Edge distribution and density in the characteristic sequence

M. E. Malliaris

Department of Mathematics
University of Chicago
5734 S. University Avenue
Chicago, IL 60637

Abstract

The characteristic sequence of hypergraphs $\langle P_n : n < \omega \rangle$ associated to a formula $\varphi(x; y)$, introduced in [5], is defined by $P_n(y_1, \ldots y_n) = (\exists x) \bigwedge_{i \leq n} \varphi(x; y_i)$. We continue the study of characteristic sequences, showing that graph-theoretic techniques, notably Szemerédi's celebrated regularity lemma, can be naturally applied to the study of model-theoretic complexity via the characteristic sequence. Specifically, we relate classification-theoretic properties of φ and of the P_n (considered as formulas) to density between components in Szemerédi-regular decompositions of graphs in the characteristic sequence. In addition, we use Szemerédi regularity to calibrate model-theoretic notions of independence by describing the depth of independence of a constellation of sets and showing that certain failures of depth imply Shelah's strong order property SOP_3; this sheds light on the interplay of independence and order in unstable theories.

Key words: Unstable theories, independence property, Szemerédi regularity

1. Introduction

The characteristic sequence $\langle P_n : n < \omega \rangle$ is a tool for studying the combinatorial complexity of a given formula φ, Definition 2.2 below. It follows from [4], [5] that the Keisler order [2] localizes to the study of φ-types and

Email address: mem@math.uchicago.edu (M. E. Malliaris)

specifically of characteristic sequences. However, this article will not focus on ultrapowers.

The analysis of [5] established that characteristic sequences are essentially trivial when the ambient theory T is NIP, Theorem 2.10 below. In this paper, we turn to the study of characteristic sequences in the presence of the independence property. The framework of characteristic sequences allows us to bring a deep collection of graph-theoretic structure theorems to bear on our investigations. Notably, the classic model-theoretic move of polarizing complex structure into rigid and random components (e.g. Shelah's isolation of the independence property and the strict order property in unstable theories) is accomplished here by the application of Szemerédi's Regularity Lemma, §4 Theorem B below. Because the Regularity Lemma describes a possible decomposition of *any* sufficiently large graph, it can be applied here to understand how arbitrarily large subsets of P_1 generically interrelate.

In Sections 3-5, we investigate how classical properties of T affect the density δ attained between arbitrarily large ϵ-regular subsets $A, B \subset P_1$ (after localization) in the sense of Szemerédi regularity, where the edge relation is given by P_2. The picture we obtain is as follows. When φ is stable, by Theorem 2.10, the density (after localization) is always 1. When φ is simple unstable, after localization, there will be an infinite number of missing edges but we can say something strong about their distribution: $(*)$ the density between arbitrarily large ϵ-regular pairs must tend towards 0 or 1 as the graphs grow (indeed, here simplicity is sufficient but not necessary). In the simple unstable case, a finer function counting the number of edges omitted over finite subgraphs of size n is meaningful, and we give a preliminary description of its possible values in Theorem 3.11. In Section 5, we use model theory to relate the property $(*)$ of having arbitrarily large ϵ-regular subsets of P_1 with edge density bounded away from 0 and 1 to the phenomenon of instability in the characteristic sequence, which is strictly more complex than failure of simplicity. In Section 6 we refine this phenomenon by defining and investigating the compatible and empty order properties. On the level of theories, the compatible order property characterizes the model-theoretic rigidity property SOP_3, which is known to imply maximality in the Keisler order by [7].

In the other direction, in Section 7 we use Szemerédi regularity to bring to light a subtle model-theoretic failure of randomness, by considering the "depth of independence" of a constellation of infinite sets. In the language of Definition 7.2, we show that theories which are I^{n+1}_n but not I^{n+1}_{n+1} for some $n > 2$, are SOP_3. This is a result about the fine structure of the classic SOP/IP distinction, illustrating the tradeoff between a weaker notion of strict order (SOP_3) and a stronger notion of independence (I^{n+1}_{n+1}) in unstable theories.

2. Preliminaries

The following conventions will be in place throughout the paper.

Convention 2.1. (Conventions)

1. *If a variable or a tuple is written x or a rather than $\overline{x}, \overline{a}$, this does not necessarily imply that $\ell(x), \ell(a) = 1$.*
2. *Unless otherwise stated, T is a complete theory in the language $\mathcal{L}$.*
3. *A set is k-*consistent *if every k-element subset is consistent, and it is k-*inconsistent *if every k-element subset is inconsistent.*
4. $\varphi_\ell(x; y_1, \ldots y_\ell) := \bigwedge_{i \leq l} \varphi(x; y_i)$
5. *$\mathcal{P}_{\aleph_0}(X)$ is the set of all finite subsets of X.*
6. *ϵ, δ are real numbers, with $0 < \epsilon < 1$ and $0 \leq \delta \leq 1$.*
7. *Let G be a symmetric binary graph. We present graphs model-theoretically, i.e. as sets of vertices on which certain edge relations hold. Throughout this paper $R(x, y)$ is a binary edge relation, which will sometimes (we will clearly say when) be interpreted as P_2.*
8. *A graph is a simple graph: no loops and no multiple edges. Definition 2.2 below implies that $\forall x(P_1(x) \rightarrow P_2(x, x))$, but we will, by convention, not count loops when taking P_2 as R.*
9. *Given a graph G, with symmetric binary edge relation $R(x, y)$:*
   - *$|G|$ is the size of G, i.e. the number of vertices.*
   - *$e(G)$ is the number of edges of G.*
   - *$\hat{e}(G)$ is the number of edges omitted in G.*
   - *An* empty graph *is a graph with no edges between distinct elements. In the case where the language contains more than one edge relation, write R_0-*empty graph *to mean that there are no R_0-edges between distinct elements.*
   - *A* complete graph *is a graph with all edges, i.e. in which $x, y \in G, x \neq y \implies R(x, y)$.*

- *We will use the word "subgraph" in the model-theoretic sense, corresponding to the graph-theoretic notion of induced subgraph. That is, X is a* subgraph *of G if X is a substructure of G in the graph language, i.e. for any vertices x, y in X and any graph edge relation R in the language, we require $R(x, y)$ in X iff $R(x, y)$ in G. This will occasionally require some translation, as for instance in Corollary 4.4.*
- *The* degree *of a vertex is the number of edges which contain it.*
- *The* complement *G' of a graph G is given by: for $x \neq y$, $G' \models R(x, y) \iff G \models \neg R(x, y)$.*
  (Further conventions are at the end of the next item.)

10. *Write (X, Y) to indicate a graph whose set of vertices has been partitioned into two disjoint sets X, Y. Call such a graph a* 2-partite graph. ***Whereas "bipartite" is often used to mean that each of the components X, Y is itself an empty graph, the term "2-partite" does not assume this to be the case.*** *Rather, we present a graph as a 2-partite graph to indicate that the edges under analysis are those between elements $x \in X$ and $y \in Y$. More precisely:*
    - *$e(X, Y)$ is the number of edges between elements $x \in X$ and $y \in Y$ (edges between elements $x, x' \in X$ or $y, y' \in Y$ are not counted).*
    - *$\hat{e}(X, Y)$ is the number of edges omitted between elements $x \in X$ and $y \in Y$.*
    - *The* density *of a finite 2-partite graph (X, Y) is $\delta(X, Y) := e(X, Y)/|X||Y|$ when $|X|, |Y| \neq 0$, and 0 otherwise.*
    - *An* empty pair *is a pair of vertices x, y with $\neg R(x, y)$.*
    - *An* infinite empty pair *is (X, Y) such that $|X| = |Y| \geq \aleph_0$ and for all $x \in X$, $y \in Y$, we have $\neg R(x, y)$.*
    - *A* complete 2-partite graph *is (X, Y) such that for all $x \in X$ and $y \in Y$, $R(x, y)$ holds.*
    - *When a graph is presented as a 2-partite graph (X, Y), we will suppose its complement $(X, Y)'$ only disagrees with (X, Y) on edges between $x \in X, y \in Y$. That is, $(X, Y)'$ contains an edge between $x \in X$ and $y \in Y$ iff the original graph (X, Y) does not, but (X, Y) and $(X, Y)'$ agree on edges between $x, x' \in X$ or $y, y' \in Y$.*

We will make extensive use of the classification-theoretic dividing lines of stability, simplicity, the independence property, and the strict order property; see, for instance, [6], Chapter II, sections 2-4 and [7]. A theory or a formula is NIP, also called dependent, if it does not have the independence property; see, for instance, [10].

We now turn to definitions. The characteristic sequence of hypergraphs was introduced in [5] as a tool for studying the complexity of a given formula φ. Let us set the stage by briefly reviewing some of the results obtained there.

Definition 2.2. (Characteristic sequences) *Let T be a first-order theory and φ a formula of the language of T.*

- *For $n < \omega$, $P_n(z_1, \ldots z_n) := \exists x \bigwedge_{i \leq n} \varphi(x; z_i)$.*
- *The* characteristic sequence *of φ in T is $\langle P_n : n < \omega \rangle$.*
- *Write $(T, \varphi) \mapsto \langle P_n \rangle$ for this association.*
- *We assume that $T \vdash \forall y \exists z \forall x (\varphi(x; z) \leftrightarrow \neg\varphi(x; y))$, i.e. by varying the parameters we can obtain any positive or negative instance of φ. If this does not already hold for some given φ, replace φ with*

$$\theta(x; y, z, w) = \begin{cases} \varphi(x; y) & \textit{if} \;\; z{=}w \\ \neg\varphi(x; y) & \textit{otherwise} \end{cases}$$

Convention 2.3. As the characteristic sequence is definable in T, its first-order properties depend only on the theory and not on the model of T chosen. Throughout this paper, we will be interested in whether certain, possibly infinite, configurations appear as subgraphs of the P_n. By this we will always mean *whether or not it is consistent with T* that such a configuration exists when P_n is interpreted in some sufficiently saturated model. Thus, without loss of generality the formulas P_n will often be identified with their interpretations in some monster model.

In the characteristic sequence, complete graphs and empty graphs have model-theoretic meaning.

Observation 2.4. *Fix T, φ and $M \models T$ and suppose $(T, \varphi) \mapsto \langle P_n \rangle$.*

1. *The following are equivalent, for a set $A \subset M$:*

(a) *A is a positive base set.*
(b) *The set $\{\varphi(x;a) : a \in A\}$ is a consistent partial type.*

2. *The following are equivalent, for a set $A \subset P_1$:*
   (a) *A is a P_n-empty graph for some n.*
   (b) *$\{\varphi(x;a) : a \in A\}$ is 1-consistent but n-inconsistent (Convention 2.1(3)).*

   Note that if A is infinite, compactness then implies some instance of φ divides.

Characteristic sequences give a natural context for studying the complexity of φ-types, which are complete P_∞-graphs by the previous observation. Let us fix some notation:

Definition 2.5. *Fix T, φ, $M \models T$ and $(T, \varphi) \mapsto \langle P_n \rangle$.*

1. *A* positive base set *is a set $A \subset P_1$ such that $A^n \subset P_n$ for all $n < \omega$.*
2. *The sequence $\langle P_n \rangle$ has* support k *if: $P_n(y_1, \dots y_n)$ iff P_k holds on every k-element subset of $\{y_1, \dots y_n\}$. The sequence has* finite support *if it has support k for some $k < \omega$. Note that support k implies support $k+1$. For our purposes here, it is usually not important to know whether k is minimal.*
3. *For $k \geq 2$, say that $A \subset P_1$ is a P_k-complete graph if $A^k \subset P_k$. If A is a P_k-complete graph for all $k \geq 2$, say that it is P_∞-complete.*
4. *The elements $a_1, \dots a_k \in P_1$ are a $k-$*point extension *of the P_r-complete graph A just in case $Aa_1, \dots a_k$ is also a P_r-complete graph, where $r \in \mathbb{N}^{\geq 2} \cup \{\infty\}$ is given. Unless otherwise specified, $r = \infty$.*

Remark 2.6. *The following are equivalent:*

1. *$\langle P_n \rangle$ has finite support.*
2. *φ does not have the finite cover property.*

Localization is a definable restriction of the predicates P_n of a certain useful form which eliminates some of the combinatorial noise around a positive base set A under analysis. Definability ensures that Convention 2.3 applies when asking whether certain configurations are present in some localization. By way of motivation, consider the following simple example.

Example 2.7. *Suppose $\mathcal{L}$ contains equality and a binary relation E, let T be the theory of an equivalence relation with infinitely many infinite classes,* *and let* $\varphi(x; y, z, w)$ *be "*xEy*" if* $z = w$ *and "*$\neg xEy$*" otherwise. Then the characteristic sequence of* φ *has support* 2*, because any* k *triples* $(y_i, z_i, w_i) \in P_1$ *will each assert the existence of an* x *which is or is not equivalent to* y_i*, and the ultimate consistency of these assertions depends on the consistency of every pair. If we consider the graphs* P_n *in some model* M*, the* P_∞*-complete subsets of* P_1 *correspond to consistent partial* φ*-types, either the type of an element in some given each equivalence class and or that of an element not in any of the equivalence classes. If* $a, b \in P_1$ *are in distinct maximal* P_∞*-complete graphs, they are inconsistent, i.e.* $\neg P_2(a, b)$ *(notice we are suppressing that these are tuples, i.e.* $a \in M^3$*).*

Now suppose we would like to analyze some partial φ*-type of the form* $\{\varphi(x; a) : a \in A\}$*. So* $A \subset M^3$*,* $A \subset P_1$ *and by definition* A *is a* P_∞*-complete graph. By stability, this is a definable type, which in our context corresponds to the following picture. Choose some* $a_0 \in A$ *and consider the restriction of* P_1 *given by* $X := P_1(y) \wedge P_2(y, a_0)$*. Now* $A \subset X$ *and moreover* X *is a* P_∞*-complete graph which, to belabor the point, is definable with parameters from* P_1 *in the graph language, i.e. by conjoining a positive instance of one of the formulas* P_n*. This motivates the slightly more general definition (2.8) of a* localization *of* P_1 *around the positive base set* A*.*

Once outside the stable case, types need not have definable extensions, and it may be too much to hope that some definable restriction (more precisely, a localization) of P_1 around a given positive base set A will itself be a complete graph. The analysis of [5] shows that the classification-theoretic complexity of φ is reflected by the graph-theoretic complexity of the finite graphs which "persist" in the vicinity of positive base sets in the characteristic sequence of φ, where a graph Y is said to be "persistent" around A if any localization containing A also contains a copy of Y. For a formal discussion of persistence, see [5], Section 4.2.

Definition 2.8. (Localization, Definition 5.1 of [5] in the case $n = 1$)
Fix a characteristic sequence $(T, \varphi) \to \langle P_n \rangle$ *and interpret the predicates* P_n *in some (monster) model* $M \models T$*. Let* $A \subset P_1$ *be a positive base set for* φ*, and let* $B \subset P_1$ *be a finite set of parameters, with* A, B *possibly empty. A* localization P_1^f of the predicate $P_1(y)$ around the positive base set A with parameters from B *is a definable subset of* P_1 *given by a function* $f : m \to \omega \times \mathcal{P}_{\aleph_0}(B)$ *where* $m < \omega$ *and:*

- *writing $f(i) = (r_i, \beta_i)$, where $\beta_i = b^i_1, \dots b^i_{r_i}$, we have:*

$$P^f_1(y) := \bigwedge_{i \leq m} P_{r_i+1}(y, b^i_1, \dots b^i_{r_i})$$

- *In any model of T containing A and B, $A \subseteq P^f_1$.*
- *For each $\ell < \omega$, there exists a P_ℓ-complete graph $C_\ell \subseteq P^f_1$. (If A is infinite, this is automatically satisfied. If not, this condition ensures that although we have restricted the parameter set of φ, the restriction still contains infinite consistent partial types.)*

When analyzing a given formula φ, we will often write "after localization, [X holds]" to mean "for any positive base set A in the parameter space of φ, there exists a localization of P_1 which contains A in which [X holds]".

A brief digression on the interest of localization may be in order. Many classical dividing lines in classification theory have the form: either there is good behavior everywhere, or *there exists* an indicator of complexity, e.g. an instance of the order property or of the independence property. But how are these indicators of complexity distributed, say, in the vicinity of a type under analysis? How many φ-ordered sequences (say) might there be and how do these configurations interact with each other and with the rest of the parameter space of φ? Which configurations will occur in *any* localization around a given positive base set, and which can be avoided by a judicious restriction of the parameter set P_1? Localization arguments thus reveal dividing lines of a different sort: to be on the "wild" side of a line seen by localization means that the indicators of complexity are everywhere in the vicinity of *some* positive base set, because they cannot be avoided. When localization arguments recognize known classification-theoretic dividing lines, the alignment of the classical and the new characterizations is of interest. Let us mention several such results.

The first is that many instances of the order property in the characteristic sequence are not essential, i.e. they disappear after localization, unless φ is quite complex: having the tree property is necessary but not sufficient. If no partition of $\{y_1, \dots y_k\}$ into object and parameter variables has been specified, to say that a symmetric formula $R(y_1, \dots y_k)$ does not have the order property means here that *none* of the formulas $R(y_1; y_2, \dots y_k)$, $R(y_1, y_2; y_3, \dots y_k)$, $\dots$, $R(y_1, \dots y_{k-1}; y_k)$ have the order property.

Conclusion 2.9. (Conclusion 5.10 of [5]) *Suppose T is simple, $(T, \varphi) \mapsto \langle P_n \rangle$. Then for any $n < \omega$ and any positive base set A, there is a localization around A in which the formulas $P_2(y_1, y_2), \ldots P_n(y_1, \ldots y_n)$ do not have the order property.*

§5 below will illuminate this curious result.

The second result is that it is possible, for any positive base set A and any given $n < \omega$, to find a localization of P_1 around A which is a P_n-complete graph, precisely when φ does not have the independence property. [Recall Convention 2.3.]

Theorem 2.10. (rewording of Theorem 6.17 of [5]) *Let φ be a formula of T and $\langle P_n \rangle$ its characteristic sequence.*

1. *Suppose $X \subseteq P_1$ is a localization and that φ does not have the independence property on parameters in X. Then for each positive base set $A \subset X$ and each $n < \omega$, there is a further localization $Y \subset X$ such that $A \subset Y$ and $Y^n \subset P_n$, i.e. Y is a P_n-complete graph.*
2. *Suppose $X \subseteq P_1$ is a localization and that φ has the independence property on parameters in X. Then for all $n < \omega$, there are elements $z_1, \ldots z_n \in X$ such that $\neg P_n(z_1, \ldots z_n)$, i.e. X is not a P_n-complete graph for any n.*

Notice that given a formula $\varphi(x; y)$ with the independence property and a stable formula $\psi(x; y)$, the merged formula $\theta(x; y, z, w)$ which is $\varphi(x; y)$ if $z = w$ and $\psi(x; y)$ otherwise will, by construction, have a characteristic sequence which is not uniformly complex. For some positive base sets $A \subset P_1$, it may well be possible to find a localization containing A which is a complete P_n-graph, while Theorem 2.10(2) says that as long as θ has the independence property over parameters in a given localization, that localization cannot be a complete P_n-graph for any $n \geq 2$. So in the course of analyzing a type, which appears as a positive ‘base set A, we continually localize around A until one of two things happens: either most of the ambient complexity of φ drops away and A is revealed to be e.g. a stable type, or else we see that φ maintains its level of classification-theoretic complexity however we attempt to zoom in around A. Subsequent sections consider this second case.

One can check that *if* φ has the independence property then there will always be *some* positive base set around which missing edges are persistent; consider a complete P_∞-subgraph of the array described in Claim 3.8.

Convention 2.11. *Suppose* $(T, \varphi) \mapsto \langle P_n \rangle$ *and let* $X \subset P_1$ *be a localization. To say that* φ *is stable (resp. unstable) on* X *means that* φ *does not (resp. does) have the order property on parameters from* X. *Likewise, we say that* φ *is simple (or has the tree property) on* X *if it does not (does) have the tree property on parameters from* X, *and that* φ *has the independence property on* X *if it has the independence property on parameters from* X. *Note that by Convention 2.3, this means asking whether it is* consistent *for these configurations to occur over parameters from* X, *as* X *is a definable set.*

Simplicity can be characterized similarly. Recall that a formula is simple if it does not have the tree property.

Theorem 2.12. (rewording of Theorem 6.24 of [5]) *Let* φ *be a formula of* T *and* $\langle P_n \rangle$ *its characteristic sequence.*

1. *Suppose* $X \subseteq P_1$ *is a localization and that* φ *does not have the tree property on parameters in* X. *Then for each positive base set* $A \subset X$ *and each* $n < \omega$, *there is a further localization* $Y \subset X$ *and an integer* k *such that* $A \subset Y$ *and for all* $Z \subset Y$ *with* Z *a* P_n*-empty graph,* $|Z| < k$.
2. *Suppose* $X \subseteq P_1$ *is a localization and that* φ *has the tree property on parameters in* X. *Then for all* $n, k < \omega$ *there is* $Z \subset X$ *such that* $|Z| > k$ *and* Z *is a* P_n*-empty graph.*

This should not come as a surprise to those familiar with D-rank. Recall Observation 2.4(2). In the simple case, the localization corresponds to choosing a finite sequence of forking extensions of the partial type A so that no further n-dividing is possible.

3. Counting functions on simple φ

Throughout this section, we consider the binary edge relation P_2 from the characteristic sequence of φ. The notation and vocabulary follow Convention 2.1. If φ remains simple unstable on any localization around a given positive base set A (Convention 2.11), Theorems 2.10 and 2.12 give lower and upper bounds on the number of missing P_2-edges. So there is some complexity, but it is not yet of a manageable form. A key move in the study of unstable theories was Shelah's proof that the presence of complexity, i.e. the order property, meant the presence of either something uniformly random (the independence property) or uniformly rigid (the strict order property SOP). In our case, insight into global behavior of the many missing edges will come from Szemerédi's regularity lemma, Theorem B, after some preliminary observations.

Observation 3.1. *Suppose φ is stable. Then after localization, for any two disjoint finite $X, Y \subset P_1$, $\delta(X, Y) = 1$. On the other hand, if φ is simple unstable then P_1 contains elements y, z with $\neg P_2(y, z)$.*

Proof. Theorem 2.10(1) says that when φ is stable, after localization P_1 is a P_2-complete graph, so a fortiori there are no edges omitted between disjoint components. The second clause is Theorem 2.10(2). □

Definition 3.2. *Define $\alpha : \omega \to \omega$ by putting, for each $n \in \omega$,*

$$\alpha(n) := \max \quad \{\hat{e}(G) : G \subset P_1, |G| = n\}$$

i.e. the largest number of P_2-edges omitted over an n-size subset of P_1. When we are given some localization $X \subseteq P_1$, α is understood to be computed on $G \subset X$.

Claim 3.3. *Suppose φ is simple, i.e., φ does not have the tree property. Then after localization, for all sufficiently large n, $\alpha(n) < \frac{n(n-1)}{2}$.*

Proof. The maximum possible value $\frac{n(n-1)}{2}$ of any $\alpha(n)$ is attained on a P_2-empty graph, on which $x \neq y \implies \neg P_2(x, y)$. Apply Theorem 2.12 which says that when φ does not have the tree property then we have, after localization, a uniform finite bound k on the size of a P_2-empty graph $X \subset P_1$. So the function α is eventually strictly below the maximum. □

Corollary 3.4. *The function $\alpha(n)$ is meaningful, i.e. after localization for all sufficiently large n*

$$\frac{n(n-1)}{2} > \quad \alpha(n) \quad \geq 0$$

precisely when φ is simple, and moreover $\alpha(n) > 0$ precisely when φ is unstable on the given localization.

Since the goal is to bound the number of possible inconsistencies, we will be mainly interested in the nondegenerate case of a simple unstable formula which remains unstable in all localizations around some given positive base set (i.e. $\alpha(n) > 0$). So let us define:

Definition 3.5. *Suppose $(T, \varphi) \mapsto \langle P_n \rangle$. The formula φ is* eventually *simple unstable if for some positive base set $A \subset P_1$ there is a localization X with $A \subset X \subset P_1$ such that φ is simple on X but φ remains unstable on every localization Y with $A \subset Y \subset X$.*

Convention 3.6. *Throughout this section, "if φ is eventually simple unstable, then after localization, $\alpha(n)$ =..." is understood to mean "either $\alpha(n) = 0$, or ...". We will not explicitly consider the trivial case, but it may happen that localizing around* some *positive base set renders the formula stable.*

With some care we can restrict the range from Corollary 3.4 further. A famous theorem of Turán says:

Theorem A. (Turán, [3]:Theorem 2.2) *If G_n is a graph with n vertices and*

$$e(G) > \left(1 - \frac{1}{k-1}\right)\frac{n^2}{2}$$

then G_n contains a complete subgraph on k vertices.

Definition 3.7. *$X = \langle a_i^t : t < 2, i < \omega \rangle$ where $a_i^t \in P_1$ for all i, t is an* $(\omega, 2)$-array *if for all $n < \omega$,*

$$P_n(a_{i_1}^{t_1}, \ldots a_{i_n}^{t_n}) \iff (\forall j, \ell \leq n)\,(i_j = i_\ell \implies t_j = t_\ell)$$

Claim 3.8. (Claim 4.5 of [5]) *The following are equivalent, for a formula φ with characteristic sequence $\langle P_n \rangle$:*

1. *φ has the independence property.*
2. *$\langle P_n \rangle$ has an $(\omega, 2)$-array.*

Observation 3.9. *Suppose that $\langle P_n \rangle$ has an $(\omega, 2)$-array. Then $\alpha(n) \geq \left\lfloor \frac{n}{2} \right\rfloor$.*

Corollary 3.10. *When φ is eventually simple unstable, then after localization*

$$\left(1 - \frac{1}{k-1}\right)\frac{n^2}{2} \geq \alpha(n) \geq \left\lfloor \frac{n}{2} \right\rfloor$$

Proof. If φ is simple unstable on some localization X, φ has the independence property and so X contains an $(\omega, 2)$-array; thus the righthand side is Observation 3.9. For the lefthand side, let $k > 1$ be the uniform finite bound on the size of a P_2-empty graph from Theorem 2.12, and apply Turán's theorem to the complement of this graph. □

At the end of Section 4 we will give a proof of the following:

Theorem 3.11. *When φ is eventually simple unstable, then after localization, either*

$$\left(1 - \frac{1}{1-k}\right)\frac{n^2}{2} \;\geq\; \alpha(n) \;\geq\; \frac{n^2}{4} \qquad \text{or} \qquad \mathcal{O}(n^2) > \alpha(n) \geq \left\lfloor \frac{n}{2} \right\rfloor$$

where k is the integer given in the proof of Corollary 3.10.

The proof will follow from Theorem 4.13 below, which will show more, namely that for φ eventually simple unstable, either $\mathcal{O}(n^2) > \alpha(n)$ or there exists an infinite empty pair in P_1. In other words, if we cannot find two disjoint infinite sets of instances of φ such that no pair of instances from distinct sets is consistent, then the overall number of inconsistencies between instances of φ is relatively small.

Our strategy is going to be to show that in the absence of such an "infinite empty pair" we can repeatedly partition sufficiently large graphs into many pieces of roughly equal size in such a way that, at each stage, the bulk of the omitted edges must occur inside the (eventually, much smaller) pieces. The main tool will be Theorem B below.

4. Szemerédi regularity

We begin with a review of Szemerédi's celebrated *regularity lemma*. Recall that ϵ, δ are real numbers, $0 < \epsilon < 1$ and $0 \leq \delta \leq 1$, following Convention 2.1.

Definition 4.1. ([9], [3]) *The finite 2-partite graph (X, Y) is ϵ-regular if for every $X' \subset X$, $Y' \subset Y$ with $|X'| \geq \epsilon|X|$, $|Y'| \geq \epsilon|Y|$, we have: $|\delta(X,Y) - \delta(X', Y')| < \epsilon$.*

The regularity lemma says that sufficiently large graphs can always be partitioned into a fixed finite number of pieces X_i of approximately equal size so that almost all of the pairs (X_i, X_j) are ϵ-regular.

Theorem B. (Szemerédi's Regularity Lemma [3], [9]) *For every ϵ, m_0 there exist $N = N(\epsilon, m_0)$, $m = m(\epsilon, m_0)$ such that for any graph X, $N \leq |X| < \aleph_0$, for some $m_0 \leq k \leq m$ there exists a partition $X = X_1 \cup \cdots \cup X_k$ satisfying:*

- $||X_i| - |X_j|| \leq 1$ *for $i, j \leq k$*

- *All but at most ϵk^2 of the pairs (X_i, X_j) are ϵ-regular.*

Remark 4.2. *The original or "old" version of the regularity lemma was stated for 2-partite graphs: given a 2-partite graph X, Y, we may partition each of X, Y into at most m pieces of approximately equal size so that almost all of the pairs (X_i, Y_j) are ϵ-regular. This version will be useful in Section 5.*

One important consequence is that we may, approximately, describe large graphs G as random graphs where the edge probability between x_i and x_j is the density $d_{i,j}$ between components X_i, X_j in some Szemerédi-regular decomposition. We will need a definition.

Definition 4.3. [3] (The reduced graph)

1. *Let $G = X_1, \ldots X_k$ be a partition of the vertex set of G into disjoint components. Given parameters ϵ, δ, define the* reduced graph $R(G, \epsilon, \delta)$ *to be the graph with vertices x_i $(1 \leq i \leq k)$ and an edge between x_i, x_j just in case the pair (X_i, X_j) is ϵ-regular of density $\geq \delta$.*
2. *Let $R(t)$ be the graph with k components $X_1, \ldots X_k$, each with t vertices, such that $e(X_i) = 0$, and $\delta(X_i, X_j) = 1$ if there is an edge between x_i and x_j in R and 0 otherwise. So $R(t)$ is the "full" graph of height t with reduced graph R.*

The following lemma (called the "Key Lemma" in [3]) says that sufficiently small subgraphs of the reduced graph must actually occur in the original graph G. *Note that in the statement of the following theorem, "subgraph" is used in the graph-theoretic sense; see the discussion following, in particular Corollary 4.4.*

Theorem C. (Key Lemma, [3]:Theorem 2.1) *Given $\delta > \epsilon > 0$, a graph R, and a positive integer m, let G be any graph whose reduced graph is R, and let H be a subgraph of $R(t)$ with h vertices and maximum degree $\Delta > 0$. Set $d = \delta - \epsilon$ and $\epsilon_0 = d^\Delta/(2 + \Delta)$. Then if $\epsilon \leq \epsilon_0$ and $t - 1 \leq \epsilon_0 m$, then $H \subset G$. Moreover the number of copies of H in G is at least $(\epsilon_0 m)^h$.*

As noted above, the statement of the Key Lemma mentions two subgraphs: "$H \subset R[t]$" and "$H \subset G$", and in both cases graph-theoretic, i.e. not necessarily induced subgraph, is meant. For our purposes, it will be important to know that the second, "$H \subset G$", has the model-theoretic meaning, i.e. is an induced subgraph. We will also not need the full strength of the first, "$H \subset R[t]$," which amounts to "any graph on the given vertex set": rather, it will suffice to have the result for graphs H' defined on some subset of the vertices of $R[t]$ which satisfy: for all x_i^1, x_i^2 in the same component of $R[t]$ and x_j^3 in a different component, there is an edge between x_j^3 and x_i^1 iff there is an edge between x_j^3 and x_i^2. That is, edges are uniform between components. Call such H' *uniform subgraphs* of $R[t]$.

We will therefore use the following modification of the Key Lemma without further comment:

Corollary 4.4. (induced-subgraph Key Lemma) *In the statement of the Key Lemma, by replacing "$H \subset R[t]$ with "H a uniform subgraph of $R[t]$" and assuming the threshold density d is bounded away from* 0 *and* 1*, we may assume that in the penultimate sentence H is an* induced *subgraph of G. [We will not use the final sentence about number of copies.]*

Proof. Suppose first that for some fixed ϵ, δ that $X_1, \ldots X_k$ are equally sized components of a graph G and for $i \neq j$, each pair (X_i, X_j) is ϵ-regular with density δ. The reduced graph (for $d = \delta$) will be complete, so if G is large enough relative to ϵ, δ, any complete graph on no more than k vertices will occur as an induced subgraph of G. Moreover, for $d = 1 - \delta$ the reduced graph of the complement of G (where edges contained within components remain the same) is complete so if G is large enough relative to $\epsilon, 1 - \delta$, any empty graph on no more than k vertices will also occur as an induced subgraph of G.

More generally, given any graph C on k vertices $z_1, \ldots z_k$, construct a graph G_C with the same vertex set as G, satisfying: there is an edge between x, y in G_C iff

- (1) x, y are both in the same component X_i
- (2) $x \in X_i, y \in X_j$ for $i \neq j$ and there is an edge between z_i, z_j in C
- (3) $x \in X_i, y \in X_j$ for $i \neq j$, there is no edge between z_i, z_j in C and there is an edge between x, y in G

That is, G_C agrees with G except when there is no edge between z_i, z_j in C: if this happens, replace (X_i, X_j) with its complement. Let $d = \min(\delta, 1 - \delta)$. Then the reduced graph of G_C is complete, guaranteeing the existence of a complete graph on k vertices in G_C, which corresponds to an isomorphic copy of C on those same vertices in G. ☐

Note that it is only possible to control the existence or nonexistence of edges between *regular* components of density bounded away from 0 and 1. If the notation is familiar, a slightly cleaner statement of the case $t = 1$ is:

Theorem D. (from Gowers [1]) *For every $\alpha > 0$ and every k there exists $\epsilon > 0$ with the following property. Let $V_1, \dots V_k$ be sets of vertices in a graph G, and suppose that for each pair (i, j) the pair (V_i, V_j) is ϵ-regular with density δ_{ij}. Let H be a graph with vertex set $(x_1, \dots x_k)$ and let $v_i \in V_i$ be chosen uniformly at random, the choices being independent. Then the probability that for all i, j $v_i v_j$ is an edge of G iff $x_i x_j$ is an edge of H differs from $\Pi_{x_i x_j \in H} \delta_{ij} \Pi_{x_i x_j \notin H} (1 - \delta_{ij})$ by at most α.*

We now work towards a proof of Theorem 3.11.

Convention 4.5. (Interstitial edges, $b_{\epsilon,\ell}$, $N_{\epsilon,\ell}$)

1. *Let G be a graph and let $G = X_1 \cup \dots \cup X_n$ be a decomposition into disjoint components, for instance as given by Theorem B. Call any edge between vertices $x \in X_i, z \in X_j$, $i \neq j$ an* interstitial edge.
2. *Let $b_{\epsilon,\ell}$ denote the upper bound on the necessary number of components, given by the regularity lemma as a function of ϵ, ℓ (so the value of m in Theorem B).*
3. *Write* $(\epsilon, \ell)^*$-decomposition *to denote any Szemerédi-regular decomposition* into k components*, for any $\ell \leq k \leq b_{\epsilon,\ell}$.*
4. *Let $N_{\epsilon,\ell}$ denote the threshold size given by the regularity lemma as a function of ϵ, ℓ, such that any graph X with $|X| > N_{\epsilon,\ell}$ admits an $(\epsilon, \ell)^*$-decomposition.*

Remark 4.6. On Definition 4.5(2)-(4): As Corollary 4.8(3) suggests, for the purposes of our asymptotic argument it is usually sufficient to know that the number of components fluctuates in a certain fixed range, as given by the Regularity Lemma.

We now apply this analysis to the characteristic sequence of a given formula φ. By "subgraph" we mean model-theoretic, i.e. induced subgraph. The Key Lemma shows that if for arbitrarily small ϵ there are arbitrarily large ϵ-regular pairs whose density remains bounded away from 0 and 1, we may extract an empty pair:

Lemma 4.7. *Suppose that for some $\eta \in (0, \frac{1}{2})$, for all $\epsilon > 0$ and all $N \in \mathbb{N}$ there exist disjoint subsets $X_N, Y_N \subset P_1$, $|X_N| = |Y_N| \geq N$ such that (X_N, Y_N) is ϵ-regular with density $\delta \in (0 + \eta, 1 - \eta)$. Then P_1 contains an infinite empty pair.*

Proof. Apply the Key Lemma to each complement graph $(X_N, Y_N)'$, which is still regular and whose density remains bounded away from 0 and 1. For each $t < \omega$, for all N sufficiently large and ϵ sufficiently small relative to the given bound $1 - \eta$ and the given maximum degree t, the lemma ensures that $(X_N, Y_N)'$ contains a complete 2-partite graph with t vertices in each part. The bound ensures that we can freely choose ϵ and N. Note that the construction remains agnostic on whether edges hold between elements $x, x' \in X_N$ or $y, y' \in Y_N$. □

Lemma 4.8. *Suppose that P_1 does not contain an infinite empty pair.*

1. *There is a function $f : (0, 1) \times \omega \to (0, 1)$ which approaches 1 as $\epsilon \to 0$ and $N \to \infty$ and such that if (X, Y) is an ϵ-regular pair with $|X| = |Y| = N$ then $\delta(X, Y) \geq f(\epsilon, N)$.*
2. *There is a function $g : ((0, 1) \times \omega) \times \omega \to (0, 1)$, which is defined on all $((\epsilon, \ell), n)$ for which $n \geq N_{\epsilon,\ell}$, and which approaches 1 as (ϵ, ℓ) stays fixed and $n \to \infty$, such that if $|X| = n$ then the density between any two regular components in an $(\epsilon, \ell)^*$-decomposition of X is at least $g((\epsilon, \ell), n)$.*
3. *For every constant $c > 0$, and for all $\epsilon_0 > 0$, there exist $0 < \epsilon < \epsilon_0$ and for each such ϵ, cofinally many $\ell < \omega$ such that: for all n sufficiently large and all graphs X with $|X| = n$, the number of missing interstitial edges in any $(\epsilon, \ell)^*$-decomposition of X is strictly less than cn^2.*

Proof. (1) This restates Lemma 4.7: the density cannot remain bounded away from 0 and 1, and if the density approaches 0, extracting an empty pair becomes even easier. In other words, for any $d \in (0, 1)$, there must be some pair (N_d, ϵ_d) such that for all $n > N_d, \epsilon < \epsilon_d$ any ϵ-regular pair of size n will have density greater than d.

(2) The regularity lemma provides a decomposition in which all components are approximately the same size (± 1), so the density of each ϵ-regular pair will be at least $f(\epsilon, \frac{n}{b_{\epsilon,\ell}})$.

It remains to prove (3). For the moment, let ϵ, ℓ be arbitrary and suppose that $|X| > N_{\epsilon,\ell}$. Then $|X| = n$ admits an ϵ-regular decomposition into k-many pieces, each of size approximately $m = \frac{n}{k}$, where

$$(\dagger) \qquad \ell \leq k \leq \ell' := b_{\epsilon,\ell}$$

Writing $\delta := g((\epsilon,\ell), \frac{n}{\ell'})$, the contribution of the interstitial edges is at most:

$$\epsilon k^2 m^2 + (1-\epsilon)(k)^2 (1-\delta) m^2$$

where the term on the left assumes the irregular pairs are empty (all missing), and the term on the right counts the expected number of interstitial edges missing from the regular pairs. By ($\dagger$), this in turn is bounded by:

$$\begin{aligned} &\leq \epsilon(\ell')^2 m^2 + (1-\epsilon)(\ell')^2 (1-\delta) m^2 \\ &\leq \epsilon(\ell')^2 \left(\frac{n}{l}\right)^2 + (1-\epsilon)(\ell')^2 (1-\delta) \left(\frac{n}{l}\right)^2 \\ &\leq n^2 \left(\frac{\ell'}{\ell}\right)^2 \Big(\epsilon + (1-\epsilon)(1-\delta)\Big) \end{aligned}$$

Thus our claim will hold whenever $\epsilon + (1-\epsilon)(1-\delta) < c(\frac{\ell}{\ell'})^2$. To obtain this, choose $\epsilon > 0$ sufficiently small (say, less than half the right-hand side). Notice that for any given ℓ, $\frac{\ell}{\ell'}$ will be less than 1; the only other place ℓ appears is in $\delta = g((\epsilon,\ell), \frac{n}{\ell'})$. By (2), for any fixed (ϵ,ℓ), g approaches 1 as $n \to \infty$. So it suffices to choose n large for the $(1-\delta)$ term to be sufficiently small. □

Lemma 4.7(3) says that for any constant c, the number of missing interstitial edges eventually falls below cn^2. We can leverage this fact to show that there must be comparatively few missing edges of any kind.

Definition 4.9. *(Successive decompositions)*

1. *Let G be a finite graph and $1 \leq t < \omega$. Say that G admits an $(\epsilon,\ell)^*$-decomposition to depth t if:*
   (1) *There is an $(\epsilon,\ell)^*$-decomposition of G.*
   (2) *Each of the components from the decomposition at stage (1) admit an $(\epsilon,\ell)^*$-decomposition.*
   ⋮

(t) *Each of the components from a decomposition at stage (t-1) admit an $(\epsilon, \ell)^*$-decomposition.*

2. *The components obtained at stage t are called* terminal components. *The components obtained at all other stages are called non-terminal components.*
3. *Say that the $(\epsilon, \ell)^*$-decomposition to depth t* respects the constant c *if for each of the non-terminal components X, the number of missing interstitial edges in any $(\epsilon, \ell)^*$-decomposition of X is strictly less than $c|X|^2$.*

Remark 4.10. *Given c, ϵ, ℓ, n satisfying Lemma 4.7(3), choose N such that $\frac{N}{(\ell')^t} > n$, where $\ell' = b_{\epsilon,\ell}$. Then for any graph G with $|G| > N$, any $(\epsilon, \ell)^*$-decomposition of G to depth t respects the constant c.*

Lemma 4.11. *Fix a constant $c \in (0, 1)$ and suppose G admits an $(\epsilon, \ell)^*$-decomposition to depth t. Let $|G| = n$, and suppose all of the terminal components in this decomposition are empty graphs. Then the total number of omitted edges from all the terminal components is at most $\frac{n^2}{\ell^t}$.*

Proof. To avoid aggregious indexing, let us work from the bottom up. Suppose we are given a component X_{t-1} from stage $t-1$, that is, X_{t-1} admits an $(\epsilon, \ell)^*$-decomposition whose components are the terminal components. The cardinality of X_{t-1} will be given by $\frac{n}{k_1 \cdots k_{t-1}}$ for *some* sequence of integers with $\ell \leq k_i \leq \ell'$ for all $1 \leq i \leq t-1$. Suppose that the $(\epsilon, \ell)^*$-decomposition of X_{t-1} has k_t components. Then the number of missing edges contributed by terminal components in X_{t-1} is no more than:

$$\left(\frac{n}{k_1 \cdots k_{t-1} k_t}\right)^2 k_t \quad = \quad \frac{n^2}{{k_1}^2 \cdots {k_{t-1}}^2 k_t}$$

Now we step back a level. The component X_{t-1} was itself one of k_{t-1} members of an $(\epsilon, \ell)^*$-decomposition of some prior stage component X_{t-2}. Let us acknowledge this by renaming X_{t-1} as $X_{t-1,1}$ and k_t as $k_{t,1}$. That is, the components of the decomposition of X_{t-2} are $X_{t-1,1}, \ldots X_{t-1,k_{t-1}}$, and the terminal components contained in $X_{t-1,i}$ contribute at most $\frac{n^2}{{k_1}^2 \cdots {k_{t-1}}^2 k_{t,i}}$ missing edges to the total count. Now the edges missing from all terminal components in X_{t-2} is at most

$$\frac{n^2}{{k_1}^2 \cdots {k_{t-1}}^2 k_{t,1}} + \cdots + \frac{n^2}{{k_1}^2 \cdots {k_{t-1}}^2 k_{t,k_{t-1}}}$$

By assumption, each of the integers $k_{t,i}$ satisfy $\ell \leq k_{t,i} \leq \ell'$, so we may replace each of them by ℓ. This gives a further upper bound

$$\left(\frac{n^2}{{k_1}^2 \cdots {k_{t-1}}^2 \ell}\right) k_{t-1} \quad = \quad \frac{n^2}{{k_1}^2 \cdots {k_{t-2}}^2 k_{t-1} \ell}$$

Continuing, we find that if in a component X_{t-r} of depth $t-r$, the contribution of missing edges from terminal components contained in X_{t-r} is at most

$$\frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 k_{t-r+1} \ell^{r-1}}$$

then writing X_{t-r-1} for the enveloping component at the immediately prior stage, and once again renaming X_{t-r} as $X_{t-r,1}$ and k_{t-r+1} as $k_{t-r+1,1}$, a bound on missing edges from terminal components contained in X_{t-r-1} is given by

$$\begin{aligned}
\frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 k_{t-r+1,1} \ell^{r-1}} + \cdots + \frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 k_{t-r+1,k_{t-r}} \ell^{r-1}} \\
&\leq \frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 \ell \ell^{r-1}} + \cdots + \frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 \ell \ell^{r-1}} \\
&\leq \left(\frac{n^2}{{k_1}^2 \cdots {k_{t-r}}^2 \ell^r}\right) k_{t-r} \\
&\leq \frac{n^2}{{k_1}^2 \cdots {k_{t-r-1}}^2 k_{t-r} \ell^r}
\end{aligned}$$

When $r = t$, the component under consideration is the entire graph, and we obtain the bound $\frac{n}{\ell^t}$ as desired. □

Lemma 4.12. *Fix a constant $c \in (0,1)$ and suppose G admits an $(\epsilon, \ell)^*$-decomposition to depth t which respects the constant c. Let $|G| = n$ and $1 \leq m \leq t-1$. Then the total number of omitted edges which occur as interstitial edges at stage m of the decomposition of G is at most $c\frac{n^2}{\ell^{m-1}}$.*

Proof. Essentially the same proof as that of Lemma 4.11. The differences are first, that the length of the induction is shorter by one, and second that rather than taking as basic units the terminal components, we take as basic units the components at stage m of the decomposition, which adds a factor of c. More precisely, let X_m be any such component and suppose as usual that it is one of finitely many components of a prior decomposition of X_{m-1}. By

Lemma 4.7(3) and the hypothesis that the successive decompositions respect c, we know that there are no more than $c|X|^2$ interstitial edges missing in any $(\epsilon, \ell)^*$-decomposition of X_m. In other words, the number of interstitial edges omitted in decompositions of the stage m components contained in X_{m-1} is no more than

$$c\left(\frac{n}{k_1 \cdots k_{m-1}}\right)^2 k_{m-1}$$

Compare the first displayed equation of the previous lemma. By applying that proof, it is straightforward to inductively combine these "basic" counts, by replacing the appropriate family of partition numbers k_i with ℓ at each inductive step as previously described, to obtain a bound of $\frac{c}{\ell^{m-1}}$ on missing edges which occur as interstitial edges at stage m across the whole graph. □

We are now prepared to prove:

Theorem 4.13. *Suppose φ is simple on some given localization $X \subseteq P_1$. If there does not exist an infinite empty pair $Y, Z \subset X$, then on X, $\alpha(n) < \mathcal{O}(n^2)$.*

Proof. Given a positive real constant $c_0 > 0$, choose c, k, t such that $0 < c < 1$, $2 < k, t \in \mathbb{N}$ and $c_0 > 2c + \frac{1}{k^t}$. Fix a pair (ϵ, ℓ) such that $\ell > k$ and (ϵ, ℓ) is one of the cofinally many pairs described in Lemma 4.8(3) for the constant c. By Remark 4.10, we may assume that all sufficiently large graphs G admit an $(\epsilon, \ell)^*$-decomposition of to depth t which respects the constant c.

We now apply the two previous lemmas to bound the number of missing edges in G. Note that the point of the decomposition is that any edges must occur either as interstitial edges at some stage of the decomposition or else occur in some terminal component. Applying the bounds obtained in Lemma 4.11 and Lemma 4.12 gives that for all sufficiently large n:

$$\begin{aligned}
\alpha(n) &< cn^2\left(1 + \frac{1}{\ell} + \frac{1}{\ell^2} + \cdots + \frac{1}{\ell^{t-1}}\right) + \left(\frac{n^2}{\ell^t}\right) \\
&< n^2\left(\frac{\ell c}{\ell - 1} + \frac{1}{\ell^t}\right) \\
&< \left(2c + \frac{1}{\ell^t}\right)n^2 \quad < \quad c_0 n^2
\end{aligned}$$

by summing the convergent series. We have shown that for any constant c_0 and for all n sufficiently large, $\alpha(n) < c_0 n^2$. This completes the proof. □

Proof. (of Theorem 3.11) This is now an immediate corollary of Corollary 3.10 and Theorem 4.13, $\frac{n^2}{4}$ being the number of edges omitted in an empty pair. □

Remark 4.14. *Theorem 4.13, and thus Theorem 3.11, are more natural than might appear. On one hand, as Szemerédi regularity deals with density, it cannot (in this formulation) give precise information about edge counts below $\mathcal{O}(n^2)$. On the other, the random graph contains many infinite empty pairs, for instance $(\{(a,z) : z \in M, z \neq a\}, \{(y,a) : y \in M, y \neq a\})$ when $\varphi(x;y,z) = xRy \wedge \neg xRz$. One could imagine a future use for such theorems in suggesting ways of decomposing the parameter spaces of simple formulas into parts whose structure resembles random graphs (with many overlapping empty pairs) and parts whose structure is more cohesive, indicated by $\alpha(n) < \mathcal{O}(n^2)$.*

5. Order and genericity

Conclusion 2.9 shows a lag between the classification-theoretic complexity of φ and that of the formulas in its characteristic sequence: for a class of unstable theories strictly containing the simple theories, and for each n, after localization P_n will be stable. This section gives a first explanation for this phenomenon, relating instability of P_2 to the complexity of the interaction between pairs of arbitrarily large P_∞-complete graphs (base sets for types) in what might be called "generic position."

Much of the technology around the regularity lemma is built to extract configurations. To avoid appeal to machinery (and to be clear that the subgraphs involved are induced), let us extract the order property explicitly.

Observation 5.1. *Let T be a theory in a language containing a symmetric binary relation R. Suppose that for some $0 < \delta < 1$ and for all ϵ, n with $0 < \epsilon < 1$, $n \in \mathbb{N}$ there exists a 2-partite R-graph (X,Y), $|X| = |Y| \geq n$, such that (X,Y) is ϵ-regular with density d, where $|d - \delta| < \epsilon$. Then R has the order property.*

Proof. It suffices to show that for arbitrarily small ϵ_0 and arbitrarily large k_0 there is a Szemerédi-regular decomposition of X and of Y into k_0 pieces each such that all but $k_0(\epsilon_0)^2$ of the pairs X_i, Y_i are ϵ_0-regular with density near some given δ. Then the Key Lemma implies, roughly speaking, that we may think of the reduced graph as a random graph with edge probability δ and that any configuration which occurs in such a random graph with positive probability will occur in our original graph R. (See Corollary 4.4.)

The subtlety is to ensure that the densities of the regular pairs are all approximately the same. Given ϵ_0, k, let k_0, N_0 be the number of components and threshold size, respectively, given by the regularity lemma. Choose ϵ so that $\frac{1}{k_0} > \epsilon$ and $n > N_0$. Let (X, Y) be the ϵ-regular pair of size at least n and density near δ, given by hypothesis.

By regularity applied to the 2-partite graph (X, Y) (Remark 4.2), $n > N_0$ means that there is a decomposition $X = \cup_{i \leq k_0} X_i$, $Y = \cup_{i \leq k_0} Y_i$ into disjoint pieces of near equal size and that all but $\epsilon_0 (k_0)^2$ of the pairs (X_i, Y_j) are ϵ_0-regular. However any one of these regular pairs (X_i, Y_j) will satisfy $|X_i|, |Y_j| = n/k_0 > \epsilon n$, so $|d(X_i, Y_j) - d(X, Y)| = |d(X_i, Y_j) - \delta \pm \epsilon| < \epsilon$ and $|d(X_i, Y_j) - \delta| < 2\epsilon$, as desired. □

Remark 5.2. *In the case where we can assume that each of the partitioned graphs (X, Y) mentioned in the previous proof have the property that X and Y are each P_∞-complete graphs, we may conclude that there is a sequence $\langle a_i b_i : i < \omega \rangle$ on which R has the order property and such that each of $A := \bigcup_i a_i$ and $B := \bigcup_i b_i$ are P_∞-complete graphs.*

A key dividing line in classification theory is Shelah's strict order property, usually called *SOP* (not to be confused with the more recent *strong order properties* SOP_n, Definition 7.5). For the purposes of analyzing the characteristic sequence, it is usually most interesting to consider theories without strict order, because of the characterization given in Theorem 2.10.

Definition 5.3. ([6] Definition 4.3 p. 69) *The formula $\varphi(x; y)$ has the* strict order property*, or SOP, if there exists an indiscernible sequence $\langle a_i : i < \omega \rangle$ on which $\exists x (\neg \varphi(x; a_j) \land \varphi(x; a_i)) \iff j < i$.*

The main step in Shelah's classic proof that any unstable theory which does not have the independence property must have the strict order property can be characterized as follows:

Theorem E. (Shelah) *Let c be a finite set of parameters and $\langle a_i : i < \omega \rangle$ a c-indiscernible sequence. For $n < \omega$, any formula $\theta(x; \overline{z})$ and relations $R(x; y)$, $R_1, \dots R_n$ where $\ell(y) = \ell(a_i)$ and $R_i \in \{R(x; y), \neg R(x; y)\}$ for $i \leq n$, if*

$$i_1 < \dots < i_n \implies \exists x \left(\theta(x; c) \land R_1(x; a_{i_1}) \land \dots \land R_n(x; a_{i_n}) \right)$$

then either

- $\exists x \Big(\theta(x;c) \wedge R_1(x;a_{i_{\sigma(1)}}) \wedge \cdots \wedge R_n(x;a_{i_{\sigma(n)}})\Big)$ *for any permutation* $\sigma : n \rightarrow n$

- *some formula of* T *has the strict order property.*

The idea is to express the permutation σ as a sequence of swaps of successive elements (in the sense of the order $<$), and use the first instance, if any, where the swap produces inconsistency to obtain a sequence witnessing strict order. For details, see [6], Theorem II.4.7, pps. 70–72.

The subtlety of the lemma below is to obtain not just the independence property but a 2-partite random graph. See Definition 7.2 for a definition of "2-partite random graph."

Lemma 5.4. *Suppose that* $R(x;y)$ *has the order property. If* T *does not have the strict order property, then there exist infinite disjoint sets* A, B *on which* R *is a 2-partite random graph.*

Proof. We first fix a template. Let M be a countable model of the theory of a 2-partite random graph with two sorts P, Q and a single binary edge relation $E(x;y)$ with $E(x;y) \implies P(x) \wedge Q(y)$. Let $\langle x_i : i < \omega \rangle$, $\langle y_i : i < \omega \rangle$ be an enumeration of P and Q, respectively.

Now let $\langle a_i b_i : i < \omega \rangle$ be an indiscernible sequence on which R has the order property, i.e. $R(a_i, b_j) \iff i < j$. Suppose that for every $i < \omega$ we could find an element c_i such that for all $j < \omega$, $R(c_i, b_j) \iff E(x_i, y_j)$ in the template. Then setting $C := \bigcup_{i<\omega} c_i$, $B := \bigcup_{j<\omega} b_j$, (C, B) is a 2-partite random graph for R.

So it remains to show that any finite subset p of the type $p_i(x) \in S(B)$ of any such c_i is consistent. Let η, ν be disjoint finite subsets of ω, and let $p(x) = \bigwedge_{j\in\eta} R(x;b_j) \wedge \bigwedge_{k\in\nu} \neg R(x;b_k)$. We are now in a position to apply Theorem E; as T is $NSOP$, $p(x)$ must be consistent. □

The next definition will be most useful in the case where $R = P_2$, but we give the general statement.

Definition 5.5. *Let* T *be a given theory,* R *a binary relation symbol in the language of* T *and suppose that* T *implies* R *is symmetric.*

1. *Call a density* $0 \leq \delta \leq 1$ attainable *for* R *w.r.t.* T *if for all* ϵ *there exists a sequence* $\langle S^\delta_\epsilon = \langle (X_i, Y_i) : i < \omega \rangle$ *of finite 2-partite* R*-graphs in some model of* T *such that for all* $n < \omega, \epsilon > 0$ *there is* $N < \omega$ *such that for all* $i > N$,

- $|X_i| = |Y_i| \geq n$,
- (X_i, Y_i) *is* ϵ*-regular with density* d_i*, where* $|d_i - \delta| < \epsilon$.

Attainable densities exist, e.g. $\frac{1}{2}$*: consider subgraphs of an infinite random 2-partite graph.*

2. *Say that* R asymptotically realizes *the density* δ*, with respect to* T*, if for all* N, ϵ *there exists a 2-partite* R*-graph* (X, Y) *in some model* $M \models T$ *with* $|X| = |Y| \geq N$ *such that* (X, Y) *is* ϵ*-regular with density* d*, where* $|d - \delta| < \epsilon$.
3. *In the special case where* $R = P_2$ *and the* X, Y *can be chosen so that* X *and* Y *are both* P_∞*-complete graphs, say that* P_2 asymptotically realizes δ on complete graphs.

Lemma 5.6. *Assume the ambient theory* T *does not have the strict order property. Then the following are equivalent for a symmetric binary relation* $R(x, y)$ *in the language of* T*:*

1. *For some* $0 < \delta < 1$*,* R *asymptotically realizes* δ*.*
2. *For any attainable* $0 < \delta < 1$*,* R *asymptotically realizes* δ*.*
3. R *has the order property.*

Proof. (1) → (3) Graph theory, i.e., Observation 5.1.

(2) → (1) This condition is not vacuous, as attainable densities exist.

(3) → (2) Model theory, i.e., suppose that R has the order property but T does not have the strict order property. Then Lemma 5.4 gives infinite disjoint sets A, B on which R is a 2-partite random graph. Given an infinite 2-partite random graph, we can construct finite subgraphs of any attainable density. □

In other words, regularity plus compactness implies that density bounded away from $0, 1$ allows us to eventually construct any 2-partite graph, and so, a fortiori, construct the order property. Model theory implies that the order property is enough to reverse the argument, i.e. to obtain a 2-partite random graph.

Corollary 5.7. *Assume* T *does not have the strict order property, and* $(T, \varphi) \mapsto \langle P_n \rangle$*. Then the following are equivalent:*

1. *After localization,* P_2 *does not have the order property.*

2. *After localization, the density of any sufficiently large P_2-regular pair (X, Y) must approach either 0 or 1.*
More precisely, there exists $f : \mathbb{N} \times (0,1) \to [0, \frac{1}{2}]$ increasing as $n \to \infty$, $\epsilon \to 0$ such that if $X, Y \subset P_1$, $|X|, |Y| \geq n$ and (X, Y) is ϵ-regular, then either $d(X, Y) < f(n, \epsilon)$ or $d(X, Y) > 1 - f(n, \epsilon)$.

Proof. (1) → (2) Suppose that we can localize, i.e., restrict the parameter set of φ so that on the restricted set $X \subset P_1$, P_2 does not have the order property. Then P_2 cannot asymptotically realize any attainable density δ on this set X, lest it come under the scope of Lemma 5.6. (2) is the statement that for any given δ, Definition 5.5 eventually does not apply.

(2) → (1) Suppose that in every localization $X \subset P_1$, P_2 has the order property. Then by Lemma 5.6, P_2 asymptotically realizes some attainable density δ on parameters in X, and therefore (2) fails. □

Corollary 5.8. *If T is simple, then any characteristic sequence associated to one of its formulas satisfies the equivalent conditions of Corollary 5.7.*

Proof. Conclusion 2.9. □

Remark 5.9. *The class of theories satisfying the equivalent conditions of Corollary 5.7 strictly contains the simple theories. Example 3.6 of [5] gives a formula with the tree property whose P_2 does not have the order property. This is essentially T^*_{feq} from [8]; basic examples of TP_2 will work.*

Remark 5.10. *Any formula with SOP_2, also called TP_1, has the order property in P_2. For SOP_2, see [8]. However, the next section suggests that more precise order properties may be useful.*

6. Two kinds of order property

When P_2 has the order property, this says something about the manner in which the family of instances of φ interacts. We obtain a deeper picture if we bring more of the weight of the characteristic sequence to bear on our definitions. If the order property for P_2 occurs between two sets A, B each of which are empty graphs, this is a statement about the interaction of (by compactness) two dividing sequences; whereas if A, B are complete graphs, it is a statement about the interaction of two types.

In this section we investigate the "empty" and "compatible" order properties, and show that on the level of theories, the second is equivalent to

SOP_3, Conclusion 6.15 below. This is surprising because there are signs in the literature that SOP_3 is a robust indicator of complexity for a theory; see Remark 6.16 below.

Definition 6.1. (Two kinds of order property) *Let $\langle P_n \rangle$ be the characteristic sequence of φ.*

1. *φ has the n-*compatible order property*, for some $n < \omega$ (or $n = \infty$) if there exist $\langle a_i, b_i : i < \omega \rangle$ such that for all $m \leq n$ (or $m < \omega$), $P_{2m}(a_{i_1}, b_{j_1}, \dots a_{i_m}, b_{j_m})$ iff* $\max\{i_1, \dots i_m\} < \min\{j_1, \dots j_m\}$.

(1)′ *When the sequence has support 2 this becomes: there exist $\langle a_i, b_i : i < \omega \rangle$ such that $P_2(a_i, a_j)$, $P_2(b_i, b_j)$ for all i, j and $P_2(a_i, b_j)$ iff $i < j$.*

2. *φ has the n-*empty order property*, for some $n \in \omega$, if: there exist $\langle a_i, b_i : i < \omega \rangle$ such that (i) $P_2(a_i; b_j)$ iff $i < j$ and (ii) $\neg P_n(a_{i_1}, \dots a_{i_n})$, $\neg P_n(b_{i_1}, \dots b_{i_n})$ hold for all $i_1, \dots i_n < \omega$.*

Let us briefly justify not focusing on a natural third possibility, the "semi-compatible order property," in which the elements $\langle a_i : i < \omega \rangle$ are an empty graph and the elements $\langle b_i : i < \omega \rangle$ are a positive base set.

Claim 6.2. *There is a formula in the random graph which has the semi-compatible order property.*

Proof. Choose two distinguished elements 0, 1 (this can be coded without parameters). Define $\psi(x; y, z)$ to be $x = y$ if $z = 0$, xRy otherwise. Then on any sequence of distinct elements $\langle a_i b_i : i < \omega \rangle \subset M$ which witness the order property ($a_i R b_j \iff i < j$), we have additionally that

$$\exists x \left(\psi(x; a_i, 0) \land \psi(x; b_j, 1) \right) \iff \exists x \left(x = a_i \land xRb_j \right) \iff i < j$$

so P_2 has the order property on the sequence $\langle (a_i, 0), (b_i, 1) : i < \omega \rangle$. On the other hand, $\exists x (x = a_i \land x = a_j) \iff i = j$, so the row of elements $(a_i, 0)$ is a P_2-empty graph. Finally, $\exists x (xRb_i \land xRb_j)$ always holds, by the axioms of the random graph; so the row of elements $(b_j, 1)$ is a P_∞-complete graph. □

Claim 6.3. *There is a formula in a simple rank 3 theory which has the 2-empty order property.*

Proof. Let T be the theory of two crosscutting equivalence relations, E and F, each with infinitely many infinite classes and such that each intersection $\{x : E(a, x) \land F(x, b)\}$ is infinite. Let P be a unary predicate such that

- $(\forall x, y)(E(x,y) \land F(x,y) \implies P(x) \iff P(y))$
- For all $n < \omega$ and $y_1, \ldots y_k, y_{k+1}, \ldots y_n$ elements of distinct E-equivalence classes, there exists z such that $i \leq k \implies (\forall x)(E(x,y_i) \land F(x,z) \implies P(x))$ and $k < i \leq n \implies (\forall x)(E(x,y_i) \land F(x,z) \implies \neg P(x)))$

Define

$$\psi(x; y, z, w) = \begin{cases} E(x,y) & \text{if z=w} \\ F(x,y) \land P(y) & \text{otherwise} \end{cases}$$

As usual, write $\psi(x; y, 0)$ for the first case and $\psi(x; y, 1)$ for the second. Let $\langle a_i, b_i : i < \omega \rangle$ be a sequence of elements chosen so that $(\forall x)(E(x, a_i) \land F(x, b_j) \implies P(x))$ iff $i < j$. Then it is easy to see ψ has the 2-empty order property on the sequence $\langle (a_i, 0), (b_i, 1) : i < \omega \rangle$. □

Remark 6.4. *Assuming $MA + 2^{\aleph_0} > \aleph_1$, Shelah has constructed an ultrafilter on ω which saturates (small) models of the random graph, but not of theories with the tree property* ([6] Theorem VI.3.9). *This is a strong argument for the "semi-compatible order property" being less complex: it cannot, by itself, imply maximality in the Keisler order, whereas we will see that the ∞-compatible order property does. It may still be that persistence, in the sense of [5], of any order property in P_2 creates complexity.*

We return to the study of the compatible order property.

Convention 6.5. *When more than one characteristic sequence is being discussed, write $P_n(\varphi)$ to indicate the nth hypergraph associated to the formula φ. Recall that φ_ℓ is shorthand for $\bigwedge_{1 \leq i \leq \ell} \varphi(x; y_i)$.*

The following general principle will be useful.

Lemma 6.6. *Suppose that we have a sequence $C := \langle c_i : i \in \mathbb{Z} \rangle$ and a formula $\rho(x; y, z)$ such that:*

1. $\exists x \rho(x; c_i, c_j) \iff i < j$
2. $\exists x \left(\bigwedge_{\ell \leq n} \rho(x; c_{i_\ell}, c_{j_\ell}) \right)$ *just in case* $\max\{i_1, \ldots i_n\} < \min\{j_1, \ldots j_n\}$

Then ρ has the ∞-compatible order property.

Proof. By compactness, it is enough to show that there are elements $\langle \alpha_i, \beta_i : i < n \rangle$ witnessing a fragment of the ∞-compatible order property of size n.

Define $\alpha_1 \ldots \alpha_n, \beta_1, \ldots \beta_n$ as follows. Remark 6.7 provides a picture.

- $\alpha_i := c_{2i-1}c_{4n-2i+1}$, $1 \leq i \leq n$
- $\beta_i := c_{-2i}c_{2i}$, $1 \leq i \leq n$

Then $P_1(\alpha_i)$, $P_1(\beta_i)$ for $1 \leq i \leq n$ by (1). For all $1 \leq k, r \leq n$ with $r + k = m$, condition (2) says that $P_m(\alpha_{j_1}, \dots \alpha_{j_k}, \beta_{i_1}, \dots \beta_{i_r})$ iff

$$\max\{2\ell : \ell \in \{i_1, \dots i_r\}\} < \min\{2s - 1 : s \in \{j_1, \dots j_k\}\}$$

that is, iff $\max\{i_1, \dots i_r\} < \min\{j_1, \dots j_k\}$, so we are done. □

Remark 6.7. *The ∞-compatible order property describes an interaction between two P_∞-complete graphs, i.e. consistent types. The hypotheses (1)-(2) of Lemma 6.6 are enough to allow a weak description of intervals. That is, we choose the sequences α_i, β_i to each describe a concentric sequence of intervals (each α_i, β_i corresponds to a set of matching parentheses) along the sequence $\langle c_i \rangle$:*

$$\leftarrow [-[-[-[-]-]-]-] - - \cdots - -(-(-(-(-)-)-)-) \rightarrow$$

which we can interlace to obtain ∞-c.o.p. by judicious choice of indexing:

$$\leftarrow [-[-[-[-(-]-(-]-(-]-(-]-)-)-)-) \rightarrow$$

In this picture, the enumeration of the αs ( ), would proceed from the outmost pair to the inmost and the enumeration of the βs [ ] from inmost to outmost.

Definition 6.8. *Given a characteristic sequence $\langle P_n \rangle$ and some set $A \subset P_1$, say that $\langle P_n \rangle$* has support k on A *if for all $r > k$ and all $\{a_1, \dots a_r\} \subseteq A$, $P_r(a_1, \dots a_r)$ iff P_k holds on every k-element subset of $\{a_1, \dots a_r\}$.*

Claim 6.9. *Suppose that φ has the strict order property, i.e. there is an infinite sequence $\langle c_i : i < \omega \rangle$ on which $\exists x(\neg\varphi(x; c_i) \wedge \varphi(x; c_j)) \iff i < j$. Then $\neg\varphi(x; y) \wedge \varphi(x; z)$ has the ∞-compatible order property.*

Proof. By compactness, we may assume that the sequence $\langle c_i \rangle$ is indiscernible. Writing $\rho(x; y, z) = \neg\varphi(x; y) \wedge \varphi(x; z)$,

- $\exists x \rho(x; c_i, c_j) \iff i < j$, by definition of strict order;
- $\exists x(\rho(x; c_i, c_j) \wedge \rho(x; c_k, c_\ell)) \iff i, k < j, \ell$

Furthermore, the characteristic sequence $P_\infty(\rho)$ has support 2 on $\langle c_i \rangle$ (Definition 6.8), so condition (2) of Lemma 6.6 is also satisfied. Apply Lemma 6.6. $\square$

Example 6.10. *The theory T of the generic triangle-free graph with edge relation R has the ∞-c.o.p. Consider $\varphi(x; y, z) = xRy \wedge xRz$. (The negative instances could be added but are not necessary.) Then:*

- $P_1((y, z)) \iff \neg yRz$.
- $P_2((y, z), (y', z'))$ *iff* $\{y, y', z, z'\}$ *is an empty graph.*
- *The sequence has support 2, as the only problems come from a single new edge:* $P_n((y_1, z_1), \dots (y_n, z_n))$ *iff*

$$\exists x \left(\bigwedge_{i \leq n} xRy_i \wedge \bigwedge_{j \leq n} xRz_j \right) \text{ that is, if } \bigcup_i y_i \cup \bigcup_j z_j \text{ is a } P_2\text{-empty graph.}$$

Let $\langle a_i, b_i : i \in \mathbb{Z} \rangle$ be a sequence witnessing the 2-empty *order property with respect to the edge relation R, say $a_i R b_j$ iff $j \leq i$. Then $\exists x(xRa_i \wedge xRb_j)$ iff $i < j$, i.e. $(a_i, b_j) \in P_1$ iff $i < j$. Also, $\exists x(xRa_i \wedge xRb_j \wedge xRa_k \wedge xRb_\ell)$ if, in addition, $i, k < j, \ell$. Apply Lemma 6.6.*

Finally, we tie the compatible order property to SOP_3, a model-theoretic rigidity property. SOP_3 will be important in the next section; the general definition is Definition 7.5, but an equivalent definition is the following. Remember that, by convention, $a_i, x, \dots$ need not be singletons.

Definition 6.11. ([8]:Fact 1.3) *T has SOP_3 iff there is an indiscernible sequence $\langle a_i : i < \omega \rangle$ and $\mathcal{L}$-formulas $\varphi(x; y), \psi(x; y)$ such that:*

1. $\{\varphi(x; y), \psi(x; y)\}$ *is contradictory.*
2. *there exists a sequence of elements $\langle c_j : j < \omega \rangle$ such that*
   - $i \leq j \implies \varphi(c_j; a_i)$
   - $i > j \implies \psi(c_j; a_i)$
3. *if $i < j$, then $\{\varphi(x; a_j), \psi(x; a_i)\}$ is contradictory.*

Lemma 6.12. *Suppose that $\theta(x; y)$ has SOP_3 in the sense of Definition 6.11. Let $\varphi_r = \varphi$, $\psi_\ell = \psi$ be the formulas from Definition 6.11. Then $\rho(x; y, z) := \varphi_r(x; y) \wedge \psi_\ell(x; z)$ has the ∞-compatible order property.*

Remark 6.13. *This is an existential assertion, and it is straightforward to check that it remains true if we modify ρ to include the corresponding negative instances.*

Proof. (of Lemma) Let $A := \langle a_i : i < \mathbb{Q} \rangle$ be an infinite indiscernible sequence from Definition 6.11. Then

$$P_1((a_i, a_j)) \iff \exists x \left(\varphi_r(x; a_i) \wedge \psi_\ell(x; a_j)\right) \iff i < j$$

by the choice of φ, ψ. More generally,

$$P_n((a_{i_1}, a_{j_1}), \ldots (a_{i_n}, a_{j_n})) \iff \exists x \left(\bigwedge_{t \leq n} \varphi_r(x; a_{i_t}) \wedge \bigwedge_{t \leq n} \psi_\ell(x; a_{j_t}) \right)$$

which, again applying Definition 6.11, is true just in case $\max\{i_1, \ldots i_n\} < \min\{j_1, \ldots j_n\}$. We now apply Lemma 6.6. □

Lemma 6.14. *Suppose $\theta(x; y)$ has the ∞-compatible order property. Then the formula $\varphi(x; y, z) := \theta(x; y) \wedge \neg\theta(x; z)$ has SOP_3.*

Proof. Let $\langle d_i b_i : i < \omega \rangle$ be a sequence witnessing the ∞-compatible order property; this will play the role of the sequence $\langle a_i : i < \omega \rangle$ from Definition 6.11. In the notation of that Definition, let $\varphi(x; y, z) := \theta(x; y) \wedge \neg\theta(x; z)$ and $\psi(x; y, z) := \theta(x; z)$. We check the conditions.

(1) Clearly $\{\varphi(x; y, z), \psi(x; y, z)\}$ is inconsistent.

(3) When $i > j$, $\{\varphi(x; d_i b_i), \psi(x; d_j b_j)\} = \{\theta(x; d_i) \wedge \neg\theta(x; b_i), \theta(x; b_j)\}$ is inconsistent because $\neg P_2(d_i, b_j)$.

Finally, for $1 \leq j < \omega$ let $p_j(x) = \{\theta(x; d_i) : 1 \leq i \leq j\} \cup \{\theta(x; b_\ell) : j < \ell < \omega\}$. The ∞-c.o.p. implies $P_n(d_1, \ldots d_j, b_{j+1}, \ldots b_n)$ for all $n < \omega$, so p_j is consistent. However, $i < j \implies \neg P_2(b_i, d_j)$ so $p_j(x) \vdash \neg\theta(x; b_i)$ for each $1 \leq i \leq j$. Choosing $c_j \models p_j$ for each $j < \omega$ gives (2). □

Conclusion 6.15. *The following are equivalent for a theory T:*

1. *T contains a formula with the ∞-compatible order property.*
2. *T contains a formula with SOP_3.*

Proof. See the two previous lemmas. □

Remark 6.16. *Applying Shelah's theorem that any theory with SOP_3 is maximal in the Keisler order [7], [8], we conclude that if T contains a formula φ with the ∞-compatible order property, then T is maximal in the Keisler order. For more on Keisler's order, see [4].*

7. Calibrating randomness

In this final section, we observe and explain a discrepancy between the model-theoretic notion of an infinite random k-partite graph and the finitary version given by Szemerédi regularity, showing essentially that a class of infinitary k-partite random graphs which do not admit reasonable finite approximations must have the strong order property SOP_3.

7.1. A seeming paradox

Observation 7.1. *Let T be the theory of the generic triangle-free graph, with edge relation R. Then it is consistent with T that there exist disjoint infinite sets X, Y, Z such that each pair $(X, Y), (Y, Z), (X, Z)$ is a 2-partite random graph.*

Proof. The construction has countably many stages. At stage 0, let $X_0 = \{a\}, Y_0 = \{b\}, Z_0 = \{c\}$ where a, b, c have no R-edges between them. At stage $i + 1$, let X_{i+1} be X_i along with $2^{|Y_i|+|Z_i|}$-many new elements:

1. for each subset $\tau \subset Y_i$, a new element x_τ such that for $y \in Y$, $x_\tau R y \iff y \in \tau$, however $\neg x_\tau R x$ for any x previously added to X_{i+1}.
2. for each subset $\nu \subset Z_i$, a new element x_ν such that for $z \in Z$, $x_\nu R z \iff z \in \nu$, with x_ν likewise R-free from previous elements of X_{i+1}.

Y_{i+1}, Z_{i+1} are defined symmetrically. As we are working in the generic triangle-free graph, in order that the the construction be able to continue, it is enough that the sets X_i, Y_i, Z_i are each empty graphs, i.e., at no point do we ask for a triangle.

To finish, set $X = \bigcup_i X_i$, $Y = \bigcup_i Y_i$, $Z = \bigcup_i Z_i$. Each pair is a 2-partite random graph, as desired. □

But recall:

Theorem F. *(weak version of Key Lemma, Theorem C) Fix $1 > \delta > 0$ and a binary edge relation R. Then there exist $\epsilon' = \epsilon'(\delta), N' = N'(\epsilon', \delta)$ such that:* if *$\epsilon < \epsilon'$, $N > N'$, X, Y, Z are disjoint finite sets of size at least N, and each of the pairs $(X, Y), (Y, Z), (X, Z)$ is ϵ-regular with density δ,* then *there exist $x \in X, y \in Y, z \in Z$ so that x, y, z is an R-triangle.*

Obviously, we cannot have an R-triangle in the generic triangle-free graph. Nonetheless each of the pairs (X, Y) in Observation 7.1 manifestly has finite subgraphs of any attainable density.

The difficulty comes when we try to choose finite subgraphs $X' \subset X, Y' \subset Y, Z' \subset Z$ so that the densities of all three pairs are *simultaneously* near the same $\delta > 0$. If (X', Y') and (Y', Z') are reasonably dense, (X', Z') will be near 0. Put otherwise, we may choose elements of X independently over Y, and independently over Z, but not both at the same time.

The constructions below generalize this example, and give a way of measuring the "depth" of independence in a constellation of sets $X_1, \dots X_n$, where any pair (X_i, X_j) is a 2-partite random graph. The example of the generic triangle-free graph is paradigmatic: we shall see that a bound on the depth of independence will produce the 3-strong order property SOP_3.

7.2. Constellations of independence properties.

Definition 7.2. *Fix a formula $R(x; y)$.*

1. *Let A, B be disjoint sets of k- and n-tuples respectively, where $k = \ell(x), n = \ell(y)$. Then A is* independent over B *with respect to R just in case for any two finite disjoint $\eta, \nu \subset B$, there exists $a \in A$ such that $b \in \eta \rightarrow R(a; b)$ and $b \in \nu \rightarrow \neg R(a; b)$.*
2. *Let $A_1, \dots A_k$ be disjoint sets (of m-tuples, where $m = \ell(x) = \ell(y)$). Then A_1 is* independent over *$A_2, \dots A_k$ with respect to R just in case A_1 is independent over $B := \bigcup_{2 \leq i \leq k} A_i$ in the sense of (2).*
3. *If there exist disjoint infinite sets A, B such that A and B are each independent over the other wrt R, then $R(x; y)$ is a* 2-partite random graph *on A, B. Often we will not name A, B explicitly and simply say $R(x; y)$ is a 2-partite random graph.*
4. *$R(x; y)$ is I^m_k, for some $2 \leq k \leq m$, if there exist disjoint infinite sets $\langle A_i : i < m \rangle$ such that for any distinct $i_1, \dots i_k < \omega$, A_{i_1} is independent over $\bigcup_{2 \leq j \leq k} A_{i_j}$ w.r.t. R. Note that k refers to the* depth *of the independence, and not the size of the finite disjoint η, ν.*

Remark 7.3. *The statement that R is I^{n+1}_n with respect to a background theory T is expressible as an infinitary partial type.*

Proof. We will build p as a type in the variables $\{x^i_j : i < \omega, 0 \leq j \leq n\}$ in the language with equality and the binary edge relation R. Note that the partition into clusters is not part of the language, and the type will not specify the edge relations between variables with the same subscript. At stage 0, let $p_0 := \{x^0_0 = x^0_0\}$. At stage $t+1$, suppose $t \equiv m \pmod{n+1}$. The partial type p_t will mention at most finitely many variables with subscript $j \neq m$: call this finite set of variables $V_{t,m}$. We construct p_{t+1} in finitely many stages. Set $p_{t+1,0} := p_t$. Denote by $h(t+1,i)$ the smallest integer h such that x^h_m is not mentioned in $p_{t+1,i}$. Enumerate the subsets $V_{t,m,i} \subseteq V_{t,m}$, and let

$$p_{t+1,i+1} := p_{t+1,i} \cup \{R(x_m^{h(t+1,i)}, v) : v \in V_{t,m,i}\} \cup \{\neg R(x_m^{h(t+1,i)}, v) : v \notin V_{t,m,i}\}$$

Let $p_{t+1} := \bigcup_i p_{t+1,i}$, completing the inductive step. Finally, let $p := \bigcup_{t<\omega} p_t$. □

Observation 7.4. *Let $R(x;y)$ be a symmetric formula. The following are equivalent.*

1. *R is I^ω_ω.*
2. *There is an infinite subset of the monster model on which R is a random graph. (Certainly this need not be definable or interpretable in any way).*

Definition 7.5. (Shelah, [7]:Definition 2.5) *For $n \geq 3$, the theory T* has *SOP_n if there is a formula $\varphi(x;y)$, $\ell(x) = \ell(y) = k$, $M \models T$ and a sequence $\langle a_i : i < \omega \rangle$ with each $a_i \in M^k$ such that:*

1. *$M \models \varphi(a_i, a_j)$ for $i < j < \omega$*
2. *$M \models \neg \exists x_1, \dots x_n (\bigwedge \{\varphi(x_m, x_k) : m < k < n$ and $k = m+1$ mod $n\})$*

Compare Definition 6.11 above.

Theorem G. (Shelah, [7]: (1) is Claim 2.6, (2) is Theorem 2.9)

1. *For a theory T, $SOP \implies SOP_{n+1} \implies SOP_n$, for $n \geq 3$ (not necessarily for the same formula).*
2. *If T is a complete theory with SOP_3, then T is maximal in the Keisler order.*

The novelty of the following argument is not the result that the generic triangle-free graph has SOP_3, which is known by [7], Claim 2.8(2); Example 6.10 and Lemma 6.12 above give an alternative proof. Rather, it illustrates the key ideas from the more elaborate construction of Theorem 7.7.

Example 7.6. *Let T be the generic triangle-free graph, with edge relation R. Then R is I^3_2 but not I^3_3, and T has SOP_3.*

Proof. Let us prove the final clause (for the rest see Observation 7.1 and the discussion following).

Suppose A, B, C are disjoint infinite sets witnessing I^3_2. Let us construct a sequence of triples $\langle a_i, b_i, c_i : i < \omega \rangle$ such that, for $i < \omega$,

- For all $j \leq i$, $b_i R a_j$.
- For all $j \leq i$, $c_i R b_j$.
- For all $j \leq i$, $a_{i+1} R c_j$.

Let $\gamma_i := (a_i b_i c_i)$ and $S := \langle \gamma_i : i < \omega \rangle$. In other words, we construct a helix of elements which approximate the forbidden configuration in the following sense. The elements fall into three clusters, A_0, A_1, A_2, and given elements x_i, x_j with $x_i \in A_i$, $x_j \in A_j$ and $i > j$, the edge between x_i, x_j agrees with the forbidden configuration except when $j = i + 1$ modulo 3.

Define a binary relation $<_\ell$ on triples by:

$$(x, y, z) \leq_\ell (x', y', z') \iff ((xRy' \wedge yRz' \wedge zRx'))$$

While $<_\ell$ need not be a partial order on the model, it does linearly order the sequence S by construction. Looking towards Definition 6.11, let us define two new formulas (the variables t stand for triples):

- $\varphi(t_0; t_1, t_2) = t_1 <_\ell t_2 <_\ell t_0$
- $\psi(t_0; t_1, t_2) = t_0 <_\ell t_1 <_\ell t_2$

Let us check that these formulas give SOP_3. For condition (1), notice that $\varphi(t_0; t_1, t_2), \psi(t_0; t_1, t_2)$ means that $(x_0, y_0, z_0) <_\ell (x_1, y_1, z_1) <_\ell (x_2, y_2, z_2) <_\ell (x_0, y_0, z_0)$. Then $x_i R y_j$, $y_j R z_k$, $z_k R x_i$ which gives a triangle, contradiction.

It is straightforward to satisfy (2) by compactness (e.g. by choosing S codense in a larger indiscernible sequence).

Finally, for condition (3), suppose $i < j$ but $\varphi(t; \gamma_i), \psi(t; \gamma_j)$ is consistent, where $t = (x, y, z)$. This means that $(x, y, z) <_\ell (a_i, b_i, c_i) <_\ell (a_j, b_j, c_j) <_\ell (x, y, z)$ (where the middle $<_\ell$ comes from the behavior of $<_\ell$ on the sequence S). As in condition (1), this gives a triangle, contradiction. $\square$

We now extend this idea to a much larger engine for producing enough rigidity for SOP_3 from a forbidden configuration.

Theorem 7.7. *Suppose that for some $2 \leq n < \omega$, the formula R of T is I^{n+1}_{n} but not I^{n+1}_{n+1}. Then T is SOP_3.*

Proof. The construction is arranged into four stages.

Step 1: Finding a universally forbidden configuration G.

Let $p(X_0, \dots X_n)$ be the infinitary type given by Remark 7.3 which describes $n+1$ infinite sets X_i which are I^{n+1}_{n+1}. By hypothesis, R is not I^{n+1}_{n+1}, so p is not consistent with T. Let G be a finite inconsistent subset in the variables $V_G = \{x^i_j : 1 \leq i \leq h, 0 \leq j \leq n\}$, and described by the edge map $E_G : \{((i,j),(i',j')) : i, i' \leq h, j \neq j' \leq n\} \rightarrow \{0,1\}$. As the inconsistency of p is a consequence of T, G will be a universally forbidden configuration:

$$T \vdash \neg(\exists x^1_0, \dots x^h_n) \left(\bigwedge_{i,i' \leq h,\ j \neq j' \leq n} R(x^i_j, x^{i'}_{j'}) \iff E((i,j),(i',j')) = 1 \right) \quad (1)$$

Note that the configuration remains agnostic on edges between elements in the same column, in keeping with the definition of I^m_ℓ.

In what follows G will appear as a template which we shall try to approximate using I^{n+1}_{n}. Here are the vertices of G arranged as they will be visually referenced (the edges are not drawn in):

$$\begin{array}{lllll}
x^h_0 & & x^h_k & & x^h_n \\
\vdots & & \vdots & & \vdots \\
x^\rho_0 & & x^\rho_k & & x^\rho_n \\
\vdots & & \vdots & & \vdots \\
x^1_0 & \dots & x^1_k & \dots & x^1_n
\end{array}$$

Figure 1: Vertices of the forbidden configuration G, arranged in columns. When comparing this configuration to an array whose rows are indexed modulo h, the superscript of the top column becomes 0.

Step 2: Building an array A of approximations to G.

Let $A_0, \dots A_n$ be disjoint infinite sets witnessing I^{n+1}_{n} for R. As in Example 7.6, we will use elements from these columns A_i to build an array $A = \langle a^\rho_i : 1 \leq \rho < \omega, 0 \leq i \leq n \rangle$. Fixing notation,

- $a^\rho_0, \dots a^\rho_n$ is called a *row*.
- $\mathrm{Col}(i) = \{j : j \neq i, i+1 \pmod{n+1}\}$ is the set of column indices associated to the column index i.
- Define a relation on pairs of indices (β for "before"):

$$\beta((t', i'), (t, i)) \iff_{def}$$
$$\Big((t' < t \wedge i' \in \mathrm{Col}(i)) \vee (t' = t \wedge i' < i)\Big)$$

Claim 7.8. *We may build the array A to satisfy:*

1. *For all ρ, $a^\rho_k \in A_k$.*
2. *For any ρ', ρ, k, k' such that $\beta((\rho', k'), (\rho, k))$,*

$$a^\rho_k \; R \; a^{\rho'}_{k'} \iff E_G((r, k), (r', k')) = 1$$

where $r \equiv \rho \pmod h$, $r' \equiv \rho' \pmod h$.

Proof. We choose elements in a helix $(a^1_0, a^1_1, \dots a^1_n, a^2_0, a^2_1, \dots)$ in such a way that $\beta((\rho', k'), (\rho, k))$ implies that $a^{\rho'}_{k'}$ is chosen before a^ρ_k.

When the time comes to choose a^ρ_k, we look for an element of A_k which satisfies Condition (2) of the Claim, that is, which, by Condition (1), realizes a given R-type over disjoint finite subsets of the columns A_i ($i \in \mathrm{Col}(k)$). Speaking informally, as we go around the circle of clusters, a shadow follows us which is not as long as we would like (i.e. it does not go n columns back) but it is next best, i.e. it goes $n-1$ columns back. The condition of I^{n+1}_n says exactly that we can choose an element in the cluster at hand which will exactly match the forbidden configuration with respect to any elements already defined which are covered by this shadow. More formally, as $(A_0, \dots A_n)$ was chosen to be I^{n+1}_n and $|\mathrm{Col}\, k| = n-1$, an appropriate a^ρ_k exists. □

Step 3: Defining the relation $<_\ell$, which has no pseudo-$(n+1)$-loops.

We now define a binary relation $<_\ell$ on m-tuples, where $m = h(n+1)$. Fix the enumeration of these tuples to agree with the natural interpretation as blocks B_ℓ of h consecutive rows in the array A (see Figure 7.2). That is,

$$
\begin{array}{lccccccc}
 & & & & \vdots & & & \\
 & & & & \vdots & & & \\
\hline
 & a_0^{\ell_n h+h} & & \dots & & a_k^{\ell_n h+h} & & \mathbf{a}_{\mathbf{n}}^{\ell_{\mathbf{n}}\mathbf{h}+\mathbf{h}} \\
 & \vdots & & & & \vdots & & \vdots \\
B_{\ell_n} = & a_0^{\ell_n h+\rho} & & & & a_k^{\ell_n h+\rho} & & \mathbf{a}_{\mathbf{n}}^{\ell_{\mathbf{n}}\mathbf{h}+\rho} \\
 & \vdots & & & & \vdots & & \vdots \\
 & a_0^{\ell_n h+1} & & & & a_k^{\ell_n h+1} & & \mathbf{a}_{\mathbf{n}}^{\ell_{\mathbf{n}}\mathbf{h}+\mathbf{1}} \\
\hline
 & & & & \vdots & & & \\
\hline
 & a_0^{\ell_k h+h} & & \dots & & \mathbf{a}_{\mathbf{k}}^{\ell_{\mathbf{k}}\mathbf{h}+\mathbf{h}} & & a_n^{\ell_k h+h} \\
 & \vdots & & & & \vdots & & \vdots \\
B_{\ell_k} = & a_0^{\ell_k h+\rho} & & & & \mathbf{a}_{\mathbf{k}}^{\ell_{\mathbf{k}}\mathbf{h}+\rho} & & a_n^{\ell_k h+\rho} \\
 & \vdots & & & & \vdots & & \vdots \\
 & a_0^{\ell_k h+1} & & \dots & & \mathbf{a}_{\mathbf{k}}^{\ell_{\mathbf{k}}\mathbf{h}+\mathbf{1}} & \dots & a_n^{\ell_k h+1} \\
\hline
 & & & & \vdots & & & \\
\hline
 & a_0^{2h} & \mathbf{a}_{\mathbf{1}}^{\mathbf{2h}} & & & a_k^{2h} & & a_n^{2h} \\
 & \vdots & \vdots & & & \vdots & & \vdots \\
B_1 = & a_0^{h+\rho} & \mathbf{a}_{\mathbf{1}}^{\mathbf{h}+\rho} & & & a_k^{h+\rho} & & a_n^{h+\rho} \\
 & \vdots & \vdots & & & \vdots & & \vdots \\
 & a_0^{h+1} & \mathbf{a}_{\mathbf{1}}^{\mathbf{h}+\mathbf{1}} & & & a_k^{h+1} & \dots & a_n^{h+1} \\
\hline
 & \mathbf{a}_{\mathbf{0}}^{\mathbf{h}} & & & & a_k^{h} & & a_n^{h} \\
 & \vdots & & & & \vdots & & \vdots \\
B_0 = & \mathbf{a}_{\mathbf{0}}^{\rho} & & & & a_k^{\rho} & & a_n^{\rho} \\
 & \vdots & & & & \vdots & & \vdots \\
 & \mathbf{a}_{\mathbf{0}}^{\mathbf{1}} & & & \dots & a_k^{1} & \dots & a_n^{1}
\end{array}
$$

Figure 2: Elements of the array A, arranged in blocks of h rows. The boldface refers to Step 4 of the proof, when a proposed witness to G is assembled from the ith columns of blocks B_i in a pseudo-$(n+1)$-loop.

write the variables $Y := \langle y_i^t : 1 \leq t \leq h, 0 \leq i \leq n\rangle$, $Z := \langle z_{i'}^{t'} : 1 \leq t' \leq h, 0 \leq i' \leq n\rangle$. Define:

$$\mathbf{Y} <_\ell \mathbf{Z} \iff_{(def)}$$
$$\bigwedge_{1\leq t',t\leq h,\ 0\leq i,i'\leq n} (i' \in \mathrm{Col}(i)) \implies \left(z_i^t\ R\ y_{i'}^{t'} \iff E_G((t,i),(t',i')) = 1\right)$$

Let B be a partition of the array A into blocks B_k ($k < \omega$) each consisting of h consecutive rows, so $B_k := \langle a_t^r : 0 \leq t \leq n, kh+1 \leq r \leq (kh)+h\rangle$, for each $k < \omega$ (see Figure 7.2). By Claim 7.8, for any $i, j < \omega$, $i < j \implies B_i <_\ell B_j$.

Definition 7.9. *A* pseudo-$(n+1)$-loop *is a sequence* W_i $(0 \leq i \leq n)$ *such that for some* m, $1 \leq m < n$*:*

$$\left(\bigwedge_{(0<j<i\leq n)} W_j <_\ell W_i\right) \wedge \left(\bigwedge_{1\leq j\leq m} W_0 <_\ell W_j\right) \wedge \left(\bigwedge_{m<j\leq n} W_j <_\ell W_0\right) \quad (2)$$

Although $<_\ell$ is not symmetric, notice that:

Remark 7.10. *Let* $X_0, \ldots X_n$ *be tuples of variables of uniform length* m *and suppose* S *is a symmetric* $2m$*-ary relation. Suppose that*

$$\left(\bigwedge_{(0<j<i\leq n)} S(X_j, X_i)\right) \wedge \left(\bigwedge_{1\leq j\leq m} S(X_0, X_j)\right) \wedge \left(\bigwedge_{m<j\leq n} S(X_j, X_0)\right)$$

Then for all $0 \leq i < j \leq n$, $S(X_i, X_j)$.

Claim 7.11. *Pseudo-*$(n+1)$*-loops in* $<_\ell$ *are inconsistent with* T.

Proof. Suppose it were consistent with T to have blocks of variables $W_0 \ldots W_n$ which form a pseudo-$(n+1)$-loop. Write $W_k(i) = \{w_i^{hk+1}, \ldots w_i^{hk+h}\}$ for the ith column of block W_k. Figure 7.2 gives the picture, where the elements a are replaced by variables w and the blocks B_i become W_i.

Notice that the asymmetric relation $<_\ell$ between columns $W_i(i), W_i(j)$ gives rise to a *symmetric* relation between those same columns, namely the relation which expresses "the edges between elements of $W_i(i)$ and those of $W_j(j)$ agree exactly with the edges which occur between the ith and jth columns in the forbidden configuration."

More formally, set $W_G = W_0(0) \cup W_1(1) \cup \cdots \cup W_n(n)$. This can be visualized as the boldface columns in Figure 7.2. By definition of $<_\ell$, the pseudo-(n+1)-loop (2) implies that whenever

$$((j \in \mathrm{Col}(i)) \wedge ((0 < j < i \leq n) \vee (j = 0 \wedge i \leq m) \vee (m < j \wedge i = 0)))$$

we will have:

$$\left(\forall\ w^t_k \in W_i(i),\ w^{t'}_{k'} \in W_j(j)\right) \left(w^t_k\ R\ w^{t'}_{k'} \iff E_G((t,k),(t',k')) = 1\right)$$

In a pseudo-(n+1)-loop, given any distinct indices $0 \leq i < j \leq n$, either $W_i(i) <_\ell W_j(j)$ or vice versa. In either case, edges between vertices in $W_i(i)$ and those in $W_j(j)$ will agree with the forbidden configuration. By Remark 7.10, W_G has the forbidden configuration, which is a contradiction. $\square$

Step 4: Obtaining SOP_3.

Step 3 showed that our array A of approximations had a certain rigidity, which we can now identify as SOP_3. Following Definition 6.11, let us define $\varphi_r(x; y_1, \ldots y_n)$ and $\psi_\ell(x; y_1, \ldots y_n)$, where the the variables are blocks, and the subscripts "ℓ" and "r" are visual aids: the element x goes to the left of the elements y_i under ψ, and to their right under φ.

That is, we set:

- $\varphi_r(x; y_1, \ldots y_n) =$

$$\bigwedge_{1 \leq i \neq j \leq n} y_i <_\ell y_j \ \wedge \ \bigwedge_{1 \leq i \leq n} y_i <_\ell x$$

- $\psi_\ell(x; y_1, \ldots y_n) =$

$$\bigwedge_{1 \leq i \leq n} x <_\ell y_i \ \wedge \ \bigwedge_{1 \leq i \neq j \leq n} y_i <_\ell y_j$$

Now let us verify that the conditions of Definition 6.11 hold. Let B be the sequence of blocks defined in Step 3, and assume without loss of generality that $B = \langle B_k : k < \omega \rangle$ is indiscernible and moreover is dense and codense in some indiscernible sequence B'. Let $A = \langle A_i : i < \omega \rangle$ be an indiscernible sequence of n-tuples of elements of B.

1. $\{\varphi_r(x; y_1, \dots y_n), \psi_\ell(x; y_1, \dots y_n)\}$ is contradictory because it gives rise to a pseudo-$(n+1)$-loop.
2. By construction, for any $k < \omega$, the type

$$\{\psi_\ell(x; A_j) : j \leq k\} \cup \{\varphi_r(x; A_i) : k < i\}$$

is consistent, because we have shown that $<_\ell$ linearly orders B, thus also B'. Choose the desired sequence of witnesses to be elements in the indiscernible sequence B' which are interleaved with B.
3. Suppose we have $\{\varphi_r(x; A_j), \psi_\ell(x; A_i)\}$ for some $i < j$, or in other words:

$$\{\varphi_r(x; B_{j_1}, \dots B_{j_n}), \psi_\ell(x; B_{i_1}, \dots B_{i_n})\} \text{ where } \{i_1, \dots i_n\} < \{j_1, \dots j_n\}$$

Then $x <_\ell B_{i_1} <_\ell \cdots <_\ell B_{i_n} <_\ell B_{j_1} <_\ell \cdots <_\ell B_{j_n} <_\ell x$ is a pseudo-$(2n+1)$-loop (remember that $<_\ell$ holds between any increasing pair of elements of B by construction). Thus a fortiori we have a pseudo-$(n+1)$-loop, contradicting the conclusion of Step 3.

We have shown that the theory T has SOP_3, so we finish. □

Acknowledgements

I thank my advisor Thomas Scanlon, and Leo Harrington, for many stimulating conversations. Thanks to Laci Babai for a copy of the helpful survey [3]. I am grateful to Scanlon's NSF grant for funding a trip to the ICM in Madrid where I first learned of Szemerédi's work. Finally, I would like to acknowledge the anonymous referee's thorough and thoughtful reading, which significantly improved the article's presentation.

References

[1] Gowers, “Hypergraph Regularity and the multidimensional Szemerédi Theorem.” Ann. of Math. (2) 166 (2007), no. 3, 897–946.

[2] Keisler, “Ultraproducts which are not saturated.” Journal of Symbolic Logic, 32 (1967) 23–46.

[3] Komlós and Simonovits, “Szemerédi’s Regularity Lemma and its Applications in Graph Theory,” Combinatorics, Paul Erdös is Eighty, vol. 2, Budapest (1996) 295–352.

[4] Malliaris, “Realization of φ-types and Keisler’s order,” Annals of Pure and Applied Logic 157 (2009) 220–224.

[5] Malliaris, “The characteristic sequence of a first-order formula,” to appear, Journal of Symbolic Logic.

[6] Shelah, *Classification Theory and the number of non-isomorphic models*, revised edition. North-Holland, 1990.

[7] Shelah, “Toward classifying unstable theories,” Annals of Pure and Applied Logic 80 (1996) 229–255.

[8] Shelah and Usvyatsov, “More on SOP_1 and SOP_2,” Annals of Pure and Applied Logic 155 (2008), no. 1, 16–31.

[9] Szemerédi, “On Sets of Integers Containing No k Elements in Arithmetic Progression.” Acta Arith. 27, 199-245, 1975a.

[10] Usvyatsov, “On generically stable types in dependent theories.” Journal of Symbolic Logic 74 (2009), 216-250.