Preservation Theorems Through the Lens of Topology

Aliaume Lopez
Université Paris-Saclay, ENS Paris-Saclay, CNRS, LSV, Gif-sur-Yvette, France
aliaume.lopez@ens-paris-saclay.fr

Abstract

In this paper, we introduce a family of topological spaces that captures the existence of preservation theorems. The structure of those spaces allows us to study the relativisation of preservation theorems under suitable definitions of surjective morphisms, subclasses, sums, products, topological closures, and projective limits. Throughout the paper, we also integrate already known results into this new framework and show how it captures the essence of their proofs.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Mathematics of computing → Discrete mathematics; Mathematics of computing → Point-set topology

Keywords and phrases Preservation theorem, Pre-spectral space, Noetherian space, Spectral space

Digital Object Identifier 10.4230/LIPIcs.CSL.2021.32

Related Version A full version of the paper is available at `https://arxiv.org/abs/2007.07879`.

Acknowledgements I thank Jean Goubault-Larrecq and Sylvain Schmitz for their help and support in writing this paper.

1 Introduction

In classical model theory, *preservation theorems* characterise first-order definable sets enjoying some semantic property as those definable in a suitable syntactic fragment [6, Section 5.2]. A well-known instance is the Łoś-Tarski Theorem [37, 28]: a first-order sentence φ is preserved under extensions on all structures – i.e., $A \models \varphi$ and A is an induced substructure of B imply $B \models \varphi$ – if and only if it is equivalent to an existential sentence.

A major roadblock for applying these results in computer science is that preservation theorems generally do not relativise to classes of structures, and in particular to the class of all finite structures (see the discussions in [31, Section 2] and [25, Section 3.4]). In fact, the only case where a classical preservation theorem was shown to hold on all finite structures is Rossman's Theorem [32]: a first-order sentence is preserved under homomorphisms on all finite structures if and only if it is equivalent to an existential positive sentence. This long-sought result has applications in database theory, where existential positive formulæ correspond to unions of conjunctive queries (also known as select-project-join-union queries and arguably the most common database queries in practice [1]). For instance, it is related in [12, Theorem 17] to the existence of homomorphism-universal models (as constructed by chase algorithms) for databases with integrity constraints, in [38, Theorem 3.4] to a characterisation of schema mappings definable via source-to-target tuple-generating dependencies, and in [18, Corollary 4.14] to the naïve evaluation of queries over incomplete databases under open-world semantics. These applications would benefit directly from preservation theorems for more restricted classes of finite structures or for other semantic properties – corresponding to other classes of queries and other semantics of incompleteness – and this has been an active area of research [5, 4, 9, 22, 17]. Like Rossman's result, these proofs typically rely on careful model-theoretic arguments – typically using Ehrenfeucht-Fraïsse games and locality – and each new attempt at proving a preservation theorem seemingly needs to restart from scratch.

29th EACSL Annual Conference on Computer Science Logic (CSL 2021).
Editors: Christel Baier and Jean Goubault-Larrecq; Article No. 32; pp. 32:1–32:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In this paper, we develop a general topological framework for investigating preservation theorems, where preservation theorems, both old and new, can be obtained as byproducts of topological constructions.

As pointed out in the literature, the classical proofs of preservation theorems fail in the finite because the Compactness Theorem does not apply. As we will see in Section 2, one can reinterpret in topological terms the two applications of the Compactness Theorem in the classical proofs of preservation theorems like the Łoś-Tarski Theorem. Here, the topology of interest has the sets of structures closed under extension as its open sets, and one application of the Compactness Theorem shows that the definable open sets are compact (Claim 2.2) while the other shows that the sets definable by existential sentences form a base for the definable open sets (Claim 2.1). In Section 3, we capture these two ingredients in general through the definitions of *logically presented pre-spectral spaces* and *diagram bases* which lead to a generic preservation theorem (Theorem 3.4): under mild hypotheses – which are met in all the preservation results over classes of finite structures in the literature – preservation holds if and only if the space under consideration is logically presented pre-spectral.

The benefit of this abstract, topological viewpoint, is that preservation results can now be proven by constructing new logically presented pre-spectral spaces from known ones.

Here, the topological core of our definition is the one of *pre-spectral* spaces, which generalise both Noetherian spaces and spectral spaces [19, 13]; see Section 4. From this point onwards the use of the word *stability* will always be used to describe *closure under some operations* and will never be used as the Model Theoretic notion of *stability*, this choice is motivated by the fact that in topology *closure* has a specific meaning, and we already are using the word *preservation* to describe *preservation theorems.* To some extent, we can rely on the stability of spectral spaces under various topological constructions to investigate the same constructions for pre-spectral spaces. We focus however in the paper on the *logically presented* pre-spectral spaces, which is where the main difficulty lies when attempting to prove preservation over classes of finite structures, and for which stability must take the logical aspect into account. Accordingly, Section 5 shows the stability of logically presented pre-spectral spaces under typical constructions: under a carefully chosen notion of morphisms, under subclasses provided a sufficient condition is met, and under finite sums and finite products.

Where the topological viewpoint really shines is when it comes to stability for various kinds of "limits" of classes of structures enjoying a preservation property. We show in Section 6 that the limit of a *single* class of structures, when it can be construed as the *closure* in a suitable topology of a logically presented pre-spectral space, is also logically presented pre-spectral. This allows us to show that Rossman's Theorem – i.e., homomorphism preservation in the finite – extends to the class of structures with the finite model property, and also extends to countable unions of finite structures (the latter was also shown in [30, Chapter 10]). In Section 7, we show that the limit of a *family* of pre-spectral spaces, when built as a *projective limit*, is also pre-spectral. We use this to show that Rossman's proof of homomorphism preservation in the finite can be re-cast in our framework as building exactly such a projective limit.

Due to space constraints, detailed proofs and additional examples will be found in the full paper.

2 Preservation Theorems

In this section, we revisit classical preservation theorems, whose proofs can be found in many books such as [6, Section 5.2]. We will recall the needed definitions, and illustrate the proof techniques in order to highlight the two ingredients that motivate our definitions of pre-spectral spaces and diagram bases later in Section 3.

2.1 Classical Preservation Theorems

Notations. A σ-structure A over a finite relational signature σ (without constants) is given by a domain $|A|$ and, for each symbol $R \in \sigma$ of arity n, a relation $\mathbf{R}^A \subseteq |A|^n$; A is finite if $|A|$ is finite. The binary symbol "=" will always be interpreted as equality, and will not be explicitly listed in our signatures. We write $\mathrm{Struct}(\sigma)$ for the set[1] of all the σ-structures and $\mathrm{Fin}(\sigma)$ for the finite ones. We assume the reader is familiar with the syntax and semantics of first-order logic over σ. We write $\mathsf{FO}[\sigma]$ for the set of first-order sentences over σ. For such a sentence φ, we write $[\![\varphi]\!]_X \triangleq \{A \in X \mid A \models \varphi\}$ for its set of models over a class of structures $X \subseteq \mathrm{Struct}(\sigma)$; by extension, we let $[\![\mathsf{F}]\!]_X \triangleq \{[\![\varphi]\!]_X \mid \varphi \in \mathsf{F}\}$ denote the collection of F-definable subsets of X for a fragment F of $\mathsf{FO}[\sigma]$.

Abstract Preservation. A preservation theorem over a class of structures $X \subseteq \mathrm{Struct}(\sigma)$ shows that first-order sentences enjoying some semantic property are equivalent to sentences from a suitable a syntactic fragment. More precisely, one can model a semantic property as a collection $\mathcal{O} \subseteq \wp(X)$ of "semantic observations" and consider a fragment $\mathsf{F} \subseteq \mathsf{FO}[\sigma]$: we will say that X has the $(\mathcal{O}, \mathsf{F})$ *preservation property* if

1. for all $\psi \in \mathsf{F}$, $[\![\psi]\!]_X \in \mathcal{O}$, and,
2. for all $\varphi \in \mathsf{FO}[\sigma]$ such that $[\![\varphi]\!]_X \in \mathcal{O}$, there exists $\psi \in \mathsf{F}$ such that $[\![\varphi]\!]_X = [\![\psi]\!]_X$.

In this definition, item 1 is usually proven by a straightforward induction on the formulæ in F, and the challenge is to establish item 2. Item 2 is also where *relativisation* to a subset $Y \subseteq X$ might fail, because a set $U \notin \mathcal{O}$ might still be such that $U \cap Y \in \{V \cap Y \mid V \in \mathcal{O}\}$, and thus there might be new first-order sentences enjoying the semantic property and requiring an equivalent sentence in F.

Put more succinctly, X has the $(\mathcal{O}, \mathsf{F})$ preservation property if

$$\mathcal{O} \cap [\![\mathsf{FO}[\sigma]]\!]_X = [\![\mathsf{F}]\!]_X \,. \tag{1}$$

This formulation explicitly shows how a semantic condition (the left-hand side in (1)) is matched with a syntactic one (the right-hand side). As preservation is of interest beyond first-order logic [20, 15, 10, 17], we will say in full generality that a set X equipped with a lattice $\mathcal{L}$ of sets definable in the logic of interest has the $(\mathcal{O}, \mathcal{L}')$ *preservation property* if

$$\mathcal{O} \cap \mathcal{L} = \mathcal{L}' \tag{2}$$

In the rest of this paper we will assume that $\mathcal{O}$ contains $\emptyset$, contains X, is closed under finite intersections and arbitrary unions. This is equivalent to $\mathcal{O}$ being a collection of open sets and defining a *topology* on X.

---

[1] In order to work over sets instead of proper classes and thereby avoid delicate but out-of-topic foundational issues, every σ-structure in this paper will be assumed to be of cardinality bounded by some suitable infinite cardinal. In particular, the Löwenheim-Skolem Theorem justifies that this is at no loss of generality when working with first-order logic.

Table 1 Classical preservation theorems and their relativisations to the finite case.

preservation theorem	quasi-ordering $\leq$	fragment F	holds in Fin(σ)
homomorphism	$\to$	EPFO	yes [32]
Tarski-Lyndon	$\subseteq$	$\mathsf{EPFO}^{\neq}$	no [3]
Łoś-Tarski	$\subseteq_i$	EFO	no [36, 21, 11]
dual Lyndon	$\twoheadleftarrow$	NFO	no [2, 34]

Monotone Preservation. In a number of cases, which are especially relevant in the applications to database theory mentioned in the introduction [12, 18], the semantic property of interest is a form of monotonicity for some quasi-ordering $\leq$ of Struct(σ). We say that a sentence φ is *monotone* in $X \subseteq \text{Struct}(\sigma)$ if $[\![\varphi]\!]_X$ is *upwards-closed*, meaning that if $A \in [\![\varphi]\!]_X$ and B is a σ-structure in X such that $A \leq B$, then $B \in [\![\varphi]\!]_X$. In terms of abstract preservation, this corresponds to choosing $\mathcal{O}$ as the collection of upwards-closed subsets of X, which is also known as the *Alexandroff topology* and is denoted by $\tau_{\leq}$.

The quasi-ordering $\leq$ in question is typically defined through some class of homomorphisms. Recall that there is a *homomorphism* between two σ-structures A and B, noted $A \to B$, if there exists $f\colon |A| \to |B|$ such that, for all relation symbols R of σ and all tuples $(a_1, \ldots, a_n) \in \mathbf{R}^A$, $(f(a_1), \ldots, f(a_n)) \in \mathbf{R}^B$. When f is injective, this entails that A is (isomorphic to) a (not necessarily induced) *substructure* of B and we write $A \subseteq B$; when f is furthermore *strong* – meaning that for all R and $(a_1, \ldots, a_n) \in |A|^n$, $(f(a_1), \ldots, f(a_n)) \in \mathbf{R}^B$ implies $(a_1, \ldots, a_n) \in \mathbf{R}^A$ – , this entails that A is (isomorphic to) an *induced substructure* of B and we write $A \subseteq_i B$; finally, we write $A \twoheadrightarrow B$ when f is surjective.

Table 1 summarises what is known about monotone preservation theorems. In this table, EFO denotes the set of existential first-order sentences, NFO the set of negative ones (namely negative atoms closed under $\vee$, $\wedge$, $\exists$, and $\forall$), EPFO the set of existential positive ones, and $\mathsf{EPFO}^{\neq}$ the set of existential positive ones extended with atoms of the form $x \neq y$ (interpreted as inequality). Note that Lydon's Theorem, which states that a first-order sentence closed under surjective homomorphisms on all structures is equivalent to a positive one, is presented in Table 1 in its *dual* form with inverse surjective homomorphisms and negative sentences. For all these fragments F and associated quasi-orderings $\leq$, the fact that $[\![\mathsf{F}]\!]_X \subseteq \tau_{\leq}$ is mostly straightforward.

2.2 The Łoś-Tarski Theorem in Topological Terms

We propose now to inspect the proof of the Łoś-Tarski Theorem on a finite relational signature σ, as found for instance in [6, Theorem 3.2.2] or [24, Section 5.4]. We work here with the collection $\mathcal{O} \triangleq \tau_{\subseteq_i}$ of upwards-closed subsets of $X \triangleq \text{Struct}(\sigma)$ for $\subseteq_i$ (this is the Alexandroff topology of the quasi-order $\subseteq_i$) and the fragment $\mathsf{F} \triangleq \mathsf{EFO}[\sigma]$. The Łoś-Tarski Theorem corresponds to the following instantiation of (1):

$$\tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\text{Struct}(\sigma)} = [\![\mathsf{EFO}[\sigma]]\!]_{\text{Struct}(\sigma)} . \tag{3}$$

The proof of the Łoś-Tarski Theorem can be decomposed into two steps, here corresponding to the upcoming claims 2.1 and 2.2, and each invoking the Compactness Theorem. When translated in topological terms, the first shows that EFO defines a *base* for the definable open sets, while the second shows that definable open sets are *compact*.

"Syntactic" Base. Recall that a *base* $\mathcal{B}$ of a topology τ is a collection of open sets such that every open set of τ is a (possibly infinite) union of elements from $\mathcal{B}$. Equivalently, $\mathcal{B}$ is a base of a topology τ whenever $\forall U \in \tau, \forall A \in U, \exists V \in \mathcal{B}, A \in V \subseteq U$. A *subbase* is a collection of open sets such that every open set of τ is a (possibly infinite) union of finite intersections of elements of the subbase. The topology $\langle \mathcal{O} \rangle$ *generated* by a collection $\mathcal{O}$ of sets is the smallest topology containing those sets; $\mathcal{O}$ is then a subbase of $\langle \mathcal{O} \rangle$.

We first prove a weaker version of Equation (3) by proving the equality on the generated topologies. Because $[\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ and $[\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ are lattices, those generated topologies can be seen as generated by infinite disjunctions of sentences in $\mathsf{FO}[\sigma]$ (resp. $\mathsf{EFO}[\sigma]$).

▷ Claim 2.1. The topologies generated by $\tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ and $[\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ are the same, i.e., $\left\langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \right\rangle = \left\langle [\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \right\rangle$.

Proof. First of all, any sentence in $\mathsf{EFO}[\sigma]$ defines an upwards-closed set for $\subseteq_i$, and moreover $\mathsf{EFO}[\sigma] \subseteq \mathsf{FO}[\sigma]$, hence $\left\langle [\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \right\rangle \subseteq \left\langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \right\rangle$.

For the converse inclusion, it suffices to show that $\mathsf{EFO}[\sigma]$ defines a base of the topology $\langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \rangle$. Consider for this a monotone sentence $\varphi \in \mathsf{FO}[\sigma]$ and a structure A such that $A \models \varphi$. Following the classical proofs (e.g., [6, Theorem 3.2.2] or [24, Corollary 5.4.3]), define $\hat{A}$ as the expansion of A with one additional constant c_a for each $a \in |A|$, interpreted by $c_a^{\hat{A}} \triangleq a$. The *diagram* $\mathrm{Diag}(A)$ of A is the set of all quantifier-free sentences over this extended signature that hold in $\hat{A}$. For a structure $\hat{B} \in \mathrm{Struct}(\sigma \cup \{c_a\}_{a \in A})$, we write B for its reduct in $\mathrm{Struct}(\sigma)$ obtained by removing the constants $\{c_a\}_{a \in A}$.

Let $T \triangleq \mathrm{Diag}(A) \cup \{\neg\varphi\}$, and consider $\hat{B} \in \mathrm{Struct}(\sigma \cup \{c_a\}_{a \in A})$ such that $\hat{B} \models T$. Because $\hat{B} \models \mathrm{Diag}(A)$, by construction $A \subseteq_i B$ (in particular, the sentence $\neg(c_a = c_b)$ belongs to $\mathrm{Diag}(A)$ for all $a \neq b$ in $|A|$), and thus $B \models \varphi$ because φ is monotone, and finally $\hat{B} \models \varphi$ because the constants c_a do not occur in φ. Therefore, $\hat{B} \models \varphi \wedge \neg\varphi$, which is absurd: the theory T is inconsistent, and by the Compactness Theorem for first-order logic, there exists a finite conjunction ψ_0 of sentences in $\mathrm{Diag}(A)$, which is already inconsistent with $\neg\varphi$.

Let ψ_A be the existential closure of the formula obtained by replacing each symbol c_a with a variable x_a in ψ_0; note that ψ_A is an existential sentence. By construction, $A \models \psi_A$, and if $B \models \psi_A$, there exists an interpretation of the constants $\{c_a\}_{a \in A}$ allowing to build an expansion $\hat{B}$ such that $\hat{B} \models \psi_0$. As we just saw that the implication $\psi_0 \implies \varphi$ is valid, $\hat{B} \models \varphi$, and since no constant symbol occurs in φ, $B \models \varphi$.

To conclude, for any open set $U \in \langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \rangle$ and for any $A \in U$, there exists a monotone sentence φ such that $A \in [\![\varphi]\!]_{\mathrm{Struct}(\sigma)}$, and we have proven that there exists $[\![\psi_A]\!]_{\mathrm{Struct}(\sigma)} \in [\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ such that $A \in [\![\psi_A]\!]_{\mathrm{Struct}(\sigma)} \subseteq [\![\varphi]\!]_{\mathrm{Struct}(\sigma)} \subseteq U$. Therefore, $[\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$ is a base of $\langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \rangle$. ◁

Compactness. The second step relies on the compactness of the sets $[\![\varphi]\!]_{\mathrm{Struct}(\sigma)}$ for monotone sentences φ. Recall that a subset K is *compact* in a topological space τ if, for any *open cover* $(U_i)_{i \in I}$ of K – i.e., a collection of open sets such that $K \subseteq \bigcup_{i \in I} U_i$ – , there exists a finite subset $I_0 \subseteq I$, such that $K \subseteq \bigcup_{i \in I_0} U_i$ (beware that this definition is also called *quasi-compact* in the literature, because we do not require any separation property here). If $\tau = \langle \mathcal{O} \rangle$, by *Alexander's Subbase Lemma*, K is compact if and only if, from every open cover of K using only sets from $\mathcal{O}$, we can extract a finite open cover of K. As open compact sets play a key role in this paper, we introduce here the notation $\mathcal{K}^{\circ}(X) \triangleq \{U \in \tau \mid U \text{ is compact}\}$. When the topology τ is not clear from the context, we shall write $\mathcal{K}^{\circ}(X, \tau)$.

▷ Claim 2.2. Every monotone sentence defines a compact open subset in the topology $\langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \rangle$, i.e., $\tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \subseteq \mathcal{K}^\circ(\mathrm{Struct}(\sigma), \langle \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \rangle)$.

Proof. Consider a monotone sentence $\varphi \in \mathsf{FO}[\sigma]_{\mathrm{Struct}(\sigma)}$. Let $(U_i)_{i \in I}$ be an open cover of $[\![\varphi]\!]_{\mathrm{Struct}(\sigma)}$. By Alexander's Subbase Lemma, we can assume that for each $i \in I$, $U_i = [\![\varphi_i]\!]_{\mathrm{Struct}(\sigma)}$ for some monotone sentence φ_i. Consider the theory $T \triangleq \{\neg\varphi_i \mid i \in I\} \cup \{\varphi\}$. Because $(U_i)_{i \in I}$ is an open cover, this theory has no models. By the Compactness Theorem for first order logic, there exists a finite set I_0 such that $T_0 \triangleq \{\neg\varphi_i \mid i \in I_0\} \cup \{\varphi\}$ is not satisfiable, proving that $(U_i)_{i \in I_0}$ is an open cover of $[\![\varphi]\!]_{\mathrm{Struct}(\sigma)}$. ◁

▶ Remark 2.3 (Compact sets in $\tau_\leq$). As we will often deal with the Alexandroff topology $\tau_\leq$ of a quasi-order $(X, \leq)$, it is worth noting that $U \in \tau_\leq$ is compact if and only if it is the upward closure $U = \uparrow F$ of some finite subset $F \subseteq_{\mathrm{fin}} X$; this is equivalent to saying that U has finitely many minimal elements up to $\leq$-equivalence [19, Exercise 4.4.22]. Thus Claim 2.2 states that any monotone sentence has finitely many $\subseteq_i$-minimal models in Struct(σ).

Proof of the Łoś-Tarski Theorem. A simple structural induction on the formulæ shows that $[\![\mathsf{EFO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)} \subseteq \tau_{\subseteq_i} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$. Regarding the converse inclusion in Equation (3), consider a sentence $\varphi \in \mathsf{FO}[\sigma]$ defining an open set in $\tau_{\subseteq_i}$. By Claim 2.1, there exists a family $(\psi_i)_{i \in I}$ of existential sentences such that $[\![\varphi]\!]_{\mathrm{Struct}(\sigma)} = \bigcup_{i \in I} [\![\psi_i]\!]_{\mathrm{Struct}(\sigma)}$. By Claim 2.2, there is a finite set $I_0 \subseteq_{\mathrm{fin}} I$ for which the equality still holds. Because $\mathsf{EFO}[\sigma]$ is a lattice, this proves the existence of an existential sentence $\psi \triangleq \bigvee_{i \in I_0} \psi_i$ such that $[\![\varphi]\!]_{\mathrm{Struct}(\sigma)} = [\![\psi]\!]_{\mathrm{Struct}(\sigma)}$. ◀

The two properties singled out in claims 2.1 and 2.2 are of different nature. Claim 2.2 really holds for any topology τ and not only for the Alexandroff topology $\tau_{\subseteq_i}$, as opposed to Claim 2.1. Moreover, Claim 2.1 appears to be the most involved one here, but is often easily proven on classes of finite structures.

3 Pre-spectral Spaces and Diagram Bases

Following the two-step decomposition of the proof of the Łoś-Tarski Theorem given in Section 2.2, we define in this section *logically presented pre-spectral spaces* and *diagram bases*, before showing in Theorem 3.4 how they characterise when a preservation theorem holds.

3.1 Pre-spectral Spaces

As a preliminary step toward our main definition, let us first propose a definition of topological spaces (X, τ) where the compact open sets form a *bounded sublattice* of $\wp(X)$ (by which we mean that $\emptyset$ and X belong to the lattice) that generates the topology.

▶ **Definition 3.1** (Pre-spectral space). *A topological space* (X, τ) *is a* pre-spectral space *whenever* $\mathcal{K}^\circ(X)$ *is a bounded sublattice of* $\wp(X)$ *that generates* τ*, i.e.,* $\langle \mathcal{K}^\circ(X) \rangle = \tau$.

The name "pre-spectral" comes from the theory of *spectral* spaces [13], for which the definition is almost identical (see Section 4.2). Pre-spectral spaces will allow us to tap into the rich topological toolset that has been developed for spectral spaces.

Logical presentations. As seen in Claim 2.2, the topology of interest in a preservation theorem is generated by combining a topological space (X, τ) with a bounded sublattice $\mathcal{L}$ of subsets of X, which will be called the *definable* subsets of X. Let us write $\langle X, \tau, \mathcal{L} \rangle$ for the topological space $(X, \langle \tau \cap \mathcal{L} \rangle)$. The following definition is then a direct generalisation of the statement of Claim 2.2.

▶ **Definition 3.2** (Logically presented pre-spectral space). *Let (X, τ) be a topological space and $\mathcal{L}$ be a bounded sublattice of $\wp(X)$. Then $\langle X, \tau, \mathcal{L}\rangle$ is a* logically presented pre-spectral space *(a* lpps*) if its definable open subsets are compact, i.e., if $\tau \cap \mathcal{L} \subseteq \mathcal{K}^\circ(X)$.*

Whenever σ is a finite relational signature, $X \subseteq \mathrm{Struct}(\sigma)$ for a topological space (X, τ) and $\mathcal{L} = [\![\mathsf{FO}[\sigma]]\!]_X$, we denote it by $\langle X, \tau, \mathsf{FO}[\sigma]\rangle$ for simplicity; e.g., $\langle \mathrm{Struct}(\sigma), \tau_{\subseteq_i}, \mathsf{FO}[\sigma]\rangle$ is a lpps by Claim 2.2.

As $\tau \cap \mathcal{L}$ is closed under finite intersection, any open set in $\langle \tau \cap \mathcal{L}\rangle$ is a union of sets from $\tau \cap \mathcal{L}$, thus any compact open set in $\mathcal{K}^\circ(X)$ is a finite union of sets from $\tau \cap \mathcal{L}$. As $\tau \cap \mathcal{L}$ is also closed under finite unions, this shows the inclusion $\mathcal{K}^\circ(X) \subseteq \tau \cap \mathcal{L}$. Thus, in a lpps, $\mathcal{K}^\circ(X) = \tau \cap \mathcal{L}$ is a bounded lattice and any lpps is indeed a pre-spectral space. Conversely, $\langle X, \tau, \mathcal{K}^\circ(X)\rangle$ is well-defined whenever (X, τ) is a pre-spectral space; in this case $\langle X, \tau, \mathcal{K}^\circ(X)\rangle$ is a lpps and it equals (X, τ) (they have the same points and opens).

Beware however that $(X, \langle \tau \cap \mathcal{L}\rangle) = \langle X, \tau, \mathcal{L}\rangle$ being pre-spectral does not entail that it is a lpps; see Remark 3.6 at the end of the section. While pre-spectral spaces capture the topological core behind Claim 2.2 with a simple definition, the logically presented ones are the real objects of interest as far as preservation theorems are concerned, and most of the technical difficulties arising in the remainder of the paper will be concerned with those.

3.2 Diagram Bases

Regarding Claim 2.1, we simply turn the statement of the claim into a definition, which is typically instantiated with $\mathcal{L} = [\![\mathsf{FO}[\sigma]]\!]_X$ and $\mathcal{L}' = [\![\mathsf{F}]\!]_X$ for a fragment F of $\mathsf{FO}[\sigma]$.

▶ **Definition 3.3** (Diagram base). *Let (X, τ) be a topological space, and $\mathcal{L}$ be a bounded sublattice of $\wp(X)$. Then $\mathcal{L}' \subseteq \mathcal{L}$ is a* diagram base *of $\langle X, \tau, \mathcal{L}\rangle$ if $\langle \tau \cap \mathcal{L}\rangle = \langle \mathcal{L}'\rangle$.*

In particular, if $\mathsf{F} \subseteq \mathsf{FO}[\sigma]$ is stable under finite conjunction, this means that any definable open set in X can be written as an infinite disjunction of F-definable sets. Over $\mathrm{Struct}(\sigma)$, this was the "difficult" step in the classical proof of the Łoś-Tarski Theorem. When $X \subseteq \mathrm{Fin}(\sigma)$, this becomes considerably simpler: for every fragment F in Table 1 and any finite structure A, there exists a *diagram* sentence ψ_A^{F} in F such that $A \leq B$ if and only if $B \models \psi_A^{\mathsf{F}}$ for the corresponding quasi-ordering. Therefore, if φ is monotone and $A \in [\![\varphi]\!]_X$, then $A \in [\![\psi_A^{\mathsf{F}}]\!]_X \subseteq [\![\varphi]\!]_X$, showing that $[\![\mathsf{F}]\!]_X$ is a base of $\langle \tau_\leq \cap [\![\mathsf{FO}[\sigma]]\!]_X\rangle$.

3.3 A Generic Preservation Theorem

We have already seen in the proof of the Łoś-Tarski Theorem why logically presented pre-spectral spaces with a diagram base yield preservation. The following theorem also proves the converse direction, under mild hypotheses on $\mathcal{L}'$: $\mathcal{L}'$ must be a lattice and must define compact sets in X for the topology generated by $\mathcal{L}'$. We usually instantiate the theorem with $X \subseteq \mathrm{Struct}(\sigma)$, $\mathcal{L} = [\![\mathsf{FO}[\sigma]]\!]_X$, and $\mathcal{L}' = [\![\mathsf{F}]\!]_X$ where F is a fragment of $\mathsf{FO}[\sigma]$.

▶ **Theorem 3.4** (Generic preservation). *Let τ be a topology on X, $\mathcal{L}$ a bounded sublattice of $\wp(X)$, and $\mathcal{L}'$ a sublattice of $\mathcal{L}$. The following are equivalent:*

1. *X has the $(\tau, \mathcal{L}')$ preservation property and $\mathcal{L}'$ defines only compact sets for the topology $\langle \mathcal{L}'\rangle$.*
2. *$\langle X, \tau, \mathcal{L}\rangle$ is a lpps and $\mathcal{L}'$ defines a diagram base of it.*

Proof. We prove the two implications separately.

1 $\implies$ 2 Assume that X has the $(\tau, \mathcal{L}')$ preservation property. Consider a set $U \in \mathcal{L} \cap \tau$: by the preservation property, $U \in \mathcal{L}'$. This already shows that $\mathcal{L}'$ defines a diagram base of $\langle X, \tau, \mathcal{L}\rangle$. Hence $\langle \mathcal{L}'\rangle = \langle \tau \cap \mathcal{L}\rangle$. Since $U \in \mathcal{L}'$, U is compact in $\langle \mathcal{L}'\rangle$, which means that U is compact in X. Therefore X is a lpps.

2 $\implies$ 1 Assume that $\mathcal{L}'$ defines a diagram base of $\langle X, \tau, \mathcal{L}\rangle$. If $U \in \tau \cap \mathcal{L}$, then it can be written as a possibly infinite union of elements in $\mathcal{L}'$. Also assume that $\langle X, \tau, \mathcal{L}\rangle$ is a lpps: then by compactness, U can be written as a finite union of elements in $\mathcal{L}'$, hence as a single element in $\mathcal{L}'$ since $\mathcal{L}'$ is a lattice. This proves that X has the $(\tau, \mathcal{L}')$ preservation property. Finally, sets in $\mathcal{L}'$ define compact sets in $\langle \mathcal{L}'\rangle$ because it is precisely the topology of $\langle X, \tau, \mathcal{L}\rangle$. ◀

The additional hypotheses on $\mathcal{L}'$ in items 1 and 2 above are somewhat at odds. Asking for $\mathcal{L}'$ to define a diagram base is asking for $\langle \mathcal{L}'\rangle$ to have enough sets, but asking for $\mathcal{L}'$ to only define compact sets is asking for $\langle \mathcal{L}'\rangle$ not to contain too many sets.

▶ Remark 3.5 (Generic monotone preservation). The condition that F must define compact sets in X in Theorem 3.4.1 is actually mild. Consider the preservation results from Table 1 for a fragment F and $\tau = \tau_{\leq}$ the Alexandroff topology of the associated quasi-ordering $\leq$. Assume that X is a $\leq$-downwards-closed subset of $\mathrm{Struct}(\sigma)$ – this is the setting of the known preservation results for classes of finite structures [5, 4, 32, 9, 22].

Observe that, in each case, $[\![\psi]\!]_{\mathrm{Struct}(\sigma)}$ for a sentence $\psi \in \mathsf{F}$ has finitely many $\leq$-minimal models up to $\leq$-equivalence. Because X is downwards-closed, $[\![\psi]\!]_X$ has the same finitely many $\leq$-minimal models. Thus, by Remark 2.3, $[\![\psi]\!]_X$ is compact in $\tau_{\leq}$, and since $[\![\mathsf{F}]\!]_X \subseteq \tau_{\leq}$, it is also compact in the topology generated by $[\![\mathsf{F}]\!]_X$.

In the case of $X = \mathrm{Fin}(\sigma)$, this downward closure condition is fulfilled and F defines a base, thus $(\tau_{\leq}, \mathsf{F})$ preservation holds if and only if $\langle \mathrm{Fin}(\sigma), \tau_{\leq}, \mathsf{FO}[\sigma]\rangle$ is a lpps.

Theorem 3.4 is a generic relationship between pre-spectral spaces and preservation theorems. The downward closure hypothesis in Remark 3.5 is necessary for the equivalence between the preservation property and pre-spectral spaces to hold, as will be shown later in Example 4.2.

▶ Remark 3.6. For each of the fragments F and associated quasi-orderings $\leq$ of Table 1, $\langle \mathrm{Fin}(\sigma), \tau_{\leq}, \mathsf{FO}[\sigma]\rangle = (\mathrm{Fin}(\sigma), \langle \tau_{\leq} \cap [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Fin}(\sigma)}\rangle)$ is a pre-spectral space. Indeed, by Remark 2.3, any compact open K from $\mathcal{K}^\circ(\mathrm{Fin}(\sigma))$ is the upward closure $K = \uparrow F$ of a finite set $F \subseteq_{\mathrm{fin}} \mathrm{Fin}(\sigma)$, thus $K = [\![\bigvee_{A \in F} \psi_A^{\mathsf{F}}]\!]_{\mathrm{Fin}(\sigma)}$, which shows that $\mathcal{K}^\circ(\mathrm{Fin}(\sigma)) \subseteq [\![\mathsf{F}]\!]_{\mathrm{Fin}(\sigma)}$. As any $\psi \in \mathsf{F}$ has finitely many $\leq$-minimal models in $\mathrm{Fin}(\sigma)$, $\mathcal{K}^\circ(\mathrm{Fin}(\sigma)) \supseteq [\![\mathsf{F}]\!]_{\mathrm{Fin}(\sigma)}$, and since F defines a base, $\langle \mathrm{Fin}(\sigma), \tau_{\leq}, \mathsf{FO}[\sigma]\rangle$ is pre-spectral. However, by Remark 3.5 and the non-preservation results of [36, 21, 3, 2, 34], $\langle \mathrm{Fin}(\sigma), \tau_{\subseteq_i}, \mathsf{FO}[\sigma]\rangle$, $\langle \mathrm{Fin}(\sigma), \tau_{\subseteq}, \mathsf{FO}[\sigma]\rangle$, and $\langle \mathrm{Fin}(\sigma), \tau_{\twoheadleftarrow}, \mathsf{FO}[\sigma]\rangle$ are not lpps: the condition $\tau \cap \mathcal{L} \subseteq \mathcal{K}^\circ(X)$ is crucial in order to derive preservation results.

Another way of reaching the topological definitions of this section is to consider a folklore result employed in several proofs of preservation theorems over classes of finite structures for fragments F of EFO [32, 4, 5, 7]: if X is downwards-closed for $\leq$, a monotone sentence φ is equivalent to a sentence from F if and only if it has finitely many $\leq$-minimal models in X (up to $\leq$-equivalence). By Remark 2.3, this says that $[\![\varphi]\!]_X$ is compact, while the folklore result itself is essentially using the fact that F defines a base.

4 Related Notions

Pre-spectral spaces generalise two notions arising from order theory, topology, and logics: Noetherian spaces and spectral spaces.

4.1 Well-Quasi-Orderings and Noetherian Spaces

A topological space in which all subsets are compact, or, equivalently, all open subsets are compact, is called *Noetherian* [19, Section 9.7]. A Noetherian space (X, τ) and a bounded sublattice $\mathcal{L}$ of $\wp(X)$ always define a lpps $\langle X, \tau, \mathcal{L}\rangle$. A related notion, considering a quasi-order instead of a topology, leads to the well-known notion of *well-quasi-orders* [26]: a quasi-order is a well-quasi-order if and only if its Alexandroff topology is Noetherian [19, Proposition 9.7.17]. Thus, if $(X, \leq)$ is a well-quasi-order and $\mathcal{L}$ is a bounded sublattice of $\wp(X)$, then $\langle X, \tau_{\leq}, \mathcal{L}\rangle$ is a lpps.

Applications of Noetherian Spaces to Preservation. Let us denote by $\mathcal{G}$ the class of finite simple undirected graphs and by $\sigma_{\mathcal{G}}$ the signature with a single binary edge relation E; then the induced substructure ordering $\subseteq_i$ coincides with the induced subgraph ordering over $\mathcal{G}$.

▶ **Example 4.1** (Finite graphs of bounded tree-depth)**.** Recall that the *tree-depth* $\mathrm{td}(G)$ of a graph G is the minimum height of the comparability graphs F of partial orders such that G is a subgraph of F [30, Chapter 6]. Let $\mathcal{T}_{\leq n}$ be the set of finite graphs of tree-depth at most n ordered by the induced substructure relation $\subseteq_i$. This is a well-quasi-order [14], thus $\langle \mathcal{T}_{\leq n}, \tau_{\subseteq_i}, \mathsf{FO}[\sigma_{\mathcal{G}}]\rangle$ is a lpps, and therefore $\mathcal{T}_{\leq n}$ enjoys the $(\tau_{\subseteq_i}, \mathsf{EFO}[\sigma_{\mathcal{G}}])$-preservation property by Theorem 3.4.

▶ **Example 4.2** (Finite cycles)**.** Consider the class $\mathcal{C} \subseteq \mathcal{G}$ of all finite simple cycles. As is well known, $(\mathcal{C}, \subseteq_i)$ is not a well-quasi-order because any two different cycles are incomparable for the induced substructure ordering [14]. In particular, every singleton is an open set: $(\mathcal{C}, \tau_{\subseteq_i})$ is actually a topological space with the *discrete topology*, and its only compact sets are the finite sets: $\langle \mathcal{C}, \tau_{\subseteq_i}, \mathsf{FO}[\sigma_{\mathcal{G}}]\rangle$ is not a lpps.

By standard locality arguments, for any sentence φ, there exists a finite threshold n_0 on the size of cycles, above which φ is either always true or always false (see the full paper for details). Let τ_n be the topology over $\mathcal{C}$ generated by the definable co-finite sets and the definable sets containing only cycles of size at most n. This is a variation of the *co-finite topology*, and is also Noetherian. Hence, $\langle \mathcal{T}_{\leq n}, \tau_{\subseteq_i}, \mathsf{FO}[\sigma_{\mathcal{G}}]\rangle$ is a lpps, and as $\mathsf{EFO}[\sigma_{\mathcal{G}}]$ defines a diagram base of it, we can apply Theorem 3.4 to deduce preservation. Now, given a monotone sentence φ, either φ has finitely many models or it has co-finitely many. In both cases, this sentence defines an open set in τ_n for some n that is definable in $\mathsf{EFO}[\sigma_{\mathcal{G}}]$. Thus the set of finite cycles has the $(\tau_{\subseteq_i}, \mathsf{EFO}[\sigma_{\mathcal{G}}])$ preservation property.

The previous example shows that the closure condition of Remark 3.5 was necessary, by proving that a space of structures can enjoy a preservation theorem while not defining a lpps.

Relativisation. The following proposition shows that, if we are looking for classes of structures where preservation theorems *always* relativise, then we should endow them with a Noetherian topology.

▶ **Proposition 4.3.** *Let (X, τ) be a pre-spectral space such that for all $Y \subseteq X$, Y with the induced topology is pre-spectral. Then X is Noetherian.*

Proof. Consider any subset Y of X: by assumption, Y is pre-spectral, hence compact in the induced topology, hence compact in (X, τ). ◀

4.2 Spectral Spaces

Spectral spaces are a class of topological spaces appearing naturally in the study of logics and algebra as a generalisation of the Stone Duality theory. Throughout this section we refer to two books and keep the notations consistent with them [19, 13]. A closed subset F of a topological space X is *irreducible* whenever F is non-empty and is not the disjoint union of two non-empty closed sets. The *closure* of a set Y in a space X is the smallest closed set containing Y and is denoted by $\overline{Y}^X$ or $\overline{Y}$ when X is clear from the context. A topological space X is *sober* whenever any irreducible closed subset F is the closure of exactly one point $x \in X$, which translates formally to $\exists x \in X, \overline{\{x\}} = F$ and $\forall y \in X, \overline{\{y\}} = F \Rightarrow y = x$. A *spectral space* is a pre-spectral space that is sober [13, Definition 1.1.5].

When a space (X, τ) is not sober, it is possible to build a *sobrified* version of this space as follows [19, Definition 8.2.17]: $\mathcal{S}(X)$ is the set of irreducible closed sets of X, and the topology is generated by the sets $\Diamond U \triangleq \{F \in \mathcal{S}(X) \mid F \cap U \neq \emptyset\}$ where U is an open set of X. It can be shown that this construction leads to a sober space, is idempotent up to homeomorphism, and constructs the *free* sober space over X [19, Theorem 8.2.44]. This leads to the following correspondence between pre-spectral spaces and spectral spaces.

▶ **Fact 4.4** (Spectral versus pre-spectral)**.** *A space X is pre-spectral if and only if $\mathcal{S}(X)$ is spectral.*

The connection with spectral spaces is of particular interest, because the sobrification functor gives a tool to translate result from the rich theory of spectral spaces to pre-spectral spaces which will be extensively used in Section 5.

5 Basic Closure Properties

To study preservation theorems, we not only want to ensure that the space is pre-spectral, but also to see that the lattice of compact open sets is obtained through a restriction of the logic. Therefore, one of our main concerns with closure properties is to characterise the lattice of compact sets, which must use properties of the definable sets and cannot rely solely on topological constructions.

5.1 Morphisms

Spectral Maps. Let us first introduce the notion of morphism between pre-spectral spaces, inherited from the case of spectral spaces [13, Definition 1.2.2]. A map $f\colon (X, \tau) \to (Y, \theta)$ is a *spectral map* whenever it is continuous and the pre-image of a compact-open set of Y is a compact-open set of X. We will write **PreSpec** for the category of pre-spectral spaces and spectral maps.

▶ **Fact 5.1.** *The image of a pre-spectral space through a spectral map is pre-spectral.*

A crucial role of spectral maps is to guard the definition of *pre-spectral subspaces*, mimicking the one of *spectral subspaces* [13, Section 2.1]. A pre-spectral subspace is not only a subset where the induced topology happens to be pre-spectral, but has the additional property that the *inclusion map* is a spectral map.

Logical Maps. In the case of a lpps, a map $f\colon \langle X, \tau, \mathcal{L}\rangle \to \langle Y, \theta, \mathcal{L}'\rangle$ is a *logical map* whenever it is continuous and the pre-image of a definable open set of Y is a definable open set of X. A map between logically defined pre-spectral spaces is logical if and only if it is spectral, since compact open subsets and definable open subsets coincide in that case. However, the use of logical maps is to prove that some spaces are pre-spectral by transferring logical properties rather than topological ones.

▶ **Fact 5.2.** *The image of a lpps $\langle X, \tau, \mathcal{L}\rangle$ through a logical map is a lpps.*

Of particular interest are the logical maps obtained through syntactic constructions. Let us define an FO-*interpretation* $f\colon X \to Y$ where $X \subseteq \mathrm{Struct}(\sigma_1)$ and $Y \subseteq \mathrm{Struct}(\sigma_2)$ through "relation" formulæ ρ_R for all $R \in \sigma_2$, where ρ_R has as many free variables as the arity of R, and an additional "domain" formula $\delta \in \mathsf{FO}[\sigma_1]$ with one free variable. The image of a σ_1-structure $A \in X$ is the σ_2-structure $f(A)$ with domain $|f(A)| \triangleq \{a \in |A| \mid A \models \delta(a)\}$ and such that $(a_1, \dots, a_n) \in \mathbf{R}^{f(A)}$ if and only if $A \models \rho_R(a_1, \dots, a_n)$. This is a simple model of logical interpretations: many different notions can be found in the literature [8].

An FO-interpretation $f\colon X \to Y$ allows to transfer logical properties from one class of structures to another: if $\varphi \in \mathsf{FO}[\sigma_2]$ is a formula on the structures of Y, then there exists a formula $f^{-1}(\varphi) \in \mathsf{FO}[\sigma_1]$ such that $A \models f^{-1}(\varphi)(\vec{a})$ if and only if $f(A) \models \varphi(f(\vec{a}))$ [24, Section 4.3]; thus, the pre-image of a definable set is definable.

▶ **Fact 5.3.** *An* FO*-interpretation is a logical map if and only if it is continuous.*

This provides us with a proof scheme to show that a space $\langle Y, \tau_2, \mathsf{FO}[\sigma_2]\rangle$ is a lpps: first, build a lpps $\langle X, \tau_1, \mathsf{FO}[\sigma_1]\rangle$, then build a FO-interpetation that is surjective and continuous from X to Y, and conclude that Y is a lpps. This is used for instance by [30, Corollary 10.7] to show that the class of all *p subdivisions* of finite graphs enjoys homomorphism preservation (using a slightly more general notion of FO-interpretations).

5.2 Relativisation

Preservation theorems do not relativise in general, but the stronger notion of being pre-spectral shows that non-trivial sufficient conditions for relativisation exists. However, unlike the theory of spectral spaces, there is not yet a full characterisation of the pre-spectral subsets of a pre-spectral space; see the full paper for a discussion.

▶ **Proposition 5.4** (Sufficient condition for relativisation)**.** *Let $\langle X, \tau, \mathsf{FO}[\sigma]\rangle$ be a lpps, Y be a Boolean combination of compact-open subsets of X, and θ be the topology induced by τ on Y. Then $\langle Y, \theta, \mathsf{FO}[\sigma]\rangle$ is a lpps.*

Proof. It suffices to prove that any definable open set U of Y is the restriction to Y of some definable open set of X. This stronger hypothesis is stable under finite unions and finite intersections, thus we only need to deal with the cases where Y is a definable open of X or the complement of one.

Let us first consider the case where Y is a definable open set of X. Then $U = U \cap Y$ is the restriction to Y of an open definable set of X. Let us next consider the case where Y is a definable closed set of X. Remark that $V \triangleq U \cup (X \setminus Y)$ is an open set of X, and is still definable. Therefore $U = V \cap Y$ with V a definable open set of X. ◀

5.3 Disjoint Unions and Products

Rather than using an already existing pre-spectral space and considering sub-spaces to build new smaller ones, it can be a rather efficient method to combine existing spaces to build bigger spaces. However, to build preservation theorems out of these constructions, it is necessary to represent those them as spaces of structures over some relational signature, which will be the role of definitions 5.5 and 5.7.

▶ **Definition 5.5** (Logical sum). *Let* $(\langle X_i, \tau_i, \mathsf{FO}[\sigma_i]\rangle)_{i\in I}$ *be a family of spaces. The* logical sum $\langle X, \tau, \mathsf{FO}[\sigma]\rangle$ *is defined as follows:*

1. *The signature* σ *is the disjoint union of the signatures* $(\sigma_i)_{i\in I}$*.*
2. *The set* X *is the union (disjoint by construction)* $\bigcup_{i\in I} f_i(X_i)$ *where, for all* $i \in I$, $f_i : X_i \to \mathrm{Struct}(\sigma)$ *is defined by* $|f_i(A)| \triangleq |A|$ *and* $(a_1, \ldots, a_n) \in \mathbf{R}^{f_i(A)}$ *if and only if* $R \in \sigma_i$ *and* $(a_1, \ldots, a_n) \in \mathbf{R}^A$*.*
3. *The topology* τ *is generated by the sets* $f_i(U)$ *where* $U \in \tau_i$ *and* $i \in I$*.*

The logical sum space is a simple translation of the topological sum space, which leads to the following result (see the full paper for a proof).

▶ **Proposition 5.6** (Stability under finite logical sum). *Let* $(\langle X_i, \tau_i, \mathsf{FO}[\sigma_i]\rangle)_{i\in I}$ *be a finite family of lpps. The logical sum of those spaces is a lpps homeomorphic to the sum of those spaces in* **PreSpec**.

In the case of products, a sentence over a product is not simply obtained by projecting on each component. This is handled in our proof of Proposition 5.8 in the full paper by reducing the first-order theory of the product to the first-order theories of its components thanks to Feferman-Vaught decompositions [16, 29].

▶ **Definition 5.7** (Logical product). *Let* $(\langle X_i, \tau_i, \mathsf{FO}[\sigma_i]\rangle)_{i\in I}$ *be a family of spaces. The* logical product $\langle X, \tau, \mathsf{FO}[\sigma]\rangle$ *is defined as follows:*

1. *The signature* σ *is the disjoint union of the signatures* $(\sigma_i)_{i\in I}$ *with additional unary predicates* ε_i *for each* $i \in I$*.*
2. *The set* X *is the image of* $\prod_{i\in I} X_i$ *through the map* $f \colon \prod_{i\in I} X_i \to \mathrm{Struct}(\sigma)$ *that associates to each* $(A_i)_{i\in I}$ *the disjoint union of the structures* A_i *with* ε_i *true on the structure* A_i *for* $i \in I$*.*
3. *The topology* τ *generated by the sets* U *such that* $f^{-1}(U)$ *is an open set of* $\prod_{i\in I}(X_i, \tau_i)$*.*

▶ **Proposition 5.8** (Stability under finite logical product). *Let* $(\langle X_i, \tau_i, \mathsf{FO}[\sigma_i]\rangle)_{i\in I}$ *be a finite family of lpps. The logical product of those spaces is a lpps homeomorphic to the product of the spaces* X_i *in* **PreSpec**.

6 Logical Closure

Consider a set Z equipped with a bounded sublattice $\mathcal{L}$ of $\wp(Z)$. In this section, we provide a way to consider the closure of a space $X \subseteq Z$ in a suitable topology so that if X is a lpps, then its closure also is. Let us write $\tau_{\mathcal{L}} \triangleq \langle \mathcal{L} \cup \{U^c \mid U \in \mathcal{L}\}\rangle$ for the topology generated by the sets of $\mathcal{L}$ and their complements. We call the closure $\overline{X}$ of X in $(Z, \tau_{\mathcal{L}})$ its *logical closure*.

We show in the full paper that lpps are stable under logical closures. For $X \subseteq Z$ and a sublattice $\mathcal{L}$ of $\wp(Z)$, we write $\mathcal{L}_X \triangleq \{U \cap X \mid U \in \mathcal{L}\}$ for the lattice induced by X.

▶ **Proposition 6.1** (Stability under logical closure). *Let* $X \subseteq Y \subseteq \overline{X}$ *and* τ *be a topology on* Y*. If* $\langle X, \tau_X, \mathcal{L}_X\rangle$ *is a lpps for the topology* τ_X *induced by* τ *on* X*, then so is* $\langle Y, \tau, \mathcal{L}_Y\rangle$*. If* $\mathcal{L}'$ *is a sublattice of* $\mathcal{L}$ *and* ${\mathcal{L}'}_X$ *is a diagram base of* X *then* ${\mathcal{L}'}_Y$ *is a diagram base of* Y*.*

Applications of logical closures. We now show that Proposition 6.1 allows to restate known preservation theorems and derive new ones. We consider the case where $Z = \mathrm{Struct}(\sigma)$ and $\mathcal{L} = [\![\mathsf{FO}[\sigma]]\!]_{\mathrm{Struct}(\sigma)}$, and we write τ_{FO} for the topology $\tau_{\mathcal{L}}$.

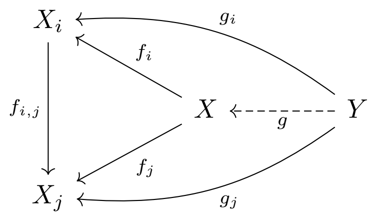

Figure 1 The commutative diagram of a projective system.

Let us define $\mathrm{FMP}(\sigma) \subseteq \mathrm{Struct}(\sigma)$ as the set of structures whose first-order theory satisfies the *finite model property*: any definable subset of $\mathrm{FMP}(\sigma)$ has a finite model. We prove that homomorphism preservation can be lifted from $\mathrm{Fin}(\sigma)$ (where it holds by Rossman's Theorem) to $\mathrm{FMP}(\sigma)$ in Corollary 6.2. To our knowledge this is a new result. This follows from Proposition 6.1 and the fact that $\mathrm{FMP}(\sigma)$ is the closure of $\mathrm{Fin}(\sigma)$ in the topology τ_{FO} (see the full paper for the proof).

▶ **Corollary 6.2** (Homomorphism preservation for structures with the finite model property)**.** $\mathrm{FMP}(\sigma)$ *has the* $(\tau_{\rightarrow}, \mathsf{EPFO}[\sigma])$ *preservation property.*

Let $\mathrm{Fin}^{\uplus}(\sigma)$ be the set of countable disjoint unions of finite structures over a finite relational signature σ. We state in Corollary 6.3 another consequence of Rossman's Theorem and Proposition 6.1, using the fact $\mathrm{Fin}(\sigma) \subsetneq \mathrm{Fin}^{\uplus}(\sigma) \subsetneq \mathrm{FMP}(\sigma) = \overline{\mathrm{Fin}(\sigma)}$; the same result was first shown by Nešetřil and Ossona de Mendez in [30, Theorem 10.6].

▶ **Corollary 6.3** (Homomorphism preservation for countable unions of finite structures)**.** $\mathrm{Fin}^{\uplus}(\sigma)$ *has the* $(\tau_{\rightarrow}, \mathsf{EPFO}[\sigma])$ *preservation property.*

7 Limits of Projective Systems

A natural construction in the category of topological spaces is the *projective limit*, and the category **Spec** of spectral spaces and spectral maps is closed under this construction [13, Corollary 2.3.8]. As an illustration, we show in Section 7.2 that $\langle \mathrm{Fin}(\sigma), \tau_{\rightarrow}, \mathsf{FO}[\sigma] \rangle$ is the projective limit of a system of Noetherian spaces, which provides an alternative understanding of Rossman's Theorem [32]. In fact, as we show in Section 7.3, any pre-spectral space is the limit of a projective system of Noetherian spaces.

7.1 Projective Systems

A *projective system* $\mathcal{F}$ in a category $\mathbf{C}$ assigns to each element i of a directed partially ordered set I an object X_i and to each ordered pair $i \leq j$ a so-called *bonding map* $f_{i,j} \colon X_i \to X_j$ so that, for all $i, j, k \in I$ with $k \leq j \leq i$, we have $f_{i,i} = \mathrm{id}_{X_i}$ and $f_{j,k} \circ f_{i,j} = f_{i,k}$. The *projective limit* of a projective system $\mathcal{F}$ is an object X with maps $f_i \colon X \to X_i$ *compatible* with the system $\mathcal{F}$, which means that, for all $i \geq j$, $f_{i,j} \circ f_i = f_j$. Moreover, X satisfies a universal property: whenever $\{g_i \colon Y \to X_i\}_{i \in I}$ is a family of maps compatible with $\mathcal{F}$, there exists a unique map $g \colon Y \to X$ such that $g_i = f_i \circ g$ for all $i \in I$.

Unfortunately, there exists projective systems in **PreSpec** that do not have limits, as can be witnessed by a slight adaptation of [35, Example 3]. Let us introduce here the category of topological spaces and continuous maps, denoted by **Top**. A projective system in **PreSpec** is a projective system of topological spaces in **Top**. A projective system in **PreSpec** always has a limit when considered as a projective system in **Top**; we give a sufficient condition for this space to be the limit in **PreSpec** (see the full paper for the proof).

▶ **Lemma 7.1** (Transfer of projective limits). *Let $\mathcal{F}$ be a projective system of pre-spectral spaces in* **PreSpec**. *If $\{f_i \colon X \to X_i\}_{i\in I}$ is the limit of $\mathcal{F}$ in* **Top** *where the maps $f_i \colon X \to X_i$ are spectral, then it is the limit of $\mathcal{F}$ in* **PreSpec**. *Moreover, $\mathcal{K}^\circ(X) = \bigcup_{i\in I} \left\{f_i^{-1}(V) \mid V \in \mathcal{K}^\circ(X_i)\right\}$.*

7.2 Application to the Homomorphism Preservation Theorem

Throughout this section, we fix a finite relational signature σ and a downwards-closed subset X of $\mathrm{Fin}(\sigma)$ for the homomorphism ordering $\to$, i.e., X is co-homomorphism closed. We will see how Rossman's Theorem can be explained as the existence of a projective limit.

n-Homomorphisms. Let us define the *tree-depth* $\mathrm{td}(A)$ of a finite structure A as the tree-depth $\mathrm{td}(\mathcal{G}(A))$ of its associated *Gaifman graph* $\mathcal{G}(A)$ [27, Definition 4.1]. Following the idea of the original proof in [32, Section 3.2], we are going to use quasi-orders that are *coarser* than the homomorphism quasi-order, and refine those progressively. For every $n \in \mathbb{N}$, we define $A \to_n B$ if for every structure C of tree-depth at most n, $C \to A$ implies $C \to B$. Note that on finite structures, $A \to B$ if and only if $A \to_{\mathrm{td}(A)} B$. Then the intersection of all the $\to_n$ relations is $\to$. Let us consider the corresponding Alexandroff topologies: $X \triangleq \langle X, \tau_\to, \mathsf{FO}[\sigma]\rangle$ and for $n \in \mathbb{N}$, let $X_n \triangleq \langle X, \tau_{\to_n}, \mathsf{FO}[\sigma]\rangle$.

Rossman's Lemma. In his paper [32], Rossman provides a function $\rho\colon \mathbb{N} \to \mathbb{N}$ and relates indistinguishability in the fragment $\mathsf{FO}_n[\sigma]$ of first-order logic with quantifier rank at most n to $\rho(n)$-homomorphism equivalence [32, Corollary 5.14]. We state this result in a self-contained manner below (see also [30, Theorem 10.5]).

▶ **Lemma 7.2** (Rossman's Lemma [32]). *There exists $\rho\colon \mathbb{N} \to \mathbb{N}$ such that, for all $n \in \mathbb{N}$, if $\varphi \in \mathsf{FO}_n[\sigma]$ is closed under homomorphisms, then it is closed under $\rho(n)$-homomorphisms.*

Rossman's Lemma is the combinatorial heart of Rossman's Theorem, so the developments in this section are only meant to show how the pre-spectral framework can capture his arguments translating the technical statement from Lemma 7.2 into a proof of homomorphism preservation in the finite.

Projective System. We are now ready obtain X as a limit of a projective system in **PreSpec**. We are going to exploit Lemma 7.2 through the definition of the topological spaces $Y_n \triangleq \langle X, \tau_\to, \mathsf{FO}_n[\sigma]\rangle$ for all n. We will use the following consequence of Rossman's Lemma (see the full paper for a proof).

▷ Claim 7.3. $\forall n \geq 1, \mathcal{K}^\circ(Y_n) \subseteq \mathcal{K}^\circ\big(X_{\rho(n)}\big) \subseteq \mathcal{K}^\circ(X)$.

The following theorem was famously first shown by Rossman in [32, Corollary 7.1]. A more recent proof in [33] uses lower bounds from circuit complexity. Similar results were shown in [30, Section 10.7] when assuming essentially the same statement as Lemma 7.2; in fact, carefully unwrapping the hypotheses of the *topological preservation theorem* of [30, Theorem 10.3] leads to the very definition of a projective system.

▶ **Theorem 7.4.** *Let σ be a finite relational signature and X be a non-empty downwards-closed subset of $\mathrm{Fin}(\sigma)$ for $\to$. Then X has the $(\tau_\to, \mathsf{EPFO}[\sigma])$ preservation property.*

Proof. Consider the projective system $\mathcal{F} \triangleq \{\mathrm{id}_{i,j} : Y_i \to Y_j\}_{i \leq j \in I}$ indexed by $I \triangleq \mathbb{N} \setminus \{0\}$. Each space Y_i is Noetherian for all $i \in I$ because $\mathsf{FO}_i[\sigma]$ contains finitely many non-equivalent sentences, hence Y_i contains finitely many open sets. Hence $\mathcal{K}^\circ(Y_i) = \tau_\to \cap [\![\mathsf{FO}_i[\sigma]]\!]_X$. Also, the maps $\mathrm{id}_{i,j}$ are spectral and $\mathcal{F}$ is a projective system in **PreSpec**. Claim 7.3 shows that the identity map $\mathrm{id}_i : X \to Y_i$ is a spectral map for all $i \in I$.

Assume that $\{g_i : Z \to Y_i\}_{i \in I}$ is a collection of morphisms in **Top** such that $\forall i \geq j \in I, g_j = \mathrm{id}_{i,j} \circ g_i$. Since $\mathrm{id}_{i,j}$ is the identity map, all the maps $(g_i)_{i \in I}$ are equal. In particular, one can build $g : Z \to X$ defined by any one of them. Let us show that g is a continuous map. If U is a definable open set of X, then U is a definable open set in Y_n for some n, hence $g^{-1}(U) = g_n^{-1}(U)$ is open. Since X has a base of definable open sets, this proves that g is continuous.

Assume that g' is an other continuous map making the diagram commute. As I is non empty, consider some $i \in I$, we have $g_i = \mathrm{id}_i \circ g = \mathrm{id}_i \circ g'$. Since f_i is the identity map we conclude $g = g'$.

We have shown that X is the limit of $\mathcal{F}$ in **Top**. Since the maps $\mathrm{id}_i : X \to X_i$ are spectral, Lemma 7.1 shows that X is a pre-spectral space such that $\mathcal{K}^\circ(X) = \bigcup_{i \in I} \mathcal{K}^\circ(Y_i) = \tau_\to \cap [\![\mathsf{FO}[\sigma]]\!]_X$. In particular, X is a lpps. As X is downwards-closed, by Remark 3.5 it has the $(\tau_\to, \mathsf{EPFO}[\sigma])$-preservation property. ◀

7.3 Completeness

We are now going to prove that any pre-spectral space can be obtained as a solution to a projective system of pre-spectral spaces, showing that the proof method of the previous sub-section is in some sense complete. In fact, this system is going to contain only Noetherian spaces (see the full paper). It is analoguous to the fact that any spectral space is a projective limit of finite T_0 spaces [23, Proposition 10].

▶ **Proposition 7.5** (Pre-spectral spaces are limits of Noetherian spaces)**.** *Let* (X, τ) *be a pre-spectral space, there exists a projective system of Noetherian spaces in* **PreSpec** *such that* X *is the limit of this projective system.*

8 Concluding Remarks

In this paper, we have introduced a general framework for preservation results, mixing topological and model-theoretic notions. The key notion here is the one of *logically presented pre-spectral spaces*, which requires the (topological) compactness of the definable sets of interest. This definition captures simultaneously the classical proofs of preservation theorems over the class of all structures (we detailed the case of the Łoś-Tarski Theorem in Section 2.2) and all the known preservation results over classes of finite structures in the literature (see Remark 3.5). Our approach is comparable to the one adopted in the *topological preservation theorem* of [30, Theorem 10.3], in that we employ topological concepts to present a generic preservation theorem; however we believe our formulation to be considerably simpler and more flexible.

We have developed a mathematical toolbox for working with logically presented pre-spectral spaces, allowing to build new spaces from known ones. Besides relatively mundane stability properties under suitable notions of morphisms, subspaces, finite sums, and finite products – which still required quite some care in order to account for first-order definability – , we have shown that more exotic constructions through topological closures or projective limits of topological spaces could also be employed. Those last two constructions give an alternative viewpoint on Rossman's proof of homomorphism preservation over the class of finite structures (Theorem 7.4), and a new homomorphism preservation result over the class of structures with the finite model property (Corollary 6.2).

References

1 Serge Abiteboul, Richard Hull, and Victor Vianu. *Foundations of Databases*. Addison-Wesley, 1995.

2 Miklós Ajtai and Yuri Gurevich. Monotone versus positive. *Journal of the ACM*, 34(4):1004–1015, 1987. doi:10.1145/31846.31852.

3 Miklós Ajtai and Yuri Gurevich. Datalog vs first-order logic. *Journal of Computer and System Sciences*, 49(3):562–588, 1994. doi:10.1016/S0022-0000(05)80071-6.

4 Albert Atserias, Anuj Dawar, and Martin Grohe. Preservation under extensions on well-behaved finite structures. *SIAM Journal on Computing*, 38(4):1364–1381, 2008. doi:10.1137/060658709.

5 Albert Atserias, Anuj Dawar, and Phokion G. Kolaitis. On preservation under homomorphisms and unions of conjunctive queries. *Journal of the ACM*, 53(2):208–237, 2006. doi:10.1145/1131342.1131344.

6 Chen Chung Chang and H. Jerome Keisler. *Model Theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics*. Elsevier, 1990.

7 Yijia Chen and Jörg Flum. Forbidden induced subgraphs and the Łoś-Tarski Theorem. Preprint, 2020. URL: https://arxiv.org/abs/2008.00420.

8 Bruno Courcelle. Monadic second-order definable graph transductions: a survey. *Theoretical Computer Science*, 126(1):53–75, 1994. doi:10.1016/0304-3975(94)90268-2.

9 Anuj Dawar. Homomorphism preservation on quasi-wide classes. *Journal of Computer and System Sciences*, 76(5):324–332, 2010. doi:10.1016/j.jcss.2009.10.005.

10 Anuj Dawar and Stephan Kreutzer. On Datalog vs. LFP. In *Proceedings of ICALP'08*, volume 5126 of *Lecture Notes in Computer Science*, pages 160–171, 2008. doi:10.1007/978-3-540-70583-3_14.

11 Anuj Dawar and Abhisekh Sankaran. Extension preservation in the finite and prefix classes of first order logic. Preprint, 2020. URL: https://arxiv.org/abs/2007.05459.

12 Alin Deutsch, Alan Nash, and Jeffrey B. Remmel. The chase revisited. In *Proceedings of PODS'08*, pages 149–158, 2008. doi:10.1145/1376916.1376938.

13 Max Dickmann, Niels Schwartz, and Marcus Tressl. *Spectral Spaces*, volume 35 of *New Mathematical Monographs*. Cambridge University Press, 2019.

14 Guoli Ding. Subgraphs and well-quasi-ordering. *Journal of Graph Theory*, 16:489–502, 1992. doi:10.1002/jgt.3190160509.

15 Tomás Feder and Moshe Y. Vardi. Homomorphism closed vs. existential positive. In *Proceedings of LICS'03*, pages 311–320, 2003. doi:10.1109/LICS.2003.1210071.

16 Solomon Feferman and Robert Vaught. The first order properties of products of algebraic systems. *Fundamenta Mathematicae*, 47(1):57–103, 1959. doi:10.4064/fm-47-1-57-103.

17 Diego Figueira and Leonid Libkin. Pattern logics and auxiliary relations. In *Proceedings of CSL-LICS'14*, pages 40:1–40:10, 2014. doi:10.1145/2603088.2603136.

18 Amélie Gheerbrant, Leonid Libkin, and Cristina Sirangelo. Naïve evaluation of queries over incomplete databases. *ACM Transactions on Database Systems*, 39(4):1–42, 2014. doi:10.1145/2691190.2691194.

19 Jean Goubault-Larrecq. *Non-Hausdorff Topology and Domain Theory*, volume 22 of *New Mathematical Monographs*. Cambridge University Press, 2013.

20 Martin Grohe. Existential least fixed-point logic and its relatives. *Journal of Logic and Computation*, 7(2):205–228, 1997. doi:10.1093/logcom/7.2.205.

21 Yuri Gurevich. Toward logic tailored for computational complexity. In *Computation and Proof Theory, Proceedings of LC'84*, volume 1104 of *Lecture Notes in Mathematics*, pages 175–216. Springer, 1984. doi:10.1007/BFb0099486.

22 Frederik Harwath, Lucas Heimberg, and Nicole Schweikardt. Preservation and decomposition theorems for bounded degree structures. In *Proceedings of CSL-LICS'14*, pages 49:1–49:10, 2014. doi:10.1145/2603088.2603130.

23 Melvin Hochster. Prime ideal structure in commutative rings. *Transactions of the American Mathematical Society*, 142:43–60, 1969. doi:10.1090/S0002-9947-1969-0251026-X.

24 Wilfrid Hodges. *A shorter model theory*. Cambridge University Press, 1997.

25 Phokion G. Kolaitis. Reflections on finite model theory. In *Proceedings of LICS'07*, pages 257–269, 2007. doi:10.1109/LICS.2007.39.

26 Joseph B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, 13(3):297–305, 1972. doi:10.1016/0097-3165(72)90063-5.

27 Leonid Libkin. *Elements of finite model theory*. Springer, 2012.

28 Jerzy Łoś. On the extending of models (I). *Fundamenta Mathematicae*, 42(1):38–54, 1955. doi:10.4064/fm-42-1-38-54.

29 Johann A. Makowsky. Algorithmic uses of the Feferman–Vaught Theorem. *Annals of Pure and Applied Logic*, 126(1–3):159–213, 2004. doi:10.1016/j.apal.2003.11.002.

30 Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*. Springer, 2012.

31 Eric Rosen. Some aspects of model theory and finite structures. *Bulletin of Symbolic Logic*, 8(3):380–403, 2002. doi:10.2178/bsl/1182353894.

32 Benjamin Rossman. Homomorphism preservation theorems. *Journal of the ACM*, 55(3):15:1–15:53, 2008. doi:10.1145/1379759.1379763.

33 Benjamin Rossman. An improved homomorphism preservation theorem from lower bounds in circuit complexity. *ACM SIGLOG News*, 3(4):33–46, 2016. doi:10.1145/3026744.3026746.

34 Alexei P. Stolboushkin. Finitely monotone properties. In *Proceedings of LICS'95*, pages 324–330, 1995. doi:10.1109/LICS.1995.523267.

35 Arthur H. Stone. Inverse limits of compact spaces. *General Topology and its Applications*, 10(2):203–211, 1979. doi:10.1016/0016-660X(79)90008-4.

36 William W. Tait. A counterexample to a conjecture of Scott and Suppes. *Journal of Symbolic Logic*, 24(1):15–16, 1959. doi:10.2307/2964569.

37 Alfred Tarski. Contributions to the theory of models. I. *Indagationes Mathematicae (Proceedings)*, 57:572–581, 1954. doi:10.1016/S1385-7258(54)50074-0.

38 Balder ten Cate and Phokion G. Kolaitis. Structural characterizations of schema-mapping languages. In *Proceedings of ICDT'09*, pages 63–72, 2009. doi:10.1145/1514894.1514903.