THESIS WORK:
“PERSISTENCE AND REGULARITY IN UNSTABLE MODEL THEORY”
MARYANTHE MALLIARIS

Historically one of the great successes of model theory has been Shelah’s stability theory: a program, described in [17], of showing that the arrangement of first-order theories into complexity classes according to a priori set-theoretic criteria (e.g. counting types over sets) in fact pushes down to reveal a very rich and entirely model-theoretic structure theory for the classes involved: what we now call stability, superstability, and \( \omega \)-stability, as well as the dichotomy between independence and strict order in unstable theories. The success of the program may be measured by the fact that the original set-theoretic criteria are now largely passed over in favor of definitions which mention ranks or combinatorial properties of a particular formula.

Because of this shift, Keisler’s 1967 order (defined below) may strike the modern reader as an anachronism. It too seeks to coarsely classify first-order theories in terms of a more set-theoretic criterion, the difficulty of producing saturated regular ultrapowers, but its structure has remained largely open. Partial results from the 70s suggest a mine of perhaps comparable richness, one which has remained largely inaccessible to current tools.

Keisler’s criterion of choice, saturation of regular ultrapowers, is natural for two reasons. First, when the ultrapower is regular, the degree of its saturation depends only on the theory and not on the saturation of the index models. Second, ultrapowers are a natural context for studying compactness, and Keisler’s order can be thought of as studying the fine structure of compactness by asking: what families of consistent types are realized or omitted together in regular ultrapowers? Thus the relative difficulty of realizing the types of \( T_1 \) versus those of some \( T_2 \) in regular ultrapowers gives a measure of the combinatorial complexity of the types each \( T_i \) is able to describe.

**Definition 1.** (Keisler’s order [7]) \( T_1 \leq T_2 \) if for all infinite \( \lambda \), \( \mathcal{D} \) regular on \( \lambda \), \( M_1 \models T_1, M_2 \models T_2 \), we have: if \( (M_2)^\lambda / \mathcal{D} \) is \( \lambda^+ \)-saturated then \( (M_1)^\lambda / \mathcal{D} \) is \( \lambda^+ \)-saturated.
Shelah in the 1970s gave a beautiful and surprising series of results showing deep links between Keisler’s order and the underlying structure of first-order theories. His dividing lines will be familiar to model theorists who have not worked on ultrapowers:

**Theorem A.** (Shelah [17]) *In the Keisler order we have: $T_1 < T_2 < \ldots \leq T_s$, where:*

1. $T_1$ is the set of countable theories without the finite cover property, which form the minimum Keisler equivalence class.
2. $T_2$ is the set of countable theories which are stable but have fcp, which form the second Keisler equivalence class.
3. $T_s$ is the maximum class, which is known to exist and to include theories with the strict order property.
4. and the intermediate structure of the unstable $\ldots$, as well as the question of determining the boundary of the maximum class, remains open.

Notice the coarseness of the order. Stability is a classic model-theoretic frontier, but the finite cover property crosscuts all of its usual refinements. Recent work of Shelah [18] and Shelah and Usvyatsov [19] has shown that $SOP_3$, a weakening of strict order, is sufficient for maximality; however, the identity of the maximal class, as well as the structure of the order on unstable theories without $SOP_3$, has remained open.

Notice also that stability, fcp and strict order are all properties of formulas. In the first chapter of this thesis we show that this is paradigmatic: the Keisler order reduces to the study of types in a single formula ([12]). In other words, the combinatorial barriers to saturation are contained in the parameter spaces of the formulas of $T$. This mirrors the crucial move of stability theory in reducing questions of a priori infinitary combinatorics to properties of formulas. But proof itself suggests the importance of a new kind of combinatorial structure.

Namely, we associate to each formula $\varphi$ a countable sequence of hypergraphs, called the “characteristic sequence,” which describe incidence relations on the parameter space of $\varphi$. We then begin the investigation of the model-theoretic complexity of $\varphi$ in terms of the graph-theoretic complexity of its characteristic sequence, that is, the distribution and recurrence of complex configurations around the base set of a $\varphi$-type under analysis.
**Definition 2.** The characteristic sequence \( \langle P_n : n < \omega \rangle \) associated to a formula \( \varphi \) of \( T \) is given by: for \( n < \omega \), \( P_n(z_1, \ldots, z_n) := \exists x \land_{i \leq n} \varphi(x; z_i) \). Write \((T, \varphi) \mapsto \langle P_n \rangle\).

This move is a natural consequence of the localization result for ultrapowers described above. Classification theory typically isolates particular configurations which signal complexity (the order property, the independence property...); an interest in saturation of ultrapowers shifts the emphasis onto understanding how the many fragments of configurations are distributed in the parameter space of the formula and how they cluster into larger constellations, into constellations of constellations, etc. Once observed and made precise, this relation of questions of “presence” as seen in the formula \( \varphi \) to questions of “persistence” as seen in the hypergraphs is an interesting structural issue beyond the context of ultrapowers.

We apply the characteristic sequence to the analysis of consistent partial \( \varphi \)-types, which correspond to complete \( P_\infty \)-graphs, i.e. sets \( A \subseteq M \) such that \( A^n \subseteq P_n \) for all \( n \). A first goal is to definably restrict the predicate \( P_1 \) around \( A \) so that the localized graph is as “uncomplicated” as possible. A combinatorial configuration will be called *persistent* around \( A \) if it appears in every finite localization around the complete graph \( A \) under analysis. We give natural characterizations of stability and simplicity in terms of persistence.

We next restrict attention to some fixed localization and consider what the complexity of configurations there imply for \( T \). This provides a second motivation for characteristic sequences: linking classification theory for \( \varphi \) to structural issues of distributions of edges in the characteristic sequence of hypergraphs is potentially quite powerful, because as properties like edge density, randomness, and regularity of the graphs are shown to give meaningful model-theoretic information about \( \varphi \), this opens up the possibility of using a deep collection of structure theorems for graphs, for instance Szemerédi-type regularity lemmas [20], to give model-theoretic information. In the notation of [9],

**Definition 3.** ([20], [9]) Fix \( 0 < \epsilon < 1 \), and write \( \delta(X, Y) \) for the edge density \( e(X, Y)/|X||Y| \).

The finite bipartite graph \((X, Y)\) is \( \epsilon \)-regular if for every \( X' \subseteq X \), \( Y' \subseteq Y \) with \(|X'| \geq \epsilon |X|, |Y'| \geq \epsilon |Y| \), we have: \( |\delta(X, Y) - \delta(X', Y')| < \epsilon \).

**Theorem B.** (Szemerédi [20]) For every \( 0 < \epsilon < 1 \), \( m_0 \in \mathbb{N} \) there exist \( N = N(\epsilon, m_0) \), \( m = m(\epsilon, m_0) \) such that: for any graph \( X \), \(|X| \geq N \), for some \( m_0 \leq k \leq m \) there exists a partition \( X = X_1 \cup \cdots \cup X_k \) satisfying:
\[
\|X_i - X_j\| \leq 1 \text{ for } i, j \leq k
\]

- All but at most \(\epsilon k^2\) of the pairs \((X_i, X_j)\) are \(\epsilon\)-regular.

Analogous lemmas for hypergraphs exist, e.g. [5], though the issue of how to extend regularity to hypergraphs is a subtle one [6].

The organizing principle is the question of how subsets of the parameter space can generically interrelate, i.e., what densities can occur between sufficiently large \(\epsilon\)-regular pairs \(A, B \subseteq P_1\), in the sense of Szemerédi. We obtain an interesting picture. When the formula is stable, after localization the density must always be 1. In a class including simple theories, after localization the density must approach either 0 or 1. We may assume NSOP as strict order is already Keisler-maximal; with this hypothesis, we characterize the property that \(P_1\) contains large disjoint \(\epsilon\)-regular sets of any reasonable density \(\delta\) in terms of instability of \(P_2\), in the sense of model theory, and obtain several corollaries.

In a slightly more technical interlude, we observe a gap between the kind of bipartite randomness given by model theory (i.e. the independence property) and that given by Szemerédi regularity. This gap has to do with the way in which the finite subgraphs approximate the infinite. We formalize this gap and use it to describe a general principle: what might be called “the depth of independence” of an infinite \(k\)-partite graph. We show that graphs which are partially, but not fully, independent in this sense give rise to SOP\(_3\). This gives a new motivation for the property, which is known to imply maximality in the Keisler order.

We conclude with several arguments necessary to apply the technology of the characteristic sequence to the analysis of types in ultrapowers.

References


