U$ are C^1-submanifolds of $\mathbb{R}^n$ which are transversal at a.
(6) For a field K, $\mathbb{G}_m(K)$ is the multiplicative group of K and $\mathbb{G}_a(K)$ is the additive group of K

We finish this section with some easy consequences of strong minimality.

Fact 2.11. f *is not almost* $\mathbb{R}$*-affine.*

Proof. If f is $\mathbb{R}$-affine then the structure $\mathbb{C}_f$ is definable in the structure $\mathbb{R}_{\mathrm{vect}}$, whose universe is $\mathbb{R}$ and whose $\emptyset$-definable subsets of $\mathbb{R}^n$ are $\mathbb{R}$-linear subspaces. Then $\mathbb{R}_{\mathrm{vect}}$ is locally modular (see 2.7(4)). But then $\mathbb{C}_f$ is also locally modular, as is any strongly minimal structure definable in a locally modular one, see 6.3 on p. 182 of [21]. □

By strong minimality we know that f has finite fibers (otherwise it would be almost constant) and the size of fibers is uniformly bounded. Indeed, if $a \in \mathbb{C}$ is generic then $d := |f^{-1}(a)| = |f^{-1}(a')|$ for all but finitely many $a' \in \mathbb{C}$. We call d the *degree* of f. We will show in Section 4 that f admits a topological degree, and that the two notions agree up to sign (a posteriori, having proved the main theorem, we know that they agree unconditionally).

Fact 2.12. *Assume $U \subseteq \mathbb{C}$ is open and f is C^1 on U. Let X denotes the set of critical points of f (i.e. those $c \in \mathbb{C}$ such that $\det(f'(c)) = 0$) on U. Then $\dim(X) \leqslant 1$.*

Proof. If not, we get by 2.8(5) an open $V \subseteq U$ such that for each $c \in V$, f is critical at c. But then, $\dim(f(V)) \leqslant 1$, hence f has an infinite fiber. Therefore, f is almost constant and $\mathbb{C}_f$ is locally modular, a contradiction. □

One of the most useful characterizations of non-locally modular strongly minimal structures is the existence of a large definable family of essentially distinct plane curves. In the next lemma we show that in $\mathbb{C}_f$ we have a natural choice of such a family:

Lemma 2.13. *For each $a \in \mathbb{C} \setminus \{0\}$ and each open $U \subseteq \mathbb{C}$, $f(x+a) - f(x)$ is not constant on U.*

Proof. Assume there existed $b \in \mathbb{C}$ such that $f(x+a) - f(x) = b$ on U. Strong minimality would imply $f(x+a) - f(x) = b$ for almost all $x \in \mathbb{C}$. So for almost all x we have

$$f(x+2a) - f(x) = f(x+a+a) - f(x+a) + f(x+a) - f(x) = 2b.$$

Inductively, for each $n \in \mathbb{N}$, $f(x+na) - f(x) = nb$ for almost all $x \in \mathbb{C}$. So by 2.9(3), there is a cofinite set $X \subseteq \mathbb{C}$ such that for all $c \in X$ there exists $l(c)$ such that $f(x+c) - f(x) = l(c)$ for almost all $x \in \mathbb{C}$.
Claim l is an almost $\mathbb{R}$-linear $\mathbb{C}_f$-definable function.

Proof of Claim. By the uniformity part in the definition of strong minimality, l is a $\mathbb{C}_f$-definable function from X to $\mathbb{C}$. Take any $c \in X$ and $d \in \mathbb{C}$ such that $\mathrm{RM}(d/c) = 1$. Then $\mathrm{RM}(d+c/c) = \mathrm{RM}(d/c) = 1$, so $d+c, d \in X$ and we have for almost all $x \in \mathbb{C}$:

$$l(d+c) = f(x+d+c) - f(x) = f(x+d+c) - f(x+d) + f(x+d) - f(x) = l(c) + l(d).$$

Therefore, l is "generically-additive", but we still need to find an actual additive function coinciding almost everywhere with l. Let L denote the graph of l and consider the following set:

$$S := \{(v,w) \in \mathbb{C}^2 \mid ((v,w) + L) \sim L\}.$$

By 2.9(3), S is $\mathbb{C}_f$-definable and it is clearly a subgroup of $(\mathbb{C},+) \times (\mathbb{C},+)$.
Let us take $(v,w) \in S$ such that $v \in X$ and choose $a \in \mathbb{C}$ generic over $\{v,w\}$. Then $(v,w) + (a, l(a)) \in L$ and $v + a \in X$. Therefore,

$$l(v) + l(a) = l(v+a) = l(a) + w,$$

so $(v,w) \in L$. Therefore, S is the graph of an additive function $\bar{l}$ which coincides with l almost everywhere.
By 2.8(3), there is an $\mathcal{R}$-definable, open $U \subseteq \mathbb{C}$ such that $\dim(\mathbb{C} \setminus U) \leqslant 1$ and $\bar{l}$ is continuous on U. For any $a \in \mathbb{C}$, if we take $b \in U$ such that $\dim(b/a) = 2$, then $\dim(b+a/a) = 2$, so $b + a \in U$ by 2.8(5). Hence $U + U = \mathbb{C}$. Now, it is an easy exercise (left to the reader) to show that an additive function on $\mathbb{C}$ which is continuous on an open set $U \subseteq \mathbb{C}$ satisfying $U + U = \mathbb{C}$ is $\mathbb{R}$-linear. □

By the Claim we can assume that l is $\mathbb{R}$-linear. For a generic $a \in \mathbb{C}$ and for almost all $x \in \mathbb{C}$, we have:

$$f(a+x) - l(x) = f(a).$$

Therefore, f is almost $\mathbb{R}$-affine, contradicting 2.13. $\square$

Let $l_{a,c}$ denote the graph of $f(x+a)+c$. We get 4 corollaries (the first one is immediate) which will be of use:

Corollary 2.14. *$\mathcal{F} := \{l_{a,c} : a, c \in \mathbb{C}\}$ is a normalized 2-dimensional family of plane curves, i.e. $(a,c) \neq (a',c')$ implies $l_{a,c} \nsim l_{a',c'}$.*

Corollary 2.15. *Assume $S \subset \mathbb{C}^2$ is strongly minimal, definable over $\emptyset$ and $a, c \in \mathbb{C}$ such that $\mathrm{RM}(a,c) > 0$. Then $S \nsim l_{a,c}$.*

Proof. Assume not and take

$$X := \{(x,y) \in \mathbb{C}^2 \mid S \sim l_{x,y}\}.$$

By 2.9(3), X is $\mathbb{C}_f$-definable over $\emptyset$. Since $(a,c) \in X$ and $\mathrm{RM}(a,c) > 0$, X is infinite. In particular, there are $(a,c) \neq (a',c')$ with $l_{a,c} \sim S, l_{a',c'} \sim S$. Hence $l_{a,c} \sim l_{a',c'}$ contradicting 2.14. $\square$

Corollary 2.16. *For any open $U \subseteq \mathbb{C}$, $\dim(f'(U)) \in \{1,2\}$.*

Proof. Since f' is $\mathcal{R}$-definable and $\dim(U) = 2$, $\dim(f'(U)) \leqslant 2$. If $f'(U)$ is finite, then without loss f' is constant on U. For simplicity, we may assume that $0 \in U$, so choose any $c \in U$ and consider the function $f_c(x) := f(x) - f(x+c)$. We get an open $V \subseteq U$ such that $f_c|_V$ is constant, which contradicts 2.13. $\square$

Corollary 2.17. *For any open $U \subseteq \mathbb{C}$, no partial derivative of a coordinate function of f is constantly 0 on U.*

Proof. Assume not. Let $f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$. Without loss there is $\alpha \in \mathbb{R}$ such that

$$\frac{\partial f_1}{\partial x}(a) = \alpha \frac{\partial f_1}{\partial y}(a), \quad \forall a \in U.$$

Take an open $U' \subset U$ and $\varepsilon > 0$ such that

$$U' + B_{2\varepsilon}(0) \subseteq U.$$

For any $(v,w) \in U'$ consider the function

$$h_{v,w} : (-2\varepsilon, 2\varepsilon) \to \mathbb{R}, \quad h_{v,w}(\gamma) = f_1(v+\gamma, w-\alpha\gamma).$$

By the Chain Rule, $h'_{v,w}$ is constantly 0, so $h_{v,w}$ is constant.
Consider now the following $\mathbb{C}_f$-definable function

$$f_{\alpha,\varepsilon}(x,y) := f(x+\varepsilon, y-\alpha\varepsilon) - f(x,y).$$

and let $g : \mathbb{C} \to \mathbb{R}$ be its first coordinate function.
Take $(v,w) \in U'$. Since $h_{v,w}$ is constant, we have

$$g(v,w) = h_{v,w}(\epsilon) - h_{v,w}(0) = 0.$$

Therefore, g is constant on U', hence

$$\dim(f_{\alpha,\varepsilon}(U')) \leqslant 1.$$

Therefore, $f_{\alpha,\varepsilon}$ has an infinite fiber, so it is almost constant contradicting 2.13. $\square$

3. TRANSVERSALITY

In this section we prove that strongly minimal subsets of $\mathbb{C}^2$ are almost closed. The idea of the proof is taken from [19] where for a strongly minimal $X \subset \mathbb{C}$ the finiteness of its frontier is proved through an investigation of the possible ways X can meet complex lines. In the present context we do not have enough $\mathbb{C}_f$-definable complex lines, so we replace them with the family of plane curves, $\mathcal{F}$, introduced in the previous section. The main problem in the proof is showing that for any fixed strongly minimal $X \subset \mathbb{C}$ there are enough plane curves in our family which intersect X at a smooth point (of the curve and X).

Consider a family $\{f_a^b : \mathbb{C} \to \mathbb{C}\}_{a\in\mathbb{C}}$ of $\mathbb{C}_f$-definable functions such that for each $a \neq a' \in \mathbb{C}$ we have $f_a^b \not\sim f_{a'}^b$ and $f_a^b(b_1) = b_2$, where $b = (b_1, b_2) \in \mathbb{C}^2$ is fixed. We call $\mathrm{graph}(f_a^b)$ the *f-line through b parameterized by a* and denoted it l_a^b. We will usually denote the graph of a function g by Γ_g.

Consider the relation

$$\Phi^b(s_1, s_2; a) \Leftrightarrow f_a^b(s_1) = s_2.$$

We say that a point $s = (s_1, s_2) \in \mathbb{C}^2$ is *b-good* if $\Phi^b(s_1, s_2; x) \neq \emptyset$ and for each $a \in \mathbb{C}$ such that $\Phi^b(s_1, s_2; a)$, there is a neighborhood $U \ni (s, a)$ such that $\Phi^b \cap U$ is the graph of a C^1-function which is a submersion at s. In other words, if s is a b-good point, we have a C^1 function $\phi : U \to \mathbb{C}$ such that for any s' close enough to s, $\phi(s')$ is a parameter for an f-line through b and s'. We say that a point is *bad* if it is not good. Let B^b denote the set of all bad points for our family.

We will now investigate the set of bad points of the the family

$$f_a^b(x) := f(a + x) - f(a + b_1) + b_2.$$

Lemma 3.1. *For each b, $\dim(B^b) \leqslant 3$. Moreover, we can assume that*

$$B^{(b_1,b_2+b_2')} = B^{(b_1,b_2)} + (0, b_2').$$

Proof. Let $E := \{(s_1, s_2) \in \mathbb{C}^2 : \Phi^b(s_1, s_2, x) = \emptyset\}$. By strong minimality and Lemma 2.13 $\dim E \leq 2$. It will, therefore, be enough to show that for almost all $s_1 \in \mathbb{C}$ the set

$$B_{s_1}^b := \{s_2 \in \mathbb{C} \mid (s_1, s_2) \in B^b \setminus E\}$$

is at most 1-dimensional.
Fix s_1. Consider the function

$$H : \mathbb{C}^3 \to \mathbb{C}, \quad H(s_1, s_2, a) = f(s_1 + a) - f(a + b_1) + b_2 - s_2.$$

Let A be the set of $a \in \mathbb{C}$ such that there is an open $U \ni (s_1, s_2, a)$ such that $H|_U$ is C^1. By Fact 2.8(4) A is of codimension 1 at most (note that this does not depend on s_2). By the implicit function theorem, (s_1, s_2) is good if for each a such that $s_2 = f(a + s_1) - f(a + b_1) + b_2$, we have $a \in A$ and

$$0 \neq \det(\frac{\partial}{\partial a} H(s_1, s_2, a)) = \det(f'(s_1 + a) - f'(b_1 + a)).$$

Consider the function $g(x) := f(s_1 + x) - f(b_1 + x) + b_2$. We have just proved that

$$B_{s_1}^b \subseteq g(\mathbb{C} \setminus A) \cup g(\{a \in A \mid \det(g'(a)) = 0)\}).$$

Since the latter set has dimension at most 1 (except the case $s_1 = 0$, which we can ignore), the result follows.
The moreover part is given by the form of g and E. $\square$

Now, given a strongly minimal $S \subset \mathbb{C}^2$, and $b \in \mathbb{C}^2$ we cannot hope to be able to say much on the possible intersections of S with f-lines through b, if S meets B^b in a large set. Therefore, we first move S so that its intersection with B^b is of small dimension for a suitable choice of b. For any $a \in \mathbb{C}$, let us define

$$F_a : \mathbb{C}^2 \to \mathbb{C}^2, \quad F_a(x,y) = (x, f(y+a)).$$

Lemma 3.2. *Assume $S \subset \mathbb{C}^2$ is strongly minimal and with infinite frontier. Then, there are $c \in \mathrm{fr}(S)$ and $a \in \mathbb{C}$ with $\dim(a,c) = 3$ such that $\dim(F_a(S) \cap B^{F_a(c)}) \leqslant 1$ and $F_a(c) \in \mathrm{fr}(F_a(S))$.*

Proof. We may assume that S is $\emptyset$-definable. Take $c = (c_1, c_2) \in \mathrm{fr}(S)$ generic (so $\dim(c) = 1$) and let

$$A := \{a \in \mathbb{C} \mid F_a(c) \in F_a(S)\},$$

Claim: A is finite.

Proof of Claim. Assume not. Let S_2 denotes the section of S over c_2. Since S is not linear, S_2 is finite. Since $c \notin S$, $c_1 \notin S_2$.
We get that for each $a \in A$, there is $d \in S_2$ such that

$$f(a + c_1) - f(a) = f(a + d) - f(a).$$

Therefore, there is a fixed $d \in S_2$ such that for an infinite set X the functions $f(x+c_1)$ and $f(x+d)$ coincide on X, which is impossible by 2.13, since $c_1 \neq d$. $\square$

Let

$$Y := \{a \in \mathbb{C} \setminus A \mid F_a \text{ is } C^1 \text{ at } c \text{ and } \det(F_a'(c)) \neq 0\}.$$

Y is clearly of codimension at most 1.
Since $F_a(c) = (c_1, f(a + c_2))$, we have by the moreover part of 3.1

$$\begin{aligned}\dim(F_a(S) \cap B^{F_a(c)}) &= \dim[F_a(S) \cap (B^c + (0, f(a+c_2) - c_2))] \\ &= \dim[(F_a(S) - (0, f(a+c_2) - c_2)) \cap B^c].\end{aligned}$$

Consider the function

$$\bar{F}_a(x,y) = (x, f(a+y) - f(a+c_2) + c_2).$$

We clearly have

$$\bar{F}_a(S) = F_a(S) - (0, f(a+c_2) - c_2).$$

It is easy to check that $\bar{F}_c(S) \cap \bar{F}_d(S)$ is finite for all $c \neq d$, hence

$$\dim\{a \in Y \mid \dim(\bar{F}_a(S) \cap B^{(c)}) = 2\} \leqslant 1$$

(otherwise, we would have a 2-dimensional family $(\bar{F}_a(S) \cap B^c)_{a \in Y_0}$ of 2-dimensional sets inside a 3-dimensional set $B^{F_a(c)}$ yielding some infinite intersection of $\bar{F}_{a_1}(S)$ and $\bar{F}_{a_2}(S)$).
Take $a \in Y$ such that $\dim(a,c) = 3$ and $\dim(F_a(S) \cap B^c) \leqslant 1$.
By the definition of Y and the inverse function theorem, there is an open set $U \ni c$ such that $F_a|_U$ is a diffeomorphism. For a topological space Z, the following is trivial:

$(*)$ For any $C \subset D \subset Z, z \in D \setminus C$, $z \in \mathrm{fr}(C)$ if and only if $z \in \mathrm{fr}_D(C)$.

Therefore, since $c \in \mathrm{fr}(S)$ and $c \in U$, we get $c \in \mathrm{fr}_U(S \cap U)$. Since $F_a|_U$ is a homeomorphism, we get

$$F_a(c) \in \mathrm{fr}_{F_a(U)}(F_a(S \cap U)).$$

Since $c \notin A$, we get $F_a(c) \notin F_a(S \cap U)$. We obtain by $(*)$ that $F_a(c) \in \mathrm{fr}(F_a(S))$. $\square$

Theorem 3.3. *Let $S \subset \mathbb{C}^2$ be strongly minimal. Then $\mathrm{fr}(S)$ is finite.*

Proof. Assume not. Without loss S is defined over $\emptyset$. Take a, c as given by 3.2 and let $b := F_a(c)$ (F_a from the statement of 3.2). Since F_a has finite fibers, $\mathrm{acl}(a, b) = \mathrm{acl}(a, c)$, so $\dim(b/a) = 1$. Replacing S with $F_a(S)$ we may assume by 3.2 that $b \in \mathrm{fr}(S)$, $\dim(b) = 1$ and $\dim(S \cap B^b) \leqslant 1$.
Consider the f-line l_a^b through b. Recall that l_a^b is the graph of the function $f_a^b(x) := f(a + x) + c_a^b$ ($c_a^b = -f(a + b_1) + b_2$, but it will not be important for us).
Claim 1

$$\mathrm{RM}(c_a^b, a) = 2.$$

Proof of Claim 1. Note that $l_a^b \cap \mathrm{fr}(S)$ is finite, since otherwise, we would have an open $U \subseteq \mathbb{C}$ and a 2-dimensional family of infinite sets $(l_a^b \cap \mathrm{fr}(S))_{a \in \mathbb{C}}$ inside a 1-dimensional set $\mathrm{fr}(S)$, yielding infinite intersection of some f-lines.
Since $\mathrm{fr}(S)$ is $\emptyset$-definable, we get that $\dim(b/c_a^b, a) = 0$. If $\mathrm{RM}(c_a^b, a) = 1$, then in particular $\dim(c_a^b/a) = 0$, so $\dim(b/a) = 0$, which contradicts the choice of a. $\square$

Take any $s \in S \cap l_a^b$. Note that $s \in \mathrm{acl}(a, b)$ because $S \cap l_a^b$ is finite and definable over a, b. On the other hand, there are finitely many f-lines through b going through s. Therefore, $a \in \mathrm{acl}(s, b)$. Hence

$$\dim(s/b) = \dim(a/b) = 2.$$

Claim 2

$$S \pitchfork_s l_a^b$$

(in the sense of 2.10).

Proof of Claim 2. Since $s \notin B^b$, there is $U \ni s$ and a C^1 parameter choice function $\Phi : U \to \mathbb{C}$ such that $\Phi(s) = a$. By the definition of Φ, $\Phi^{-1}(a) \subseteq l_a^b$. Let $V := T_s(l_a^b) \cap T_s(S)$.
Assume S is not transversal to l_a^b at s, i.e. $\dim V \geqslant 1$. Since Φ is a submersion at s, $\Phi^{-1}(a)$ is smooth and 2-dimensional around s (so it locally coincides with l_a^b around s). We have

$$\ker \Phi'(s) = T_s(\Phi^{-1}(s)) = T_s(l_a^b).$$

Therefore, there is a real curve $\gamma : \mathbb{R} \to S$ such that $\gamma(0) = s$ and $\gamma'(0) \in V$, $\Phi \circ \gamma$ is constant on U_0 a neighborhood of $0 \in \mathbb{R}$. Therefore, Φ is constantly a on $\gamma(U_0)$ which is infinite. Hence l_a^b intersects S at infinitely many points. By 2.15, $\mathrm{RM}(a, c_a^b) = 0$ which contradicts Claim 1. $\square$

We can now conclude as in [19]. Consider a subfamily of the family of f-lines from 2.14

$$\mathcal{F} := (l_{a,c})_{c \in \mathbb{C}}, \quad l_{a,c} := \Gamma_{f(a+x)+c}.$$

Which are merely translations of l_a^b (so there is no fixed point now through which they all pass). By Claim 1, l_a^b is a $\mathbb{C}_f$-generic f-line in $\mathcal{F}$. Therefore, there are cofinitely many $c \in \mathbb{C}$ such that

$$n := |l_a^b \cap S| = |l_{a,c} \cap S|.$$

Consider now $l_a^b \cap S = \{s_1, \ldots, s_n\}$. By Claim 2, there is $\varepsilon > 0$ such that if $|c' - c| < \varepsilon$, then $l_{a,c'}$ still intersects $l_{b'}$ near each s_i. But there are infinitely many

c' as above such that $l_{c'}$ intersects S near b as well (since $b \in \mathrm{fr}(S)$), contradicting the choice of n. □

4. TOPOLOGY

4.1. **The topological properties of f.** Because even rational functions are only continuous on the Riemann sphere, this will be the natural domain for us to work in. We will use the finiteness of the frontier to show that allowing the value ∞, f is an almost continuous, proper open map.

The key to the proof is Theorem 3.3 which will be used repeatedly to obtain various finiteness results. First we apply the theorem to show that f is locally injective on a cofinite set.

Definition 4.1. We say that c is a *non-injective point* of f if for any open $U \ni c$, $f|_U$ is not injective. Let $\mathcal{N}(f)$ denote the set of non-injective points of f.

Fact 4.2. *$\mathcal{N}(f)$ is finite.*

Proof. Let

$$S := \{(x, y) \in \mathbb{C}^2 : f(x) = f(y) \wedge x \neq y\}.$$

Clearly, $\mathrm{RM}(S) = 1$ and $\mathrm{fr}(S)$ is contained in the diagonal. For $c \in \mathbb{C}$, we have $(c, c) \in \mathrm{fr}(S)$ if and only if $c \in \mathcal{N}(f)$. But $\mathrm{fr}(S)$ is finite by 3.3, so we are done. □

Let S^2 denote the 2-sphere. We view it as a 1-point compactification of $\mathbb{C}$, so $S^2 = \mathbb{C} \cup \{\infty\}$. Our first task is to show that $\infty \in S^2$ can not be obtained as a limit infinitely often. Let us define

$$L := \{a \in \mathbb{C} \mid \exists a_n \to a,\ \lim(f(a_n)) = \infty\}.$$

We need a lemma which gives a partial description of the topological properties of f.

Lemma 4.3. *There is a finite $F \subset \mathbb{C}$ such that f is continuous and open on $\mathbb{C} \setminus (L \cup F)$.*

Proof. Let F be the union of $\mathcal{N}(f)$ and the projection to the first coordinate of $\mathrm{fr}(\Gamma_f)$. By 3.3 and 4.2, F is finite. It is clear from our choices that f is continuous and locally injective on $\mathbb{C} \setminus (L \cup F)$.
By Brouwer's Invariance of Domain Theorem [2], f is open on $\mathbb{C} \setminus (L \cup F)$. □

Fact 4.4. *L is finite.*

Proof. The proof is similar to that of Theorem 3.3 but considerably simpler. Assume L is infinite. Then, L is clearly $\mathcal{R}$-definable and by 2.8(3) of dimension 1. By 2.8(3) again, $f|_L$ is continuous on a cofinite subset of L, so we may assume that $f|_L$ is continuous everywhere. We will also assume $L \cap F = \emptyset$, where F is as in the conclusion of 4.3.
Claim 1: There are $a, a' \in L$ and $b \in \mathbb{C}$ such that $\dim(a, a', b) = 4$ and for each $\varepsilon > 0$:

$$f(B_\varepsilon(a)) \cap (f(B_\varepsilon(a')) + b) \text{ is unbounded.}$$

Proof of Claim 1. Let $N > 0$ and

$$U_{a,b} := \mathrm{int}[(f(B_\varepsilon(a)) + b) \cap (\mathbb{C} \setminus B_N(0))].$$

By 4.3, $f(B_\varepsilon(a) \setminus (L \cup F))$ is open and, since $f|_L$ is continuous and $a \in L \setminus F$, it is unbounded. Therefore $(U_{a,b})_{(a,b)\in L\times\mathbb{C}}$ is a family of non-empty open subsets of $\mathbb{C}$ parameterized by a 3-dimensional set, so

$$X := \{(a,b) \in l \times \mathbb{C} \mid \dim\{(a',b') \in l \times \mathbb{C} \mid U_{a,b} \cap U_{a',b'} \neq \emptyset\} < 3\}$$

is at most 2-dimensional. Hence, if we take any $a, a' \in L$ and $b_1, b' \in \mathbb{C}$ such that $\dim(a, a', b_1, b') = 6$, we get that

$$(f(B_\varepsilon(a)) + b_1) \cap (f(B_\varepsilon(a')) + b')$$

is unbounded, since it intersects $\mathbb{C} \setminus B_N(0)$ for every $N > 0$. Hence, we can take $b := b_1 - b'$. $\square$

Take a, a', b as in Claim 1. Since $\dim(a') = \dim(a) = 1$, we get $\dim(a'/b, a) = 1$, so a' is $\mathbb{C}_f$-generic over a, b. Let

$$d := \max\{\|y\| \mid \exists x \in \mathbb{C},\ (x,y) \in \Gamma_{f(a+x)} \cap \Gamma_{f(a'+x)+b}\}.$$

By Claim 1, for each $\varepsilon > 0$, there are $x_0, x_1 \in B_\varepsilon(0)$ such that

$$y = f(a + x_0) = f(a' + x_1)$$

and $\|y\| > d$. Taking $a'' := x_1 - x_0$, we get some x such that

$$(*) \qquad (x,y) \in \Gamma_{f(a+x)} \cap \Gamma_{f(a''+x)+b}.$$

We need now a transversality claim similar to Claim 2 in the proof of 3.3.

Claim 2: Let $g, h : \mathbb{C} \to \mathbb{C}$ be $\mathbb{C}_f$-definable over A, not almost affine and $b \in \mathbb{C}$ such that $\dim(b/A) = 2$. Assume also $h - g$ is not affine, then for each $a \in \mathbb{C}$ such that $h(a) = g(a) + b$, h and g are C^1 at a and

$$\Gamma_h \pitchfork_{(a,g(a))} \Gamma_{g(x)+b}.$$

Proof of Claim 2. Without loss $A = \emptyset$. Because $h - g$ is not affine and b is (in particular) $\mathbb{C}_f$-generic there exists $a \in \mathbb{C}$ such that $h(a) = g(a)+b$ so $\mathrm{acl}(a) = \mathrm{acl}(b)$, so $\dim(a) = 2$. Hence a is a C^1-point of both h and g and $\det(h-g)'(a) \neq 0$. Since $(h - g - b)' = (h - g)'$ and transversality of graphs is equivalent to injectivity of the difference of derivatives, the result follows. $\square$

Taking

$$A = \{a, a'\},\ h(x) = f(a + x),\ g(x) = f(a' + x),$$

we get by Claim 2

$$\Gamma_{f(a+x)} \pitchfork_s \Gamma_{f(a'+x)+b}$$

for every $s \in S$ (S is the intersection of the two graphs).

We may choose a'' so close to a' such that, by Claim 2, for every $s \in S$ we get

$$B_\varepsilon(s) \cap \Gamma_{f(a+x)} \cap \Gamma_{f(a''+x)+b} \neq \emptyset$$

for some $\varepsilon > 0$ such that $B_\varepsilon(s) \cap B_\varepsilon(s') = \emptyset$ for $s \neq s' \in S$. In particular the norm of the second coordinate of these intersection points is bounded by d.
But for y obtained in $(*)$ above $\|y\| > d$ so

$$|\Gamma_{f(a+x)} \cap \Gamma_{f(a''+x)+b}| > |\Gamma_{f(a+x)} \cap \Gamma_{f(a'+x)+b}|,$$

Since a'' was arbitrary, there we get infinitely many points with the same property. This contradicts the $\mathbb{C}_f$-genericity of a' over a, b. $\square$

It seems that from this point on the assumption that f is definable in an o-minimal field $\mathcal{R}$ is used very mildly, except for the fact that f is C^1 on a large set. We will try to point out all the places where o-minimality is used. It is not clear to us whether o-minimality is crucial for the conclusion of the proof.

We need a simple topological lemma:

Fact 4.5. *Let $D = B\setminus\{a\}$ be a punctured disc, Y a topological space and $f : D \to Y$ a continuous function. Then the set*

$$E := \{e \in Y : (a, e) \in \mathrm{fr}(\Gamma_f)\}$$

is connected.

Proof. Suppose not and let U, V be open and disjoint subsets of Y such that $E = (U \cap E) \cup (V \cap E)$. Consider $\tilde{D} := f^{-1}(U) \cup f^{-1}(V)$. By continuity, $\tilde{D}$ is open, and by the definition of E, $\tilde{B} := \tilde{D} \cup \{a\}$ is open in B, since for any sequence $B \ni x_n \to a$ we know that $\lim f(x_n) \in E$ so $x_n \in \tilde{D}$ for n big enough.
B has a basis of open discs, so there must be a disc C centered at a which is contained in $\tilde{B}$. Therefore $C \setminus \{a\} \subseteq \tilde{D}$ and $C \setminus \{a\} \nsubseteq f^{-1}(U), f^{-1}(V)$, which contradicts the connectedness of $C \setminus \{a\}$. □

Gathering all this information we easily get:

Proposition 4.6. *Let A be the set of points at which f is not continuous. Then A is finite and there is a continuous function $\hat{f} : \mathbb{C} \to S^2$ which coincides with f on $\mathbb{C} \setminus A$.*

Proof. By 3.3, 4.4 and 4.5, for each $a \in \mathbb{C}$,

$$\hat{f}(a) := \lim_{x\to a} f(x)$$

exists in S^2. For $a \in \mathbb{C} \setminus A$, $\hat{f}(a) = f(a)$. If $a \in A$, then either a is contained in the projection of the frontier of the graph of f, which is finite by 3.3, or $\hat{f}(a) = \infty$, so $a \in L$ which is finite by 4.4. □

Since there is no harm in changing f on a finite set, from now on we replace f with $\hat{f}$. Therefore f is a continuous function into S^2. We call each point $a \in \mathbb{C}$ such that $f(a) = \infty$, a *pole* of f.

Proposition 4.7. *f is an open map.*

Proof. By Brouwer's Invariance of Domain Theorem ([2]), a locally injective function $f : V \to \mathbb{R}^2$ is open for any open $V \subseteq \mathbb{R}^2$. Hence for any $a \notin \mathcal{N}(f)$ there exists a neighborhood V_a such that $f(V_a)$ is open. So we focus on $a \in \mathcal{N}(f)$. Without loss, $a = 0 = f(0)$. Since $\mathcal{N}(f)$ is finite we can find an open disc V around 0 such that $\mathrm{cl}(V) \cap \mathcal{N}(f) = \{0\}$. We may also assume that V is so small that $f^{-1}(0) \cap \mathrm{cl}(V) = \{0\}$. Denote $D := V \setminus \{0\}$.

By the Invariance of Domain $f|_D$ is open so $\partial f(D) \subseteq f(\partial D)$. By our choice of V we know that $0 \notin f(\partial D \setminus \{0\})$ and as the latter is closed, 0 is an isolated point of $\partial f(D)$. Hence for a small enough ball $B \ni 0$ we get that $B \cap \partial f(D) = \{0\}$. Therefore $(B \setminus \{0\}) \cap f(D)$ is non-empty and clopen in $B \setminus \{0\}$, so (since $B \setminus \{0\}$ is connected) $B \setminus \{0\} \subseteq f(D)$ and $B \subseteq f(V)$. So $0 \in \mathrm{int}(f(V))$, which is what we needed. □

We can now show that f is proper, i.e.:

Proposition 4.8. *f continuously extends to S^2.*

Proof. We need to show that $\lim_{x\to\infty} f(x)$ exists. By 4.5, it is enough to show that

$$E := \{e \in \mathbb{C} : (\infty, e) \in \operatorname{fr}_{S^2\times S^2} \Gamma_f\}$$

is finite.
Since $\{z : |f^{-1}(z)| \neq d\}$ is finite (d is the degree of f), it will suffice to show that there does not exists $e \in E$ such that $|f^{-1}(e)| = d$. Assume the contrary, and let $e_1, \ldots, e_d$ be distinct such that $f(e_i) = e$ for $i \leqslant d$. Let $U_i \ni e_i$ be pairwise disjoint open sets. Since $f : \mathbb{C} \to \mathbb{C}$ is open, $V := f(U_1) \cap \ldots \cap f(U_d)$ is an open set which is non-empty, as it contains e. But $e \in E$ implies that $f^{-1}(V)$ is unbounded. Therefore, the set

$$A := f^{-1}(V) \setminus U_1 \cup \ldots \cup U_d$$

is infinite. But for any $a \in A$ and $i \leqslant k$, $f(a)$ has a preimage by f in U_i, therefore each element of the infinite set $f(A)$ has at least $d + 1$ preimages by f, a contradiction. □

Question 4.9. In the presence of a $\mathbb{C}_f$-definable algebraically closed field Peterzil and Starchenko [19] use similar arguments to conclude that g is C^1 in all but finitely many points. Can such a result be obtained already at this stage?

Before we proceed we note that what we have proved for f is also valid for any (non-almost constant) $\mathbb{C}_f$-definable function $g : \mathbb{C} \to \mathbb{C}$, i.e. any such function is almost continuous, open and with finitely many non-injective points as a function from S^2 to S^2. Therefore, even when the complex addition does not extend to S^2, we can still add continuous $\mathbb{C}_f$-definable functions from S^2 to S^2, understanding e.g. $f(x+1) - f(x)$ as the continuous correction of a $\mathbb{C}_f$-definable function. Whenever we say that a $\mathbb{C}_f$-definable function is not injective, we mean that its continuous correction is not. Since a $\mathbb{C}_f$-definable function can have finitely many "accidental" values, non-injectivity does not say much before this correction. Note however, that if a $\mathbb{C}_f$-definable function has at least 2 poles, then it is necessarily non-injective in the right sense described above.

Fact 4.10. *If f is injective, then $g(x) := f(x + 1) - f(x)$ is not.*

Proof. By 2.13, g is not almost constant. Therefore the image of g is cofinite, hence g is onto as a continuous function on the compact space S^2. Since f is injective it has only one pole at $a \in S^2$, say. If $a \in \mathbb{C}$, then g has poles at a and $a - 1$, so it is not injective.
Assume $a = \infty$. Then g has no poles on $\mathbb{C}$, so $g(\mathbb{C}) = \mathbb{C}$. In particular, there is $c \in \mathbb{C}$ such that $g(c) = 0$. But then $f(c + 1) = f(c)$, so f is not injective, a contradiction. □

In the case f was injective (e.g. if $f(x) = 1/x$), we replace f with the g above, so we can assume f is not injective. The reader may wonder if we do not lose any information by replacing $\mathbb{C}_f$ with this reduct $\mathbb{C}_g$. However, the course of the proof is to find a $\mathbb{C}_f$-definable field (so replacing f with g is fine) and then show that all the extra structure must be already definable in that field. Hence, there is no danger in replacing f with g, as long as $\mathbb{C}_g$ is not locally modular, which is guaranteed e.g. if g is continuous and not injective. We will often make such replacements.

In a similar way, we may also assume that $f(\infty)=\infty$. If not, there exists some $a'\in\mathbb{C}$ such that $f(a')=\infty$. Replacing f with $g:=f(f(x)-f(\infty)+a')$ we get a function with the desired property. Not that if f was not injective clearly so is g.

From now on whenever we consider a $\mathbb{C}_f$-definable function $g:S^2\to S^2$ which is not almost affine we will implicitly assume that g has the same topological properties as f, i.e. we will always assume that g is continuous, open, non-injective with finite fibers and with finitely many non-injective points. The following result, which combines $\S VIII.4.5$ and $\S VIII.6.1$ of [26] will be crucial in our analysis of such functions:

Theorem 4.11. *Assume that $g:S^2\to S^2$ has all the above properties, then:*
(i) g is a ramified covering, i.e. for each $c\in S^2$ there exists a neighborhood $V\ni c$ and homeomorphisms $\phi:V\to B_1(0),\psi:g(V)\to B_1(0)$ such that $\phi(c)=\psi(g(c))=0$ and there is $k\geqslant 1$ such that the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "g"] \arrow[d, "\phi"'] & g(V) \arrow[d, "\psi"] \\ B_1(0) \arrow[r, "g_k"] & B_1(0), \end{tikzcd}$$

for $g_k(z)=z^k$. The number k is called the multiplicity *of g at c and denoted* $\mathrm{mlt}_c\,g$.
(ii) For $y\in S^2$, the number

$$\deg(g):=\sum_{c\in g^{-1}(y)}\mathrm{mlt}_c\,g$$

does not depend on y and is called the degree *of g.*

Remark 4.12. Note that the above theorem implies in particular that the size of a generic fiber of f is maximal (which can be also inferred easily from the openness of f). In particular, since we know that f is not injective, we get that $\deg(f)>1$.

Question 4.13. By Stoïlow's theorem (see, e.g. [26]) if g is a ramified covering of S^2, say, there exists a unique complex structure on S, denoted $\$$, such that

$$g:(S^2,\$)\to(S^2,\text{Riemann sphere structure})$$

is a holomorphic function. Moreover, $(S^2,\$)$ is biholomorphic to the Riemann sphere. Can those two facts be used to shorten the proof of our main result?

Definition 4.14. We call a point $y\in S^2$ a *ramification point* of g if $|g^{-1}(y)|<\deg(g)$.

By strong minimality we obtain immediately that the set of ramification points is finite. We have a relation between non-injective points and ramification points.

Fact 4.15. *$y\in S^2$ is a ramification point of g if and only if there is $c\in g^{-1}(y)$, which is a non-injective point of g.*

Proof. By 4.11(ii), y is a ramification point of g if and only if there is $c\in g^{-1}(y)$ such that $\mathrm{mlt}_c\,g>1$. Such c is clearly a non-injective point of g. $\square$

We prove that non-injective points of non-injective functions exist.

Lemma 4.16. *If $\deg g>1$, and $g(\infty)=\infty$ then g has at least 2 ramification points. Therefore, there is $c\in\mathbb{C}$ which is a non-injective point of g and not a pole.*

Proof. The proof is actually an application of the Hurwitz formula which is valid in our context. However, we need a very weak form of it, so we show the easy calculation.
Let $d := \deg(g)$, R be the set of ramification points of g, $r := |R|$ and $X := S^2 \setminus g^{-1}(R)$. Since $f|_X$ is d-to-1, we get $E(X) = dE(S^2 \setminus R)$, where E is the Euler characteristic. Therefore

$$d(2-r) = dE(S^2 \setminus R) = E(X) \leqslant 2 - r.$$

Since $d > 1$, $d(2-r) \leqslant 2-r$ implies $r > 1$.
By 4.15, we have 2 non-injective points of g on which g attains different values, so one of these values is from $\mathbb{C}$ and since $g(\infty) = \infty$ we get what we wanted. □

By the last lemma, we can add to all our assumptions about f that it has a non-injective point at 0 and that $f(0) = 0$ (replacing f with $f(x) - f(0)$, since we know 0 is not a pole).

4.2. **Degree theory.** We want to investigate the way multiplicities vary in continuous families, which we will do using the notion of topological degree. We start with some notation. For a simple closed curve C, let C^0 denote the interior of the region bounded by C and $\bar{C} := C \cup C^0$. For a continuous function $h : B \to \mathbb{C}$ from a bounded open subset of $\mathbb{C}$ and $y \in \mathbb{C} \setminus \partial B$, let $d(h, B, y)$ denote the topological degree as in [5]. We need a lemma about the relation between the topological degree and the degree we have defined earlier.

Lemma 4.17. *Let B, y be as above and $h : B \to \mathbb{C}$ be an $\mathcal{R}$-definable ramified covering. Then*

$$\sum_{x \in h^{-1}(y)} \mathrm{mlt}_x(h) = \pm d(h, B, y).$$

Proof. By Theorem 2.9(1) of [5], we have

$$\sum_{x \in h^{-1}(y)} i(h, x, y) = d(h, B, y),$$

where $i(h, x, p)$ is the index of h with respect to x, p defined in 2.8 of [5].

Since $d(h, B, y)$ is continuous with y and the smooth points of h are dense, it will be enough to show the equality for $y \in \mathbb{C}$ such that h is C^1 on $h^{-1}(y)$ and $h'(x_i) \neq 0$ for all $x_i \in h^{-1}(y)$. Since h is locally injective on a co-finite set, by 3.2 of [20] h' has constant sign (whenever it is defined). Therefore $i(h, x, y)$ has constant sign. Because we restrict to those points where $f' \neq 0$ we get that $i(h, x, y) = \mathrm{sgn}(h'(y))$ and $\mathrm{mlt}_x(h) = \pm i(h, x, y)$, with the sign independent of y. Hence the desired equality. □

In [20] o-minimality is used to prove that $\mathrm{sgn}(h')$ is constant. It seems, however, that o-minimality is not really needed for the present context, besides the fact that C^1-points are dense (see also the proof of Theorem IX(2.1) of [26]).

Remark 4.18. We know a posteriori that any $\mathbb{C}_f$-definable function *is* orientation preserving, so the topological degree and our degree do, in fact, agree (not only up to sign). But we are unable to show it at this stage.

Let $G : \mathbb{C}^n \times S^2 \to S^2$ be a $\mathbb{C}_f$-definable continuous family of continuous functions. In the next two lemmas we gather some facts on the trajectories of non-injective points in the family

$$(g_a(x) := G(a,x))_{a\in\mathbb{C}^n}.$$

These results will be crucial in the next section. We note that both lemmas follow immediately from the Argument Principle (applied to g'_a) if g is holomorphic. It is possible that they are still folklore even in this more general topological context, but we could not find a reference, so we give the proofs instead.

The function g_0 is either orientation preserving everywhere or orientation reversing everywhere. If it is orientation preserving (resp. reversing) then for each a close enough to 0, g_a is orientation preserving (resp. orientation reversing). Since, in the proofs below we are interested only in a close enough to 0, we can assume that all our functions are orientations preserving and that we actually have the equality in the statement of 4.17 (otherwise, we just use the negative topological degree, which has all the properties we want as well).

In the next lemma we show that if 0 is a non-injective point of g_0 then for every a close enough 0, g_a has a non-injective point near 0. In other words, we cannot lose non-injective points by changing our functions within a continuous family of continuous functions.

Lemma 4.19. *Assume $0 \in \mathcal{N}(g_0)$ and 0 is not a pole. Then for each open $U \ni 0$, there is $\varepsilon > 0$ such that for each $a \in B_\varepsilon(0)$, g_a has a non-injective point in U.*

Proof. Set $g := g_0$. We may assume $g(0) = 0$. By 4.11(i), we can find a simple closed curve $0 \in C \subset U$ such that $g(C) = D$ is a simple closed curve, $g(C^0) = D^0$ and $g|_{C^0}$ is homeomorphic to z^n. Take a closed 2-ball (i.e. a set homeomorphic to a closed disc) $0 \in B \subset D^0$ such that $g|_{g^{-1}(B)\cap\bar{C}}$ is still homeomorphic to z^n. Take $\varepsilon > 0$ such that for each $|a| < \varepsilon$ we have

$$g_a(D) \cap B = \emptyset.$$

Since the topological degree is invariant under homotopy, (see D4 from the Introduction to [5]), we have (by 4.17) for each $|a| < \varepsilon$ and for each $p \in B$

$$d(g_a, C^0, p) = n.$$

Because $g|_{g^{-1}(B)\cap C}$ is homeomorphic to z^n we know that $\partial g^{-1}(B) \cap \bar{C}$ is a simple closed curve. In particular $g|_{g^{-1}(B)\cap C}$ is connected. Therefore, there is $0 < \varepsilon' < \varepsilon$ such that $g_a^{-1}(B) \cap \bar{C}$ is connected for all $|a| < \varepsilon'$. To simplify the notation we will from now on identify g_a with its restriction to $\bar{C}$. Assume g_a has no non-injective points in C^0 for some $|a| < \varepsilon'$. In particular, by 4.11(i), g_a is a local homeomorphism on C^0. Therefore $g_a^{-1}(B) = \mathrm{cl}(g_a^{-1}(B)^0)$, and $\partial g_a^{-1}(B)$ is a 1-dimensional set which by 4.11 is everywhere locally homeomorphic to an interval (since, by assumption, there are no non-injective points of g_a in C^0).

Case 1 $\partial g_a^{-1}(B)$ is connected

Then, $g_a^{-1}(B)$ is homeomorphic to a closed disc and $g_a|_{g_a^{-1}(B)}$ can not be n-to-1, therefore, by the definition of the topological degree, there are non-injective points.

Case 2 $\partial g_a^{-1}(B)$ is not connected.

Each connected components of $\partial g_a^{-1}(B)$ is not self intersecting, and therefore homeomorphic to S^1. Therefore $g_a^{-1}(B)$ is homeomorphic to a closed disc with finitely

many closed discs removed. Hence, the Euler characteristic of $g_a^{-1}(B)$ is not positive, so $g_a|_{g_a^{-1}(B)}$ is not n-to-1, and again $\mathcal{N}(g_a) \cap C^0 \neq \emptyset$ (since $d(g_a, C^0, p) = n$ for every $p \in g^{-1}(B)$).
In both cases we reach a contradiction. $\square$

In the next lemma we investigate the behavior of the local degree, as we trace it along the trajectories of non-injective points within our family G. Roughly, our statement is that if this trajectory splits at some point into several trajectories then the local degree must go down.

Lemma 4.20. *Assume $c \in \mathcal{N}(g_0)$, $g_0(c) \neq \infty$ and for all open balls $B \ni 0, C \ni c$ there exists $a \in B$ such that $|\mathcal{N}(g_a) \cap C| > 1$, then there exist $a \in \mathbb{C}$ and $d \in \mathcal{N}(g_a)$ arbitrarily close to 0 and c respectively, such that $\operatorname{mlt}_d(g_a) < \operatorname{mlt}_0(g_0)$.*

Proof. Without loss $c = 0$. Take a simple closed curve C such that $g_0|_{C^0}$ is topologically equivalent to $z \mapsto z^n$. Take a smaller closed 2-ball $D \subset C^0$ as above and $\varepsilon > 0$ such that for each $a \in B_\varepsilon(0)$ and each $x \in D$

$$d(g_a, C^0, g_a(x)) = n.$$

We may assume that ε was also chosen in such a way that for each $a \in B_\varepsilon(0)$ there are $s_1, s_2 \in D^0 \cap \mathcal{N}(g_a)$ and $g_a(s_1)$ and $g_a(s_2)$ belong to the same component of $\mathbb{C} \setminus g_a(C)$. By 4.17 and a remark before 4.19, we have:

$$n = d(g_a, C, g_a(s_i)) = \sum_{x \in g_a^{-1}(g_a(s_i)) \cap D} \operatorname{mlt}_x g_a$$

for $i = 1, 2$. So, we are done if there is $i \in \{1, 2\}$ and $s_0 \neq s_i$ such that $g_a(s_0) = g_a(s_i)$.
If not, we are in the situation of [26, VIII(5.21)] for $E = D$, $p_i = s_i$, $q_i = g_a(s_i)$. Take any closed region $B \subset g_a(D) \setminus g_a(C)$ bounded by a simple closed curve and containing $g_a(s_1)$ and $g_a(s_2)$ (it can be done since $g_a(s_1)$ and $g_a(s_2)$ belong to the same component of $\mathbb{C} \setminus g_a(C)$). By [26, VIII(5.21)], there exist components $B_1 \ni s_1$, $B_2 \ni s_2$ of $g_a^{-1}(B)$ such that $g_a|_{B_i}$ is topologically equivalent to $z \mapsto z^{n_i}$ and $g_a(B_i) = B$. Clearly $B_1 \neq B_2$, therefore $n \geqslant n_1 + n_2$, so $n_1 < n$. $\square$

5. Smoothness

The main result of this section is a weak version of the Cauchy-Riemann theorem. We show that if g is a $\mathbb{C}_f$-definable function with the usual properties and g is not locally injective at some C^1 point c then $g'(c) = 0$. Rather surprisingly, proving that such functions existed (namely, ones with smooth non-injective points) turned out to be one of the hardest parts of the proof.

Lemma 5.1. *Let $g : S^2 \to S^2$ be a ramified covering and $g(0) = 0$. Then $0 \in \mathcal{N}(g)$ if and only if there are an open set $U \ni 0$ and a natural number $k > 1$ such that for some $\varepsilon > 0$,*

$$g^{-1}([-\varepsilon, \varepsilon] \times \{0\}) \cap U = \bigcup_{i=1}^{k} C_i,$$

where C_i's are distinct arcs (i.e. each C_i is homeomorphic to a closed interval) such that for $i \neq j$, we have $C_i \cap C_j = \{0\}$ and $0 \in \operatorname{Int}(C_i)$ for all i.

Proof. For the right-to-left direction is enough to see that the above condition implies that g is not a local homeomorphism at 0.
For the left-to-right direction, note that the complex mapping $h_k(z) := z^k$ on $B_1(0)$ has the corresponding property for any arc J with $0 \in \mathrm{int}(J)$. By 4.11(i), there exists $U \ni 0$ and homeomorphisms $h_1 : U \to B_1(0)$, $h_2 : B_1(0) \to g(U)$ such that $g = h_1 \circ h_k \circ h_2$ on $B_1(0)$. Necessarily $h_1(0) = h_2(0) = 0$. The result follows, if we take $J := h_2^{-1}([-\varepsilon, \varepsilon] \times \{0\})$. □

From the above lemma we get.

Proposition 5.2. *Let $g : S^2 \to S^2$ be a ramified covering, $g(0) = 0$ and $0 \in \mathcal{N}(g)$. Assume that g is C^1 at 0 then $g'(0) = 0$.*

The proof follows immediately from the more general statement:

Lemma 5.3. *Let $g : S^2 \to S^2$ be a ramified covering, $g(0) = 0$ and $0 \in \mathcal{N}(g)$. For all but finitely many $a \in S^1$ if the partial derivative $\frac{\partial g}{\partial \alpha}$ exists and is continuous at 0 then $\frac{\partial g}{\partial \alpha}(0) = 0$.*

Proof. Let g_2 be the second coordinate function of g (locally around 0). It is enough to focus on g_2. It will be enough to prove that for each $\varepsilon > 0$ for cofinitely many directions α there is $c \in B_\varepsilon(0)$ where $\frac{\partial g_2}{\partial \alpha}(c) = 0$.
Let U, $C_1, \ldots, C_k$ be as provided by 5.1 so $g_2|_{C_i} = 0$. Fix a direction α such that (possibly after shrinking U) $\frac{\partial g_2}{\partial \alpha}$ exists and is continuous on U. Let $L_\alpha \subset \mathbb{R}^2$ be the line through 0 given by α. By the mean value theorem it will be enough to find $d \in \mathbb{C}$ such that

$$|(L_\alpha + d) \cap (\bigcup_{i=1}^{k} C_i) \cap B_\varepsilon(0)| > 1.$$

Take any parameterizations $\tau_i : [-1, 1] \to C_i$ such that $\tau(0) = 0$. By Fact 2.8(5) for any i the limits

$$c_i^+ := \lim_{t \to 0^+} \tau_i'(t), \quad c_i^- := \lim_{t \to 0^-} \tau_i'(t)$$

exist. We are done if L_α meets C_1 and C_2 “both-sided transversally” at 0, i.e. if $\alpha \notin \{c_1^+, c_1^-, c_2^+, c_2^-\}$, which is a finite set. □

In the last proof we used o-minimality in order to assure that the (one-sided) tangents to the curves C_i existed at 0. We are not sure whether o-minimality can be avoided:

Question 5.4. Is 5.2 true without the o-minimality assumption?

To apply the above proposition, we need to have a $\mathbb{C}_f$-definable continuous function with a non-injective point at which the function is C^1. Toward this end we will use the degree theory developed in the previous section.
We introduce some notation. For a function $h : \mathbb{C} \to \mathbb{C}$, we denote by $\mathcal{S}(h)$ the set of all points $z \in \mathbb{C}$ such that h is not C^1 at z. By definition, $\mathcal{S}(h)$ is closed. If h is $\mathcal{R}$-definable, then $\mathcal{S}(h)$ is also $\mathcal{R}$-definable and (by 2.8(3)) $\dim \mathcal{S}(h) \leqslant 1$.
We fix the continuous $\mathbb{C}_f$-definable family of continuous functions

$$f_a(x) := f(a + x) + f(x).$$

If $\mathcal{N}(f_a) \nsubseteq \mathcal{S}(f_a)$ for some $a \in \mathbb{C}$ we are done. So we will now investigate the bad case that $\mathcal{N}(f_a) \subseteq \mathcal{S}(f_a)$ for all a. Note that $\mathcal{S}(f_a) \subseteq \mathcal{S}(f) \cup (\mathcal{S}(f) - a)$.

Note that $\mathcal{S}(f)$ is relatively small (i.e. at most 1-dimensional). In the next lemma, we show that we can replace $\mathcal{S}(f)$ with a finite set.

Lemma 5.5. *If $\mathcal{N}(f_a) \subseteq \mathcal{S}(f) \cup (\mathcal{S}(f) - a)$ for all $a \in \mathbb{C}$, then there is a finite $S \subset \mathbb{C}$ such that for each a, we have $\mathcal{N}(f_a) \subseteq S \cup (S - a)$.*

Proof. Let $d_n : \mathbb{C} \to \mathcal{S}(f)$ be the (partial) $\mathcal{R}$-definable function such that $d_n(a)$ is the n^{th}-smallest (in the lexicographical order) element of $\mathcal{N}(f_a) \cap (\mathcal{S}(f) \cup (\mathcal{S}(f) - a))$, if it exists.
Claim: For all n such that $\dim(\mathrm{dom}(d_n)) = 2$ there is a finite $S_n \subseteq \mathcal{S}$ such that $\dim(d_n^{-1}(\mathcal{S} \setminus S_n)) \leqslant 1$.

Proof of Claim. Assume not. Since the set of ramification points of f_a is uniformly $\mathbb{C}_f$-definable, by strong minimality (2.9(3)) we obtain

$$N := \max\{|f_a(\mathcal{N}(f_a))| \mid a \in \mathbb{C}\} < \infty.$$

Since $\dim \mathcal{S}(f) < 2$ and $\dim(\mathrm{dom}(d_n)) = 2$, there are infinitely many $x \in \mathcal{S}(f)$ such that $d_n^{-1}(x)$ is infinite. For any such x, the set

$$R_x := \{a \in \mathbb{C} \mid f_a(x) \text{ is a ramification point of } f_a\}$$

is infinite and $\mathbb{C}_f$-definable, hence cofinite.
Let

$$D := \max\{\deg(f_a) \mid a \in \mathbb{C}\}$$

and take $DN + 1$ distinct x_i such that $d_n^{-1}(x_i)$ is infinite. Then for any

$$a \in R_{x_1} \cap \ldots \cap R_{x_{DN+1}},$$

we get that f_a has at least $N + 1$ ramification points, which contradicts the choice of N. $\square$

Note that $|\mathcal{N}(f_a)|$ is uniformly bounded by N in the proof of the Claim and the uniform bound for $\deg(f_a)$. Therefore, there is $n \in \mathbb{N}$, the least such that $\dim(\mathrm{dom}(d_n)) < 2$. Let

$$X := \mathbb{C} \setminus \bigcup_{i<n} (d_i^{-1}(\mathcal{S} \setminus S_i)),\ S := \bigcup_{i<n} S_i.$$

So $\dim(X) \leqslant 1$ and S is finite. For each $a \in \mathbb{C} \setminus X$, we have $\mathcal{N}(f_a) \subseteq S \cup (S - a)$. It remains to show that it still holds for $a \in X$. Assume there is $c \in \mathcal{N}(f_a)$ such that $c \notin S \cup (S - a)$. Therefore, there is an open $U \ni c$ and $\varepsilon > 0$ such that for each $a' \in B_\varepsilon(a)$, we have

$$U \cap (S \cup (S - a')) = \emptyset.$$

By 4.19, there is $\varepsilon' < \varepsilon$ such that for each $a' \in B_{\varepsilon'}(a)$, there is a non-injective point of $f_{a'}$ in U. But then $\mathcal{N}(f_{a'}) \nsubseteq S \cap (S - a')$, which is a contradiction, since $B_{\varepsilon'}(a)$ is 2-dimensional. $\square$

Theorem 5.6. *There is a $\mathbb{C}_f$-definable function $g : \mathbb{C} \to \mathbb{C}$ such that 0 is a non-injective point of g and g is C^1 at 0.*

Proof. We may assume that for all $a \in \mathbb{C}$ we have $\mathcal{N}(f_a) \subseteq \mathcal{S}(f_a)$. Since $\mathcal{S}(f_a) \subseteq \mathcal{S}(f) \cup (\mathcal{S}(f) - a)$, we get by 5.5 a finite S such that $N(f_a) \subseteq S \cup (S - a)$ for each $a \in \mathbb{C}$. By 4.16, $\mathcal{N}(f) \cap \mathbb{C} \neq \emptyset$ so without loss (after replacing f with $f(x - a)$) we have $0 \in \mathcal{N}(f) \subseteq S$.

Take a disc D with boundary C around 0 such that $D\cap S=\{0\}$. Since $f_0=2f$, so $0\in\mathcal{N}(f_0)$, we get by 4.19 that there is $\varepsilon>0$ such that for all $a\in B_\varepsilon(0)$ we have:

$$\mathcal{N}(f_a)\cap D\neq\emptyset \text{ and } D\cap(S-a)=\{-a\}.$$

We need the following
Claim: There is $a\in B_\varepsilon(0)$ such that $|N(f_a)\cap D|>1$.

Proof. $\mathcal{N}(f_a)\cap D\subseteq\{0,-a\}$ for each $a\in B_\varepsilon(0)$. Let

$$X:=\{a\in B_\varepsilon(0)\setminus\{0\}\mid 0\in\mathcal{N}(f_a)\}.$$

By 4.19, X is an open subset of $B_\varepsilon(0)\setminus\{0\}$.
The same applies to

$$Y:=\{a\in B_\varepsilon(0)\setminus\{0\}\mid -a\in\mathcal{N}(f_a)\}.$$

By 4.19, $B_\varepsilon(0)\setminus\{0\}=X\cup Y$.
Since $f_a=f_{-a}\circ L_a$, where $L_a(x)=a+x$, we get

$$\mathcal{N}(f_a)=L_a^{-1}(\mathcal{N}(f_{-a}))=\mathcal{N}(f_{-a})+a.$$

Therefore $X=-Y$, so they are both non-empty. Since $B_\varepsilon(0)\setminus\{0\}$ is connected, $X\cap Y\neq\emptyset$. □

By 4.20, there is $a\in\mathbb{C}$ such that

$$\min\{\mathrm{mlt}_{f_a}(x)\mid x\in\mathcal{N}(f_a)\}<\min\{\mathrm{mlt}_f(x)\mid x\in\mathcal{N}(f)\}.$$

Since this minimum can not go down for ever, at some stage we get a function with the required property. □

Replacing f, if needed, we will from now on assume that f (aside from all the usual properties) is C^1 at 0 and $0\in\mathcal{N}(f)$. We will also assume that $f(0)=0$. For $c\in\mathbb{C}$ let $f_c(x):=f(x+c)$ and for $a:=(a_1,\dots,a_n)$, $n\geq 2$ denote

$$f_a:=f_{a_{n-1}}\circ\ldots\circ f_{a_1}\circ f-f_{a_n}.$$

For example, if $n=2$ we get $f_a(x)=f(f(x)+a_1)-f(x+a_2)$.
With this notation, an immediate but important application of the last theorem is:

Corollary 5.7. *There is a ball $B\subseteq\mathbb{C}^n$ with $0\in B$ such that f_a has a C^1 non-injective point for all $a\in B$. Moreover, for each a if $c\in\mathcal{N}(f_a)$ then*

$$(*)\qquad f'_{a_n}(c)=f'_{a_{n-1}}(f_{a_{n-2}}\circ\ldots\circ f_{a_1}(f(c)))\cdot\ldots\cdot f'_{a_1}(f(c))\cdot f'(c).$$

Proof. The existence of such B is immediate from 4.19 and the last theorem. The equality $(*)$ follows immediately from 5.2 using the Chain Rule. □

We denote

$$R(a;c)\iff f_a(c)\text{ is a ramification point of }f_a,$$
$$R'(a;c)\iff R(a;c)\wedge f'_a(c)=0.$$

Note that R is $\mathbb{C}_f$-definable and R' is $\mathcal{R}$-definable. We claim that:

Lemma 5.8. *Let*

$$E:=\{(y,f(y)+a_1,\dots,f_{a_{n-2}}\circ\ldots\circ f_{a_1}(f(y))+a_{n-1})\mid(\exists a_n\in\mathbb{C})R'(a;y)\}.$$

Then $\dim(E)=2n$.

Proof. Let B be as provided by Corollary 5.7 and $a_1, \ldots, a_{n-1} \in \mathbb{C}$ be such that for

$$a_{<n} := (a_1, \ldots, a_{n-1})$$

we have $\dim(a_{<n}) = 2(n-1)$ and $a_{<n}$ is in the projection of B. Let

$$R'_{a_{<n}} := \{y \in \mathbb{C} \mid (\exists x \in \mathbb{C}) R'(a_{<n}, x; y)\}.$$

Claim: $\dim(R'_{a_{<n}}) = 2$.

Proof of Claim. Let $U := \{x \in \mathbb{C} : (a_{<n}, x) \in B\}$. Then U is open and there is an o-minimally definable function $c : U \to \mathbb{C}$ such that for each $x \in U$, $R'(a_{<n}, x; c(x))$. It is enough to show that $\dim(c(U)) = 2$. Assume not.
Case 1: $\dim(c(U)) = 0$
Then, there is $y_0 \in \mathbb{C}$ such that $\dim(c^{-1}(y_0)) = 2$, hence there is an open $U \subseteq \mathbb{C}$ such that $U \subseteq c^{-1}(y_0)$. Therefore, for each $a_n \in U$, $(*)$ from 5.7 holds for $a_{<n}, a_n, y_0$ (with $a_{<n}, y_0$ fixed). But the right hand side of $(*)$ does not depend on a_n, implying that f' is constant on an open set, which contradicts 2.16.
Case 2: $\dim(c(U)) = 1$
Then, there is an infinite set $Y \subseteq c(B)$ such that for each $y \in Y$, $c^{-1}(y)$ is infinite. Since $R' \subseteq R$, we get that for each $y \in Y$ the set

$$R^y := \{x \in \mathbb{C} \mid R(a_{<n}, x, y)\}$$

is cofinite. Let

$$R_x := \{y \in \mathbb{C} \mid R(a_{<n}, x, y)\}, \quad n := \max\{|R_x| \mid x \in \mathbb{C}\}.$$

If we take $y_1, \ldots y_{n+1} \in Y$, then for any $x \in R^{y_1} \cap \ldots \cap R^{y_{n+1}}$, we have $|R^x| > n$, a contradiction. $\square$

By the claim, there are $a_n, c \in \mathbb{C}$ such that $\dim(c/a_{<n}) = 2$ and $R(a_{<n}, a_n; c)$, so

$$(c, f(c) + a_1, \ldots, f_{a_{n-2}} \circ \ldots \circ f_{a_1}(f(c)) + a_{n-1}) \in E.$$

Since clearly,

$$\dim(f(c) + a_1, \ldots, f_{a_{n-2}} \circ \ldots \circ f_{a_1}(f(c)) + a_{n-1}/c) = \dim(a_{<n}/c) = 2n - 2,$$

we get that

$$\dim(c, f(c) + a_1, \ldots, f_{a_{n-2}} \circ \ldots \circ f_{a_1}(f(c)) + a_{n-1}) = 2n,$$

which is exactly what we want. $\square$

It follows that for some $\mathbb{C}_f$-definable functions $g : \mathbb{C} \to \mathbb{C}$ the set $g'(\mathbb{C}) \subset M_2(\mathbb{R})$ has at least some of the flavor of a group:

Corollary 5.9. *Suppose g is a $\mathbb{C}_f$-definable function and $V \subseteq \mathbb{C}$ is open such that g is C^1 on V and there is $v \in V$ such that g is non-injective at v. Then for any $n > 1$, there are open $U_1, \ldots, U_n \subseteq V$ such that for $i \leqslant n$, $g'(U_i) \subseteq \mathrm{GL}_2(\mathbb{R})$ and*

$$\dim(g'(U_1) \cdot \ldots \cdot g'(U_n)) \leqslant 2.$$

Proof. Let E be as a in the previous lemma, V the set of C^1 points of g, which is open. Since $\dim(E) = 2n$ (by 5.8) and $\dim(\mathbb{C} \setminus V) = 1$, we can assume that $E \subseteq V^n$. By 2.8(5), there are open $U_i \subseteq \mathbb{C}$ such that $U_1 \times \cdots \times U_n \subseteq E$. The result follows from the definition of E, the moreover part of 5.7 and 2.12. $\square$

6. ANALITICITY

The aim of this section is to prove the converse of Proposition 5.2, namely that if $f'(c)=0$ then f is not a local homeomorphism at c. We will do it by proving that f is analytic on some open set. The proof will go as follows. As a first step we show that from 5.9 we obtain a local Lie group $X \subset \mathrm{GL}_2(\mathbb{R})$ and an open set $U \subseteq \mathbb{C}$ such that $f'(U) \subseteq XM$ for some $M \in \mathrm{GL}_2(\mathbb{R})$. We then use local Lie group theory to prove that there exists a Lie subgroup $H < \mathrm{GL}_2(\mathbb{R})$ such that $H \cap X$ is a relatively open subset of X. Then, we use the classification of maximal solvable Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$ to show that $H = \mathrm{GL}_1(\mathbb{C})$ (naturally embedded into $\mathrm{GL}_2(\mathbb{R})$). Then we use 5.9 again to show that we can take $M = I$ and therefore f is holomorphic on U.

We will try to keep the following notational conventions: U, V, W will denote (usually open) subsets of $\mathbb{C}$, a, b, c elements of $\mathbb{C}$, A, B, C, D subsets of $\mathrm{GL}_2(\mathbb{R})$, H, G Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$, X, Y local Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$ and M, N elements of $\mathrm{GL}_2(\mathbb{R})$ (with the exception of the more general 6.1 below).
In this section we omit the group multiplication symbol.

Lemma 6.1. *Let B, C, D be $k-$dimensional, $\mathcal{R}$-definable subsets of an $\mathcal{R}$-definable group G with $\dim(G) > k$. Assume that $\dim(BCD) = k$. Then there is a definable relatively open $A \subseteq B$ such that*

$$\dim(AA^{-1}AA^{-1}) = k.$$

Proof. Choose $b_0 \in B, c_0 \in C, d_0 \in D$ independent generics, and set $z_0 = b_0c_0d_0$. Since z_0 is interdefinable with d_0 over $\{b_0, c_0\}$ we get that $\dim(z_0) = k$, i.e. it is generic in BCD. By the definability of dimension and the genericity of b_0, c_0 there are $B_0 \ni b_0, C_0 \ni c_0$ relatively open such that for all $b \in B_0$ and $c \in C_0$ there is some $d \in D$ such that $z_0 = bcd$. It follows that

$$C_0^{-1}B_0^{-1}z_0 \subseteq D$$
$$B_0C_0C_0^{-1}B_0^{-1}z_0 \subseteq BCC_0^{-1}B_0^{-1}z_0 \subseteq BCDz_0$$
$$\dim(B_0C_0C_0^{-1}D_0^{-1}) = k.$$

Since $\dim(B_0C_0) = k$, we get in exactly the same way that for a relatively open $B_{00} \subseteq B_0$ and some $z_{00} \in B_0C_0$, we have $B_{00}^{-1}z_{00} \subseteq C_0$, and therefore $z_{00}^{-1}B_{00} \subseteq C_0^{-1}$. It follows that

$$B_{00}B_{00}^{-1}z_{00}z_{00}^{-1}B_{00}B_{00}^{-1} \subseteq B_{00}C_0C_0^{-1}B_{00}^{-1}.$$

Therefore, $\dim(B_{00}B_{00}^{-1}B_{00}B_{00}^{-1}) = k$ and $A := B_{00}$ works. □

It may be worth mentioning that in the stable case (and in particular in the strongly minimal case) the analogue with just 2 subsets of G would have already produced a definable group. It is unclear to us if the same holds in the o-minimal case.

Question 6.2. Can we take just 2 sets B, C in the statement of the above lemma?

It is also unclear to us whether the o-minimality assumption on top of the fact that G is a Lie group (which is the case in practice) is needed at all.

Question 6.3. Is 6.1 true if we replace "definable" by "submanifold" (so, no o-minimal theory around)?

Proposition 6.4. *Let $X \subset \mathrm{GL}_2(\mathbb{R})$ be local Lie subgroup, i.e. a submanifold such that $X = X^{-1}$, XX is a submanifold and $\dim(XX) = \dim(X)$. Assume $\dim(X) \leqslant 2$. Then, there is a Lie subgroup $H < \mathrm{GL}_2(\mathbb{R})$ such that $H \cap X$ is open in X. Moreover, H is a conjugate of either the upper-triangular group or $\mathrm{GL}_1(\mathbb{C})$.*

Proof. It is easy to see that $\mathfrak{h} := T_1(X)$ is a Lie subalgebra of $\mathfrak{gl}_2(\mathbb{R})$. Since $\dim(\mathfrak{h}) = 2$, $\mathfrak{h}$ is solvable. Hence $\mathfrak{h}$ is contained in $\mathfrak{g}$, a maximal solvable subalgebra of $\mathfrak{gl}_2(\mathbb{R})$.
We want to show that $\mathfrak{g}$ is conjugated to $\mathrm{GL}_1(\mathbb{C})$ or the algebra of upper-triangular matrices. This should be well-known but we could not find a direct reference, so we present a quick argument. All elementary facts on Lie algebras we use here can be found in [15]. Since $\mathfrak{gl}_2(\mathbb{R})$ splits into the product of its center and $\mathfrak{sl}_2(\mathbb{R})$, we can assume that $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{sl}_2(\mathbb{R})$. If $\dim(\mathfrak{g}) = 1$, then $\mathfrak{g}$ is commutative, so it must be a Cartan subalgebra of $\mathfrak{sl}_2(\mathbb{R})$. Therefore, $\mathfrak{h}$ is conjugated to $\mathfrak{so}_2(\mathbb{R})$ (which is contained in $\mathfrak{gl}_1(\mathbb{C})$) or to the subalgebra of diagonal matrices. But the latter is not maximal among solvable subalgebras of $\mathfrak{sl}_2(\mathbb{R})$. If $\dim(\mathfrak{g}) = 2$, then (by e.g. the Lie Theorem) $\mathfrak{g}$ contains a semi-simple element M. If M is not diagonalizable over $\mathbb{R}$, then the $\mathbb{R}$-span of M is conjugated to $\mathfrak{so}_2(\mathbb{R})$. But $\mathfrak{so}_2(\mathbb{R})$ is maximal in $\mathfrak{sl}_2(\mathbb{R})$ even as a Lie subalgebra (by a direct computation or e.g. exercise 1 on p.265 of [15]), a contradiction. Hence, M is diagonalizable over $\mathbb{R}$, so we may assume it is diagonal. It is easy to see that if $N \in \mathfrak{sl}_2(\mathbb{R})$ is not triangular, then the Lie algebra generated by M and N is the entire $\mathfrak{sl}_2(\mathbb{R})$, so $\mathfrak{g}$ consists of triangular matrices. Taking one more conjugation (if needed), we make sure that these matrices are upper-triangular.
We have obtained that $\mathfrak{g}$ is a Lie algebra of a Lie subgroup $H < \mathrm{GL}_2(\mathbb{R})$ and H is conjugated to either the upper-triangular group or to $\mathrm{GL}_1(\mathbb{C})$.
To finish, it is enough to notice that for local Lie subgroups X, Y, we have $T_1(X) \subseteq T_1(Y)$ if and only if $X \cap Y$ is open in X (see Theorem 74 on page 258 of [23]). $\square$

Remark 6.5. This is the place in the proof of the main theorem where we find the matrix A (as in the statement of 1.2) – it will be the conjugating matrix appearing in the moreover part of the above Proposition. This will be made clearer in the last three statements in this section and Remark 7.5.

Proposition 6.6. *There are $M, M' \in \mathrm{GL}_2(\mathbb{R})$ and an open $U \subseteq \mathbb{C}$ such that $M \circ f \circ M'$ is holomorphic on U. Moreover, we may assume that $M' = I$.*

Proof. Take $n = 9$ in 5.9 to obtain open sets $U_1, \dots U_9$ such that f is C^1 and regular on $U_1 \cup \dots \cup U_9$ and

$$\dim(f'(U_1) \dots f'(U_9)) \leqslant 2.$$

We want to be under the assumptions of 6.1. By 2.16, $\dim(f'(U_i)) \geqslant 1$ for each $i \leqslant 9$. If there is $i \leqslant 7$ such that

$$\dim f'(U_i) = \dim f'(U_{i+1}) = \dim f'(U_{i+2}) = \dim f'(U_i) f'(U_{i+1}) f'(U_{i+2})$$

we take $B = f'(U_i)$, $C = f'(U_{i+1})$, $D = f'(U_{i+2})$.
If there is no such i, we get in particular

$$\dim f'(U_1) f'(U_2) f'(U_3) = \dim f'(U_4) f'(U_5) f'(U_6) = \dim f'(U_7) f'(U_8) f'(U_9) = 2.$$

Then we take

$$B = f'(U_1) f'(U_2) f'(U_3),\ C = f'(U_4) f'(U_5) f'(U_6),\ D = f'(U_7) f'(U_8) f'(U_9).$$

By 6.1, there is a relatively open $A \subseteq B$ such that $X := AA^{-1}$ satisfies the assumptions of 6.4. Take $H < \mathrm{GL}_2(\mathbb{R})$ as given by 6.4. By shrinking A we may assume that $X \subseteq H$. For any $a \in A$, $Aa^{-1} \subseteq H$, therefore $A \subseteq Ha$.
By our choice of A, there is $i \leqslant 7$ and $b \in \mathrm{GL}_2(\mathbb{R})$ such that $A \cap f'(U_i)b$ is open in $f'(U_i)b$, so $A' := Ab^{-1}$ is open in $f'(U_i)$.
Since f is C^1 on U_i, $f'|_{U_i} : U_i \to \mathrm{GL}_2(\mathbb{R})$ is continuous, therefore $U := (f'|_{U_i})^{-1}(A')$ is open in $\mathbb{C}$.
By our choices we obtain

$$f'(U) \subseteq Ab^{-1} \subseteq Hab^{-1}.$$

We want to show now that H is a conjugate of $\mathrm{GL}_1(\mathbb{C})$. By 6.4 and since real conjugation is an isomorphism of $\mathbb{C}_f$, it is enough to show that $f'(B)$ is not contained in a single coset of the upper-triangular group. However, this follows immediately from Corollary 2.17 since if B were contained in a coset of the upper triangular matrices, it would imply that some partial derivative of f is constant on B.
The moreover part follows, since again we can replace f with its conjugate. $\square$

Strangely enough, despite our good understanding of f we were not, until this stage, able to show that f' does not have infinite fibers, which is the reason we had to use $n = 9$ (rather that $n = 3$) in the proof of the last proposition.

Our next step is to show that in fact f is holomorphic (up to conjugation) on an open set, i.e. that $M = I$ (for M as in the statement of the last proposition). First, we note that as an immediate application of the last result we do get the desired control on the dimension of $f'(B)$.

Lemma 6.7. *There is an open $W \subseteq U$ such that $f'(W)$ is an open subset of $M\,\mathrm{GL}_1(\mathbb{C})$ (U, M are from 6.6).*

Proof. By 6.6, $M^{-1} \circ f$ is holomorphic on U. In particular $\dim(M^{-1} \cdot f'(U)) = 2$, since $(M^{-1} \circ f)'$ is again holomorphic, hence open.
Therefore,

$$B := \mathrm{int}_{M\,\mathrm{GL}_1(\mathbb{C})}(f'(B)) \neq \emptyset.$$

Take $W := (f'|_U)^{-1}(B)$ (it is open, since f is C^1 on U). $\square$

Proposition 6.8. *f is holomorphic on an open $V \subseteq \mathbb{C}$.*

Proof. Take W from 6.7. We may assume that $0 \in W$. By 6.7, $f'(W) - f'(0)$ is an open subset of $MM_1(\mathbb{C})$ for some $M \in \mathrm{GL}_2(\mathbb{R})$ (we work in $M_2(\mathbb{R})$ now).
Take $0 \neq A \in f'(W) - f'(0)$ such that $-A \in f'(W) - f'(0)$. Take $a, b \in W$ such that

$$f'(a) - f'(0) = A, \quad f'(b) - f'(0) = -A.$$

Then we have

$$f'(a) + f'(b) - 2f'(0) = 0.$$

Consider the $\mathbb{C}_f$-definable function

$$g(x) := f(a + x) + f(b + x) - 2f(x).$$

Then $g'(0) = 0$ and there is an open $V \subseteq W$ such that $0 \in V$, g is C^1 on V and $g'(V) \subseteq MM_1(\mathbb{C})$. Therefore, $M^{-1}g$ is holomorphic on V and $(M^{-1}g)'(0) = 0$. Thus 0 is a non-injective point of $M^{-1}g$, hence 0 is also a non-injective point of g.

Therefore g satisfies the assumptions of Corollary 5.9 whence there are $U_1, U_2 \subseteq V$ open such that
$$\dim(g'(U_1) \cdot g'(U_2)) \leq 2.$$
By the argument as in 6.7, we have
$$\dim(g'(U_1)) = \dim(g'(U_2)) = 2.$$
Since the set of critical points of $g|_V$ can not be 2-dimensional, we can assume that
$$g'(U_1), g'(U_2) \subseteq M\,\mathrm{GL}_1(\mathbb{C}).$$
Therefore, it is enough to show:
Claim: For any open $B_1, B_2 \subseteq \mathrm{GL}_1(\mathbb{C})$ if $\dim(MB_1MB_2) = 2$ then $M \in \mathrm{GL}_1(\mathbb{C})$.

Proof of Claim. For $h \in M_2$, let $X_h := MB_1Mh$. Since
$$MB_1MB_2 = \bigcup_{h \in B_2} X_h,$$
there is $h \in B_1$ such that
$$\dim(\{h' \in B_2 \mid X_h \cap X_{h'} \neq \emptyset\}) = 2.$$
Then for any such h' we have
$$MB_1Mh \cap MB_1Mh' \neq \emptyset$$
$$B_1Mh \cap B_1Mh' \neq \emptyset$$
$$\mathrm{GL}_1(\mathbb{C})Mh = \mathrm{GL}_1(\mathbb{C})Mh'$$
$$Mh'h^{-1}M^{-1} \in \mathrm{GL}_1(\mathbb{C}).$$
Hence the group
$$H := \{h'' \in \mathrm{GL}_1(\mathbb{C}) \mid Mh''M^{-1} \in \mathrm{GL}_1(\mathbb{C})\}$$
contains an open set, so (since $\mathrm{GL}_1(\mathbb{C})$ is connected) $H = \mathrm{GL}_1(\mathbb{C})$ and $M\,\mathrm{GL}_1(\mathbb{C}) = \mathrm{GL}_1(\mathbb{C})M$.
Since $\mathrm{GL}_1(\mathbb{C})$ is self-normalizing in $\mathrm{GL}_2(\mathbb{R})$, we get $M \in \mathrm{GL}_1(\mathbb{C})$. □

7. Rationality

In this section we finish the proof of our main result. For the first time in this paper we will have to use some non-trivial model theoretic results. To obtain an infinite field interpretable in $\mathbb{C}_f$ we will use a variant of the Hrushovski-Weil group configuration theorem. We will then have to use known results of Hrushovski's to show that in fact the field is definable in $\mathbb{C}_f$ in order to conclude that it is o-minimally definably isomorphic to $\mathbb{C}$. To conclude that f is rational with respect to this field structure we will have to use the purity of the field, as proved in [19]. The following will be the key to our argument:

Definition 7.1. Let $\mathcal{M}$ be a strongly minimal structure. A *field configuration* is a collection of tuples such that $\mathrm{RM}(g_1, g_2, g_3, b_1, b_2, b_3) = 5$ and:

- $\mathrm{RM}(g_i) = 2$ and $\mathrm{RM}(b_i) = 1$ for all $1 \leq i \leq 3$.
- The g_i are pairwise independent but $g_3 \in \mathrm{acl}(g_1, g_2)$.
- $\mathrm{RM}(g_1, b_1, b_2) = \mathrm{RM}(g_2, b_2, b_3) = \mathrm{RM}(g_3, b_1, b_3) = 3$.
- There are no other dependencies.

For a proof of the theorem below see 4.5 on page 206 of [21] and 3.27(2) of [22].

Theorem 7.2 (Hrushovski)**.** *If $\mathcal{M}$ is strongly minimal and admits a field configuration then there is a strongly minimal field interpretable in $\mathcal{M}$.*

Note that if $\mathcal{K}$ is an algebraically closed field then a field configuration can be easily constructed using the action of $\mathbb{G}_m(K) \ltimes \mathbb{G}_a(K)$ on $\mathbb{G}_a(K)$. More precisely, take $g_j := (\alpha_j^i, \alpha_j^i)$ for independent generic elements of K for $i, j \in \{1,2\}$. Thinking of g_j as the affine transformation $x \mapsto \alpha_j^1 x + \alpha_j^2$ we let $g_3 = g_2 \circ g_1$. Take $b_1 \in K$ any generic element independent over g_1, g_2 and let $b_2 = g_1(b_1)$ and $b_3 = g_3(b_1)$. It is easy to check that this gives us a field configuration in the structure $\mathcal{K}$.

We will now use the analyticity of f on an open set together with the field configuration theorem to show:

Theorem 7.3. *There is a $\mathbb{C}_f$-interpretable strongly minimal field.*

Proof. Our aim is to construct a field configuration. Take V from 6.8. Without loss of generality, $0 \in V$ and let $M := f'(0)$.
We will find a field configuration of $\mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$ acting on $\mathbb{G}_a(\mathbb{C})$ within $f'(V)$ and use it to construct a similar field configuration in $\mathbb{C}_f$.
Take $a_1, a_2 \in \mathbb{C}$ $\mathcal{R}$-independent generics very close to $1 \in \mathbb{C}$ and b_1, b_2, b $\mathcal{R}$-independent generics over a_1, a_2 in a small ball around $0 \in \mathbb{C}$. Let

$$g := (a_1, b_1), h := (a_2, b_2) \in \mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$$

(note that g, h are very close to the identity of $\mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$).
We get the afore mentioned group configuration

$$\mathcal{X} := (g, h, hg, b, g \cdot b, hg \cdot b)$$

where $\cdot$ denotes the action of $\mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$ on $\mathbb{G}_a(\mathbb{C})$. Denoting, for $(a,b) \in \mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$,

$$M(a,b) = (M \cdot a, M + b),$$

(where on the right-hand side $a, b \in \mathrm{GL}_1(\mathbb{C}) \hookrightarrow \mathrm{GL}_2(\mathbb{R})$) we get that

$$M(a,b) \in f'(V)$$

for all a close enough to 1 and b close enough to 0. Therefore, we get a group configuration $M\mathcal{X} \subset f'(V)$.
We aim to show that for any six-tuple $\mathcal{Y} \subset V$ such that $f'(\mathcal{Y}) = M\mathcal{X}$, $\mathcal{Y}$ is a group configuration in $\mathbb{C}_f$.
Since $\mathcal{R}$-independence is stronger than $\mathbb{C}_f$-independence, we need to check only the algebraicity conditions. They come from either a group operation in $\mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$ or a group action of $\mathbb{G}_m(\mathbb{C}) \ltimes \mathbb{G}_a(\mathbb{C})$ on $\mathbb{G}_a(\mathbb{C})$.
We consider only the latter case, so we take $g, b, g \cdot b \in \mathcal{X}$. We have

$$g = (a_1, b_1), \quad g \cdot b = a_1 b + b_1$$

$$Mg = (Ma_1, M + b_1), \quad Mb = M + b, \quad M(g \cdot b) = M + a_1 b + b_1.$$

Take $\alpha, \beta, \gamma, \delta \in V$ such that

$$(f'(\alpha), f'(\beta)) = Mg, \quad f'(\gamma) = Mb, \; f'(\delta) = M(g \cdot b).$$

We want to show $\delta \in \mathrm{acl}(\alpha, \beta, \gamma)$.

$$\begin{aligned} f'(\delta) = M + a_1 b + b_1 &= M + M^{-1} f'(\alpha)(f'(\gamma) - M) + f'(\beta) - M \\ &= M^{-1} f'(\alpha) f'(\gamma) - M^{-1} f'(\alpha) M + f'(\beta). \end{aligned}$$

Since we have set $M = f'(0)$ and $M \in \mathrm{GL}_1(\mathbb{C})$, M commutes with $f'(\alpha)$ (this is one of the reasons we needed f to be analytic on V and not only a translate of one) and we get

$$f'(0)f'(\delta) = f'(\alpha)f'(\gamma) - f'(0)f'(\alpha) + f'(0)f'(\beta).$$

For each $\mu, \nu \in \mathbb{C}$ there is a $\mathbb{C}_f$-definable function $f_{\mu,\nu}$ which is holomorphic in a neighborhood of 0 such that

$$f_{\mu,\nu}(0) = 0, \quad f'_{\mu,\nu}(0) = f'(\mu)f'(\nu).$$

Consider the function

$$g_y := f_{\alpha,\gamma} - f_{\alpha,0} + f_{\beta,0} - f_{y,0}.$$

We will show that g_δ is not a generic function in the strongly minimal family $\{g_y : y \in \mathbb{C}\}$. Since this family is defined over (α, β, γ) this will prove that indeed $\delta \in \mathrm{acl}(\alpha, \beta, \gamma)$.

It will be enough to find infinitely many y such that $|g_y^{-1}(0)| > |g_\delta^{-1}(0)|$. Since g_δ is holomorphic (which is another reason we needed f to be analytic on V) in a neighborhood of 0, and $g'_\delta(0) = 0$, we get that g_δ has a multiple 0 at 0.

Let $\{a_1, \dots, a_k\} = g_\delta^{-1}(0)$ and without loss of generality assume that $a_1 = 0$. Let $V_i \ni a_i$ be pairwise disjoint open sets. Choose δ' close enough to δ such that $g'_{\delta'}(0) \neq 0$. Therefore $g_{\delta'}$ has a single zero at 0. We use the Argument Principle to find $c \in V_1 \setminus \{0\}$ such that $g_{\delta'}(c) = 0$.

It remains to check that for δ' close enough to δ, $|g_{\delta'}^{-1}(0) \cap V_i| > 0$ for all $i > 1$. By 4.17, $d(0, g_\delta, V_i) \neq 0$. By Axiom (D4) of the degree theory from the introduction to [5], $d(0, g_{\delta'}, V_i) \neq 0$ for δ' close enough to δ. By 4.17 again, $|g_{\delta'}^{-1}(0) \cap V_i| > 0$.

By Theorem 7.2, a strongly minimal field is interpretable in $\mathbb{C}_f$. □

It remains to prove:

Theorem 7.4. *There is $A \in \mathrm{GL}_2(\mathbb{R})$ such that $\mathbb{C}_{A \circ f \circ A^{-1}}$ is interdefinable with the field of complex numbers.*

Proof. We will need a little more model theory for this proof. Let $\mathcal{K}$ be the strongly minimal field interpretable in $\mathbb{C}_f$ by the previous theorem. Note that in particular $\dim(K) = 2$. By [22, 2.27], $(\mathbb{C}, +)$ is $\mathcal{K}$-internal, i.e. there exists a partial $\mathbb{C}_f$-definable surjective map $g : K^n \to (\mathbb{C}, +)$. Therefore the structure $(\mathbb{C}, +)$ is interpretable in the structure $\mathcal{K}_{\mathbb{C}_f}$ whose universe is K and the definable sets are those which are $\mathbb{C}_f$-definable.

By Theorem 1.1. of [16] (and $\dim(K) = 2$), there is an $\mathcal{R}$-definable isomorphism of fields $\Phi : \mathcal{K} \to \mathbb{C}$. Then $\Phi(\mathcal{K}_{\mathbb{C}_f})$ is an $\mathcal{R}$-interpretable, strongly minimal structure and the complex field is definable in $\Phi(\mathcal{K}_{\mathbb{C}_f})$. By Theorem 1.3 of [19], the structure $\Phi(\mathcal{K}_{\mathbb{C}_f})$ coincides with the complex field. Therefore, $\mathcal{K}_{\mathbb{C}_f} = \mathcal{K}$ (as structures). Hence $(\mathbb{C}, +)$ is interpretable in $\mathcal{K}$. By [22, 3.1], [22, 4.13] and Remark 3. on p. 69 of [22], $(\mathbb{C}, +)$ is $\mathbb{C}_f$-definably isomorphic to a $\mathcal{K}$-algebraic group. Since $\dim \mathbb{C} = \dim \mathcal{K} = 2$ we get that also $\dim_{\mathcal{K}}(\mathbb{C}, +) = 1$. Since $(\mathbb{C}, +)$ has no torsion, it is $\mathbb{C}_f$-definably isomorphic to $\mathbb{G}_a(\mathcal{K})$.

Using the above isomorphism we can find a $\mathbb{C}_f$-definable $\star : \mathbb{C}^2 \to \mathbb{C}$ such that $(\mathbb{C}, +, \star)$ is a field.

Claim: There is $A \in \mathrm{GL}_2(\mathbb{R})$ such that

$$A : (\mathbb{C}, +, \cdot) \cong (\mathbb{C}, +, \star).$$

Proof. Let $1_\star$ be the neutral element for $\star$ and for any $a \in \mathbb{C}$ let us denote $a_\star := a \cdot 1_\star$. For any $a \in \mathbb{C}$, we have

$$a_\star \star 2_\star = a_\star \star (2 \cdot 1_\star) = a_\star \star (1_\star + 1_\star) = a_\star + a_\star = 2 \cdot a_\star = 2 \cdot a \cdot 1_\star = (a \cdot 2)_\star.$$

We can replace 2 above with any rational number to obtain

$$\forall a \in \mathbb{C}, b \in \mathbb{Q} \quad a_\star \star b_\star = (a \cdot b)_\star \; .$$

Fix $a \in \mathbb{C}$ and consider 2 additive, $\mathbb{C}_f$-definable functions

$$\phi_a(x) := a_\star \star b_\star, \quad \psi_a(x) := (a \cdot b)_\star.$$

By 4.6, ϕ_a and ψ_a are continuous on a cofinite set. But any additive function g continuous on a cofinite set is continuous everywhere – for $c \in \mathbb{C}$ and $c_n \to c$ take N such that g is continuous at Nc, then $Nc_n \to Nc$ and $f(Nc_n) \to f(Nc)$, so $f(c_n) \to f(c)$.
Therefore ϕ_a, ψ_a are continuous and they coincide on $\mathbb{Q}$, so they coincide on $\mathbb{R}$.
Take $i^\star \in \mathbb{C}$ such that $i^\star \star i^\star = -1_\star$. By an argument as above, we get that for all $r \in \mathbb{R}$, we have $r \cdot i^\star = r_\star \star i^\star$.
Let the first column of A consists of the real and imaginary part of $1_\star$ and the second column of A consists of real and imaginary part of $i^\star$. Take any 2 complex numbers $a = x + iy, b = w + iv$. We have

$$A(a \cdot b) = A(xw - vy + i(xv + yw)) = (xw - vy)_\star + (xv + yw) \cdot i^\star$$
$$= (xw - vy)_\star + (xv + yw)_\star \star i^\star = x_\star \star w_\star - v_\star \star y_\star + (x_\star \star v_\star - y_\star \star w_\star)_\star \star i^\star$$
$$= (x_\star + i^\star \star y_\star) \star (w_\star + i^\star \star v_\star) = (x_\star + y \cdot i^\star) \star (w_\star + v \cdot i^\star) = Aa \star Ab.$$

□

By the claim the complex field is definable in $\mathbb{C}_{AfA^{-1}}$. By [19, 1.3], AfA^{-1} is also definable in the complex field. Therefore, $\mathbb{C}_{AfA^{-1}}$ is interdefinable with the complex field. □

Remark 7.5. (1) Note that in the end of the proof of 7.4, we get that AfA^{-1} is $\mathbb{C}$-rational, hence there is an open set $U \subseteq \mathbb{C}$ on which both f and AfA^{-1} are holomorphic. As in the proof of 6.8 this implies that A is multiplication by a non-zero $a \in \mathbb{C}$. Since $af(a^{-1}z)$ is rational, f is also rational and we can take $A = I$ in the statement of 7.4. Therefore 6.4 is the only place of the proof where conjugation by a real matrix takes place. Hence A from the statement of 1.2 is indeed as was mentioned in 6.5.

(2) In particular we have shown the following statement:
Let $f : \mathbb{C} \to \mathbb{C}$ be $\mathcal{R}$-definable such that $\mathbb{C}_f$ is strongly minimal and not locally modular. Assume f is holomorphic on an open set. Then f is an almost $\mathbb{C}$-rational function and $\mathbb{C}_f$ is interdefinable with the complex field.

(3) Note, however, that we still needed the Claim in the proof of 7.4, since [16, 1.1] gives only a certain $\mathcal{R}$-definable isomorphism between $(\mathbb{C}, +, \cdot)$ and $(\mathbb{C}, +, \star)$ and we want this isomorphism to be $\mathbb{R}$-linear. It is also easy to find $\star \neq \cdot$ such that $(\mathbb{C}, +, \star)$ is isomorphic to the complex field by a $\mathbb{C}$-linear map (it must be of the form $x \star y = axy$ for a fixed $a \in \mathbb{C} \setminus \{0\}$).

Let us forget now all the changes we have done to f and just take it as it was in the statement of 1.2. We easily obtain:

Corollary 7.6. *f is a conjugate of a complex constructible function.*

Note that this is the best possible description of f, since any such a function yields a strongly minimal structure $\mathbb{C}_f$.

8. O-MINIMALITY

This section is dedicated to generalizing our main result to any o-minimal (real closed) field. More precisely, assume $\mathcal{R}$ is an ω-saturated o-minimal expansion of a real closed field R. Let $K = R[i]$ be the algebraic closure of R and $f : K \to K$ an $\mathcal{R}$-definable function such that $K_f := (K, +, f)$ is strongly minimal and not locally modular. Theorem 1.2 generalizes naturally to:

Theorem 8.1. *K_f is biinterpretable with the field K.*

Before we give the details of the proof, we would like to mention that many classical geometrical and topological theorems translate naturally into our present context. Theorems such as the Mean Value Theorem, the Implicit Function Theorem (together with its relatives, the Open Mapping Theorem and the Inverse Function Theorem) as well as deeper theorems from differential geometry can be proved in the o-minimal context by investigating classical proofs and replacing the geometrical notions involved with their definable counterparts.

We do the exact same thing here. There is one significant difference (referred to in the introduction) with the classical case. Working over $\mathbb{R}$, the o-minimality assumption simplifies the proofs, but it is unclear to what extent it is actually needed. In the more general context considered here it *gives* the natural working grounds, since the pure topology on R may be totally disconnected and the usual topological notions will not be meaningful anymore. For example, in the totally disconnected case the topological dimension of every set in R is 0, contrary to the classical case where it coincides with the o-minimal one. Hence, the o-minimal dimension is the meaningful one in this context, and because it only applies to definable sets, the natural framework is that of $\mathcal{R}$-definable sets and functions.

Since the methods of translating proofs are pretty standard by now, and our original proofs were written with the abstract real closed field context in mind, we just sketch how the translation can be done section by section, focusing on two places where it may not straightforward.

Section 2 was written in the general context and since we assume $\mathcal{R}$ to be ω-saturated needs no modifications. The same applies to Section 3, having in mind the theory of smooth o-minimal manifolds as developed, e.g., by Berarducci and Otero in [1].

We focus a bit more on Section 4. The fact that f is open (Lemma 4.3) goes through by the o-minimal version of the Invariance of Domain Theorem (see [11]) and gives the finiteness of the number of poles of f as in Lemma 4.4. The (definable) connectedness of the set of possible limits of f at a point (Lemma 4.5) also goes through – replace sequences and their limits with definable curves and their limits. The proof of the continuity and properness of f need no modification.

We postpone the discussion of the transfer of crucial facts from topological analysis from [26] (notably Theorem 4.11) to the Appendix. Other parts of Section 4.1 go through. For the degree theory we use Peterzil and Starchenko's development of the theory in the o-minimal setting ([19] and [20]). This allows for a smooth adaptation of the topological Lemmas in subsection 4.2 to the present setting. Section 5 is done purely in the definable realm so needs no adaptation.

We focus now on Section 6. The corresponding Lie algebra theory for definable groups was developed by Peterzil, Pillay and Starchenko in [17] and [18]. The algebraic facts on Lie algebras we needed (classification of maximal solvable subalgebras of $\mathfrak{gl}_2(\mathbb{R})$) hold for any real closed field and the proofs from [17] work in the local case to give an analogue of Theorem 74 of [23]. The rest of Section 6 and the entire Section 7 go through, using the theory of K-holomorphic functions as developed in [19].

We conclude this section by showing that some care is needed when translating proofs from the classical case to the o-minimal one. The key result in Section 6 is the proof in Proposition 6.4 of a correspondence between local Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$ and solvable Lie subgroups of $\mathrm{GL}_2(\mathbb{R})$. Our original proof of 6.4 went along the following path – a local subgroup X produces a Lie subalgebra $\mathfrak{h}$. ¿From $\mathfrak{h}$ we can only get a *virtual* Lie subgroup of $\mathrm{GL}_2(\mathbb{R})$, namely a smooth injective map $\phi : H \to \mathrm{GL}_2(\mathbb{R})$, which allows for a simple conclusion of the argument.

However, this map need not be definable in any o-minimal expansion of $\mathbb{R}$ as the following example shows.

Example 8.2. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $T = S^1 \times S^1$ and $\mathfrak{h} := (1, \alpha)\mathbb{R}$ be a Lie subalgebra of $\mathrm{Lie}(T)$. Then $\mathfrak{h}$ corresponds to a virtual Lie subgroup $\phi : \mathbb{R} \to T$ and $\phi(\mathbb{R})$ is dense in T. In particular, $\phi(\mathbb{R})$ is not definable in any o-minimal expansion of $\mathbb{R}$.

Taking a piece of $\phi(\mathbb{R})$ which is a local Lie subgroup of T definable in an o-minimal theory, we get a definable local subgroup not corresponding to any global definable subgroup (in the definable category, the notion of definable subgroup and virtual definable subgroup clearly coincide).

So, in o-minimal theories there is no chance for a full Lie correspondence

$$\text{Lie subalgebras of } \mathrm{Lie}(G) \;\leftrightarrow\; \text{Virtual Lie subgroups of } G.$$

As was pointed out by Kobi Peterzil, there is no Lie correspondence even on the level of local subgroups in the real field, since one can not define there even locally subgroups of $\mathbb{G}_m(\mathbb{R})^2$ which are graphs of non-rational power function. However, there is a natural o-minimal expansion of $\mathbb{R}$ (by analytic functions) for which we have the full local correspondence. Hence, we find it natural to ask the following question:

Question 8.3. Assume $\mathcal{R}$ is an o-minimal expansion of a real closed field. Is there an o-minimal expansion $\mathcal{R}' \supseteq \mathcal{R}$ such that we have the correspondence

$$\text{Lie subalgebras of } \mathfrak{gl}_n(R) \;\leftrightarrow\; \mathcal{R}'\text{-definable local subgroups of } \mathrm{GL}_n(R) \text{ ?}$$

9. Appendix

We give a short overview of the results in topological analysis used in the paper, setting them in the o-minimal context. Since the classical proofs (taken from [26]) are given in a purely topological setting their adaptation to the o-minimal case is rather straightforward. For the sake of completeness, and in view of its importance to our analysis we give a proof of Theorem 4.11. Other results taken from [26] and used in the text transfer to the o-minimal context as easily. We will use the o-minimal version of the Jordan plane curve theorem (see [27]) without reference. All references to [26] are to Chapter VIII unless stated otherwise. We keep the notation and assumptions from the previous section.

The proof of the following can be found in Whyburn (4.2) and needs no adaptation to the general context:

Lemma 9.1. *Any definable open mapping of one simple definable closed curve onto another is definably homeomorphic to the mapping $z \mapsto z^k$ for $|z| = 1$ and some $k \geq 1$.*

Definition 9.2. A definable mapping satisfying the conclusion of the previous lemma will be said to be of degree k.

The following should be standard (though possibly out of style):

Definition 9.3. A set T homeomorphic to a straight line interval is called a simple *arc*. If $a_1, \dots, a_k$ are any (distinct) points in R^n we sometimes denote by $a_1 a_2 \dots a_k$ an arbitrary simple arc T through $a_1, \dots, a_k$ such that if $\phi : [c, d] \to T$ is a homeomorphism witnessing it then $\phi^{-1}(a_1) = c$, $\phi^{-1}(a_k) = d$ and $\phi^{-1}(a_i) < \phi^{-1}(a_{i+1})$ for $1 \leq i < k$.

The second Lemma we will need is:

Lemma 9.4. *Let A, B be definably homeomorphic to a definable closed disc, $g(A) = B$ a definable finite-to-one open mapping. Let $C = \partial A$ and $J = \partial B$, and assume that $g(A^o) = B^o$, $g(C) = J$ and $g \restriction C$ is of degree k. Then $\left|g^{-1}(p)\right| \leq k^2$ for all $p \in B$.*

Proof. It will be enough to show that if $a, b \in J$, $p \in B^o$ and apb is a definable simple arc spanning J (i.e. having no common points with J but a, b) then $\left|g^{-1}(p)\right| \leq m(ap) \cdot m(pb)$ where for any set S we define $m(S)$ to be the number of definable components of $g^{-1}(S)$.

If this were not true there would exist definable components K of $g^{-1}(ap)$ and H of $g^{-1}(pb)$ such that $\left|H \cap K \cap g^{-1}(p)\right| > 1$. Clearly H and K are infinite and 1-dimensional. By 1-dimensionality for any generic $(a_1, a_2) \in H$ we may assume that $a_2 \in \operatorname{acl}(a_1)$ so for any $a_1 \in \pi(H)$, the set $\{a_2 \mid (a_1, a_2) \in H\}$ is finite. Similarly for K. By the existence of Skolem functions (Fact 2.8(4)) there are finitely many definable continuous (by 2.8(1)) functions $h_1(x) \dots h_i(x)$ and $k_1(x), \dots, k_r(x)$ whose domains are intervals in R and such that

$$H = \bigcup \Gamma_{h_i},\ K = \bigcup \Gamma_{k_i}.$$

Let $a \in H$, and assume without loss that $a \in \Gamma_{h_1}$. Since $h_1(x)$ is continuous, for all $\varepsilon > 0$ small enough, $\Gamma_{h_1} \cap B_\varepsilon(a)$ is connected. By the existence of Skolem functions again, if for all $\varepsilon > 0$ there exists $b \in H \cap B_\varepsilon(a) \setminus \Gamma_{h_1}$ then there exists a definable curve $\gamma \subseteq H$ such that for every $0 < \varepsilon < \varepsilon_0$ (some ε_0 small enough),

$$\gamma(\varepsilon) \in H \cap B_\varepsilon(a) \setminus \Gamma_{h_1}.$$

It follows (by the definition of o-minimality) that for some $\varepsilon > 0$ small enough $\{\gamma(x) : x < \varepsilon\} \subseteq \Gamma_{h_i}$ for some $i \neq 1$. So by continuity of h_i, $\lim_{x \to 0} \gamma(x) \in \Gamma_{h_i}$, implying that $(\Gamma_{h_1} \cup \Gamma_{h_i}) \cap B_\varepsilon(a)$ is definably connected (because the Γ_{h_j} are definably connected and those two have a as a common point). By induction it is now easy to check that H (and similarly K) are definably locally connected, so $H \cup K$ is definably path connected. Since $H \cap K \subseteq g^{-1}(p)$ it is in particular finite. From the assumption that $|H \cap K| > 1$ it follows that $H \cup K$ contains a definable simple closed curve. So there exists a definable component T of $A \setminus (H \cup K)$ such that $T \subseteq A \setminus C$. But $\mathrm{fr}(T) \subseteq H \cup K$ so $T \cap g^{-1}(apb) = \emptyset$ and so $g(T)$ would have to be a definable component of $B \setminus (apb)$ which is impossible, since $g(T) \cap J = \emptyset$. $\square$

The next corollary is very easy:

Corollary 9.5. *Let $g(A) = B$ be as in the previous lemma, $k = \deg(g \restriction C)$.*

(1) *If $k = 1$ then g is a homeomorphism.*
(2) *For any $p_1, p_2 \in B$, $\left|g^{-1}(p_1)\right| \leq k \left|g^{-1}(p_2)\right|$. In particular, if for some $p \in B$ we have $\left|g^{-1}(p)\right| = 1$ then $\left|g^{-1}(q)\right| \leq k$ for all $q \in B$.*

We can now prove the main result of Whyburn we are using:

Proposition 9.6. *Let A, B be definably homeomorphic to a definable closed disc, $g(A) = B$ a definable finite-to-one open mapping. Let $C = \partial A$ and $J = \partial B$, and assume that $g(A^o) = B^o$, $g(C) = J$. If there exists a point $q \in B$ such that $g^{-1}(q) = \{p\}$ for some $p \in A$ then g is definably homeomorphic to $z \mapsto z^k$ on $|z| \leq 1$ for some $k \geq 1$.*

Proof. By Lemma 9.1, $g|_C$ is definably homeomorphic to $z \mapsto z^k$ on $|z| = 1$ for some $k \geq 1$. Let aqb be any definable simple arc spanning J and

$$\{a_1, \ldots, a_k\} = g^{-1}(a),\ \{b_1, \ldots, b_k\} = g^{-1}(b)$$

be cyclically ordered on C. Because g is finite-to-one $\dim(g^{-1}(apb)) = 1$. By o-minimality (as in the proof of the Lemma 9.4 above) we get that $g^{-1}(apb)$ is locally connected and since p must be in all its definable components, it must be definably connected. Thus, $g^{-1}(aq)$ contains arcs of the form $a_i q$ and $g^{-1}(qb)$ contains arcs qb_i $(i = 1, \ldots, k)$. Since $a_i p \cup a_j p$ separates b_i and b_k for $i \neq j$, and

$$(a_i p \cup a_j p) \cap (pb_i \cup pb_n) = \{p\}$$

it separates pb_i from pb_j. Whence $pb_i \cap pb_j = \{p\}$. By (2) of the last corollary,

$$g^{-1}(apb) = \bigcup_{i=1}^{k} (a_i p \cup pb_i).$$

Denote

$$C_i := a_i p \cup pb_i \cup b_i a_i,$$

where $a_i b_i$ is the arc in C satisfying $a_i b_i \cap g^{-1}(a) = \{a_i\}$. It is now straightforward to check that $g|_C$ satisfies the assumptions of the previous corollary with $k = 1$, and so is a homeomorphism. The conclusion now follows readily. $\square$

To obtain *(i)* of theorem 4.11 we still have to show that if g is a definable finite-to-one continuous open map, for every $q \in K$ there exists a small enough closed disc $D \ni q$ such that every definable component of $g^{-1}(D)$ is definably homeomorphic to a closed disc. This is the analogue of (3.1) of Whyburn, which is easily translated into the present context as well. Alternatively, this can easily be obtained using the properness of g. Part *(ii)* of 4.11 is an immediate corollary of part *(i)*, as it shows that the degree function is continuous.

References

[1] A. Berarducci and M. Otero. Intersection theory for o-minimal manifolds. *Ann. Pure Appl. Logic* **107**(1-3), 87–119 (2001).

[2] L. Brouwer. Zur invarianz des n-dimensionalen gebiets. *Mathematische Annalen* **72**, 55–56 (1912).

[3] S. Buechler. "Essential Stability Theory". Springer (1996).

[4] Z. Chatzidakis, E. Hrushovski, and Y. Peterzil. Model theory of difference fields. II. Periodic ideals and the trichotomy in all characteristics. *Proc. London Math. Soc. (3)* **85**(2), 257–311 (2002).

[5] W. Gangbo and I. Fonseca. "Degree Theory in Analysis and Applications". Oxford University Press (1995).

[6] W. Hodges. "A shorter model theory". Cambridge University Press, Cambridge (1997).

[7] E. Hrushovski. Strongly minimal expansions of algebraically closed fields. *Israel Journal of Mathematics* **79**, 129–151 (1992).

[8] E. Hrushovski. The Mordell-Lang conjecture for function fields. *J. Amer. Math. Soc.* **9**(3), 667–690 (1996).

[9] E. Hrushovski and Z. Sokolovic. Minimal subsets of differentially closed fields. unpublished.

[10] E. Hrushovski and B. Zilber. Zariski geometries. *Journal of the American Mathematical Society* **9**(1), 1–55 (1996).

[11] J. Johns. An open mapping theorem for o-minimal structures. *J. Symbolic Logic* **66**(4), 1817–1820 (2001).

[12] D. Macpherson. Notes on o-minimality and variations. In "Model theory, algebra, and geometry", vol. 39 of "Math. Sci. Res. Inst. Publ.", pp. 97–130. Cambridge Univ. Press, Cambridge (2000).

[13] D. Marker and A. Pillay. Reducts of $(\mathbb{C}, +, \cdot)$ which contain $+$. *Journal of Symbolic Logic* **55**(3), 1243–1251 (1990).

[14] G. A. Martin. Definability in reducts of algebraically closed fields. *J. Symbolic Logic* **53**(1), 188–199 (1988).

[15] A. L. Onishchik and È. B. Vinberg. "Lie groups and algebraic groups". Springer Series in Soviet Mathematics. Springer-Verlag, Berlin (1990). Translated from the Russian and with a preface by D. A. Leites.

[16] M. Otero, Y. Peterzil, and A. Pillay. On groups and rings definable in o-minimal expansions of real closed fields. *Bull. London Math. Soc.* **28**(1), 7–14 (1996).

[17] Y. Peterzil, A. Pillay, and S. Starchenko. Definably simple groups in o-minimal structures. *Trans. Amer. Math. Soc.* **352**(10), 4397–4419 (2000).

[18] Y. Peterzil, A. Pillay, and S. Starchenko. Linear groups definable in o-minimal structures. *J. Algebra* **247**(1), 1–23 (2002).

[19] Y. Peterzil and S. Starchenko. Expansions of algebraically closed fields in o-minimal structures. *Selecta Math. (N.S.)* **7**(3), 409–445 (2001).

[20] Y. Peterzil and S. Starchenko. Computing o-minimal topological invariants using differential topology. Available on the Newton institute preprint server http://www.newton.cam.ac.uk/preprints/NI05039.pdf (2005).

[21] A. Pillay. "Geometric stability theory", vol. 32 of "Oxford Logic Guides". The Clarendon Press Oxford University Press, New York (1996). Oxford Science Publications.

[22] B. Poizat. "Stable groups", vol. 87 of "Mathematical Surveys and Monographs". American Mathematical Society, Providence, RI (2001). Translated from the 1987 French original by Moses Gabriel Klein.

[23] L. Pontrjagin. "Topological groups". Princeton University Press, Princeton, N.J. (1939 1958). Translated from the Russian by Emma Lehmer, (Fifth printing, 1958).

[24] E. D. Rabinovich. "Definability of a field in sufficiently rich incidence systems", vol. 14 of "QMW Maths Notes". Queen Mary and Westfield College School of Mathematical Sciences, London (1993). With an introduction by Wilfrid Hodges.

[25] L. van den Dries. "Tame topology and o-minimal structures", vol. 248 of "London Mathematical Society Lecture Note Series". Cambridge University Press, Cambridge (1998).

[26] G. T. Whyburn. "Topological analysis". Second, revised edition. Princeton Mathematical Series, No. 23. Princeton University Press, Princeton, N.J. (1964).

[27] A. Woerheide. "Topology of Definable Sets in O-Minimal Structures". PhD thesis, University of Illinois at Urbana-Chapaign (1996).
[28] B. Zilber. The structure of models of uncountably categorical theories. In "Proc. Int. Congr. Of Mathematicians, 1983, Warszawa", vol. 1, pp. 359–368. PWN - North-Holland P. Co. (1984).
[29] B. Zilber. Non-commutative geometry and new stable structures. Available on http://www.maths.ox.ac.uk/ zilber/noncom.pdf (2005).

*Mathematical Institute, Oxford University, Oxford, UK
E-mail address: `hasson@maths.ox.ac.uk`
URL: `http://www.maths.ox.ac.uk/~hasson/`

♣Mathematical Institute, Oxford University, Oxford, UK
E-mail address: `kowalski@maths.ox.ac.uk`

†Instytut Matematyczny, Uniwersytet Wrocławski, Wrocław, Poland
E-mail address: `pkowa@math.uni.wroc.pl`
URL: `http://www.math.uni.wroc.pl/~pkowa/`