Definability in Rationals with Real Order in the Background

Yuri Gurevich* Alexander Rabinovich†

Abstract

The paper deals with logically definable families of sets (or *point-sets*) of rational numbers. In particular we are interested whether the families definable over the real line with a unary predicate for the rationals are definable over the rational order alone. Let $\phi(X,Y)$ and $\psi(Y)$ range over formulas in the first-order monadic language of order. Let Q be the set of rationals and F be the family of subsets J of Q such that $\phi(Q,J)$ holds over the real line. The question arises whether, for every ϕ, F can be defined by means of an appropriate $\psi(Y)$ interpreted over the rational order. We answer the question negatively. The answer remains negative if the first-order logic is strengthen to weak monadic second-order logic. The answer is positive for the restricted version of monadic second-order logic where set quantifiers range over open sets. The case of full monadic second-order logic remains open.

1 Introduction

We consider the monadic second-order theory of linear order. For the sake of brevity, linearly ordered sets will be called chains.

Let $\mathcal{A} = \langle \mathbf{A} < \rangle$ be a chain. A formula $\phi(t)$ with one free individual variable t defines a point-set on $\mathbf{A}$ which contains exactly the points of $\mathbf{A}$ that satisfy $\phi(t)$. As usual we identify a subset of $\mathbf{A}$ with its characteristic predicate and we will say that such a formula defines a predicate on $\mathbf{A}$.

More generally, a formula $\chi(X)$ with one free monadic predicate variable defines the set of those predicates (or the family of those point-sets) on $\mathbf{A}$ that satisfy $\chi(X)$. This family is said to be definable by $\chi(X)$ in $\mathcal{A}$. The second kind of definability is more general because a set Y can be adequately represented by the family $\{\{y\} : y \in Y\}$.

Suppose that $\mathcal{A}$ is a subchain of $\mathcal{B} = \langle \mathbf{B}, < \rangle$. With a formula $\chi(X,\ A)$ we associate the following family of point-sets (or set of predicates) $\{\mathbf{P} : \mathbf{P} \subseteq \mathbf{A} \text{ and } \chi(\mathbf{P}, \mathbf{A}) \text{ holds in } \mathcal{B}\}$ on $\mathbf{A}$. This family is said to be definable by χ in $\mathcal{A}$ with $\mathcal{B}$ in the background.

*Microsoft Research and Dept EECS, University of Michigan, One Microsoft Way, Redmond, WA 98052, USA, e.mail: gurevich@microsoft.com

†Department of Computer Science, Beverly Sackler School of Exact Sciences, Tel Aviv University, Israel 69978, e.mail: rabino@math.tau.ac.il

Note that in such a definition bound individual (respectively predicate) variables of χ range over $\mathbf{B}$ (respectively over subsets of $\mathbf{B}$). Hence, it is reasonable to expect that the presence of a background chain $\mathcal{B}$ allows one to define point sets (or families of point-sets) on $\mathbf{A}$ which are not definable inside $\mathcal{A}$.

In [3] we proved

Theorem 1.1 *For any closed subset* $\mathbf{F}$ *of the reals, a family of point-sets is definable in a subchain* $\mathcal{F} = \langle \mathbf{F}, < \rangle$ *of the reals if and only if it is definable in* $\mathcal{F}$ *with the chain of the reals in the background.*

In fact, we proved a somewhat stronger theorem, namely, there is a uniform way to translate a definition (in the closed subchains of the reals) with the reals in the background into a definition without a background. On the other hand, it was shown in [3]

Theorem 1.2 *There exists an open subset* $\mathbf{G}$ *of the reals and a family of point-sets which is definable in* $\mathbf{G}$ *with the reals in the background and is not definable in* $\mathbf{G}$*.*

Note that the notions of "definable family of point sets" and "definable family of point sets with a background" can be naturally adopted to the first-order and other languages. For example, let P, A be unary predicate names and let ϕ (respectively χ) be first-order sentence in the signature $\{<, P\}$ (respectively $\{<, P, A\}$) . Let $\mathcal{A} = \langle \mathbf{A} < \rangle$ be a chain. A family $\mathcal{F}$ of subsets of $\mathbf{A}$ is definable by ϕ in $\mathcal{A}$ if $\mathbf{P} \in \mathcal{F} \Longleftrightarrow \langle \mathbf{A} <, \mathbf{P} \rangle \models \phi$. Suppose that $\mathcal{A}$ is a subchain of $\mathcal{B} = \langle \mathbf{B}, < \rangle$. A family $\mathcal{F}$ of subsets of $\mathbf{A}$ is definable by χ in $\mathcal{A}$ with $\mathcal{B}$ in the background if $\mathbf{P} \in \mathcal{F} \Longleftrightarrow \langle \mathbf{B} <, \mathbf{P}, \mathbf{A} \rangle \models \chi$.

Analyzing the proofs in [3] one can see that Theorems 1.1 and 1.2 also hold when "definable" is replaced by "first-order definable".

We were unable to resolve the following

Problem: Is it true that a family of point-sets is definable in the chain $\mathcal{Q}$ of rationals if and only if it is definable in $\mathcal{Q}$ with the chain of reals in the background?

We have not solved the problem, but have some related results. We prove the following theorem.

Theorem 1.3 *There is a family of point sets which is first-order definable in* $\mathcal{Q}$ *with the chain of reals in the background and is not first-order definable in* $\mathcal{Q}$*.*

Then we consider the weak interpretation of monadic language of order. Under this interpretation the bound monadic variables range over finite subsets.

We prove that Theorem 1.3 holds when "first-order definable" is replaced by "definable in weak monadic logic", i.e.,

Theorem 1.4 *There is a family of point sets which is definable by a formula of weak monadic logic in* $\mathcal{Q}$ *with the chain of reals in the background and is not definable by a formula of weak monadic logic in* $\mathcal{Q}$*.*

Finally, we consider the interpretation of monadic logic in which bound variables range over the open subsets of the chain. We show that under this interpretation definability in $\mathcal{Q}$ is the same as definability in $\mathcal{Q}$ with the chain of reals in the background.

Theorem 1.5 *Under the open sets interpretation of monadic second order language, a family of point-sets is definable in $\mathcal{Q}$ if and only if it is definable in $\mathcal{Q}$ with the chain of reals in the background.*

The paper is organized as follows: In section 2 we fix notations and recall some well-known theorems. Section 3 and Section 4 deal with definability in first-order monadic logic and weak monadic logic. Some of the results here are of independent interest. In Section 5 we consider the interpretations of monadic logic in which bound variables range over the open and closed subsets.

2 Preliminaries

Notations We use $k,\ l,\ m,\ n$ for natural numbers; $\mathbf{Q}$ for the set of rational numbers, $\mathbf{R}$ for the set of reals, Z for the set of integers, N for the set of natural numbers; $\mathbf{Q}^+$ for non-negative rational numbers and $\mathbf{R}^+$ for non-negative reals. We use standard notations for ordinals, e.g., ω is the order type of natural numbers, ω^* is the order type of negative integers, ω^k is the order type of lexicographically ordered k-tuples of natural numbers. As usual in set theory, a natural number n can be viewed as a linear order, namely the initial segment $(\{0,..,n-1\},\ <)$ of the standard ordering of natural numbers.

Let $\tau = \langle <, P_1, \dots, P_n\rangle$ be a signature with binary predicate name $<$ and unary predicate names $P_1, \dots P_n$. Let $A_1 = \langle \mid A_1 \mid, <^{A_1},\ P_1^{A_1}, \dots, P_n^{A_1}\rangle$ and let $A_2 = \langle \mid A_2 \mid, <^{A_2},\ P_1^{A_2}, \dots, P_n^{A_2}\rangle$ be structures for τ.

The structure $A = A_1 + A_2$ for τ is defined as follows:

1. The universe is $\{\langle 1, a_1\rangle\ :\ a_1 \in\mid A_1 \mid\} \cup \{\langle 2, a_2\rangle\ :\ a_2 \in\mid A_2 \mid\}$.
2. $\langle i, a\rangle <^A \langle j, b\rangle$ iff $i = 1$ and $j = 2$, or $i = j$ and $a <^{A_i} b$.
3. $\langle i, a\rangle \in P_k^A$ iff $a \in P_k^{A_i}$.

It is clear that if $<^{A_1}$ and $<^{A_2}$ are linear orders on $\mid A_1 \mid$ and $\mid A_2 \mid$ respectively, then $<^A$ is a linear order on $\mid A \mid$.

Let τ be the signature as above and let A_1 be a structure for τ and let $B = \langle \mid B \mid,\ <^B\rangle$ be a chain. The structure $A = A_1 \times B$ for τ is defined as follows:

1. The universe is $\{\langle a_1, b\rangle\ :\ a_1 \in\mid A_1 \mid$ and $b \in\mid B \mid\}$.
2. $\langle a_1, b_1\rangle <^A \langle a_2, b_2\rangle$ iff $b_1 <^B b_2$, or $b_1 = b_2$ and $a_1 <^{A_1} a_2$.
3. $\langle a, b\rangle \in P_k^A$ iff $a \in P_k^{A_1}$.

It is clear that the operations $+$ and $\times$ are associative, when, as usual, one does not distinguish between isomorphic structures.

We use $\models$ for the satisfaction relation between structures and formulas. We use $\models_w$ for the satisfaction relations between structures and monadic formulas when bound monadic variables range over the finite subsets of the structures. We use $\equiv_n$ for the indiscernability by first order formulas of quantifier rank n.

We use $\approx_n^f$ for indiscernability [2]. by the n-round Ehrenfeucht games appropriate for the weak monadic second-order logic [2]. For reader's convenience we recall the definition of those games.

The playing board is composed of two structures A_1 and A_2 of the same signature. Given two structures A_1 and A_2. The game $H_n(A_1, A_2)$ is played by two players Spoiler and Duplicator . The game is played n rounds. In the i-th round Spoiler first chooses one of the structures A_{l_i} $(l_i = 1, 2)$ and then points out in the chosen structure an arbitrary finite sequence $a_{i,1}^{l_i},\ a_{i,2}^{l_i}, \dots a_{i,k_i}^{l_i}$ of elements. Then Duplicator chooses in the other structure A_{3-l_1} a sequence of k_i elements $a_{i,1}^{3-l_i},\ a_{i,2}^{3-l_i}, \dots a_{i,k_i}^{3-l_i}$. After n rounds we have $k_1 + k_2 + \dots k_n$ pairs.

$$\left.\begin{array}{c} a_{1,1}^1 \leftrightarrow a_{1,1}^2 \\ a_{1,2}^1 \leftrightarrow a_{1,2}^2 \\ \vdots \\ a_{1,k_1}^1 \leftrightarrow a_{1,k_1}^2 \end{array}\right\} \text{first move}$$

$$\vdots$$

$$\left.\begin{array}{c} a_{n,1}^1 \leftrightarrow a_{n,1}^2 \\ a_{n,2}^1 \leftrightarrow a_{n,2}^2 \\ \vdots \\ a_{n,k_n}^1 \leftrightarrow a_{n,k_n}^2 \end{array}\right\} n\text{-th move}$$

Duplicator wins if the correspondence written above is a partial isomorphism. Spoiler wins if and only if Duplicator does not win. We write $A_1 \approx_n^f A_2$ if Duplicator has a winning strategy in the game $H_n(A_1, A_2)$. It is clear that $\approx_n^f$ is an equivalence relation.

Theorem 2.1 *(Ehrenfeucht [2]) If structures A_1 and A_2 are $\approx_n^f$ equivalent, then for every weak monadic sentence ϕ with at most n quantifiers $A_1 \models_w \phi$ iff $A_2 \models_w \phi$.*

Let A be a chain and $a < b$ be elements of A. We use (a, b) for the subchain of A which consists of all elements between a and b; we use $(-\infty, a)$ (respectively (a, ∞)) for the subchain which consists of all elements which are less than a (respectively greater than a). Similar notations are used for the closed and for half open intervals of A.

Lemma 2.2 *(Ehrenfeucht [2]) Assume that A, B are chains. Assume that*

1. *For every finite sequence $a_1 < a_2 < \dots < a_n$ of elements of A there is a sequence $b_1 < b_2 < \dots < b_n$ of elements of B such that $(-\infty, , a_1] \approx_k^f (-\infty, b_1]$, and $(a_n, \infty) \approx_k^f (b_n, \infty)$, and $(a_i, a_{i+1}] \approx_k^f (b_i, b_{i+1}]$, and*

2. *for every finite sequence* $b_1 < b_2 < \ldots < b_n$ *of elements of* B *there is a sequence* $a_1 < a_2 < \ldots < a_n$ *of elements of* A *such that* $(-\infty, a_1] \approx_k^f (-\infty, b_1]$, *and* $(a_n, \infty) \approx_k^f (b_n, \infty)$, *and* $(a_i, a_{i+1}] \approx_k^f (b_i, b_{i+1}]$.

Then $A \approx_{k+1}^f B$.

The following lemma is well-known and easy to check (see [4] for similar lemmas).

Lemma 2.3 1. *If* $A_1 \equiv_n A_2$ *and* $A_1' \equiv_n A_2'$ *then* $A_1 + A_1' \equiv_n A_2 + A_2'$.

2. *If* $A_1 \equiv_n A_2$ *and* $B_1 \equiv_n B_2$ *then* $A_1 \times B_1 \equiv_n A_2 \times B_2$.

3. *If* $A_1 \approx_n^f A_2$ *and* $A_1' \approx_n^f A_2'$ *then* $A_1 + A_1' \approx_n^f A_2 + A_2'$.

4. *If* $A_1 \approx_n^f A_2$ *and* $B_1 \approx_n^f B_2$ *then* $A_1 \times B_1 \approx_n^f A_2 \times B_2$.

Lemma 2.4 $k \equiv_n m$ *for* $k, m \geq 2^n$ *(see e.g., in [1]).*

3 Definability in Monadic first-order logic of order

Lemma 3.1 $k \equiv_n \omega + \omega^*$ *for* $k \geq 2^n$.

Proof Straightforward Induction on n; use Lemma 2.4 and Theorem 2.3(1). □

Corollary 3.2 Z *is elementarily equivalent to* $Z + Z$.

Proof We will prove that $Z \equiv_n Z + Z$ for every n. Indeed,

$$Z = \omega^* + 2^n + \omega \equiv_n \omega^* + (\omega + \omega^*) + \omega = (\omega^* + \omega) + (\omega^* + \omega) = Z + Z$$

. □

Let Q_0 be the structure $\langle Q^+, <, \{0\}\rangle$ where Q^+ is the set of non-negative rationals, $<$ is the standard order on Q^+ and $\{0\}$ is the unary predicate that contains only 0.

The following is straightforward

Lemma 3.3 1. $Q_0 \times Z$ *is isomorphic to* $\langle Q, <, P_Z\rangle$ *where* Q *are the rationals,* $<$ *is the standard order on the rationals and* P_Z *is interpreted as the set of integers.*

2. $Q_0 \times (Z + Z)$ *is isomorphic to* $\langle Q, <, P\rangle$ *where* Q *are the rationals,* $<$ *is the standard order on the rationals and* P *is interpreted as a subset* $\{u_i : i \in Z\} \cup \{v_i : i \in Z\}$, *where* u_i *and* v_i *are rational numbers,* $u_i < u_{i+1}$, $v_i < v_{i+1}$ *and* $\lim_{i \to \infty} u_i = \lim_{i \to -\infty} v_i = \alpha$ *for some irrational number* α *and* $\{u_i : i \in Z\}$ *(respectively* $\{v_i : i \in Z\}$*) is unbounded in* Q *from below (respectively, from above).*

Corollary 3.4 *There is no sentence* ϕ *in the language of first-order logic of order with an additional unary relation* P *such that the following are equivalent*

1. $\langle Q, <, P\rangle \models \phi$.

2. *there exists an irrational number α such that for all $q_1 < \alpha < q_2$ there are points of P both in $(q_1,\ \alpha)$ and $(\alpha,\ q_2)$.*

Proof If such ϕ exists then (by Lemma 3.3) the structure $Q_0 \times (Z + Z)$ satisfies ϕ and structure $Q_0 \times Z$ does not satisfy ϕ. However, these structures are elementarily equivalent (by Corollary 3.2 and Lemma 2.3(2)). Contradiction. □

Lemma 3.5 *There is a sentence ϕ in the language of first-order logic of order with additional unary relations $P,\ Q$ such that $\langle R, <, Q, P\rangle \models \phi$ iff P is a subset of the set Q of rationals and there exists an irrational number α such that for all $q_1 < \alpha < q_2$ there are points of P both in $(q_1,\ \alpha)$ and $(\alpha,\ q_2)$.*

Proof Immediate. □ Finally, Theorem 1.3 follows

from Lemma 3.5 and Corollary 3.4. Indeed, the family of all point-sets P which have an irrational number α both as its left and right limit point is not definable in the set Q of rationals, but is definable in Q with the reals in the background.

4 Definability in weak monadic logic of order

Theorem 4.1 $\omega^k \approx_k^f \omega^k + \omega^k \times A$ *for every linear order A*

Proof The proof is similar to the Ehrenfeucht proof that for any ordinals α and β, if $\alpha = \beta(mod\ \omega^k)$ then $\alpha \approx_k^f \beta$ (Theorem 12 in [2]). It proceeds by the induction on k. The basis (k=1) is trivial, because both structures are infinite.

Now let us assume that the theorem is true for some k. Let Spoiler choose a sequence $a_1, \dots a_n$ on his first move. We can assume, without restriction of generality, that $a_1 < \dots < a_n$.

We will show that Duplicator can choose elements $b_1 < \dots < b_n$ in the second structure such that $(-\infty, a_1] \approx_k^f (-\infty, b_1]$, and $(a_n, \infty) \approx_k^f (b_n, \infty)$, and $(a_i, a_{i+1}) \approx_k^f (b_i, b_{i+1})$. Therefore, according to Theorem 2.2, we obtain that our structures are $\approx_{k+1}^f$ equivalent. Let us describe how Duplicator should reply to the above move of Spoiler.

If Spoiler have chosen his first move in the structure ω^{k+1} then Duplicator will choose the same elements in the first component of $\omega^{k+1} + \omega^{k+1} \times A$. The first n segments are identical. The segment (a_n, ∞) has the order type of ω^{k+1} and the segment (b_n, ∞) has the order type of $\omega^{k+1} + \omega^{k+1} \times A = \omega^k + \omega^k \times (\omega + \omega \times A)$; by the inductive hypothesis both these segments are $\approx_k^f$ equivalent to ω^k. This completes our arguments for the case when Spoiler chooses elements in ω^{k+1}.

Assume that Spoiler chooses elements in $\omega^{k+1} + \omega^{k+1} \times A$. We define a sequence of auxiliary ordinals d_i $(i = 1, \dots, n)$.

Define d_1 as a_1 if a_1 is in the first summand of $\omega^{k+1} + \omega^{k+1} \times A$; define d_1 as $\omega^k + d$ if $a_1 = \langle d, j \rangle \in \omega^{k+1} \times A$ for some $j \in A$.

For $i = 1, \ldots, n-1$ define d_{i+1} according to the following cases:

Case 1. a_i and a_{i+1} are in the first summand ω^{k+1} or there is j such that a_i and a_{i+1} are both in the same j-th summand of $\omega^{k+1} \times A$. In this case we define d_{i+1} be the order type of $(a_i, a_{i+1}]$.

Case 2. The other case. Then a_{i+1} is some element (d, j) of $\omega^{k+1} \times A$. Define $d_{i+1} = \omega^k + d$.

It is clear that

$$d_i < \omega^{k+1} \tag{1}$$

Observe that

$$d_1 \approx^f_k (-\infty, a_1] \tag{2}$$

Indeed, if a_1 is in the first summand of $\omega^{k+1} + \omega^{k+1} \times A$, then these two segments are isomorphic. If a_1 is d-th elements of j-th component of $\omega^{k+1} \times A$, then

$$(-\infty, a_1] = \omega^{k+1} + \omega^{k+1} \times (-\infty, j) + d = \omega^k + \omega^k \times (\omega + \omega \times (-\infty, j)) + d$$

and therefore, by the inductive hypothesis, $(-\infty, a_1]$ is $\approx^f_k$ equivalent to d_1. Similar arguments show that for $i = 1, \ldots, n-1$

$$d_{i+1} \approx^f_k (a_i, a_{i+1}] \tag{3}$$

Define b_i as $d_1 + \ldots + d_i$. It follows from (1) that $b_1 < b_2 \ldots < b_n < \omega^{k+1}$, and hence $(b_n, \infty) \approx^f_k \omega^k \approx^f_k (a_n, \infty)$. From (2) and (3) it follows that $(-\infty, a_1] \approx^f_k (-\infty, b_1]$ and $(a_i, a_{i+1}] \approx^f_k (b_i, b_{i+1}]$. Therefore, according to Theorem 2.2, we obtain that our structures are $\approx^f_{k+1}$ equivalent.

□

Corollary 4.2 *There is no sentence in the weak monadic logic of order that defines the set of Dedekind complete linear orders.*

Proof Observe that ω^k is Dedekind complete and $\omega^k + \omega^k \times \omega^*$ is not. Applying Theorem 4.1 and Theorem 2.1 we obtain the Corollary. □ The same arguments show

Corollary 4.3 *There is no sentence in the weak monadic logic of order that defines the set of well-ordered chains.*

Corollary 4.4 *There is no sentence ϕ in the weak monadic language of order with an additional unary relation P such that the following are equivalent.*

1. *$\langle Q, <, P \rangle \models \phi$,*
2. *there exists an irrational number α such that for all $q_1 < \alpha < q_2$ there are points of P both in $(q_1,\ \alpha)$ and $(\alpha,\ q_2)$.*

Proof Toward a contradiction, let ϕ be a sentence in a prenex normal form that expresses the above property and let k be the number of quantifiers in ϕ. Let Q_0 be as in Section 3 and let Q_1 be $\langle Q, <, \emptyset\rangle$. Observe that $Q_1 + Q_0 \times \omega^k$ does not satisfy ϕ and $Q_1 + Q_0 \times (\omega^k + \omega^k \times \omega^*)$ satisfies ϕ (both of the constructed chains are countable and dense, so they are order-isomorphic to Q). However, these structures are $\approx_k^f$ equivalent by Theorem 4.1 and Lemma 2.3. This contradicts Theorem 2.1. $\square$ This corollary together with

Lemma 3.5 imply

Theorem 4.5 *There is a family of point sets which is definable in Q with the chain of reals in the background by a weak monadic formula and is not definable in Q by a weak monadic formula.*

5 Open and Closed Sets Interpretations of Monadic Second-order Language

Let A be a linear order. An interval J of A is said to be open if (1) $sup(J)$ either does not exist or it exists, but does not belong to J and (2) the same for inf.

A subset of A is said to be open if it is the union of a family of disjoint open intervals. We will be interested in the chain of reals and the chain of rationals; on these chains the above definition of an open set is equivalent to the standard topological definition. A subset O of A is said to be perforated if it is open and for any two distinct maximal intervals I and J in O there is a non-empty open interval H such that all the points of H are between the points of I and the points of J.

Observe that

Lemma 5.1 *Every open set is a union of two disjoint perforated sets.*

Proof Let G be an open set. Let S be the relation on maximal open intervals of G defined as follows: I and J are in S iff there is no non-empty open interval H such that all the points of H are between the points of I and the points of J. It is clear that S is reflexive and symmetric. Let S^* be the transitive closure of S. Observe that S^* is an equivalence relation.

Every equivalence class C of S^* is a set of intervals. These intervals are naturally ordered. This order is either finite or else has one of the following order types: ω or ω^* or Z. In each of these cases, the order of intervals is isomorphic to a contiguous segment of integers. Fix such an isomorphism ξ. Call a member I of C odd (respectively even) if so is the number $\xi(I)$.

Let P_1 be the union of all odd intervals (in all equivalence classes of S^*), and let P_2 be the union of all even intervals. It is clear that $G = P_1 \cup P_2$ and that P_1, P_2 are disjoint perforated open sets. $\square$

First, we consider two interpretations of the language of monadic second-order logic: in the first one, the bound monadic variables range over open sets; in the second one, the bound monadic variables range over perforated sets; free monadic variables are interpreted as arbitrary sets. We use $\models_{open}$ and $\models_{\text{perf}}$ for the satisfaction relation under the first and the second interpretations respectively.

Lemma 5.2 *For every monadic formula $\phi(t_1, \ldots, t_k, X_1, \ldots, X_n)$ there is a monadic formula $\psi(t_1, \ldots, t_k, X_1, \ldots, X_n)$ such that for all $a_1, \ldots, a_k \in \mathbf{R}$ and $P_1, \ldots, P_n \subseteq \mathbf{R}$*

$$\langle \mathbf{R}, a_1, \ldots, a_k, P_1, \ldots, P_n \rangle \models_{open} \phi(t_1, \ldots, t_k, X_1, \ldots, X_n)$$
$$\textit{iff}$$
$$\langle \mathbf{R}, a_1, \ldots, a_k, P_1, \ldots, P_n \rangle \models_{perf} \psi(t_1, \ldots, t_k, X_1, \ldots, X_n)$$

Proof By Lemma 5.1, every open set is the union of two perforated sets. For every bound variable Y_i introduce two bound variables Z_i and U_i and replace in ϕ substrings "$\exists Y_i$" by "$\exists Z_i \exists U_i$" and "$t \in Y_i$" replace by "$t \in Z_i \vee t \in U_i$". $\square$

For the second reduction we introduce canonical formulas and show that every monadic formula under the perforated sets interpretation in some sense is equivalent to a canonical formula.

Below we use variables U_i which always range over sets of the form $\{a \in \mathbf{R} \ : \ a > \alpha,$ where α is irrational $\}$; these variables will be always bound.

We use variables O_i to range over perforated sets; they might be bound or free.

X_i will range over arbitrary subsets of rational numbers; X_i are always free.

Variables v_i will range over the rational numbers and they always will be bound; t_i will range over the reals and they always will be free.

Contains(O, U) is interpreted as "perforated set O contains the infimum (in $\mathbf{R}$) of U". The formula $U_i \prec U_j$ is interpreted as the infimum of U_i is less than the infimum of U_j. The formula $t \ll U$ is interpreted as t is less than the infimum of U. (*Contains* and $\prec$ are predicates over the set variables.)

Q is a unary predicate name interpreted as the set of rational numbers.

Consider the formulas constructed from

1. *Contains*(O_i, U_j).
2. $U_i \prec U_j$.
3. $t \ll U$.
4. $u \in Z$, where u is a first order variable v_i or t_j and Z is a monadic variable O_i, U_j or X_k.
5. $u \in Q$, where u is a first order variable v_i or t_j and Q is the predicate name interpreted as the set of rationals.
6. $u_i < u_j$, where u_i and u_j are individual variables.

by applying the boolean connectives, quantification over perforated set variables O_i, quantification over U_j, and quantifiers $\exists v_i \in Q$. Quantifiers of the form $\exists v_i \in \mathbf{R}$ are not allowed.

Such a formula is a C formula if all v_i are bound by $\exists v_i \in Q$. We use $\models_{\mathrm{can}}$ for the satisfaction relation of C formulas.

Lemma 5.3 *For every formula $\phi(Q, t_1, \ldots, t_k, X_1, \ldots X_n, O_1, \ldots O_m)$ of monadic second-order language of order there exists a C-formula $\phi^{tr}(Q, t_1, \ldots, t_k, X_1, \ldots X_n, O_1, \ldots O_m)$ such that for $a_1, \ldots a_k \in \mathbf{R}$ and $P_1, \ldots, P_n \subseteq \mathbf{Q}$ and perforated sets $G_1, \ldots G_m$*

$$\langle \mathbf{R}, \mathbf{Q}, a_1, \ldots, a_k, P_1, \ldots, P_n, G_1, \ldots, G_m \rangle \models_{perf} \phi(Q, t_1, \ldots, t_k, X_1, \ldots, X_n, O_1, \ldots, O_m)$$
$$\textit{iff}$$
$$\langle \mathbf{R}, \mathbf{Q}, a_1, \ldots, a_k, P_1, \ldots, P_n, G_1, \ldots, G_m \rangle \models_{can} \phi^{tr}(Q, t_1, \ldots, t_k, X_1, \ldots, X_n, O_1, \ldots, O_m)$$

Proof The proof proceeds by induction on formulas. The only non-trivial case is the quantification over the individual variables.

So assume that we know how to construct ϕ_1^{tr} and we will show how to construct the translation for $\exists t.\phi_1$.

We use $\psi\{v/t\}$ for the formula obtained by replacing all free occurrences of t in ψ by v.

The translation of $\exists t.\phi_1$ is defined as $\exists v \in Q.\phi_1^{tr}\{v/t\} \vee \exists U_t.\alpha_1$, where v (respectively U_t) is a fresh individual (respectively monadic) variable and α_1 is obtained from ϕ_1^{tr} by the following transformations:

Replace

1. " $t < w$" by "$w \in U_t$", where w is an individual variable.
2. "$v < t$" by "$v \notin U_t$".
3. "$t' < t$" by $t' \ll U_t$.
4. "$t \ll U$" by "$U_t \prec U$".
5. "$t \in X_i$" and "$t \in Q$" by "False".
6. "$t \in U$" by "$U \prec U_t$".
7. "$t \in O$" by "$Contain(O, U_t)$".

It is easy to verify that the translation indeed satisfies the Lemma. $\square$

The next lemma provides a reduction from the definability by canonical formulas in $\mathbf{Q}$ with the chain of reals in the background to the definability in $\mathbf{Q}$ under the perforated set interpretation.

Lemma 5.4 *Let $\phi(Q, X_1, \ldots, X_n)$ be a C formula (without free individual variables and free perforated set variables). There exists a formula $\psi(X_1, \ldots, X_n)$ in the language of monadic second-order logic of order such that for all $P_1, \ldots, P_n \subseteq \mathbf{Q}$*

$$\langle \mathbf{R}, \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{can} \phi(Q, X_1, \ldots, X_n) \textit{ iff } \langle \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{perf} \psi(X_1, \ldots, X_n)$$

.

Proof Observe that if $\mathbf{U}_i = \{a \in \mathbf{R} : a > \alpha_i\}$ for an irrational number α_i then

$$\langle \mathbf{Q}, <, \mathbf{Q} \cap \mathbf{U}_i\rangle \models (\forall t t'.\ t \in U \wedge t' > t \rightarrow t' \in U) \wedge \neg\exists t'.\ (\forall t.\ t > t' \leftrightarrow t \in U)$$

and

$$\mathbf{U}_j \prec \mathbf{U}_i \text{ if and only if } \langle \mathbf{Q}, <, \mathbf{Q} \cap \mathbf{U}_i, \mathbf{Q} \cap \mathbf{U}_j \models \exists t.\ t \in U_j \wedge t \notin U_i$$

Moreover, if $\mathbf{O}$ is a perforated subset of reals and $\mathbf{U} = \{a \in \mathbf{R} : a > \alpha\}$ for some irrational α then (a) $\mathbf{O} \cap \mathbf{Q}$ is a perforated subset of rationals and (b) $Contain(\mathbf{O}, \mathbf{U})$ if and only if there is an open (in $\mathbf{Q}$) interval $\mathbf{I}$ such that $\mathbf{I} \subset \mathbf{O} \cap \mathbf{Q}$ and the sets $\mathbf{I} \cap \mathbf{U}$ and $\mathbf{I} \setminus (\mathbf{I} \cap \mathbf{U})$ are non-empty.

Note also that in ϕ all individual variables range over $\mathbf{Q}$, because ϕ is a canonical formula without free individual variables.

Therefore, ψ can be constructed as follows: replace in ϕ

1. "$U_j \prec U_i$" by "$\exists t.\ t \in U_j \wedge t \notin U_i$".
2. "$\exists U$" by "$\exists U.\ (\forall t t'.\ t \in U \wedge t' > t \rightarrow t' \in U) \wedge \neg\exists t'.\ (\forall t.\ t > t' \leftrightarrow t \in U) \wedge$".
3. "$Contain(O, U)$" by "$\exists t_1 t_2.\ t_1 < t_2 \wedge t_1 \in O \wedge t_2 \in O \wedge t_1 \notin U \wedge t_2 \in U \wedge \forall t.\ t_1 < t < t_2 \rightarrow t \in O$". (Actually this is the only place where we use that O are interpreted as perforated sets.)
4. "$\exists v \in Q$" by "$\exists v$".

□ The following is immediate:

Lemma 5.5 *For every monadic formula $\phi(X_1, \ldots, X_n)$ there is a monadic formula $\psi(X_1, \ldots, X_n)$ such that for every $P_1, \ldots, P_n \subseteq \mathbf{Q}$*

$$\langle \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{perf} \phi(X_1, \ldots, X_n) \textit{ iff } \langle \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{open} \psi(X_1, \ldots, X_n)$$

From Lemmas 5.2-5.5 we obtain that under the open sets interpretation, definability in $\mathbf{Q}$ with the reals in the background is equivalent to definability in $\mathbf{Q}$.

Theorem 5.6 *For every monadic formula $\phi(Q, X_1, \ldots, X_n)$ there is a monadic formula $\psi(X_1, \ldots, X_n)$ such that for every $P_1, \ldots, P_n \subseteq \mathbf{Q}$*

$$\langle \mathbf{R}, \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{open} \phi(Q, X_1, \ldots, X_n) \textit{ iff } \langle \mathbf{Q}, P_1, \ldots, P_n\rangle \models_{open} \psi(X_1, \ldots, X_n).$$

Hence,

Theorem 5.7 *A family of point sets is definable in* $\mathbf{Q}$ *by a monadic second-order formula under the open sets interpretation if and only if it is definable in* $\mathbf{Q}$ *with the reals in the background under the open sets interpretation.*

Under the closed sets interpretation the bound monadic predicate variables range over the closed subsets.

Corollary 5.8 *A family of point sets is definable in* $\mathbf{Q}$ *by a monadic second order formula under the closed sets interpretation if and only if it is definable in* $\mathbf{Q}$ *with the reals in the background under the closed sets interpretation.*

Proof We provide a reduction between the open sets interpretation and the closed sets interpretation.

Let ϕ be a formula and let ψ be obtained from ϕ when for every bound monadic variable X and every individual variable t, "$t \in X$" is replaced by "$t \notin X$". It is obvious that a family of point sets is definable by ϕ under the open (respectively closed) sets interpretation if and only if the family is definable by ψ under the closed (respectively open) sets interpretation.

This reduction together with Theorem 5.7 give the desirable result. $\square$

References

[1] H.-D. Ebbinghaus and J. Flum (1995). *Finite Model Theory.* Springer Perspectives in Mathematical Logic.

[2] A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. Fundamenta mathematicae 49,129-141,1961.

[3] Y. Gurevich and A. Rabinovich. Definability and Undefinability with Real Order in the Background. To appear in JSL.

[4] H. Läuchli (1968). A decision procedure for the weak second order theory of linear order. In *Contributions to Mathematical Logic, Proceedings of Logic Colloquium Hanover 1966*, North-Holland.