

ROOTING ALGEBRAIC VERTICES OF CONVERGENT SEQUENCES*

(EXTENDED ABSTRACT)

David Hartman^{†‡} Tomáš Hons[†] Jaroslav Nešetřil[†]

Abstract

Structural convergence is a framework for convergence of graphs by Nešetřil and Ossona de Mendez that unifies the dense (left) graph convergence and Benjamini-Schramm convergence. They posed a problem asking whether for a given sequence of graphs (G_n) converging to a limit L and a vertex r of L it is possible to find a sequence of vertices (r_n) such that L rooted at r is the limit of the graphs G_n rooted at r_n . A counterexample was found by Christofides and Král', but they showed that the statement holds for almost all vertices r of L . We offer another perspective to the original problem by considering the size of definable sets to which the root r belongs. We prove that if r is an algebraic vertex (i.e. belongs to a finite definable set), the sequence of roots (r_n) always exists.

DOI: <https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-075>

1 Introduction

The field of graph convergence studies asymptotic properties of large graphs. The goal is to define a well-behaved notion of a limit structure that describes the limit behavior of a convergent sequence of graphs. Several different approaches are studied. The two most prominent types of convergence are defined for sequences of dense [2][7][6] and sparse graphs

*Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Synergy Grant DYNASNET, grant agreement No 810115).

[†]Computer Science Institute, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

[‡]Institute of Computer Science of the Czech Academy of Sciences, Prague, Czech Republic

[1][4]. The recently introduced notion of structural convergence by Nešetřil and Ossona de Mendez offers a generalizing framework for these cases using ideas from analysis, model theory and probability [8][9].

Structural convergence is a framework of convergence for general relational structures; however, we follow the usual approach that we restrict to the of language graphs and rooted graphs without loss of generality. Our arguments remain valid in the general case (e.g. as in [3]). The *Stone pairing* of a formula ϕ in the language of graphs and a finite graph G , denoted by $\langle \phi, G \rangle$, is the probability that ϕ is satisfied by a tuple of vertices of G selected uniformly at random (for ϕ sentence, we set $\langle \phi, G \rangle = 1$ if $G \models \phi$, and $\langle \phi, G \rangle = 0$ otherwise). A sequence of finite graphs (G_n) is said to be *FO-convergent* if the sequence $(\langle \phi, G_n \rangle)$ converges for each ϕ . The limit structure L , called *modeling*, is a graph with measure ν on a standard Borel space satisfying that all the first-order definable are sets measurable. The value $\langle \phi, L \rangle$ is defined as the measure of the set $\phi(L)$, the set of solutions of ϕ in L , using the appropriate power of the measure ν . A modeling L is a limit of an FO-convergent sequence (G_n) if $\lim \langle \phi, G_n \rangle = \langle \phi, L \rangle$ for each ϕ . A modeling limit does not exist for each convergent sequence. It is known to exist for all sequences of graphs from a class \mathcal{C} if and only if \mathcal{C} is a nowhere dense class [10].

The authors of this framework asked in [8] the following question: given a sequence (G_n) converging to a modeling L and a vertex of r of L , is there a sequence of vertices (r_n) such that the graphs G_n rooted at r_n converge to L rooted at r ? Christofides and Král' provided an example that the answer is negative in general. However, they also proved that it is possible to find such a sequence (r_n) for almost all choices of the vertex r . That is, if the root of L is chosen at random (according to the measure ν), the vertices (r_n) exist with probability 1 [3].

In this paper, we refine the original problem by considering the root r to be an *algebraic* vertex of L . That is, r belongs to a finite definable set of L . We prove that the sequence of roots (r_n) always exists under such condition. Our main result reads as follows:

Theorem 1. *Let (G_n) be an FO-convergent sequence of graphs with a modeling limit L and r be an algebraic vertex of L . Then there is a sequence (r_n) , $r_n \in V(G_n)$, such that (G_n, r_n) FO-converges to (L, r) .*

Note that Theorem 1 deals with full FO-convergence and not just convergence with respect to sentences (called *elementary convergence*), for which it is a trivial statement (see the case of $p = 0$ in Lemma 3).

2 Notation

All graphs are finite except modelings, which are of size continuum. The vertex set of a graph G is denoted by $V(G)$. We use $\mathbb{N} = \{1, 2, \dots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $[n] = \{1, 2, \dots, n\}$, $[n]_0 = [n] \cup \{0\}$. The set of formulas in p free variables in the language of graphs is denoted by FO_p and $\text{FO} = \bigcup_{p \in \mathbb{N}_0} \text{FO}_p$ is the set of all formulas. Tuples of vertices, free variables, etc. are denoted by boldface letters, e.g. $\mathbf{x} = (x_1, \dots, x_p)$. Multiset

is a set that allows multiplicities of its elements. The power set of a set X is denoted by 2^X .

Let G be an arbitrary graph and r one of its vertices. By (G, r) we denote the graph G rooted at r . Formally, considering G as a structure in the language of graphs, we add a new constant “Root” to the vocabulary and interpret it as r . We refer to the extended language as the language of rooted graphs. The set of formulas in the extended language is denoted by FO^+ . Note that $\text{FO}_p \subseteq \text{FO}_p^+$.

Let L be a modeling. A formula $\phi \in \text{FO}_p$ is algebraic in L if $\phi(L)$ is finite, where $\phi(L) = \{\mathbf{v} \in V(L)^p : L \models \phi(\mathbf{v})\}$ is the set of solutions of ϕ in L . A vertex of L is algebraic if it satisfies an algebraic formula.

3 Rooting in algebraic sets

We prove the following statement, which is equivalent to Theorem 1.

Theorem 2. *Let (G_n) be an FO-convergent sequence of graphs with a modeling limit L and $\xi(x)$ be an algebraic formula in L . Then there is a sequence (r_n) , $r_n \in V(G_n)$, and a vertex $r \in \xi(L)$ such that (G_n, r_n) FO-converges to (L, r) .*

Obviously, Theorem 2 is implied by Theorem 1. The converse follows from fact that ξ has only finitely many solutions in L and we can iteratively root them one by one until we reach r .

Fix (G_n) , L and ξ for the rest of the paper. Without loss of generality, assume that $|\xi(G_n)| = |\xi(L)|$ for each n and $\xi(L)$ is an inclusion-minimal definable set in L . We prove Theorem 2 in three steps. First, we consider a single formula ϕ in the language of rooted graphs and show that we can find the roots (r_n) and r such that $\lim \langle \phi, (G_n, r_n) \rangle = \langle \phi, (L, r) \rangle$. Then we consider an arbitrary finite collection of formulas ϕ_1, \dots, ϕ_k and construct a single formula ψ with the property that convergence of $\langle \psi, (G_n, r_n) \rangle$ to $\langle \psi, (L, r) \rangle$ implies convergence of each $\langle \phi_i, (G_n, r_n) \rangle$ to $\langle \phi_i, (L, r) \rangle$. Finally, a routine use of compactness extends the previous to all formulas, which proves the theorem.

3.1 Single formula

For a formula $\phi(\mathbf{x}) \in \text{FO}_p^+$, let $\phi^-(\mathbf{x}, y) \in \text{FO}_{p+1}$ be the formula created from ϕ by replacing each occurrence of the term “Root” by “ y ” (we assume that y does not appear in ϕ).

Lemma 3. *For a given $\phi \in \text{FO}_p^+$ there is a sequence (r_n) , $r_n \in \xi(G_n)$, and a vertex $r \in \xi(L)$ such that $\lim \langle \phi, (G_n, r_n) \rangle = \langle \phi, (L, r) \rangle$.*

Proof. If $p = 0$, then either the sentence $(\forall y)(\xi(y) \rightarrow \phi^-(y))$ or $(\forall y)(\xi(y) \rightarrow \neg \phi^-(y))$ is satisfied in L (using the assumption that $\xi(L)$ is an inclusion-minimal definable set); hence, it holds in each G_n from a certain index on. Therefore, an arbitrary choice of $r_n \in \xi(G_n)$ and $r \in \xi(L)$ meets the conclusion.

Let ν be the measure associated to the modeling L . Define $f_L : V(L)^p \rightarrow 2^{\xi(L)}$ to be the function that sends \mathbf{v} to the set $\{u \in \xi(L) : L \models \phi^-(\mathbf{v}, u)\}$. Consider the pushforward measure μ_L on $2^{\xi(L)}$ of the p -th power of ν by f_L . Viewing $2^{\xi(L)}$ as a lattice, we are mostly interested in the measure of the filter generated by atoms of $2^{\xi(L)}$. Let X^\uparrow denote the filters generated by $X \in 2^{\xi(L)}$. Observe that for $u \in \xi(L)$ we have $\mu_L(\{u\}^\uparrow) = \langle \phi, (L, u) \rangle$. Suppose that $|\xi(L)| = t$ and define an ordering $R_L = (u_1, u_2, \dots, u_t)$ such that $\mu_L(R_L) = (\mu_L(\{u_i\}^\uparrow))_{i \in [t]}$ satisfies $\mu_L(\{u_1\}^\uparrow) \geq \mu_L(\{u_2\}^\uparrow) \geq \dots \geq \mu_L(\{u_t\}^\uparrow)$. Define similarly for each n the function $f_n : V(G_n)^p \rightarrow 2^{\xi(G_n)}$, measure μ_n (as the pushforward of the uniform measure) and the vector R_n .

We claim that the sequence $(\mu_n(R_n)) \subset ([0, 1]^t, \|\cdot\|_\infty)$ converges to $\mu_L(R_L)$. Then an arbitrary choice of an index $i \in [t]$ yields the sequence (r_n) and vertex r as the i -th elements of the vectors R_n , resp. R_L .

The claim follows from the fact that the vectors $\mu_n(R_n)$ continuously depend on the values $\langle \psi_{k,\ell}, G_n \rangle$, where $\psi_{k,\ell}(\mathbf{x}_1, \dots, \mathbf{x}_k) \in \text{FO}_{k,p}$ is

$$(\exists y_1, \dots, y_\ell) \left(\bigwedge_{i=1}^l \xi(y_i) \wedge \bigwedge_{1 \leq i < j \leq \ell} y_i \neq y_j \wedge \bigwedge_{i=1}^k \bigwedge_{j=1}^\ell \phi^-(\mathbf{x}_i, y_j) \right)$$

for $\ell \in [m]_0, k \in \binom{[m]}{\ell}$ and that $\langle \psi_{k,\ell}, G_n \rangle \rightarrow \langle \psi_{k,\ell}, L \rangle$. This continuous dependency can be proved by inclusion-exclusion with a help of classical results from combinatorics and complex analysis: Girard-Newton formulas [11] and the continuous dependency of the roots of a polynomial on its coefficients [12]. □

3.2 Finite collection of formulas

In this part, we use Lemma 3 to prove an analogous statement for a finite collection of formulas.

Lemma 4. *For given formulas ϕ_1, \dots, ϕ_k there is a sequence $(r_n), r_n \in \xi(G_n)$, and a vertex $r \in \xi(L)$ such that $\lim \langle \phi_i, (G_n, r_n) \rangle = \langle \phi_i, (L, r) \rangle$ for each ϕ_i .*

Proof. Since for sentences any choice of (r_n) and r works, we assume that neither of ϕ_1, \dots, ϕ_k is a sentence.

Consider an inclusion-maximal set $I \subseteq [k]$ for which there is $v \in \xi(L)$ such that every $i \in I$ satisfies $\langle \phi_i, (L, v) \rangle > 0$, denote $|I|$ by k' . If $I = \emptyset$, we can choose (r_n) and r arbitrarily; hence, assume otherwise. For $i \in I$ set $A_i = \{\langle \phi_i, (L, u) \rangle : u \in \xi(L)\} \cap (0, 1]$. Take a vector $\mathbf{e} \in \mathbb{N}^{k'}$ of exponents with the property that for each distinct $\mathbf{a}, \mathbf{b} \in \prod_{i \in I} A_i$ we have $\prod_{i \in I} a_i^{e_i} \neq \prod_{i \in I} b_i^{e_i}$. Such a vector exists as each A_i is finite and contains only positive values. The set of *bad* choices of rational exponents that make the values for particular \mathbf{a}, \mathbf{b} coincide form a $(k' - 1)$ -dimensional hyperplane in $\mathbb{Q}^{k'}$. We can surely avoid finitely many of such hyperplanes (one for each choice of \mathbf{a} and \mathbf{b}) to find a *good* vector of positive rational exponents and scale them to integers.

Use Lemma 3 for the formula ψ of the form

$$\bigwedge_{i \in I} \bigwedge_{j=1}^{e_i} \phi_i(\mathbf{x}_{i,j}),$$

where all the tuples $\mathbf{x}_{i,j}$ are pairwise disjoint, to obtain roots (r_n) and r . In particular, we can take the vertex r such that $\langle \psi, (L, r) \rangle > 0$ (due to our choice of I).

We have $\lim \langle \phi_i, (G_n, r_n) \rangle = \langle \phi_i, (L, r) \rangle > 0$ for each $i \in I$ as

$$\langle \psi, (L, r) \rangle = \prod_{i \in I} \langle \phi_i, (L, r) \rangle^{e_i},$$

using our selection of exponents \mathbf{e} .

Also, it holds that $\lim \langle \phi_j, (G_n, r_n) \rangle = \langle \phi_j, (L, r) \rangle = 0$ for each $j \notin I$: for the formula $\chi = \bigwedge_{i \in I \cup \{j\}} \phi_i(\mathbf{x}_i)$, we have $\lim \langle \chi, (G_n, r_n) \rangle = \langle \chi, (L, r) \rangle = 0$ due to the maximality of I (this is for *any* choice of (r_n) and r). We have

$$\langle \chi, (G_n, r_n) \rangle = \prod_{i \in I \cup \{j\}} \langle \phi_i, (G_n, r_n) \rangle$$

and as for some $\varepsilon > 0$ there is n_0 such that $\langle \phi_i, (G_n, r_n) \rangle > \varepsilon$ for each $i \in I$ and $n \geq n_0$, the factor $\langle \phi_j, (G_n, r_n) \rangle$ must tend to 0. \square

We remark that the rationalization of the fact that the sequence $(\langle \phi_j, (G_n, r_n) \rangle)$ for $j \notin I$ even converge is the reason why we are proving Theorem 2 instead of Theorem 1 directly. We are using the fact that we can choose the set I (and the root r for the formula ψ) such that any rooting (r_n) makes the sequence $\langle \chi, (G_n, r_n) \rangle$ converge to 0.

4 Concluding remarks

An iterative use of Theorem 1 or 2 allows us to gain complete control over the algebraic elements as we can consider each of them separately.

We note that it is possible to root solutions of algebraic formulas with multiple free variables as the projection to each coordinate yields an algebraic set. Moreover, the natural modification of Theorem 2 remains valid for FO-convergent sequences (G_n) without a modeling limit. The proofs are analogous except that the set I in Lemma 4 is defined as an inclusion-maximal set for which there are roots (r_n) such that $\lim \langle \bigwedge_{i \in I} \phi_i(\mathbf{x}_i), (G_n, r_n) \rangle > 0$.

Besides the original problem in [8], our motivation was the study of structural convergence of sequences created via gadget construction, see [5]. Using the result of this paper, we conclude that FO-convergence is preserved if the gadgets replace only finitely many edges (under natural additional assumptions).

In the typical case, the modeling L is of size continuum and the set of algebraic vertices (which is at most countable) has measure 0. Hence, our results reveal only a negligible portion of vertices of L for which the roots (r_n) exist, which shows that there is still room for further research.

References

- [1] Itai Benjamini and Oded Schramm. “Recurrence of Distributional Limits of Finite Planar Graphs”. In: *Electronic Journal of Probability* 6.none (Jan. 2001). DOI: 10.1214/ejp.v6-96.
- [2] Christian Borgs et al. “Counting Graph Homomorphisms”. In: *Topics in Discrete Mathematics*. Springer Berlin Heidelberg, 2006, pp. 315–371. ISBN: 978-3-540-33700-3. DOI: https://doi.org/10.1007/3-540-33700-8_18.
- [3] Demetres Christofides and Daniel Král’. “First-Order Convergence and Roots”. In: *Combinatorics, Probability and Computing* 25.2 (Feb. 2015), pp. 213–221. DOI: 10.1017/s0963548315000048.
- [4] Gábor Elek. “Note on limits of finite graphs”. In: *Combinatorica* 27.4 (July 2007), pp. 503–507. DOI: 10.1007/s00493-007-2214-8.
- [5] David Hartman, Tomáš Hons, and Jaroslav Nešetřil. *Gadget construction and structural convergence*. 2022. DOI: 10.48550/ARXIV.2212.10985.
- [6] László Lovász. *Large networks and graph limits*. American Mathematical Society, 2012. ISBN: 9780821890851.
- [7] László Lovász and Balázs Szegedy. “Limits of dense graph sequences”. In: *Journal of Combinatorial Theory, Series B* 96.6 (Nov. 2006), pp. 933–957. DOI: 10.1016/j.jctb.2006.05.002.
- [8] Jaroslav Nešetřil and Patrice Ossona de Mendez. “A Unified Approach to Structural Limits and Limits of Graphs with Bounded Tree-Depth”. In: *Memoirs of the American Mathematical Society* 263.1272 (Jan. 2020). DOI: 10.1090/memo/1272.
- [9] Jaroslav Nešetřil and Patrice Ossona de Mendez. “Cluster analysis of local convergent sequences of structures”. In: *Random Structures & Algorithms* 51.4 (Apr. 2017), pp. 674–728. DOI: 10.1002/rsa.20719.
- [10] Jaroslav Nešetřil and Patrice Ossona de Mendez. “Existence of Modeling Limits for Sequences of Sparse Structures”. In: *The Journal of Symbolic Logic* 84.02 (Mar. 2019), pp. 452–472. DOI: 10.1017/jsl.2018.32.
- [11] Richard P. Stanley and Sergey Fomin. *Enumerative Combinatorics*. Cambridge University Press, Jan. 1999. DOI: 10.1017/cbo9780511609589.
- [12] Hassler Whitney. *Complex Analytic Varieties*. Addison Wesley Longman Publishing Group, 1972, p. 399. ISBN: 9780201086539.