

DEFINABLE ISOMORPHISM PROBLEM

KHADIJEH KESHVARDOOST, BARTEK KLIN, SŁAWOMIR LASOTA, JOANNA OCHREMIK,
AND SZYMON TORUŃCZYK

Velayat University

University of Warsaw

University of Warsaw

Universit Paris Diderot

University of Warsaw

ABSTRACT. We investigate the isomorphism problem in the setting of definable sets (equivalent to sets with atoms): given two definable relational structures, are they related by a definable isomorphism? Under mild assumptions on the underlying structure of atoms, we prove decidability of the problem. The core result is parameter-elimination: existence of an isomorphism definable with parameters implies existence of an isomorphism definable without parameters.

1. INTRODUCTION

First-order definable sets, although usually infinite, can be finitely described and are therefore amenable to algorithmic manipulation. Definable sets (we drop the qualifier *first-order* in what follows) are parametrized by a fixed underlying relational structure \mathcal{A} whose elements are called *atoms*.

Example 1.1. Let \mathcal{A} be a countable set $\{\underline{1}, \underline{2}, \underline{3}, \dots\}$ equipped with the equality relation only; we shall call this structure the *pure set*. Let

$$V = \{ \{a, b\} \mid a, b \in \mathcal{A}, a \neq b \},$$

$$E = \{ (\{a, b\}, \{c, d\}) \mid a, b, c, d \in \mathcal{A}, a \neq b \wedge a \neq c \wedge a \neq d \wedge b \neq c \wedge b \neq d \wedge c \neq d \}.$$

Both V and E are definable sets (over \mathcal{A}), as they are constructed from \mathcal{A} using (possibly nested) set-builder expressions with first-order guards ranging over \mathcal{A} . In general, we allow finite unions in the definitions, and finite tuples (as above) are allowed for notational convenience. Precise definitions are given in Section 2. The pair $G = (V, E)$ is also a definable set, in fact, a definable graph. It is an infinite Kneser graph (a generalization of the famous Petersen graph): its vertices are all two-element subsets of \mathcal{A} , and two such subsets are adjacent iff they are disjoint.

Key words and phrases: the graph isomorphism problem, definable sets, homogeneous structures.

The graph G is \emptyset -*definable*: its definition does not refer to any particular elements of \mathcal{A} . In general, one may refer to a finite set of parameters $S \subseteq \mathcal{A}$ to describe an S -*definable* set. For instance, the set $\{a \mid a \in \mathcal{A}, a \neq \underline{1} \wedge a \neq \underline{2}\}$ is $\{\underline{1}, \underline{2}\}$ -definable. Definable sets are those which are S -definable for some finite $S \subseteq \mathcal{A}$. \square

We remark that in the pure set \mathcal{A} (or, more generally, in an *effective homogeneous* relational structure \mathcal{A}), every first-order formula is effectively equivalent to a quantifier-free formula. In the sequel we shall assume \mathcal{A} to be effective homogeneous, and thus, as long as complexity issues are ignored and decidability is the only concern, we can safely restrict to quantifier-free formulas. In consequence, the first-order theory of \mathcal{A} is decidable.

Although definable relational structures (over a finite signature) correspond (up to isomorphism) to first-order interpretations well-known from logic and model theory [9], we prefer to use a different definition since standard set-theoretic notions directly translate into this setting. For example, a definable function $f : X \rightarrow Y$ is simply a function whose domain X , codomain Y , and graph $\Gamma(f) \subseteq X \times Y$ are definable sets. A relational structure is definable if its signature, universe, and interpretation function that maps each relation symbol to a relation on the universe, are definable. Finally, a definable isomorphism between definable structures over the same signature is a definable bijective mapping between their universes that preserves and reflects every relation in the signature. Likewise one introduces, e.g., definable homomorphisms. All hereditarily finite sets (finite sets, whose elements are finite, and so on, recursively) are definable, and every finite relational structure over a finite signature is (isomorphic to) a definable one.

Contribution. The classical *isomorphism problem* asks whether two given *finite* structures are isomorphic. In this paper, we consider its counterpart in the setting of definable sets (the problem is called *definable isomorphism problem* in the sequel): given two definable structures \mathbb{A}, \mathbb{B} over the same definable signature Σ , all over the same fixed structure \mathcal{A} , are they related by a definable isomorphism? Note that definable structures can be meaningfully considered as input to a computational problem since they are finitely described with the set-builder notation and first-order formulas in the language of \mathcal{A} . The structure \mathcal{A} is considered here in a parametric manner, not as a part of input: every structure \mathcal{A} induces a different decision problem.

As our main result we prove, under a mild assumption on the structure \mathcal{A} , decidability of the definable isomorphism problem. The key technical difficulty is to show that every two S -definable structures related by a definable isomorphism are also related by an S -definable one. (When $S = \emptyset$ this is parameter-elimination: existence of an isomorphism defined with parameters enforces existence of one defined without parameters.) Having this, the problem reduces to testing whether two S -definable structures are related by an S -definable isomorphism, which in turn reduces to the first-order satisfiability problem in \mathcal{A} .

As witnessed by Example 2.4 below, existence of an isomorphism does not guarantee existence of a definable one. Therefore we do not solve the *isomorphism problem* for definable structures, which asks whether two given definable structures are isomorphic. In fact, decidability status of the latter problem remains an intriguing open question, even for \mathcal{A} the pure set.

Motivation and related work. This paper is part of a programme aimed at generalizing classical decision problems (for instance the homomorphism problem, studied recently in [10, 11]), and computation models such as automata [4], Turing machines [5] and programming

languages [3, 6, 12, 13], to sets with atoms. For other applications of sets with atoms (called there *nominal sets*) in computing, see [16].

Isomorphism testing is at the core of many decision problems in combinatorics and logic. In case of finite graphs it is well known to be solvable in NP, and since recently in quasi-polynomial time [1]. Whether it can be solved in P is still an extremely challenging open question, and only special cases are shown so far, e.g. [14].

2. PRELIMINARIES

Throughout the paper, fix a countable relational structure \mathcal{A} , called *atoms*. We overload the notation and use the symbol \mathcal{A} both for the relational structure, and for the set of its element, hoping that this does not lead to confusion. We assume that the vocabulary of \mathcal{A} is finite. We shall now formally introduce the notion of definable sets over \mathcal{A} , following [10, 11].

Definable sets. An *expression* is either a variable from some fixed infinite set, or a formal finite union (including the empty union \emptyset) of *set-builder expressions* of the form

$$\{e \mid a_1, \dots, a_n \in \mathcal{A}, \phi\}, \quad (2.1)$$

where e is an expression, a_1, \dots, a_n are pairwise different (bound) variables, and ϕ is a first-order formula over the signature of \mathcal{A} which may use variables. The variables a_1, \dots, a_n are bound. Free variables in (2.1) are those free variables of e and of ϕ which are not among a_1, \dots, a_n .

For an expression e with free variables V , any valuation $\text{val} : V \rightarrow \mathcal{A}$ defines in an obvious way a value $X = e[\text{val}]$, which is either an atom or a set, formally defined by induction on the structure of e . We then say that X is a *definable set over \mathcal{A}* , and that it is *defined* by e with val . When a structure of atoms \mathcal{A} is obvious from the context, we simply speak of *definable sets* without explicitly specifying \mathcal{A} . Note that one set X can be defined by many different expressions. Finally, observe that the family of definable sets is hereditary: every element of a definable set is a definable set, or an atom $a \in \mathcal{A}$.

Sometimes we want to emphasize those atoms that appear in the image of the valuation $\text{val} : V \rightarrow \mathcal{A}$. For any finite set $S \subseteq \mathcal{A}$ of atoms with $\text{val}(V) \subseteq S$ we say that S *supports* $X = e[\text{val}]$, or that S is a *support* of X , or that X is *S -definable*. Clearly, an S -definable set is also T -definable whenever $S \subseteq T$.

As syntactic sugar, we allow atoms to occur directly in set expressions (these atoms we call *parameters*). For example, what we write as the $\{\underline{1}\}$ -definable set $\{a \mid a \in \mathcal{A}, a \neq \underline{1}\}$ is formally defined by the expression $\{a \mid a \in \mathcal{A}, a \neq b\}$, together with a valuation mapping b to $\underline{1}$. With this syntactic sugar, a definable set is determined by a sole expression e , without valuation.

As a notational convention, when writing set-builder expressions (2.1) we omit the formula ϕ when it is trivial, and omit the enumeration $a_1, \dots, a_n \in \mathcal{A}$ when $n = 0$. This allows us, in particular, to write singletons, like $\{\underline{1}\}$.

Remark 2.1. *To improve readability, it will be convenient to use standard set-theoretic encodings to allow a more flexible syntax. In particular, ordered pairs and tuples can be encoded e.g. by Kuratowski pairs, $(x, y) = \{\{x, y\}, \{x\}\}$. We will also consider as definable infinite families of symbols, such as $\{R_x : x \in X\}$, where R is a symbol and X is a definable set. Formally, such a family can be encoded as the set of ordered pairs $\{R\} \times X$, where the symbol R is represented by some \emptyset -definable set, e.g. \emptyset or $\{\emptyset\}$. Here we use the fact that definable sets (over fixed atoms \mathcal{A}) are closed under Cartesian products.*

Definable relational structures. Any object in the set-theoretic universe (a relation, a function, a logical structure, etc.) may be definable. For example, a definable relation on X, Y is a relation $R \subseteq X \times Y$ which is a definable set of pairs, and a definable function $X \rightarrow Y$ is a function whose graph is definable. Along the same lines, a definable relational signature is a definable set of *symbols* Σ , partitioned into definable sets $\Sigma = \Sigma_1 \uplus \Sigma_2 \uplus \dots \uplus \Sigma_l$ according to the arity of symbols. We say that σ has *arity* r if $\sigma \in \Sigma_r$, and $l \in \mathbb{N}$ is thus the maximal arity of a symbol in Σ .

For a signature Σ , a definable Σ -structure \mathbb{A} consists of a definable universe A and a definable interpretation function which assigns a relation $\sigma^{\mathbb{A}} \subseteq A^r$ to each relation symbol $\sigma \in \Sigma$ of arity r . (We denote structures using blackboard font, and their universes using the corresponding symbol in *italics*). More explicitly, such a structure can be represented by the tuple $\mathbb{A} = (A, I_1, \dots, I_l)$ where $I_r = \{(\sigma, a_1, \dots, a_r) \mid \sigma \in \Sigma_r, (a_1, \dots, a_r) \in \sigma^{\mathbb{A}}\}$ is a definable set for $r = 1, \dots, l$. It is not difficult to see that the interpretation $\sigma^{\mathbb{A}}$ of every fixed symbol $\sigma \in \Sigma$ is definable.

Example 2.2. The graph G from Example 1.1 is a definable (over \mathcal{A} the pure set) structure over a finite signature Σ containing a single binary relation symbol. To give an example of a definable structure over an infinite definable signature, extend G to a structure \mathbb{A} by infinitely many unary predicates representing the neighborhoods of each vertex of G . To this end, define the signature $\Sigma = \{E\} \cup \{N_v \mid v \in V\}$, where $V = \{\{a, b\} \mid a, b \in \mathcal{A}, a \neq b\}$ is the vertex set of G and N is a symbol (cf. Remark 2.1). The interpretation of N_v is specified by the set $I_1 = \{(N_v, w) \mid (v, w) \in E\}$ (where E is defined by the expression from Example 1.1). \square

Representing the input. Definable relational structures can be input to algorithms, as they are finitely presented by expressions defining the signature, the universe, and the interpretation function. If the input is an S -definable set X , defined by an expression e with parameters $a_1, \dots, a_n \in S$, then we also need to represent the tuple a_1, \dots, a_n of atoms. For example, in the case of pure set \mathcal{A} these elements can be represented as numbers.

Definable isomorphism problem. Let atoms \mathcal{A} be fixed. Recall that a definable function $f : X \rightarrow Y$ is a function whose graph $\Gamma(f) \subseteq X \times Y$ is a definable sets. A definable isomorphism between definable structures \mathbb{A}, \mathbb{B} over the same signature Σ is a definable bijective function $h : A \rightarrow B$ between their universes that preserves and reflects every relation in the signature: for every $\sigma \in \Sigma$ of arity r and every r -tuple $a_1, \dots, a_r \in A$ of elements of \mathbb{A} , $(a_1, \dots, a_r) \in \sigma^{\mathbb{A}}$ if, and only if $(h(a_1), \dots, h(a_r)) \in \sigma^{\mathbb{B}}$. Likewise one can also introduce definable homomorphisms, embeddings, etc.

Remark 2.3. As argued in [11], definable structures over finite signatures coincide, up to definable isomorphism, with first-order interpretations with parameters in \mathcal{A} , in the sense of model theory [9]. \square

We focus in this paper on the following family of decision problems (note that the structure of atoms \mathcal{A} is fixed, and not part of input, and hence every choice of \mathcal{A} yields a different decision problem):

Problem: DEFINABLE-ISOMORPHISM(\mathcal{A})

Input: A definable signature Σ and two definable Σ -structures \mathbb{A} and \mathbb{B} .

Decide: Is there a definable isomorphism from \mathbb{A} to \mathbb{B} ?

Example 2.4. Imposing the definability requirement on isomorphisms clearly does matter. Let \mathcal{A} be the pure set again, and consider the following two \emptyset -definable graphs, each of them being an infinite clique:

$$V_1 = \{a \mid a \in \mathcal{A}\} = \mathcal{A} \quad V_2 = \{\{a, b\} \mid a, b \in \mathcal{A}, a \neq b\},$$

$$E_1 = \{\{a, b\} \mid a \neq b\} \quad E_2 = \{\{\{a, b\}, \{c, d\}\} \mid a, b, c, d \in \mathcal{A}, a \neq b \wedge c \neq d \wedge \{a, b\} \neq \{c, d\}\},$$

where $\{a, b\} \neq \{c, d\}$ is a shorthand for $\neg((a = c \wedge b = d) \vee (a = d \wedge b = c))$. The two graphs are clearly isomorphic. On the other hand there is no *definable* isomorphism between them, simply because there is no definable bijection between V_1 and V_2 , as we will argue in Example 4.5 in Section 4. \square

3. DECIDABILITY OF THE DEFINABLE ISOMORPHISM PROBLEM

As before, let atoms \mathcal{A} be an arbitrary fixed countable relational structure over a finite vocabulary. Let's start by recalling some terminology in order to formulate the assumptions we are going to impose on the structure of atoms.

Recall that a relational structure \mathcal{A} is *homogeneous* if every isomorphism of its two finite induced substructures extends to an automorphism of the whole structure \mathcal{A} . Equivalently, a homogeneous structure is the Fraïssé limit of a class of finite structures closed under induced substructures and amalgamation (see [8, 15] for details). Among good properties of homogeneous structures, the one most relevant in the sequel is *quantifier elimination*: every first-order formula is equivalent to a quantifier-free one. We will assume homogeneity of \mathcal{A} .

Our second assumption is *effectiveness* of the structure \mathcal{A} : it is decidable, for a given finite relational structure \mathbb{A} over the same vocabulary as \mathcal{A} , whether \mathbb{A} *embeds* into \mathcal{A} (i.e., whether \mathbb{A} is isomorphic to an induced substructure of \mathcal{A}). Effectiveness of \mathcal{A} , combined with homogeneity, implies that the quantifier elimination is effective and, in consequence, the first order theory of \mathcal{A} is decidable.

Automorphism of \mathcal{A} we call *atom automorphisms*. Those atoms automorphisms π which fix all elements of S , i.e., $\pi(a) = a$ for every $a \in S$, we call *atom S -automorphisms*. The structure \mathcal{A} we call *dense* if for every finite $S \subseteq \mathcal{A}$, there is an embedding $H : \mathcal{A} \rightarrow \mathcal{A} - S$ such that every automorphism of $H(\mathcal{A})$ extends to an atom S -automorphism. When \mathcal{A} is homogeneous, the image $H(\mathcal{A})$ – seen as an induced substructure of \mathcal{A} – being the isomorphic copy of \mathcal{A} , is therefore homogeneous too.

We say that the structure \mathcal{A} *admits least supports* if every definable set x over \mathcal{A} has the least (with respect to inclusion) support. Equivalently: the supports of x are closed under intersections (recall that supports are finite). The least support of x we denote by $\text{supp}(x)$ and refer to as *the support of x* . Here are all our assumptions:

- \mathcal{A} is homogeneous;
- \mathcal{A} is effective;
- \mathcal{A} is dense;
- \mathcal{A} admits least supports.

Many structures \mathcal{A} satisfy all the above assumptions. Here are few examples:

- the pure set;
- the dense total order (\mathbb{Q}, \leq) of rational numbers;
- the universal (random) graph (the Fraïssé limit of all finite graphs [8]);

- the universal partial order (the Fraïsse limit of all finite partial orders).

Theorem 3.1. *DEFINABLE-ISOMORPHISM(\mathcal{A}) is decidable whenever the structure of atoms \mathcal{A} is homogeneous, effective, dense, and admits least supports.*

We note that Theorem 3.1 holds for *any* definable signature, possibly infinite. A key fact which we use to prove Theorem 3.1 is the following lemma:

Lemma 3.2 (Parameter elimination). *Suppose that the structure of atoms \mathcal{A} is homogeneous, dense, and admits least supports. Let $T \subseteq \mathcal{A}$ be a finite set of atoms, and let \mathbb{A}, \mathbb{B} be two T -definable relational structures over a T -definable signature Σ . If \mathbb{A} and \mathbb{B} are related by a definable isomorphism, then they are also related by a T -definable one.*

In rough words, the lemma says that in a definition of an isomorphism between T -definable structures, one can eliminate parameters outside of T , maybe at a price of modifying the isomorphism.

Necessity of assumptions. We briefly comment on the choice of assumptions in Theorem 3.1 and in Lemma 3.2. Homogeneity is a standard assumption that makes definable sets orbit-finite (cf. Section 4) and effectiveness is a standard way of making definable sets algorithmically tractable (see e.g. [7]). These two assumptions alone are not sufficient for our purposes, in view of the following counterexample:

Example 3.3. Consider, as the structure \mathcal{A} of atoms, the full bipartite graph, with both parts infinite countable. This structure is effective and homogeneous, but does not satisfy Lemma 3.2. Indeed, consider two \emptyset -definable two-element sets (by E we denote the edge relation in \mathcal{A}):

$$A = \{ \{ b \mid \neg b E a \} \mid a \in \mathcal{A} \} \quad \text{and} \quad B = \{ \emptyset, \{ \emptyset \} \}.$$

The two elements of the first one are just the two parts of the bipartite graph \mathcal{A} ; the two elements of the other one are two arbitrarily chosen \emptyset -definable sets. We consider A and B as the relational structures over the empty signature. Clearly, A and B are related by a bijection; on the other hand, there is no \emptyset -definable bijection between them. Indeed, according to Lemma 4.2 in Section 4 below, any such \emptyset -definable bijection would commute with all atom automorphisms, which is in contradiction with the fact that atom automorphisms sometimes swap the elements of A , but never the elements of B . \square

The assumption that \mathcal{A} is dense is a crucial property which we apply a number of times when proving Lemma 3.2, and we do not envisage proving the result without this assumption. On the other hand, we suspect that admitting least supports can be relaxed. However, as homogeneous structures typically admit least supports, one would not gain much by dropping this assumption, while a new layer of technical complication would be necessarily introduced.

Lemma 3.2 implies Theorem 3.1. We now sketch the proof of Theorem 3.1, using Lemma 3.2. By the latter lemma, DEFINABLE-ISOMORPHISM(\mathcal{A}) reduces to testing whether the given T -definable Σ -structures \mathbb{A}, \mathbb{B} are related by a T -definable isomorphism. In turn, as we show now, testing of the latter condition reduces to evaluation of first-order formulas in \mathcal{A} . Let the given structures be $\mathbb{A} = (A, I_1, \dots, I_l)$ and $\mathbb{B} = (B, J_1, \dots, J_l)$. We follow the lines of the proof of Thm. 12 in [10]; in particular, we build on the following fact (cf. Lem. 13 in [10]):

Lemma 3.4. *For any finite set $T \subseteq \mathcal{A}$ of atoms, a T -definable set X has finitely many T -definable subsets, and expressions defining them can be computed effectively from an expression defining X .*

Indeed, for each definable set X represented by a single set-builder expression of the form (2.1), replace ϕ by each (up to equivalence) quantifier-free formula ψ with the same free variables, such that $\psi \implies \phi$.

To verify existence of a T -definable isomorphism from A to B , apply Lemma 13 to $X = A \times B$ and for every T -definable subset $R \subseteq A \times B$, test the validity of the first-order formula

$$\forall a \in A \exists! b \in B R(a, b) \wedge \forall b \in B \exists! a \in A R(a, b)$$

ensuring that R is the graph of a bijection; and for every $i = 1, \dots, l$, test the validity of the formula

$$\forall \sigma \in \Sigma_i \forall a_1, \dots, a_i \in A \forall b_1, \dots, b_i \in B \bigwedge_{1 \leq j \leq i} R(a_j, b_j) \implies (I_i(\sigma, a_1, \dots, a_i) \iff J_i(\sigma, b_1, \dots, b_i))$$

ensuring that the function is an isomorphism. Evaluation of first-order formulas of the above form reduces to evaluation of first order formulas in \mathcal{A} , see [6, 10] for further details.

4. DEFINABLE SETS VIA ACTION OF ATOM AUTOMORPHISMS

For the proof of Lemma 3.2 it will be more convenient to take a different perspective on definable sets, namely via action of atom automorphisms. This view emphasises that definable sets are always *orbit-finite* (cf. Lemma 4.3 below). In this section we provide the necessary definitions and properties that will be useful in the proof of Lemma 3.2 in the next section. All further missing details can be found in [2].

Definable sets contain, as elements, either other definable sets or atoms $a \in \mathcal{A}$. The group of atom automorphisms acts naturally on such sets, by renaming all atoms appearing as elements, as elements of elements, etc. The action preserves definable sets: a definable set is mapped to a definable set. By πx we denote the result of the action of an atom automorphism π on a definable set x . For instance, consider the pure set \mathcal{A} as atoms and the atom automorphism π that swaps $\underline{0}$ with $\underline{1}$, and $\underline{3}$ with $\underline{4}$, and preserves all other atoms. Then

$$\pi \{a \mid a \in \mathcal{A}, a \neq \underline{1} \wedge a \neq \underline{2}\} = \{a \mid a \in \mathcal{A}, a \neq \underline{0} \wedge a \neq \underline{2}\} \quad \pi \{\underline{0}, \underline{1}, \underline{2}\} = \{\underline{0}, \underline{1}, \underline{2}\}.$$

The action defines a partition of all definable sets into *orbits*: x and x' are in the same orbit if $\pi x = x'$ for some atom automorphism π . In the same vein, for every finite set $S \subseteq \mathcal{A}$, the action of the subgroup of atom S -automorphisms defines a finer partition of all definable sets into *S -orbits* (\emptyset -orbits are just orbits). In particular, the set \mathcal{A} of atoms is itself also partitioned in S -orbits.

By inspecting the syntactic form of definable sets, one easily verifies the following basic fact:

Lemma 4.1. *Every S -definable set is closed under the action of atom S -automorphisms on its elements, i.e., is a union of S -orbits.*

We will intensively use the following direct conclusion from Lemma 4.1:

Lemma 4.2. *Every S -definable function h commutes with atom S -automorphisms: for every atom S -automorphism π , $h\pi = \pi h$.*

Finite unions of S -orbits we call S -orbit finite sets. Homogeneity of \mathcal{A} guarantees finiteness in the statement of Lemma 4.1:

Lemma 4.3. *Assume that the structure of atoms \mathcal{A} is homogeneous. Then every S -definable set is S -orbit finite.*

For an S -orbit O and an S -definable set X , if $O \subseteq X$ we say that O is an S -orbit inside X ; on the other hand, whenever $x \in O$ we say that O is the S -orbit of x , and write $O = \text{orbit}_S(x)$. The converse of Lemma 4.2 is also true when \mathcal{A} is homogeneous:

Lemma 4.4. *Assume that the structure of atoms \mathcal{A} is homogeneous. Then every function h that commutes with atom S -automorphisms, with $\text{Dom}(h)$ and $\text{Codom}(h)$ being S -definable, is itself S -definable.*

Example 4.5. Lemma 4.2 can be used to prove the claim formulated in Example 2.4 in Section 2: there is no definable bijection between the following two sets

$$V_1 = \{a \mid a \in \mathcal{A}\} = \mathcal{A} \quad V_2 = \{\{a, b\} \mid a, b \in \mathcal{A}, a \neq b\}.$$

Suppose the contrary, and let $f : V_1 \rightarrow V_2$ be an S -definable bijection. Consider any $a, b, c \in \mathcal{A} - S$ with $f(a) = \{b, c\}$, and any atom S -automorphism π that swaps b and c and does not preserve a , say $\pi(a) = a'$. Note that such π always exists, e.g., when $a = b$ then $a' = c$. By Lemma 4.2 we obtain: $f(a') = f\pi(a) = \pi f(a) = \pi\{b, c\} = \{b, c\} = f(a)$, which is in contradiction with bijectivity of f . \square

We will also need some basic properties involving the least supports. Denote by $\text{supp}(x)$ the least support of x . First, we note that the support function commutes with atom automorphisms:

Lemma 4.6. *For every definable set x and atom automorphism π , we have $\text{supp}(\pi x) = \pi \text{supp}(x)$.*

Proof. Indeed, the inclusion $\text{supp}(\pi x) \subseteq \pi \text{supp}(x)$ follows by the fact that $\text{supp}(\pi x)$ is the least support of πx while $\pi \text{supp}(x)$ still supports πx . The other inclusion $\pi \text{supp}(x) \subseteq \text{supp}(\pi x)$ follows from the first one: replace in the first one x by πx , and π by π^{-1} , obtaining $\text{supp}(\pi^{-1} \pi x) \subseteq \pi^{-1} \text{supp}(\pi x)$, and then apply π to both sides. \square

The cardinality of $\text{supp}(x)$ we call the *dimension* of x . As a conclusion from Lemma 4.6:

Corollary 4.7. *Every two elements of the same \emptyset -orbit have necessarily the same dimension.*

Moreover, the action of an atom automorphism π on x depends only on the restriction of π to $\text{supp}(x)$:

Lemma 4.8. *If atom automorphisms π, π' coincide on $\text{supp}(x)$, then $\pi x = \pi' x$.*

Finally, we observe the relationship between the least support of a function, its argument and value:

Lemma 4.9. *Let f be a definable function and let $a \in \text{Dom}(f)$. Then $\text{supp}(f(x)) \subseteq \text{supp}(f) \cup \text{supp}(x)$.*

5. PROOF OF LEMMA 3.2

In this section we prove Lemma 3.2. Consider two T -definable structures \mathbb{A} and \mathbb{B} related by an S -definable isomorphism f and assume, w.l.o.g., that $S \supseteq T$. We are going to modify suitably the isomorphism in order to obtain a possibly different one which will be T -definable. For the sake of readability we conduct the proof in the special case, under the following assumptions:

- $T = \emptyset$,
- the signature contains just one binary symbol.

Thus we assume the structures \mathbb{A}, \mathbb{B} to be \emptyset -definable directed graphs. The proof adapts easily to the general case, as discussed at the end of this section.

Consider therefore two \emptyset -definable directed graphs $\mathbb{A} = (A, E)$ and $\mathbb{B} = (B, F)$, where A and B are sets of nodes, and $E \subseteq A \times A$ and $F \subseteq B \times B$ are sets of directed edges, together with an S -definable isomorphism $f : A \rightarrow B$ of graphs. Assume w.l.o.g. that the node sets A and B are disjoint. By Lemma 4.3, the set A of nodes of \mathbb{A} , being itself \emptyset -definable, splits into finitely many \emptyset -orbits. Likewise the set B of nodes of \mathbb{B} splits into finitely many \emptyset -orbits, but a priori it is not clear whether the numbers of \emptyset -orbits inside A and B are equal. As a side conclusion of our proof, it will be made clear that they really are.

By density of \mathcal{A} we know that the structure \mathcal{A} embeds into $\mathcal{A} - S$. Fix in the sequel an embedding $H : \mathcal{A} \rightarrow \mathcal{A} - S$ such that every automorphism of $H(\mathcal{A})$ extends to an atom S -automorphism. Atoms in $H(\mathcal{A})$ we call *S -independent*. Along the same lines, a node $x \in A \cup B$ we call *S -independent* if $\text{supp}(x) \subseteq H(\mathcal{A})$. By homogeneity and density of atoms we have:

Claim 5.1. *Every \emptyset -orbit inside $A \cup B$ contains an S -independent node.*

(The claim, as well as few other claims formulated below, will be proved below once the proof of Lemma 3.2 is finished.)

We are going to construct an \emptyset -definable bijection $h : A \rightarrow B$ (which will be later shown to be an isomorphism); to this aim we will define inductively a sequence of partial bijections

$$h_i : A \rightarrow B,$$

for $i = 0, 1, \dots, m$, where m is the number of \emptyset -orbits inside A (or B , as made explicit below), such that the domain $\text{Dom}(h_{i+1})$ of every h_{i+1} extends the domain $\text{Dom}(h_i)$ of h_i by one \emptyset -orbit inside A . The required bijection will be $h = h_m$. The order of adding \emptyset -orbits into the domain of h will be relevant for showing h to be an isomorphism.

We start easily by taking as h_0 the empty function. For the induction step, suppose that h_n is already defined. Among the remaining \emptyset -orbits of A and B , i.e., among those which are not included in $\text{Dom}(h_n) \cup \text{Codom}(h_n)$, choose an orbit O whose elements have maximal dimension (cf. Corollary 4.7). W.l.o.g. assume that $O \subseteq A$ (if $O \subseteq B$, change the roles of \mathbb{A} and \mathbb{B} , and replace f and h_n by their inverses). Choose an arbitrary S -independent node $x_0 \in O$. We observe that the S -orbit of x_0 does not depend on the choice of x_0 :

Claim 5.2. *All S -independent nodes in O belong to the same S -orbit.*

We call this S -orbit $M \subseteq O$ the *starting S -orbit* inside O , and all its elements *starting nodes*.

Define the sequence $x_0, y_0, x_1, y_1, x_2, \dots$ of nodes, alternating between nodes of \mathbb{A} and nodes of \mathbb{B} , by the following equalities:

$$h_n(x_{i+1}) = y_i = f(x_i). \quad (5.1)$$

In words, every next y_i is obtained from x_i by applying f ; and every next x_{i+1} is obtained from y_i by applying $(h_n)^{-1}$. Clearly, the latter application is well-defined only when $y_i \in \text{Codom}(h_n)$. We thus stop generating the sequence when $y_i \notin \text{Codom}(h_n)$ for the first time. We need however to prove that this will eventually happen, i.e., that the sequence is finite:

Claim 5.3. $y_l \notin \text{Codom}(h_n)$, for some $l \geq 0$.

(In particular, when $n = 0$ the claim necessarily holds for $l = 0$.) The S -orbit M' of y_l we also call the starting S -orbit inside the \emptyset -orbit O' of y_l (\mathbb{A} and \mathbb{B} are treated symmetrically here). The number l we will call the length of the starting orbit M , and $n + 1$ we will call its order; we write $\text{order}(M) = n + 1$, and $\text{length}(M) = l$. By abuse of notation, we will also assign the same order and length to the \emptyset -orbit O , and to every element of O .

We are now ready to define the bijection h_{n+1} with $\text{Dom}(h_{n+1}) = \text{Dom}(h_n) \cup O$. It agrees with h_n on $\text{Dom}(h_n)$; on the orbit O , we define h_{n+1} by extending the mapping $x_0 \mapsto y_l$ to the whole O . We claim that there is a (unique) \emptyset -definable bijection between the \emptyset -orbit O of x_0 and the \emptyset -orbit O' of y_l that maps x_0 to y_l :

Claim 5.4. *The set of pairs $\{(\pi x_0, \pi y_l) \mid \pi \text{ is an atom automorphism}\}$ is an \emptyset -definable bijection between O and O' .*

This completes the induction step of the definition of h .

Before proving that h is an isomorphism, we formulate a concise equality describing h ; the equality will be useful later. The definition of h_{n+1} in the induction step does not depend on the choice of S -independent node $x_0 \in M$. To see this, observe that Claim 5.3 holds, with the same value l , for every other choice of $x_0 \in M$; indeed, both f and $(h_n)^{-1}$ are S -definable, and hence by Lemma 4.2 they preserve the relation of belonging to the same S -orbit; thus, irrespectively which S -independent node $x_0 \in M$ is chosen, it will always belong to the same S -orbit (by Claim 5.2), and thus the node y_l will always belong to the same S -orbit inside B . Thus by Lemma 4.2 and by the very definition of h we have, for every starting node $x \in A$, the following equality:

$$h(x) = \left[f \circ (h^{-1} \circ f)^l \right](x)$$

where $l = \text{length}(x)$; or equivalently (as f and h are bijections)

$$(h^{-1} \circ f)^{l+1}(x) = x. \quad (5.2)$$

(In the special case of starting nodes x of $\text{order}(x) = 1$ (recall that 1 is the minimal possible order) we have $\text{length}(x) = 0$, and thus the above equality reduces to $h(x) = f(x)$. Thus h agrees with h on the starting S -orbit of order 1.)

Now we embark on proving that h is an isomorphism, i.e., that for every two nodes $x, x' \in A$,

$$(x, x') \in E \quad \text{if, and only if} \quad (h(x), h(x')) \in F. \quad (5.3)$$

The proof is by induction on the orders of x and x' . Let $n = \text{order}(x)$ and $n' = \text{order}(x')$. For the induction step, suppose that (5.3) holds for all pairs y, y' of starting nodes with $\langle m, m' \rangle = \langle \text{order}(y), \text{order}(y') \rangle$ pointwise strictly smaller than $\langle n, n' \rangle$. Thus we assume that

the claim holds for y, y' whenever $m \leq n$, $m' \leq n'$, but either $m < n$ or $m' < n'$. Let $l = \text{length}(x)$ and $l' = \text{length}(x')$.

We first prove (5.3) in the special case when both x and x' are starting nodes. (Recall that in the special case $n = n' = 1$ we have $h(x) = f(x)$ and $h(x') = f(x')$, and thus the claim follows since f is an isomorphism.) Consider two sequences of nodes,

$$x_0, y_0, x_1, y_1, \dots, \quad \text{and} \quad x'_0, y'_0, x'_1, y'_1, \dots$$

where $x_0 = x$ and $x'_0 = x'$, alternating between nodes of \mathbb{A} and \mathbb{B} , determined by the equalities analogous to (5.1):

$$h(x_{i+1}) = y_i = f(x_i) \quad h(x'_{i+1}) = y'_i = f(x'_i).$$

Analogously as before, the sequence is obtained by alternating applications of f and h^{-1} . In particular, $x_{i+1} = h^{-1}(f(x_i))$ and $x'_{i+1} = h^{-1}(f(x'_i))$. Consider the smallest $k > 0$ such that $x_k = x$ and $x'_k = x'$. Observe that equality (5.2) guarantees that such k exists, for instance $k = (l+1)(l'+1)$ works. Note that the equality (5.2) implies also that

$$h(x) = y_{k-1} \quad \text{and} \quad h(x') = y'_{k-1}. \quad (5.4)$$

Recall that f is an isomorphism, hence for every $i \geq 0$ we have:

$$(x_i, x'_i) \in E \quad \text{if, and only if} \quad (y_i, y'_i) \in F. \quad (5.5)$$

Using the induction assumption we want to prove additionally

$$(x_{i+1}, x'_{i+1}) \in E \quad \text{if, and only if} \quad (y_i, y'_i) \in F, \quad (5.6)$$

for every i such that $0 < i+1 < k$. By the very definition of h we know that $\text{order}(x_i) \leq n$ and $\text{order}(x'_i) \leq n'$ for every $i \geq 0$. Furthermore, observe that for every $0 < i < k$ we have $\text{order}(x_i) < n$ or $\text{order}(x'_i) < n'$; indeed, the equalities $\text{order}(x_i) = n$ and $\text{order}(x'_i) = n'$, together with the bijectivity of $(h^{-1} \circ f)^i$, would imply $x_i = x$ and $x'_i = x'$. In consequence, the induction assumption (5.3) applies for every pair (x_{i+1}, x'_{i+1}) where $0 < i+1 < k$, which proves the equivalences (5.6). Combining (5.5) with (5.6) we get

$$(x, x') \in E \quad \text{if, and only if} \quad (y_{k-1}, y'_{k-1}) \in F$$

which, by (5.4), is exactly (5.3), as required.

Now we proceed to proving the claim (5.3) for arbitrary nodes $x, x' \in A$. To this aim we will use the following fact, easy to prove using homogeneity and density of \mathcal{A} :

Claim 5.5. *For every two nodes $x, x' \in A$ there is an atom automorphism π such that both πx and $\pi x'$ are starting nodes.*

For any π in Claim 5.5 we have the following sequence of equivalences:

$$\begin{aligned} (x, x') \in E & \text{ iff} \\ (\pi x, \pi x') \in E & \text{ iff} \\ (h(\pi x), h(\pi x')) \in F & \text{ iff} \\ (\pi h(x), \pi h(x')) \in F & \text{ iff} \\ (h(x), h(x')) \in F. \end{aligned} \quad (5.7)$$

The first and the last equivalence hold as the relations E and F , being \emptyset -definable, are closed under atom automorphisms (cf. Lemma 4.1); the second equivalence has been treated previously as both πx and $\pi x'$ are starting nodes; and the third equivalence holds as the

function h is \emptyset -definable (cf. Lemma 4.2). Lemma 3.2 is thus proved, once we prove the yet unproved claims, namely Claims 5.1–5.5.

Proof of Claim 5.1. Consider an \emptyset -orbit $O \subseteq A \cup B$. We will use the embedding $H : \mathcal{A} \rightarrow \mathcal{A} - S$. Take any node $x \in O$ and consider the restriction of H to the finite set $\text{supp}(x) \subseteq \mathcal{A}$. As H is an embedding, the restriction is a finite partial isomorphism which extends, by homogeneity of \mathcal{A} , to an atom automorphism, say π . Then πx is an S -independent node in O , as $\text{supp}(\pi x) = \pi \text{supp}(x) \subseteq H(\mathcal{A})$ (cf. Lemma 4.6). \square

Proof of Claim 5.2. Let $x, x' \in O$ be two S -independent nodes, thus $\text{supp}(x) \cup \text{supp}(x') \subseteq H(\mathcal{A})$. Take an atom automorphism π such that $\pi x = x'$. By Lemma 4.6 we have $\pi \text{supp}(x) = \text{supp}(x')$. The restriction of π to $\text{supp}(x)$ is a finite partial isomorphism, and hence extends to an automorphism of $H(\mathcal{A})$, which in turn extends to an atom S -automorphism τ , by density. Thus $\tau x = x'$ (we use Lemma 4.8 here) and hence x and x' are in the same S -orbit. \square

Proof of Claim 5.3. It is enough to prove that the nodes y_0, y_1, \dots belong to different S -orbits inside B (as B is S -orbit finite, so is $\text{Codom}(h_n)$, and the claim follows). To this aim consider the sequence of S -orbits:

$$\text{orbit}_S(x_0), \text{orbit}_S(y_0), \text{orbit}_S(x_1), \dots$$

and suppose, towards contradiction, that some S -orbit O repeats in the sequence, say $\text{orbit}_S(y_k) = \text{orbit}_S(y_j)$ for some indices $j < k$. Recall that A and B are disjoint, and that $y_i = f(x_i)$ and $x_{i+1} = h_n^{-1}(y_i)$, for every i . Since f is S -definable, and h_n , being \emptyset -definable, is also S -definable, by Lemma 4.2 they both (preserve and) reflect the relation of belonging to the same S -orbit; therefore $\text{orbit}_S(y_{k-\alpha}) = \text{orbit}_S(y_{j-\alpha})$ for $\alpha = 1, \dots, j$, and in consequence the first orbit O necessarily repeats in the sequence:

$$\text{orbit}_S(x_0) = \text{orbit}_S(x_i) \tag{5.8}$$

for $i = k - j > 0$. Recall that $x_i = (h_n^{-1} \circ f)^i$ to observe that the equality (5.8) is impossible: $x_i \in \text{Dom}(h_n)$ while $x_0 \notin \text{Dom}(h_n)$, and $\text{Dom}(h_n)$, being the union of \emptyset -orbits, is also the union of S -orbits. \square

Proof of Claim 5.4. Relying on Lemma 4.4 it is enough to demonstrate that the set of pairs

$$\{ (\pi x_0, \pi y_l) \mid \pi \text{ is an atom automorphism} \}$$

is (the graph of) a bijection. In other words, it is enough to prove that for every two atom automorphism π, π' , the equality $\pi x_0 = \pi' x_0$ holds if, and only if the equality $\pi y_l = \pi' y_l$ holds. This is equivalent to the following condition:

$$\text{for every atom automorphism } \pi, \pi x_0 = x_0 \text{ iff } \pi y_l = y_l. \tag{5.9}$$

Recall that $y_l = g(x_0)$, where $g = [f \circ (h_n^{-1} \circ f)^l]$; as both f and h_n are S -definable functions, their composition g is also S -definable, and hence the condition (5.9) holds for all atom S -automorphisms π . Our aim is to prove (5.9) for all atom automorphisms, using S -independence of x_0 .

Define the S -support as $\text{supp}_S(x) \stackrel{\text{def}}{=} \text{supp}(x) - S$. Let $U = \text{supp}(x_0)$. We claim that $\text{supp}(y_l) = U$ as well, i.e., y_l is S -independent too. To demonstrate this, we first observe using Lemma 4.9 that the mapping g , being S -definable, can only decrease the S -support: $\text{supp}_S(x_0) \supseteq \text{supp}_S(y_l)$. But g^{-1} , being also S -definable, has the same property

and hence $\text{supp}_S(x_0) = \text{supp}_S(y_l)$. Furthermore, x_0 is S -independent and thus satisfies $\text{supp}_S(x_0) = \text{supp}(x_0)$, which yields $U = \text{supp}_S(y_l) \subseteq \text{supp}(y_l)$. Finally, the dimension of y_l is at most equal to the dimension of x_0 (because the node x_0 has been chosen as one with the maximal dimension), i.e., to the cardinality of U , therefore we deduce the equality $U = \text{supp}(y_l)$.

We are now ready to prove (5.9); we concentrate on one implication – the other is proved similarly. Consider any atom automorphism π satisfying $\pi x_0 = x_0$. By Lemma 4.6 we know that π preserves the set U , i.e., $\pi U = U$. We claim that some atom S -automorphism π' coincides with π on U . Indeed, the restriction of π to U extends, as a finite partial isomorphism inside $H(\mathcal{A})$, to an automorphism of $H(\mathcal{A})$; and the latter extends to an atom S -automorphism π' , by density of \mathcal{A} . As π and π' coincide on $U = \text{supp}(x_0)$, by Lemma 4.8 we have $\pi' x_0 = \pi x_0 = x_0$. We can apply (5.9) to the atom S -automorphism π' , thus obtaining $\pi' y_l = y_l$. Finally, again by Lemma 4.8 applied to π and π' , coinciding on $U = \text{supp}(y_l)$, we deduce $\pi y_l = \pi' y_l$. This entails $\pi y_l = y_l$ as required. \square

Proof of Claim 5.5. Similarly as in the proof of Claim 5.1, we use the embedding $H : \mathcal{A} \rightarrow \mathcal{A} - S$. The restriction of H to the finite set $\text{supp}(x) \cup \text{supp}(x') \subseteq \mathcal{A}$ is necessarily a finite partial isomorphism which extends, by homogeneity of \mathcal{A} , to an atom automorphism π , such that the nodes πx and $\pi x'$ are both S -independent and hence, by Claim 5.2, also starting. \square

The general case. The additional assumptions imposed in the proof are not essential, and the proof easily adapts to the general case. To get rid of the assumption $T = \emptyset$, one just needs to consider T -orbits O, O' inside A, B , respectively, instead of \emptyset -orbits. One also easily gets rid of the assumption that the signature has just one binary symbol. First, to deal with (possibly infinitely) many symbols, in the inductive argument towards h being an isomorphism one treats each symbol separately. In case symbols of arity other than 2, say r , the induction is with respect to the r -tuples of orders, instead of pairs thereof. The inductive argument itself, as well as Claim 5.5, adapt easily to r -tuples. Finally, in case of a (possibly infinite) T -definable signature, one needs to modify the sequence of equalities (5.7) appropriately, in order to take into account the action of atom automorphisms on signature symbols. For a signature symbol σ , denote by $\sigma^{\mathbb{A}}, \sigma^{\mathbb{B}}$ the interpretation of σ in \mathbb{A}, \mathbb{B} , respectively. Then the sequence of equalities (5.7) is adapted as follows:

$$\begin{aligned}
(x_1, \dots, x_r) &\in \sigma^{\mathbb{A}} \quad \text{iff} \\
(\pi x_1, \dots, \pi x_r) &\in \pi(\sigma^{\mathbb{A}}) \quad \text{iff} \\
(\pi x_1, \dots, \pi x_r) &\in (\pi\sigma)^{\mathbb{A}} \quad \text{iff} \\
(h(\pi x_1), \dots, h(\pi x_r)) &\in (\pi\sigma)^{\mathbb{B}} \quad \text{iff} \\
(h(\pi x_1), \dots, h(\pi x_r)) &\in \pi(\sigma^{\mathbb{B}}) \quad \text{iff} \\
(\pi h(x_1), \dots, \pi h(x_r)) &\in \pi(\sigma^{\mathbb{B}}) \quad \text{iff} \\
(h(x_1), \dots, h(x_r)) &\in \sigma^{\mathbb{B}}.
\end{aligned}$$

We consider here the action $\pi\sigma$ on the signature symbol σ , as well as the action on the interpretation of the signature symbol, $\pi(\sigma^{\mathbb{A}})$ or $\pi(\sigma^{\mathbb{B}})$. In particular, $(\pi\sigma)^{\mathbb{A}}$ denotes the interpretation of the signature symbol $\pi\sigma$ in \mathbb{A} . As π is a T -automorphism, the first, the second, the fourth and the last equivalence follow by T -definability of \mathbb{A} and \mathbb{B} (cf. Lemma 4.1);

the third equivalence is proved previously, as $\pi x_1, \dots, \pi x_r$ are starting nodes; and the fifth equivalence holds since the function h is T -definable (cf. Lemma 4.2).

REFERENCES

- [1] L. Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In *Procs. STOC 2016*, pages 684–697, 2016.
- [2] M. Bojańczyk. Slightly infinite sets. A draft of a book available at <https://www.mimuw.edu.pl/~bojan/paper/atom-book>.
- [3] M. Bojańczyk, L. Braud, B. Klin, and S. Lasota. Towards nominal computation. In *Procs. POPL 2012*, pages 401–412, 2012.
- [4] M. Bojańczyk, B. Klin, and S. Lasota. Automata theory in nominal sets. *Log. Meth. Comp. Sci.*, 10, 2014.
- [5] M. Bojańczyk, B. Klin, S. Lasota, and S. Toruńczyk. Turing machines with atoms. In *Procs. LICS 2014*, pages 183–192, 2013.
- [6] M. Bojańczyk and S. Toruńczyk. Imperative programming in sets with atoms. In *Procs. FSTTCS 2012*, volume 18 of *LIPIcs*, 2012.
- [7] L. Clemente and S. Lasota. Reachability analysis of first-order definable pushdown systems. In *Procs. CSL’15*, pages 244–259, 2015.
- [8] R. Fraïssé. *Theory of relations*. North-Holland, 1953.
- [9] W. Hodges. *Model Theory*. Number 42 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993.
- [10] B. Klin, E. Kopczyński, J. Ochremiak, and S. Toruńczyk. Locally finite constraint satisfaction problems. In *Procs. of LICS’15*, pages 475–486, 2015.
- [11] B. Klin, S. Lasota, J. Ochremiak, and S. Toruńczyk. Homomorphism problems for first-order definable structures. In *Procs. FSTTCS 2016*, pages 14:1–14:15, 2016.
- [12] B. Klin and M. Szynwelski. SMT solving for functional programming over infinite structures. In *Procs. MSFP 2016*, pages 57–75, 2016.
- [13] E. Kopczyński and S. Toruńczyk. LOIS: an application of SMT solvers. In *Procs. SMT Workshop*, volume 1716 of *CEUR Proceedings*, pages 51–60, 2016.
- [14] E. M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *J. Comput. Syst. Sci.*, 25(1):4265, 1982.
- [15] D. Macpherson. A survey of homogeneous structures. *Discrete Mathematics*, 311(15):15991634, 2011.
- [16] A. M. Pitts. *Nominal Sets: Names and Symmetry in Computer Science*. Cambridge University Press, 2013.