*LOGIC COLLOQUIUM '82*
*G. Lolli, G. Longo and A. Marcja (editors)*



# Remarks on Tarski's problem concerning $(\mathbb{R}, +, \cdot, \exp)$

Lou van den Dries

School of Mathematics
The Institute for Advanced Study
Princeton, New Jersey 08540
U.S.A.

## INTRODUCTION

In his monograph on the elementary theory of the structure $(\mathbb{R}, +, \cdot)$ Tarski asked whether his results could be extended to the structure $(\mathbb{R}, +, \cdot, \exp)$ ([T, p. 45]). (Instead of $\exp(x) = e^x$, Tarski suggested the function $f(x) = 2^x$, but this makes little difference since exp is definable in $(\mathbb{R}, +, \cdot, f)$ as the unique function of the form $x \longmapsto f(ax)$ which is its own derivative; the axioms mentioned in [T, p. 57, note 20] for $\mathrm{Th}(\mathbb{R}, +, \cdot, f)$ are far from adequate, see e.g. [D-W].)

Before we discuss Tarski's question, let us briefly review some aspects of his work on $(\mathbb{R}, +, \cdot)$ and see what use has been made of it:

(1) Decidability of $\mathrm{Th}(\mathbb{R}, +, \cdot)$,

(2) $\mathrm{Th}(\mathbb{R}, +, \cdot)$ = theory of real closed fields,

(3) Elimination of quantifiers for $(\mathbb{R}, <, 0, 1, +, \cdot)$,

(4) Properties of definable subsets of $\mathbb{R}^n$,

(5) Properties of definable functions.

These aspects are closely related in Tarski's work, but it makes sense to discuss them separately. (1) is a nice result in its own right and quite useful in many theoretical decidability questions, but has otherwise not been important in settling open problems, as far as I know. (2) is sometimes useful in proving properties of real closed fields: in certain cases the only known proof consists of first establishing the property for the field of reals by transcendental methods and then invoking (2). (This is called Tarski's principle.) (2) and (3) combined give a trivial and improved solution of Hilbert's 17th problem, and some important generalizations, due to A. Robinson.

The central result in Tarski's work seems to me (3) as I hope to indicate in the discussion of (4) and (5) below. (Also, (1) and (2) are easy consequences of Tarski's method of establishing (3).) Concerning (4): the single most fruitful fact is the so-called Tarski-Seidenberg theorem: the image of a semialgebraic subset of $\mathbb{R}^m$ under a semialgebraic map $\mathbb{R}^m \longrightarrow \mathbb{R}^n$ is a semialgebraic subset of $\mathbb{R}^n$. Clearly this is the same as the existence of a quantifier elimination for the structure $(\mathbb{R}, <, (r)_{r \in \mathbb{R}}, +, \cdot)$ which is slightly weaker than (3). (Semialgebraic = quantifier free definable with parameters in $(\mathbb{R}, <, +, \cdot)$.) Another important property of semialgebraic sets is that they have only finitely many connected components, and that each component is also semialgebraic[1]; see [K] for a nice use of this result.

The basic fact about (5) is that a continuous semialgebraic function $\mathbb{R}^n \longrightarrow \mathbb{R}$ is bounded in absolute value by a polynomial function: this follows easily from (3); an important application occurs in [Hör, p. 276]. I cannot resist giving one other beautiful application: <u>A simple proof, due to K. McKenna, that the inverse of a bijective polynomial map $p: \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is also a polynomial map</u>. From complex analysis we know that $p^{-1}$ is holomorphic, and clearly the real valued function $z \longmapsto |p^{-1}(z)|$ is continuous and semialgebraic, hence bounded by a (real) polynomial function (identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$). Therefore, by Liouville, $p^{-1}$ is a polynomial map. Q.E.D.

How to extend all this to $(\mathbb{R}, +, \cdot, \exp)$?

It seems to me that concentrating most attention on the analogue of (1), that is, decidability of the elementary theory, is a waste of time: consider for example the perplexing problem of deciding the statements

$$p(e, e^e, e^{e^e}, \dots) = 0 ,$$

with $p \in \mathbb{Z}[X_1, X_2, X_3, \dots]$. (And this is just a tiny part of the <u>quantifier free</u> part of the theory.) A natural 'exponential' analogue of <u>'real closed field'</u> does not seem likely (but see [vdD 1] [D-W]), so I don't expect

[1] This property is not an obvious consequence of Tarski's work; see also the end of this Introduction.

an attractive analogue of (2) for our exponential structure.

More plausible problems arise in the attempt to extend (3), (4) and (5). To explain this let us go back to the result that each semialgebraic set has only finitely connected components, each semialgebraic. This follows from Collins [C][2] in which a new decision method for the reals is constructed, much more time efficient than Tarski's. But this efficiency aspect does not concern us here: we are interested in Collins' key <u>geometric</u> idea, which he calls "cylindrical decomposition"; it is partly an alternative to, partly a considerable sharpening of the notion of quantifier elimination. In (3.6) we shall define what a cylindrical decomposition of a set $X \subset \mathbb{R}^n$ is. For the moment, we only mention that such a set $X$ is the disjoint union of finitely many <u>cells</u>, a cell being a subset of $\mathbb{R}^n$ homeomorphic to a space $\mathbb{R}^m$, $m \leq n$.

The following considerations indicate that alternatives to (naive) quantifier elimination are quite welcome in the situation we are facing.

## FAILURE OF 'NAIVE' QUANTIFIER ELIMINATION

The example below shows that the elementary theory of $(\mathbb{R}, <, (r)_{r\in\mathbb{R}}, +, \cdot, \exp)$ does not admit elimination of quantifiers. In fact, much more is true:

<u>Proposition</u>. <u>Let</u> $(F_i)_{i\in I}$ <u>be any family of (total) real analytic functions,</u> $F_i : \mathbb{R}^{n_i} \longrightarrow \mathbb{R}$. <u>Then the structure</u> $(\mathbb{R}, <, (r)_{r\in\mathbb{R}}, +, \cdot, (F_i)_{i\in I})$ <u>admits quantifier elimination if and only if each</u> $F_i$ <u>is semialgebraic</u>.

(Note: Semialgebraic = definable in $(\mathbb{R}, <, (r)_{r\in\mathbb{R}}, +, \cdot)$, by Tarski.)

The key to this negative result is the following example which we shall treat before proving the proposition.

<u>Example</u>. (Osgood, see [Ł, p. 133]) Define $f : \mathbb{R} \times \mathbb{R}^{>0} \longrightarrow \mathbb{R}$ by $f(x,y) = y\cdot\exp(x/y)$. Then its graph $G(f)$ is the subset $\{(x,y,z) \mid y > 0 \wedge \exists t(z = y\cdot\exp(t) \wedge ty = x)\}$ of $\mathbb{R}^3$ which is obviously definable in $(\mathbb{R}, <, \cdot, \exp)$.

[2] The result is actually due to H. Whitney. See [Ł, p. 110] for an elegant proof.

Claim. There are no (real) analytic functions $F_1,\dots,F_k : U \longrightarrow \mathbb{R}$, $U$ an open ball in $\mathbb{R}^3$ centered at $0$, such that $G(f) \cap U$ belongs to the boolean algebra generated by the subsets $\{F_i > 0\}$, $\{F_i = 0\}$ of $U$.

Proof. The crucial facts about $f$ are: $f(\lambda x, \lambda y) = \lambda f(x,y)(\lambda > 0, y > 0)$, and $f$ is not an "algebraic" function. Suppose the functions $F_i : U \longrightarrow \mathbb{R}$, $i = 1,\dots,k$, have the property we want to refute; we may of course assume that none of the $F_i$ is identically zero. At least one of the $F_i$, say $F_1$, must vanish on $G(f) \cap U$ (otherwise there would be $c \in G(f) \cap U$ with $F_i(c) \neq 0$ for each $i$, and $G(f)$ would contain a whole neighborhood of $c \in \mathbb{R}^3$). Write $F_1 = P_0 + P_1 + P_2 + \dots$, $P_d$ a homogeneous polynomial of degree $d$.[3)] Let $(x,y,z) \in G(f) \cap U$. Then for all $0 < \lambda < 1$ we have $(\lambda x, \lambda y, \lambda z) \in G(f) \cap U$, and

$0 = F_1(\lambda x, \lambda y, \lambda z) = P_0(x,y,z) + \lambda P_1(x,y,z) + \lambda^2 P_2(x,y,z) + \dots$, hence $P_d(x,y,z) = 0$ for all $d$. Take $d$ with $P_d \neq 0$, and we see that $P_d(x,y,f(x,y)) = 0$ for all $(x,y) \in \mathbb{R} \times \mathbb{R}^{>0}$, and $f$ would be an algebraic function. $\square$

We model the proof of the proposition on the argument just given. Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ be analytic and define $f : \mathbb{R}^n \times \mathbb{R}^{>0} \longrightarrow \mathbb{R}$ by $f(x_1,\dots,x_{n+1}) = x_{n+1} F(x_1/x_{n+1},\dots,x_n/x_{n+1})$. If we assume that its graph $G(f)$ is quantifier free definable in $(\mathbb{R}, <, (r)_{r \in \mathbb{R}}, +, \cdot, (F_i)_{i \in I})$, then we derive exactly as before that there is a nonzero real polynomial $P$ in $n+2$ variables with $P(x,f(x)) = 0$ for all $x \in \mathbb{R}^n \times \mathbb{R}^{>0}$. Substituting a suitable value $\lambda > 0$ for the $n+1^{st}$ variable, we derive from this that the analytic function $F_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}$, given by $(x_1,\dots,x_n) \longmapsto \lambda F(x_1/\lambda,\dots,x_n/\lambda)$ satisfies identically an equation $Q(x, F_\lambda(x)) = 0$, $x \in \mathbb{R}^n$, $Q$ a nonzero real polynomial in $n+1$ variables. The lemma below shows that then $F_\lambda$ is a semialgebraic function, hence $F$ is also semialgebraic. $\square$

---

3) We assume here that $U$ is taken so small that this Taylor series of $F_1$ converges on $U$.

Lemma. If a continuous function $g:\mathbb{R}^n \longrightarrow \mathbb{R}$ satisfies identically an equation $Q(x,g(x)) = 0$ , $x \in \mathbb{R}^n$ , where $Q$ is a nonzero real polynomial in n+1 variables, then $g$ is semialgebraic.

Proof. From the results on pp. 106-110 of [Ł], it follows that $\mathbb{R}^n$ can be partitioned into semialgebraic subsets $A_0, A_1, \ldots, A_m$ with $A_1, \ldots, A_m$ connected such that if $x \in A_0$ , then $Q(x,Y)$ vanishes identically, while if $1 \leq i \leq m$ and $x \in A_i$ , the real roots of $Q(x,Y)$ are given by continuous functions $\rho_{i1}(x) < \ldots < \rho_{ik(i)}(x)$ of $x$ . (Obviously, these functions $\rho_{ij}:A_i \longrightarrow \mathbb{R}$ are semialgebraic.) By continuity and connectedness, $g$ must coincide on each $A_i$ $(i > 0)$ with one of the functions $\rho_{ij}$ , hence $g$ restricted to $A_1 \cup \ldots \cup A_m$ is semialgebraic. Since every $p \in A_0$ is the limit of points in $A_1 \cup \ldots \cup A_m$ , the value of $g$ at $p$ is semialgebraically determined by its values on $A_1 \cup \ldots \cup A_m$ . It follows that $g$ is semialgebraic. □

The proposition forces us to look for new ways of solving (the realistic part of) Tarski's problem.

A line of attack which seems quite promising to me goes roughly as follows: Define a k-manifold, $k \in \mathbb{N}$ , to be a real analytic submanifold $M$ of some $\mathbb{R}^n$ , $n \geq k$ , equipped with an analytic isomorphism $h_M:M \simeq \mathbb{R}^k$ . We introduce for each $k$ a class $\mathscr{M}^{(k)}$ of k-manifolds, and for each manifold $M \in \mathscr{M}^{(k)}$ an $\mathbb{R}$-algebra $\mathscr{A}(M)$ of (real) analytic functions on $M$ . The classes $\mathscr{M}^{(k)}$ and, for each $M \in \mathscr{M}^{(k)}$ , the algebra $\mathscr{A}(M)$ , are constructed in stages. (N.B., as soon as we have $M$ , we do not necessarily introduce all of $\mathscr{A}(M)$ at once.)

At stage 0 we introduce all semialgebraic k-manifolds and all semialgebraic (analytic) functions on them. To obtain new manifolds, and functions on them, as well as new functions on the manifolds already available, we use several constructions of which the following three are the most important.

Let $M$ be a k-manifold already available.

(i) If $f,g:M \longrightarrow \mathbb{R}$ are already available, then we introduce the k-manifold

$$\mathrm{graph}(f) \stackrel{\mathrm{def}}{=} \{x,f(x)) : x \in M\} \quad ,$$

the k+1-manifold $(f,g)_M \stackrel{\mathrm{def}}{=} \{(x,y) : f(x) < y < g(x)\}$ (if $f < g$ on $M$), and the k+1-manifolds $(-\infty,f)_M$, $(f,\infty)_M$, and $(-\infty,\infty)_M$, defined similarly.

(ii) If $f_0,\ldots,f_d:M \longrightarrow \mathbb{R}$ are available, then we introduce the function $f_0 + f_1 Y + \ldots + f_d Y^d$ on $M \times \mathbb{R} = (-\infty,\infty)_M$, and also its restrictions to $\mathrm{graph}(f)$, $(f,g)_M$, $(-\infty,f)_M$, $(f,\infty)_M$ (where $f,g$ are as in (i)), as well as all those analytic functions which are algebraic over the 'polynomial' functions so obtained.

(iii) Let $X_1,\ldots,X_k$ be the coordinate functions on $M$ defined by the chart $h_M:M \simeq \mathbb{R}^k$. Then we introduce as new functions on $M$ those analytic functions $f:M \longrightarrow \mathbb{R}$ which satisfy a system of differential equations

$$\frac{\partial f}{\partial X_i} = F_i(X_1,\ldots,X_k,f) \quad , \quad i = 1,\ldots,k \quad ,$$

where each $F_i$ is a function already available, e.g., via constructions (i) and (ii).

Now the delicate part is to show that the zeroset of each function in $\mathscr{A}(M)$, $M \in \mathscr{M}^{(k)}$, as well as its complement, is a disjoint union of finitely many manifolds in $\bigcup_{i \le k} \mathscr{M}^{(i)}$. This can only be proved if the construction is done in exactly the right order (the 'stages' have to be indexed by a suitably chosen well ordering), and if the correct induction hypotheses are selected. (The induction hypothesis on the "asymptotic behavior of zerosets" seems to be the crucial problem, but I have an idea as to what it should be. Since precise statements are complicated and I expect to come back to it later, I'll leave it at these vague indications.)

Remarks

(1) All functions built up from real constants, variables, and the operations $+$, $\cdot$, and exp are eventually introduced via construction (iii), but it is

interesting that it is the system of differential equations which is crucial to prove the required results for these functions (and their zerosets), and not the defining expression in terms of $+$ , $\cdot$ , and exp .[4)]

(2) If the program we just sketched works out, then one obtains a Tarski-Seidenberg theorem for a very large class of sets and functions, including all those definable in our exponential structure, as well as the result that each of these sets is a disjoint union of finitely many k-manifolds. Quite apart from its interest for Tarski's problem, it would be very desirable to obtain results in this direction, for example, in connection with Hilbert's 16th problem, see [Hov], [P].

(3) References [vdD2] and [Hov] give partial results in the spirit of the program sketched above, and led me to the present formulation.

In this paper we shall carry out a more relaxed investigation which has the advantage of introducing the notion of cylindrical decomposition in the most natural and painless way: let $\mathscr{R}$ be any expansion of $(\mathbb{R},<)$ . A subset of $\mathbb{R}^m$ is called definable if it is definable by a formula in the language of $\mathscr{R}$ . Under a convenient assumption on $\mathrm{Th}(\mathscr{R})$ - to be introduced in the next section, and obviously satisfied if $\mathscr{R} = (\mathbb{R}, <, +, \cdot)$ - we shall prove, inter alia:

(*) Each definable function $\mathbb{R} \longrightarrow \mathbb{R}$ is piecewise continuous, c.f. (2.2).

(**) Each definable subset of $\mathbb{R}^m$ is the disjoint union of finitely many cells, each of which is also definable. See (3.11). (Cell = space homeomorphic to an $\mathbb{R}^k$.)

(It is of some interest that this gives a proof, based on quite general principles, of Whitney's "finite number of components" result on semialgebraic sets.) If we assume moreover that $\mathscr{R}$ is an expansion of $(\mathbb{R}, <, +)$ , then

4) The treatment in [vdD2] could have been simplified and generalized considerably if this point had been clearer to me.

(*) and (**) can be strengthened in the sense that in (*) the function $f$ is, for each $n$, piecewise $C^n$, and in (**), one can take, given any $n$, the cells as $C^n$-submanifolds of $\mathbb{R}^m$. In an appendix we shall indicate the proof of this and make some remarks on the connection with the important notion of Whitney stratification.

## GENERAL CONVENTIONS

In this article we consider L-structures $\mathscr{A} = (A,\ldots)$ and $\mathscr{R} = (\mathbb{R},\ldots)$, L a first-order language with equality. An L-formula $\phi(v_1,\ldots,v_n)$, $n \geq 1$, is said to define the set $\{a_1,\ldots,a_n) \in A^n \mid \mathscr{A} \models \phi(a_1,\ldots,a_n)\}$ (in $\mathscr{A}$), and a set of this form is called a definable subset of $A^n$. If $\phi = \phi(v_1,\ldots,v_n,v_{n+1})$ is given and $x \in A^n$, then $\phi(x,A)$ denotes the set $\{a \in A \mid \mathscr{A} \models \phi(x,a)\}$; these sets are special cases of A-definable sets: A-definable subsets of $A^n$ generalize definable subsets of $A^n$ in that their defining formula are $L_A$-formulas, i.e., we allow parameters.

If $X \subset A^m$, $Y \subset A^n$ we identify $X \times Y$ in the usual way with the subset $\{(x_1,\ldots,x_m,y_1,\ldots,y_n) \in A^{m+n} \mid (x_1,\ldots,x_m) \in X, (y_1,\ldots,y_n) \in Y\}$ of $A^{m+n}$. A function $f:X \longrightarrow A (X \subset A^m)$ is called (A-)definable if its graph $G(f) = \{(x,f(x)) \mid x \in X\} \subset A^{m+1}$ is (A-)definable.

We let $k,\ell,m,n$ vary over elements of $\mathbb{N} = \{0,1,2,\ldots\}$ and write $\#X$ for the number of elements of $X$ if $X$ is finite; otherwise $\#X = \infty$.

## §1. Theories of finite type

(1.1) Let $(A,<)$ be a nonempty dense linear order without endpoints. For notational purposes we adjoin two 'endpoints' $-\infty$, $+\infty$, and put $-\infty < a < +\infty$ for all $a \in A$. Let $\hat{A} = A \cup \{-\infty,+\infty\}$. An interval is a subset $(a,b) \overset{\text{def}}{=} \{r \in A \mid a < r < b\}$ of $A$, where $-\infty \leq a < b \leq +\infty$. Note that $a$ and $b$ are uniquely determined by the interval. An interval is clearly infinite and $(A,<)$-definable.

(1.2) Definition. A subset $X$ of $A$ is called of finite type if $X$ is the union of a finite set and finitely many intervals.

Example: If $(A,<) = (\mathbb{R},<)$ then a set $X \subset \mathbb{R}$ is of finite type if and only if $X$ has only finitely many connected components.

(1.3) Definition. For $X \subset A$ we define the boundary of $X$ to be the set

$$\mathrm{bd}(X) = \{a \in A \mid \text{every interval containing } a \text{ contains both a point from } X \text{ and a point outside } X\} .$$

Note that for $(A,<) = (\mathbb{R},<)$ this gives the usual notion of boundary of $X$ as a subset of the topological space $\mathbb{R}$.

(1.4) Lemma. Suppose $X \subset A$ is of finite type. Then:

(1) $\mathrm{bd}(X)$ is finite;

(2) if $\mathrm{bd}(X) = \{a_1,\ldots,a_m\}$ where $a_1 < \ldots < a_m$, then, putting $a_0 = -\infty$, $a_{m+1} = \infty$, each interval $(a_i, a_{i+1})$, $i = 0,\ldots,m$, either is part of $X$ or disjoint from $X$.

Proof. Left to the reader. □

(1.5) Definition. Let $X \subset A$ be of finite type and $\{a_1,\ldots,a_m\}$ its boundary, $a_1 < \ldots < a_m$. Then the type of $X$ is the sequence $\tau_X = <\tau_1,\tau_2,\ldots,\tau_{2m+1}>$ where for each $i = 0,\ldots,m$ :

$$\tau_{2i+1} = \begin{cases} +1 & \text{if } (a_i,a_{i+1}) \subset X \\ -1 & \text{if } (a_i,a_{i+1}) \subset A \setminus X \end{cases}$$

and for each $i = 1,\ldots,m$ :

$$\tau_{2i} = \begin{cases} +1 & \text{if } a_i \in X \\ -1 & \text{if } a_i \notin X . \end{cases}$$

A type is a finite sequence $\tau = <\tau_1,\ldots,\tau_N>$, all whose terms $\tau_i$ are in $\{-1, +1\}$. $N$ is called the length of $\tau$.

(1.6) <u>In the remainder of this section we assume that the language</u> $L$ <u>contains a binary predicate symbol</u> $<$ <u>and that the</u> L-<u>structure</u> $\mathscr{A} = (A,<,\dots)$ <u>is an expansion of the dense linear ordering without endpoints</u> $(A,<)$ .

<u>Definition</u>. We call an L-theory $T$ (extending the theory of dense linear order without endpoints) <u>of finite type</u> if each model $\mathscr{A} = (A,<,\dots)$ has the property that each $A$-definable subset of $A$ is of finite type.

We call $\mathscr{A} = (A,<,\dots)$ of finite type if $\mathrm{Th}(\mathscr{A})$ is of finite type. (This condition implies that each $A$-definable subset of $A$ is of finite type, but the converse implication is probably not valid.)

(1.7) <u>Lemma</u>. For each L-formula $\phi(v_1,\dots,v_{n+1})$ there is a formula $(\mathrm{bd}\,\phi)(v_1,\dots,v_{n+1})$ such that for all $a \in A^n$ :

$$\mathrm{bd}(\phi(a,A)) = (\mathrm{bd}\,\phi)(a,A) \quad .$$

<u>Proof</u>. Clear. □

(1.8) <u>Lemma</u>. If $\mathscr{A} = (A,<,\dots)$ is of finite type and $\phi(v_1,\dots,v_{n+1})$ is an L-formula, then there is $k = k_\phi \in \mathbb{N}$ such that $\#\,\phi(a,A) \in \{0,1,\dots,k\} \cup \{\infty\}$ for all $a \in A^n$ .

<u>Proof</u>. With respect to $\mathrm{Th}(\mathscr{A})$ the 'infinitary formula' $\exists^{\infty} v_{n+1}\, \phi(v_1,\dots,v_{n+1})$ is equivalent to the formula $\exists x\, \exists y (x < y \wedge \forall v_{n+1}(x < v_{n+1} < y \longrightarrow \phi(v_1,\dots,v_{n+1})))$, and we can use a simple model theoretic compactness argument to establish the existence of a number $k$ as required. □

(1.9) <u>Proposition</u>. Let $\mathscr{A} = (A,<,\dots)$ be of finite type and $\phi(v_1,\dots,v_{n+1})$ an L-formula. Then there is a finite partition of $A^n$ into disjoint definable subsets $X_1,\dots,X_M$ , and there are distinct types $\tau(1),\dots,\tau(M)$ such that for each $i$ and $x \in X_i$ the set $\phi(x,A)$ is of type $\tau(i)$ .

<u>Proof</u>. There is $k \in \mathbb{N}$ such that for each $x \in A^n$ we have $\#(\mathrm{bd}\,\phi(x,A)) \le k$ , by (1.4), (1.7) and (1.8). Now each set $\phi(x,A)$ is of

finite type, and the type of $\phi(x,A)$ is moreover of length $\leq 2k+1$, so only finitely many types actually occur among the types of the sets $\phi(x,A)$, $x \in A^n$. Let $\tau(1),\ldots,\tau(M)$ be these types and define $X_i$ as the set of all $x \in A^n$ for which $\phi(x,A)$ has type $\tau(i)$. $\square$

Remarks

(1) For $x \in X_i$ let $\tau(i)$ have length $2m(i)+1$, and let $\mathrm{bd}\,\phi(x,A) = \{f_{i,1}(x),\ldots,f_{i,m(i)}(x)\}$ where $f_{i,1}(x) < \ldots < f_{i,m(i)}(x)$. Then the functions $f_{i,j}: X_i \longrightarrow A$ are definable in $\mathscr{A}$.

(2) The proposition will be used later in an inductive argument (see (3.10).

## §2. Piecewise continuity of definable functions

Throughout this section $\mathscr{R}$ is an expansion of $(\mathbb{R},<)$ with the property that each $\mathbb{R}$-definable subset of $\mathbb{R}$ has only finitely many connected components. ( This assumption seems to be somewhat weaker than: $\mathscr{R}$ is of finite type.)

(2.1) Lemma. If $I \subset \mathbb{R}$ is an interval and $f: I \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, then $f$ is continuous in at least one point of $I$.

Proof. Assume first that there is a subinterval $J$ of $I$ whose image under $f$ is finite. Then $J$ is the union of the finitely many sets $J \cap f^{-1}(b)$, $b \in f(J)$, so there is $b$ such that $J \cap f^{-1}(b)$ is infinite, hence the $\mathbb{R}$-definable set $J \cap f^{-1}(b)$ must contain a subinterval $J'$ of $J$. Then $f$ is constant on $J'$, hence $f$ is continuous at each point of $J'$.

In the remainder of this proof we shall assume that each subinterval of $I$ has infinite image under $f$. Inductively we shall find a descending sequence of closed segments $[a_n,b_n] \subset I$ with $0 < b_n - a_n < 1/n$, $n \geq 1$, such that $f[a_n,b_n]$ is contained in an (open) interval $J_n$ of length $< 1/n$. It is clear that then $f$ is continuous at the unique point in $\bigcap_{n=1}^{\infty} [a_n,b_n]$.
$[a_1,b_1]$ and $J_1$ are obtained as follows: for $J_1$ we take an interval of length $< 1$ contained in the $\mathbb{R}$-definable infinite set $f(I)$, and $[a_1,b_1]$ is chosen as a closed segment contained in the $\mathbb{R}$-definable infinite set

$f^{-1}(J_1)$, with $0 < b_1 - a_1 < 1$. Given $[a_n, b_n]$ we choose an interval $J_{n+1} \subset f[a_n, b_n]$ of length $< 1/n+1$; then we choose $[a_{n+1}, b_{n+1}] \subset f^{-1}(J_{n+1})$ with $0 < b_{n+1} - a_{n+1} < 1/n+1$. □

(2.2) Corollary. If $f:(a,b) \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, $a < b$, then $f$ is piecewise continuous, that is, there are $a_0 < a_1 < \ldots < a_m$, $a_0 = a$, $a_m = b$, such that $f$ is continuous on each subinterval $(a_i, a_{i+1})$.

Proof. The set of points in $(a,b)$ at which $f$ is not continuous is obviously $\mathbb{R}$-definable, but it cannot contain a whole subinterval of $(a,b)$ by the previous lemma. Hence, it must be a finite set. □

(2.3) Lemma. If $I \subset \mathbb{R}$ is an interval and $f: I \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable then there is a subinterval of $I$ on which $f$ is constant or strictly monotone.

Proof. By the previous corollary we may as well assume that $f$ is continuous. Take $a, b \in I$, $a < b$, and assume that $f$ is not constant on $[a,b]$. Then $f[a,b]$ contains a segment $[c,d]$, $c < d$, and we define $g:[c,d] \longrightarrow [a,b]$ by: $g(y) = \min\{x \in [a,b] : f(x) = y\}$. By definition $g$ is injective, and by the previous corollary $g$ is continuous on a subsegment $[c',d']$ of $[c,d]$, so $g$ is strictly monotone on $[c',d']$ and therefore $f$ is strictly monotone on the image of $[c',d']$ under $g$ (since $f(gy) = y$), and this image contains an interval. □

(2.4) Corollary. If $f:(a,b) \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, then there are $a_0 < a_1 < \ldots < a_m$, $a_0 = a$, $a_m = b$, such that for each $i < m$ $f$ is either constant on $(a_i, a_{i+1})$ or continuous and strictly monotone on $(a_i, a_{i+1})$. In particular $\lim_{x \downarrow a} f(x)$ and $\lim_{x \uparrow b} f(x)$ exist in $\hat{\mathbb{R}}$.

Proof. The set of points in $(a,b)$ which have no neighborhood on which $f$ is constant, or continuous and strictly monotone, is $\mathbb{R}$-definable, and cannot contain a whole subinterval of $(a,b)$, by the previous lemma. So it must be a finite set $\{a_1, \ldots, a_{m-1}\}$, with $a = a_0 < a_1 < \ldots < a_{m-1} < a_m = b$. It is easy to see that on each subinterval $(a_i, a_{i+1})$ the function $f$ is either constant, or continuous and strictly monotone. □

§3. Cylindrical Decomposition

(3.1) This section contains the main result, Theorem (3.7). First some conventions and definitions. Recall that $\hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ . We fix once and for all an expansion $\mathcal{R}$ of $(\mathbb{R},<)$ which is of finite type, c.f. (1.6). (So the results of Sections 1 and 2 are available to us.) The term 'definable' as applied to subsets of $\mathbb{R}^m$ and functions between subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ refers to definability by an L-formula in the structure $\mathcal{R}$ .

(3.2) Given a nonempty definable subset $X$ of $\mathbb{R}^m$ , $m \geq 0$ , we let $C(X)$ be the set of definable continuous functions $f:X \longrightarrow \mathbb{R}$ , and we put $\hat{C}(X) = C(X) \cup \{-\infty, \infty\}$ . For $f,g \in \hat{C}(X)$ we write $f < g$ (or $g > f$ ) if $f(x) < g(x)$ for all $x \in X$. (Here $-\infty$ and $\infty$ are identified with the corresponding constant functions defined on $X$ .) If $f \in C(X)$ we let $G(f) = \{(x,f(x)):x \in X\} \subset \mathbb{R}^{m+1}$ be the graph of $f$ , and if $f,g \in \hat{C}(X)$ , $f < g$ , we write $(f,g)_X$ for the set $\{(x,y)|x \in X$ and $f(x)<y<g(x)\} \subset \mathbb{R}^{m+1}$ . It is helpful to view $(f,g)_X$ as the family of intervals $(f(x),g(x))_{x \in X}$ parametrized by $X$ . Note that if $X$ is homeomorphic to $\mathbb{R}^k$ , then $G(f)$ is also homeomorphic to $\mathbb{R}^k$ , and $(f,g)_X$ is homeomorphic to $\mathbb{R}^{k+1}$ .

(3.3) By induction we define for each $n \geq 0$ and each finite sequence $(i_1,\ldots,i_n)$ of zeros and ones a collection $F(i_1,\ldots,i_n)$ of nonempty definable subsets of $\mathbb{R}^n$ .

$n = 0$: there is only the empty sequence $\emptyset$ and $F\emptyset$ has only one member, namely $\mathbb{R}^0 = \{0\}$ .

$n = 1$: $F(0)$ consists of all $\{a\}$ with $a \in \mathbb{R}$ definable in $\mathcal{R}$ .
$F(1)$ consists of all intervals of $\mathbb{R}$ whose (finite) endpoints are definable in $\mathcal{R}$ .

More generally, given the collection $F(i_1,\ldots,i_n)$ we define $F(i_1,\ldots,i_n,0)$ to consists of all graphs $G(f)$ with $f \in C(X)$ , $X \in F(i_1,\ldots,i_n)$ , and we define $F(i_1,\ldots,i_n,1)$ to consist of all sets $(f,g)_X$ where $X \in F(i_1,\ldots,i_n)$, $f,g \in \hat{C}(X)$ , $f < g$ . (Note: this generalizes the definition of $F(0)$ and $F(1)$ .)

Clearly, for given $n$, the $2^n$ collections $F(i_1,\ldots,i_n)$ are mutually disjoint. Also, if $X \in F(i_1,\ldots,i_n)$, then $X \subset \mathbb{R}^n$ is homeomorphic to $\mathbb{R}^k$, where $k = i_1 + \ldots + i_n$. We put $\dim(X) = k$. Note: $\dim(X) = n$ if and only if $X$ is open in $\mathbb{R}^n$. (All this is immediate by induction on $n$.) We let $F^{(n)}$ be the union of the $2^n$ collections $F(i_1,\ldots,i_n)$, so $F^{(n)}$ is a collection of definable subsets of $\mathbb{R}^n$.

(3.4) For $X \in F^{(n)}$ such that $\dim(X) < n$ we shall define a definable homeomorphism

$$h_X : X \simeq h(X) \quad \text{where} \quad h(X) \in F^{(n-1)} \ .$$

If $\dim(X) = \dim(hX) < n-1$, then one can repeat the same construction with $h(X)$. By induction this shows that if $\dim X = k$, then $X$ is definably homeomorphic to an open subset of $\mathbb{R}^k$ belonging to $F^{(k)}$.

In case $X$ is a graph $G(f)$ where $f \in C(Y)$, $Y \in F^{(n-1)}$ we put $Y = h(X)$ and let $h_X : X \longrightarrow Y$ be the projection map $(y,r) \longmapsto y$.

Suppose $X = (f,g)_Y$ where $Y \in F^{(n-1)}$, $f < g$, $f,g \in \hat{C}(Y)$. Then $\dim(X) < n$ implies $\dim(Y) < n-1$, so by recursion we may assume that $h_Y : Y \simeq h(Y)$ has already been defined. We get $fh_Y^{-1}$, $gh_Y^{-1} \in \hat{C}(hY)$, and we let $h(X)$ be the set $(fh_Y^{-1}, gh_Y^{-1})_{h(Y)}$. Clearly $h_Y$ lifts to a homeomorphism $h_X : X \simeq h(X) : (y,r) \longmapsto (h_Y(y),r)$.

(3.5) For each set $X$ in $F^{(n)}$ we shall define a collection $\mathrm{Dec}(X)$ of finite partitions of $X$ into sets also belonging to $F^{(n)}$. We call a member of $\mathrm{Dec}(X)$ a decomposition of $X$.

For $n = 0$ we have $X = \mathbb{R}^0$, since $F^{(0)} = \{\mathbb{R}^0\}$, and the only (finite) partition of $X$ is $\{X\}$; we let $\{X\}$ be the only member of $\mathrm{Dec}(X)$.

Let $n \geq 0$ and suppose $\mathrm{Dec}(X)$ has been defined for all $X \in F^{(n)}$. Let $X \in F^{(n)}$ and $f \in C(X)$. Then each finite partition $\mathscr{D} = \{A_1,\ldots,A_d\}$ of $X$ lifts to one of $G(f)$, namely $\{\pi^{-1}(A_1),\ldots,\pi^{-1}(A_d)\}$, where $\pi : G(f) \longrightarrow X$ is the projection on the first $n$ coordinates. Let us call this 'lifted' partition $\mathscr{D}^f$ . Then we define $\operatorname{Dec}(G(f)) = \{\mathscr{D}^f \mid \mathscr{D} \in \operatorname{Dec}(X)\}$ .

The definition of $\operatorname{Dec}((f,g)_X)$ for $f,g \in \hat{C}(X)$ , $f < g$ , is a bit more complicated. Let again $\mathscr{D} = \{A_1,\ldots,A_d\}$ be a decomposition of $X$ , i.e., $\mathscr{D} \in \operatorname{Dec}(X)$ . Then $\mathscr{D}$ induces a partition

$$\{(f|A_1,g|A_1)_{A_1},\ldots,(f|A_d,g|A_d)_{A_d}\}$$

of $(f,g)_X$ , which we denote by $\mathscr{D}^{f,g}$ .

A second kind of partition of $(f,g)_X$ is obtained as follows: let $f_0,f_1,\ldots,f_m \in \hat{C}(X)$ with $f = f_0 < f_1 < \ldots < f_m = g$ on $X$ . Then we have the partition of $(f,g)_X$ consisting of the sets $(f_i,f_{i+1})_X$ , $i = 0,\ldots,m-1$ , and the sets $G(f_i)$ , $0 < i < m$ . We denote this partition by $(f_0|f_1|\ldots|f_m)_X$ and call it a <u>proper</u> decomposition of $(f,g)_X$ . ("Proper" because it is not induced by a decomposition of $X$ .)

Suppose now that $\mathscr{D} = \{A_1,\ldots,A_d\} \in \operatorname{Dec}(X)$ and that for each $A_i$ a <u>proper</u> decomposition $\mathscr{D}_i$ of $(f|A_i,g|A_i)_{A_i}$ is given. Then $\mathscr{D}_1 \cup \ldots \cup \mathscr{D}_d$ is a partition of $(f,g)_X$ which we denote by $<\mathscr{D};\{\mathscr{D}_1,\ldots,\mathscr{D}_d\}>$ . We define $\operatorname{Dec}((f,g)_X)$ as consisting of all partitions $<\mathscr{D};\{\mathscr{D}_1,\ldots,\mathscr{D}_d\}>$ . We call $\mathscr{D}$ the <u>base</u> of the decomposition $<\mathscr{D};\{\mathscr{D}_1,\ldots,\mathscr{D}_d\}>$ .

(3.6) Let us say that a decomposition $\mathscr{D}$ of $\mathbb{R}^n$ <u>partitions</u> a set $A \subset \mathbb{R}^n$ if $A$ is the union of a subcollection of $\mathscr{D}$ , in other words, if each set in $\mathscr{D}$ is contained in $A$ or disjoint from $A$ .

The following terminology will not be used in the remainder of this section, but it might be helpful in comparing with results in [C]. A <u>cylindrical decomposition</u> is a decomposition of $\mathbb{R}^n$ , that is, a member of $\operatorname{Dec}(\mathbb{R}^n)$ , for some $n$ . A cylindrical decomposition which partitions a set $A \subset \mathbb{R}^n$ is also called a <u>cylindrical</u> <u>decomposition</u> of $A$ . With this terminology the next theorem states that for any definable sets $A_1,A_2,\ldots,A_m \subset \mathbb{R}^n$ there is a common cylindrical decomposition.

(3.7) <u>Theorem</u>. <u>Let</u> $n \geq 1$ .

(a) For each definable set $X \subset \mathbb{R}^{n-1}$ and definable $f: X \longrightarrow \mathbb{R}$ there is a finite partition of $X$ into definable sets on each of which $f$ is continuous.

(b) Given any definable subsets $A_1, \ldots, A_m$ of $\mathbb{R}^n$ there is $\mathscr{D} \in \mathrm{Dec}(\mathbb{R}^n)$ partitioning each of $A_1, \ldots, A_m$ .

The proof is by induction on $n$ . For $n = 1$ statement (a) holds trivially (since $X = \emptyset$ or $X$ is a singleton), and (b) is an easy exercise in boolean algebra using the fact that each $A_i \subset \mathbb{R}$ is of finite type. The induction steps take place in the proof of the following three lemmas.

(3.8) Lemma. Suppose $N > 1$ and statements (a) and (b) of (3.7) hold for all $n < N$ . Then statement (a) holds for $n = N$ .

Proof. Let $X \subset \mathbb{R}^{N-1}$ be definable and $f: X \longrightarrow \mathbb{R}$ definable. Using (b) for $n = N-1$ we may restrict to the case that $X \in F^{(N-1)}$ .

Suppose first that $\dim X < N-1$ . Then, by (3.4), we have a definable homeomorphism $h: X \simeq X'$ where $X' \in F^{(N-2)}$ . The function $fh^{-1}: X' \longrightarrow \mathbb{R}$ is definable, so by the induction hypothesis there is a finite partition of $X'$ into definable sets on each of which $fh^{-1}$ is continuous. If we lift this partition via $h$ we get a finite partition of $X$ into definable sets on each of which $f$ is continuous.

We are left with the case $\dim X = N-1$ . Then $X$ is open in $\mathbb{R}^{N-1}$ .

Claim. The set $Y = \{x \in X \mid f \text{ is continuous at } x\}$ is dense in $X$ .

If we accept this claim for a moment then, since $Y$ is definable, we can use (b) for $n = N-1$ to get a decomposition $\mathscr{D}$ of $\mathbb{R}^{N-1}$ which partitions both $Y$ and $X \setminus Y$ . The sets in $\mathscr{D}$ which are contained in $X$ and are open must be contained in $Y$ , so $f$ is continuous on each of those sets. The sets in $\mathscr{D}$ which are contained in $X$ but not open are of dimension $< N-1$ , and so the arguments used in the case $\dim X < N-1$ apply.

To prove the claim we take any nonempty open set $U \subset X$ . To show that $f$ is continuous in at least one point of $U$ we use exactly the same arguments as in the proof of (2.1). The role of the decreasing segments $[a_n, b_n]$ is now of course taken over by a decreasing sequence of closed balls contained in $U$, their diameter tending to $0$. (Here we let the distance on $\mathbb{R}^{N-1}$ be defined by the norm $\|(x_1,\ldots,x_{N-1})\| = \max(|x_1|,\ldots,|x_{N-1}|)$, in order that the balls are $\mathbb{R}$-definable.) $\square$

(3.9) <u>Lemma</u>. Assume that $N > 1$ and statement (b) of (3.7) holds for $n = N-1$. Then any two decompositions $\mathscr{D}_1$ and $\mathscr{D}_2$ of $\mathbb{R}^N$ have a common refinement, that is, there is a decomposition of $\mathbb{R}^N$ partitioning each of the sets in $\mathscr{D}_1 \cup \mathscr{D}_2$.

<u>Proof</u>. Note that $\mathbb{R}^N = (-\infty,\infty)_X$ where $X = \mathbb{R}^{N-1}$. This gives us the structure of decompositions of $\mathbb{R}^N$, see (3.5). So we can use the assumption that (b) holds for $n = N-1$ to reduce to the case that $\mathscr{D}_1$ and $\mathscr{D}_2$ have a common base $\mathscr{D}$, say

$$\mathscr{D}_1 = <\mathscr{D};\{\mathscr{D}_{11},\ldots,\mathscr{D}_{1d}\}> \quad ,$$

$$\mathscr{D}_2 = <\mathscr{D};\{\mathscr{D}_{21},\ldots,\mathscr{D}_{2d}\}> \quad ,$$

where $\mathscr{D} = \{B_1,\ldots,B_d\}$ is a decomposition of $\mathbb{R}^{n-1}$. Fix a set $B_i$ in $\mathscr{D}$ and say $\mathscr{D}_{1i}$, $\mathscr{D}_{2i}$ are (proper) decompositions of $(-\infty,\infty)_{B_i}$,

$$\mathscr{D}_{1i} = (-\infty|f_1|\ldots|f_p|\infty)_{B_i} \quad ,$$

$$\mathscr{D}_{2i} = (-\infty|g_1|\ldots|g_q|\infty)_{B_i} \quad ,$$

where the $f_\lambda$, $g_\mu : B_i \longrightarrow \mathbb{R}$ are definable and continuous. Now clearly $B_i$ can be partitioned into finitely many definable sets $B_{ij}$ such that, given any $B_{ij}$, $f_\lambda$, and $g_\mu$, we have: either for all $x \in B_{ij}$: $f_\lambda(x) < g_\mu(x)$,
or for all $x \in B_{ij}$: $f_\lambda(x) = g_\mu(x)$,
or for all $x \in B_{ij}$: $f_\lambda(x) > g_\mu(x)$.

By the assumption that (b) holds for $n = N-1$ there is $\mathscr{D}^* \in \mathrm{Dec}(\mathbb{R}^{N-1})$ such that $\mathscr{D}^*$ partitions each of the sets $B_{ij}$. Now it should be clear how to construct a decomposition of $\mathbb{R}^N$ with base $\mathscr{D}^*$ (and using restrictions of the $f_\lambda$ and $g_\mu$) which partitions each set in $\mathscr{D}_1 \cup \mathscr{D}_2$ . □

Note that Lemma (3.9) immediately extends to a finite collection of decompositions of $\mathbb{R}^N$ . It is this stronger form that we shall use in the final step which follows now.

(3.10) Lemma. Suppose $N > 1$ , statement (a) of (3.7) holds for $n = N$ and statement (b) holds for $n = N-1$ . Then statement (b) holds for $n = N$ .

Proof. Let definable sets $A_1,\ldots,A_m$ of $\mathbb{R}^N$ be given. Fix an $A_i$ and consider $A_i$ as lying over $\pi(A_i)$ where $\pi:\mathbb{R}^N \longrightarrow \mathbb{R}^{N-1}$ is the projection on the first $N-1$ coordinates. We are going to apply (1.9) to the defining formula for $A_i$ , taking into account Remark (1) following (1.9). We also use the hypothesis of the lemma. These considerations give us a decomposition $\mathscr{D}_i$ of $\mathbb{R}^{N-1}$ partitioning $\pi(A_i)$ and such that for each set $B \in \mathscr{D}_i$ with $B \subset \pi(A_i)$ there is $k = k(B)$ and there are definable continuous functions $f(B,1),\ldots,f(B,k):B \longrightarrow \mathbb{R}$ with $f(B,1) < \ldots < f(B,k)$ on $B$ such that $\pi^{-1}(B) \cap A_i$ is partitioned by the decomposition $\mathscr{D}_B \overset{\text{def}}{=\!=} (-\infty|f(B,1)|\ldots|f(B,k)|\infty)_{\pi^{-1}(B)}$ of $\pi^{-1}(B) = (-\infty,\infty)_B$ . It is clear that this gives us a decomposition

$$\mathscr{D}_i^* = \langle \mathscr{D};\{\mathscr{D}_B|B \in \mathscr{D}_i, B \subset \pi(A_i)\}\cup\{\{\pi^{-1}(B)\}|B \in \mathscr{D}_i ,$$

$$B \cap \pi(A_i) = \emptyset\}\rangle$$

of $\mathbb{R}^N$ which partitions $A_i$ . Now a common refinement $\mathscr{D} \in \mathrm{Dec}(\mathbb{R}^N)$ of $\mathscr{D}_1^*,\ldots,\mathscr{D}_m^*$ partitions each of $A_1,\ldots,A_m$ . □

This completes the proof of Theorem (3.7).

(3.11) Corollary. Each definable subset of $\mathbb{R}^n$ has only finitely many connected components, and each component is also definable.

<u>Proof.</u> By (3.7) each definable subset $A$ of $\mathbb{R}^n$ is partitioned by a decomposition $\mathscr{D}$ of $\mathbb{R}^n$. Now each of the finitely many sets in $\mathscr{D}$ is a (definable) cell, hence connected. Each component of $A$ is therefore a union of finitely many cells belonging to $\mathscr{D}$, and is therefore definable. $\square$

(3.12) <u>Remarks</u>

(1) Theorem (3.7) was derived under the assumption that $\mathscr{R}$ is of finite type (which is really an assumption on $\mathrm{Th}(\mathscr{R})$ ). Conversely, a weak form of (3.7) implies that $\mathscr{R}$ is of finite type. To be precise: all of the definitions and arguments in (3.2) - (3.6) make sense and go through without change for any expansion $\mathscr{R}$ of $(\mathbb{R},<)$, whether of finite type or not. Now we have the following fact.

If $\mathscr{R}$ is an expansion of $(\mathbb{R},<)$ with the property that for each definable subset $A$ of $\mathbb{R}^n$, $n \geq 1$, there is a decomposition of $\mathbb{R}^n$ partitioning $A$, then $\mathscr{R}$ is of finite type. (Hint: derive first (1.9) for $\mathscr{A} = \mathscr{R}$.)

(2) The assumption that $\mathscr{R}$ is of finite type remains valid upon expanding $\mathscr{R}$ with names for real numbers. Therefore, Theorem (3.7) also applies to $\mathbb{R}$-definable sets and functions where of course the notion of decomposition of $\mathbb{R}^n$ is suitably relativized.

(3) The condition that $\mathscr{R}$ is of finite type might be difficult to prove in a concrete case. The point of this paper is just to show what more is true if this condition holds. For 'exponentiation' one may try to use this knowledge the other way around: since I expect that $(\mathbb{R}, <, +, \cdot, \exp)$ is of finite type, but don't see any model theoretic technique to prove this, I prefer to stay inside $\mathbb{R}$ and use the analytic machinery available there to establish directly the conclusion of Theorem (3.7). The introduction sketches a (tentative) method to achieve this.

(4) $\mathscr{R}$ is of finite type if and only if each infinite $\mathbb{R}'$-definable subset of $\mathbb{R}'$, where $\mathscr{R}' = (\mathbb{R}', <, \ldots) \equiv \mathscr{R}$, has nonempty interior (i.e., contains an interval of $\mathbb{R}'$). (Exercise.)

The second part of this equivalence also makes sense for other topological-algebraic structures. Now Macintyre [Mac] established the analogue of this second half for p-adically closed fields. It would be nice to use this fact to prove an analogue of Theorem (3.7) for the field of p-adic numbers $\mathbb{Q}_p$.

(5) The relation between 'cylindrical decomposition' and 'quantifier elimination' is the following: <u>if each quantifier free definable subset of $\mathbb{R}^m$ has a quantifier free definable cylindrical decomposition, then $\mathrm{Th}(\mathscr{R})$ admits quantifier elimination</u>. This is how Collins [C] proves that $(\mathbb{R}, <, 0, 1, +, \cdot)$ has quantifier elimination.

(6) Let $\mathscr{R}$ be an expansion of $(\mathbb{R}, <, +, \cdot)$. Postulating the finite type condition is of course a rather drastic way of avoiding the Gödel phenomena that would appear if $\mathbb{N}$ were definable. In this connection I would like to know the answer to the following question: <u>is the set $\mathbb{N}$ definable whenever some infinite discrete subset of $\mathbb{R}$ is definable</u>?

APPENDIX

<u>Throughout this appendix we assume that $\mathscr{R}$ is an expansion of $(\mathbb{R}, <, +)$ with the property that each $\mathbb{R}$-definable subset of $\mathbb{R}$ is of finite type</u>, i.e., has only finitely many connected components.

Recall the notation $\hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$.

(A.1) <u>Lemma</u>. Suppose $I$ is an interval, $a \in I$, and $f: I \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable. Then both $f'(a^-) = \lim_{h \uparrow 0} \frac{f(a+h) - f(a)}{h}$ and $f'(a^+)$ (defined similarly) exist in $\hat{\mathbb{R}}$.

Proof. We shall just treat the case of $f'(a^+)$. (The case of $f'(a^-)$ is handled similarly.)

Suppose $\ell = \liminf_{h \downarrow 0} \frac{f(a+h)-f(a)}{h} < L = \limsup_{h \downarrow 0} \frac{f(a+h)-f(a)}{h}$.

Choose a rational $r$ with $\ell < r < L$. Then there are arbitrarily small $h > 0$ such that $f(a+h)-f(a) < rh$, but also arbitrarily small $h > 0$ such that $f(a+h)-f(a) \geq rh$. This situation is incompatible with the fact that the set of all $h > 0$ such that $f(a+h)-f(a) < rh$ is $\mathbb{R}$-definable, contradiction. $\square$

(A.2) Lemma. Suppose $I$ is an interval, $f: I \longrightarrow \mathbb{R}$ is continuous and $f'(a^+)$ is defined (in $\hat{\mathbb{R}}$), and $> 0$ for all $a \in I$. Then $f$ is strictly increasing, and its inverse $f^{-1}$ defined on the interval $f(I)$ has the property that $f^{-1})'(b^+)$ is defined and equal to $1/f'(a^+)$ if $f(a) = b$.

Proof. Left to the reader. (Note: the continuity assumption cannot be omitted.) $\square$

(A.3) Lemma. Suppose $I$ is an interval, $f: I \longrightarrow \mathbb{R}$ is continuous, and the maps $a \longmapsto f'(a^+)$, $a \longmapsto f'(a^-)$ are well defined, real valued, and continuous on all of $I$. Then $f$ is continuously differentiable on $I$.

Proof. It suffices to show that $f'(a^+) = f'(a^-)$ for all $a \in I$. Suppose the contrary, say $f'(a^+) > f'(a^-)$ for a certain $a \in I$. Then there is $c \in \mathbb{R}$ and an interval $J \subset I$ around $a$ such that $f'(x^+) > c > f'(x^-)$ for all $x \in J$. Then the continuous function $g: J \longrightarrow \mathbb{R}$ with $g(x) = f(x) - cx$ has the property that $g'(x^+) > 0$, $g'(x^-) < 0$ for all $x$, hence $g$ would be both strictly increasing and strictly decreasing on $J$, by the previous lemma. Contradiction. $\square$

(A.4) Lemma. If $I$ is an interval and $f: I \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, then there are only finitely many $x \in I$ such that $f'(x^+) = \pm\infty$.

Proof. Suppose for example that the $\mathbb{R}$-definable set $\{x \in I \mid f'(x^+) = \infty\}$ is infinite. So this set contains a whole interval, and for the sake of deriving a contradiction, we may as well assume that $f'(x^+) = \infty$ for all $x \in I$. By (2.2) we may further restrict to the case that $f$ is continuous, which implies that $f$ is strictly increasing on $I$, whence $f'(x^-) \geq 0$ for all $x \in I$.

After further shrinking the interval, we may also assume that we are in one of two cases:

(1) $f'(x^-) = \infty$ for all $x \in I$

(2) $f'(x^-)$ is finite for all $x \in I$, and $x \longmapsto f'(x^-)$ is continuous on $I$. (This uses again (2.2).)

In case (1) the inverse $f^{-1}$ of $f$ satisfies $(f^{-1})'(b^+) = (f^{-1})'(b^-) = 0$ for all $b$, whence $f^{-1}$ is constant, contradicting its injectiveness. In case (2) we can apply the same argument as in the proof of Lemma (A.3) to get a contradiction. $\square$

(A.5) Proposition. Suppose $f:(a,b) \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, $a < b$. Then $f$ is piecewise continuously differentiable, that is, there are $a_0 < \ldots < a_m$, $a_0 = a$, $a_m = b$ such that $f$ is continuously differentiable on each subinterval $(a_i, a_{i+1})$.

Proof. By (2.2) we may as well assume that $f$ is continuous, and by (A.4) that $f'(a^+)$ and $f'(a^-)$ are finite for all $a \in I$. But then the functions $a \longmapsto f'(a^+)$ and $a \longmapsto f'(a^-)$ are $\mathbb{R}$-definable, hence piecewise continuous, and the result follows immediately from (A.3). $\square$

(A.6) Corollary. Suppose $f:(a,b) \longrightarrow \mathbb{R}$ is $\mathbb{R}$-definable, $a < b$. Then $f$ is piecewise $C^n$, for each $n \in \mathbb{N}$. (The number of 'pieces' is allowed to increase with $n$.)

Proof. This follows from Proposition (A.5) by induction on $n$. $\square$

(A.7) Let us also mention the following interesting consequence of the hypothesis that each $\mathbb{R}$-definable subset of $\mathbb{R}$ is of finite type; we shall not use it in the rest of this appendix.

Proposition. $\mathrm{Th}(\mathscr{R})$ has definable Skolem functions. Also each definable equivalence relation on a definable subset of $\mathbb{R}^n$ has a definable set of representatives.

The proof is similar to the one given in [vdD3,(1.2) and (4.1)] for the theory of real closed fields.

(A.8) Fix a number $M \in \mathbb{N}$.
The considerations of §3 can now be refined as follows. For a definable $C^M$-submanifold $X$ of $\mathbb{R}^n$ we put $C^M(X) = \{f: X \longrightarrow \mathbb{R} \mid f \text{ is a definable } C^M\text{-function}\}$. Further we let $F_M(i_1,\ldots,i_n)$ be a class of $C^M$-submanifolds of $\mathbb{R}^n$, for each sequence $(i_1,\ldots,i_n)$ in $\{0,1\}$; the definition is by induction on $n$ and similar to the one in (3.3), but we require the 'defining' functions to be $C^M$. Put $F_M^{(n)} = \cup F_M(i_1,\ldots,i_n)$. Note that if $X \in F_M^{(n)}$ and $\dim(X) < n$, then $h(X) \in F_M^{(n-1)}$ and the homeomorphism $h: X \simeq h(X)$ is a $C^M$-isomorphism. So the $C^M$-analogue of (3.4) holds.
We now define $\mathrm{Dec}_M(X)$ for $X \in F_M^{(n)}$ as in (3.5), and we can state the following $C^M$-analogue of Theorem (3.7).

(A.9) Theorem. Suppose $\mathscr{R} = (\mathbb{R}, <, +, \ldots)$ is of finite type. Let $n \geq 1$. Then:

(a) For each definable set $X \subset \mathbb{R}^{n-1}$ and definable $f: X \longrightarrow \mathbb{R}$ there is a finite partition of $X$ into $C^M$-manifolds $\in F_M^{(n-1)}$ on each of which $f$ is a $C^M$-function.

(b) Given any definable subsets $A_1,\ldots,A_m$ of $\mathbb{R}^n$ there is $\mathscr{D} \in \mathrm{Dec}_M(\mathbb{R}^n)$ partitioning each of $A_1,\ldots,A_m$.

The proof is along the lines of the proof of (3.7) and we leave it to the reader to supply the extra work needed in the $C^M$-version of (3.8). (Hint: use (A.6), induction on $M$ and the fact that a function on an open subset of $\mathbb{R}^m$ is $C^M$ iff its partial derivatives up to order $M$ exist and are continuous.)

(A.10) This differentiable analogue of Theorem (3.7) could be useful in a proof of the following conjecture:

Let the expansion $\mathcal{R}$ of $(\mathbb{R}, <, +, \cdot)$ be of finite type and $\phi(v_1,\ldots,v_{m+n})$ an $L_{\mathbb{R}}$-formula. Then the family of $\mathbb{R}$-definable sets $(X_a)_{a \in \mathbb{R}^m}$, where

$$X_a = \{b \in \mathbb{R}^n \mid \mathcal{R} \models \phi(a,b)\} ,$$

contains only finitely many homeomorphism types. (In particular, there is a uniform bound on the number of connected components of $X_a$ as $a$ ranges over $\mathbb{R}^m$.)

(A.11) Mather [Mat, p. 218] gives a beautiful proof for the case of semialgebraic sets, i.e., for $\mathcal{R} = (\mathbb{R}, <, +, \cdot)$. The main tools in his proof are the Tarski-Seidenberg theorem and the existence of a finite Whitney stratification for each semialgebraic set. In our more general situation, the analogue of the Tarski-Seidenberg theorem is automatically true since we work in the category of all $\mathbb{R}$-definable sets. But to give a set a Whitney stratification means to partition the set into differentiable manifolds such that a certain technical condition holds, and this technical condition is unfortunately not generally satisfied by the $C^M$-decomposition of Theorem (A.9). When we look at Wall's proof in [W] that semialgebraic sets have finite Whitney stratifications, it seems nevertheless quite plausible that this proof essentially goes through in our general context, mainly because Theorem (A.9) is available.

It is remarkable that Mather obtains his homeomorphisms by integration of a semialgebraic vector field, an operation which leaves the semialgebraic context. This is one more reason to try to extend the structure $(\mathbb{R}, <, +, \cdot)$ as suggested in the Introduction.

REFERENCES

[C] G. E. Collins, Quantifier Elimination for Real Closed Fields by Cylindrical Algebraic Decomposition, Automata Theory and Formal Language, 2nd G.I. Conf., Kaiserslautern, pp. 134-183, Berlin, Springer-Verlag, 1975.

[vdD1] L. van den Dries, Exponential Rings, Exponential Polynomials, and Exponential Functions, to appear in Pacific Journal of Mathematics.

[vdD2] L. van den Dries, Analytic Hardy Fields and Exponential Curves in the Real Plane, to appear in Am. Journal of Math.

[vdD3] L. van den Dries, Algebraic Theories with definable Skolem Functions, submitted.

[D-W] B. Dahn and H. Wolter, On the theory of exponential fields, preprint.

[Hör] L. Hörmander, Linear Partial Differential Operators, Springer-Verlag, Berlin, 1963.

[Hov] A. G. Hovanskii, On a class of systems of transcendental equations, Soviet Math. Dokl. 22 (1980) 762-765.

[K] P. J. Kahn, Counting Types of Rigid Frameworks, Inv. Math. 55 (1979) 297-308.

[Ł] S. Łojasiewicz, Ensembles Semi-Analytiques, mimeographed notes, IHES, 1965.

[Mac] A. Macintyre, On definable subsets of p-adic fields, JSL 41 (1976) 605-610.

[Mat] J. Mather, Stratifications and Mappings, in: Dynamical Systems, ed. by M. Peixoto, pp. 195-232, Academic Press, 1973.

[P] C. Pugh, Hilbert's 16th problem: Limit cycles of polynomial vector fields in the plane, in: Dynamical Systems - Warwick 1974, pp. 55-57, SLN 468, 1975.

[T] A. Tarski, A Decision Method for Elementary Algebra and Geometry, 2nd ed., revised, Berkeley and Los Angeles, 1951.

[W] C. T. Wall, Regular Stratifications, in: Dynamical Systems - Warwick 1974, pp. 332-344, SLN 468, 1975.