

1

Definability in classes of finite structures

Dugald Macpherson^a and Charles Steinhorn^b

1.1 Introduction

This paper provides an overview of recent work by the authors and others on two topics in the model theory of finite structures. The point of view here differs from that usually associated with the term ‘finite model theory’, as presented for example in [21] or [46], in which the emphasis and motivation come primarily from computer science. Instead, the inspiration for this work has its origins in contemporary (infinite) model theoretic themes such as dimension, independence, and various measures of the complexity of definable sets. Each of the topics deals with classes of finite structures for first-order logic that are isolated by conditions that are drawn from these model-theoretic considerations. Moreover, in both cases, connections exist to areas in infinite model theory such as stability and simplicity theory, and o-minimality. This survey is intended for both mathematical logicians and computer scientists whose work focuses on logical aspects of the subject.

The first theme concerns *asymptotic classes* of finite structures. This subject has its origins in the model theory of finite fields, via the work of Chatzidakis, van den Dries and Macintyre [13] (see Theorem 1.2.1) and the earlier model theory of finite fields developed by Ax [4], and ultimately rests on the Lang-Weil bounds for the number of points in a finite field of an irreducible variety defined over that field. Given a first-order formula φ in the language of rings, the analysis in [13] provides estimates for the cardinality of the set defined by this formula in all fi-

^a School of Mathematics, University of Leeds, Leeds LS2 9JT England
h.d.macpherson@leeds.ac.uk

^b Vassar College, Poughkeepsie, New York 12604, USA steinhorn@vassar.edu;
partially supported by U.S. National Science Foundation Grant DMS-0070743.

nite fields in terms of two parameters, dimension and measure. Denoting the universe of a finite field by F , the cardinality estimate has the form $\mu|F|^k$, where k represents the dimension and μ the measure of the size of the set defined by φ . Asymptotic classes are, roughly speaking, classes of finite structures with a strong uniformity condition on the cardinality of definable sets that mirrors precisely that for finite fields (see [22], [25], or [48]). Indeed, finite fields, as developed in [13], provides a key example, but there are many others. Ryten [52] has shown that certain difference fields, and hence any family of finite simple groups of fixed Lie rank, fit into this framework, and Elwes [23] has established that every smoothly approximable structure admits a class of (finite) envelopes which forms an asymptotic class that witnesses smooth approximability. The uniformity properties of asymptotic classes feed through to ultraproducts of the members of any such class: they are supersimple of finite rank, with an additional ingredient, still rather mysterious, called *measure*.

The second topic, *robust classes* of finite structures, has its origins in an attempt to bring an appropriate version of o-minimality to classes of finite structures. Obstacles quickly present themselves: for example, no class of finite totally ordered structures can be an asymptotic class—see [25] or [48]. In fact, every o-minimal structure elementarily equivalent to an ultraproduct of totally ordered finite structures must be discretely ordered, and thus carry very limited structure [51]. A robust class is a directed system of finite structures *with embeddings*, such that any formula, interpreted in a structure in the class that is sufficiently large relative to the parameters appearing in the formula, assumes a constant truth value. In this setting, a non-trivial notion of o-minimality may be defined. For example, the group $(\mathbb{Q}, +, <)$, with $+$ interpreted as a ternary relation, is the direct limit of an o-minimal robust class of finite structures. The initial theory of robust classes has been developed by the authors, and in the Ph.D. thesis of Macpherson's student, R. Marshall [49]. It is our hope is that connections eventually will emerge between the two topics of this survey and current concerns of finite model theory. This seems particularly possible for robust classes; here Ehrenfeucht-Fraïssé games intervene, locality can play a role, and there are very natural, if still rudimentary, notions of complexity.

The organization of this paper is as follows. Asymptotic classes are introduced in Section 1.2, and examples are the focus of Sections 1.3, 1.4 and 1.6. Smoothly approximable structures, which provide important examples for both asymptotic classes and robust classes, are given a brief overview in Section 1.4. Section 1.5 links asymptotic classes with

contemporary infinite model theory, in particular simple theories, and introduces the notion of a measurable structure. Asymptotic classes of groups, with connections to simple theories, are treated in Section 1.6. Robust classes are introduced in Section 1.7 and Section 1.8 is devoted to examples. In the final section of the paper, Section 1.9, an ‘o-minimal’ robust approximation of the ordered group of rational numbers is presented. Notation throughout is standard; any uncommon terminology or notation is defined where it arises.

1.2 Asymptotic classes

The starting point here is the following theorem of Chatzidakis, van den Dries, and Macintyre. One considers, uniformly across finite fields, families of definable sets determined by formulas $\varphi(\bar{x}, \bar{y})$, where the \bar{y} are parameter-variables.

Theorem 1.2.1 ([13]) *Let $\varphi(\bar{x}, \bar{y})$ be a formula in the language $L_{\text{rings}} = (+, \times, -, 0, 1)$ for rings, with $\bar{x} = (x_1, \dots, x_n)$ and $\bar{y} = (y_1, \dots, y_m)$. Then there is a positive constant C , and a finite set D of pairs (d, μ) with $d \in \{0, \dots, n\}$ and μ a non-negative rational number, such that for each finite field \mathbb{F}_q and $\bar{a} \in \mathbb{F}_q^m$,*

$$|\varphi(\mathbb{F}_q^n, \bar{a})| - \mu q^d \leq Cq^{d-(1/2)} \quad (*)$$

for some $(d, \mu) \in D$.

Furthermore, for each $(d, \mu) \in D$, there is a formula $\varphi_{(d, \mu)}(\bar{x})$ which defines in each finite field \mathbb{F}_q the set of tuples \bar{a} such that $(*)$ holds.

Each such pair (d, μ) may be understood as providing a finite combinatorial version of the *dimension* d and measure μ of those definable sets to which the pair corresponds. For formulas which define absolutely irreducible varieties—without the μ —this is the result of Lang-Weil [45]. A ‘near model completeness’ result of Kiefe [40], coming rapidly out of Ax’s work [4], asserts that every formula $\varphi(\bar{x})$ is equivalent, uniformly across finite fields, to a boolean combination of formulas $\exists y g(\bar{x}, y) = 0$ where $g(\bar{X}, Y) \in \mathbb{Z}[\bar{X}, Y]$. This suggests why the above theorem should hold: one can reduce definable sets to sets built from finite-to-one projections of varieties. The details are intricate.

Theorem 1.2.1 suggests that one might consider arbitrary classes of finite structures satisfying asymptotic uniformities in the spirit of the

theorem. Note that there are possible natural weakenings of the conditions. First, one could weaken the error term $Cq^{d-(1/2)}$. There also is perhaps no reason to require μ to be rational. Additionally, and most importantly, the class of finite fields is in a sense 1-dimensional: any formula uniformly picking out an arbitrarily large subset of the field (in affine 1-space) picks out a positive fraction of the field. Thus, for example, Theorem 1.2.1 answers a question of Felgner (that in fact inspired the paper), showing that \mathbb{F}_q is not uniformly definable in \mathbb{F}_{q^2} . One could easily consider the universe F of a structure to be N -dimensional if all definable subsets of F are roughly of size $\mu|F|^{d/N}$ for $d \in \{0, 1, \dots, N\}$ and μ a constant. These considerations lead to the following definition of Elwes [22] of an *N -dimensional asymptotic class* of finite structures. For the initially considered concept, 1-dimensional asymptotic classes, see [48]. A more extensive survey of asymptotic classes than provided here, with more emphasis on the infinite limits, may be found in [25].

Definition 1.2.2 (Elwes, [23]) Let $N \in \mathbb{N}$, and let \mathcal{C} be a class of finite L -structures, where L is a finite language. Then we say that \mathcal{C} is an *N -dimensional asymptotic class* if the following hold.

(i) For every L -formula $\varphi(\bar{x}, \bar{y})$ where $l(\bar{x}) = n$ and $l(\bar{y}) = m$, there is a finite set of pairs $D \subseteq (\{0, \dots, Nn\} \times \mathbb{R}^{>0}) \cup \{(0, 0)\}$ and for each $(d, \mu) \in D$ a collection $\Phi_{(d, \mu)}$ of pairs of the form (M, \bar{a}) where $M \in \mathcal{C}$ and $\bar{a} \in M^m$, so that $\{\Phi_{(d, \mu)} : (d, \mu) \in D\}$ is a partition of $\{(M, \bar{a}) : M \in \mathcal{C}, \bar{a} \in M^m\}$, and

$$|\varphi(M^n, \bar{a})| - \mu|M|^{\frac{d}{N}} = o(|M|^{\frac{d}{N}})$$

as $|M| \rightarrow \infty$ and $(M, \bar{a}) \in \Phi_{(d, \mu)}$.

(ii) Each $\Phi_{(d, \mu)}$ is \emptyset -definable, that is to say $\{\bar{a} \in M^m : (M, \bar{a}) \in \Phi_{(d, \mu)}\}$ is uniformly \emptyset -definable across \mathcal{C} .

We may write D_φ for D in the definition above, and call $\{\Phi_{(d, \mu)} : (d, \mu) \in D\}$ a *(definable) asymptotic partition*. We define $h(\varphi(M^n, \bar{a}))$ to be the pair

$$(\text{Dim}(\varphi(M^n, \bar{a})), \text{Meas}(\varphi(M^n, \bar{a}))),$$

which equals (d, μ) if $(M, \bar{a}) \in \Phi_{(d, \mu)}$, except that if $d = \mu = 0$ we work with the convention that $\text{Dim}(\varphi(M^n, \bar{a})) = -1$.

The o -notation in (i) here means that for every $\varepsilon > 0$ there is $Q \in \mathbb{N}$ such that for all $M \in \mathcal{C}$ with $|M| > Q$ and all $\bar{a} \in M^m$, where $(M, \bar{a}) \in \Phi_{d, \mu}$, we have

$$|\varphi(M^n, \bar{a})| - \mu|M|^{\frac{d}{N}} < \varepsilon|M|^{\frac{d}{N}}.$$

We call \mathcal{C} a *weak asymptotic class* if \mathcal{C} satisfies the asymptotic criterion (i) for all φ , but the $\Phi_{(d,\mu)}$ are not assumed to be definable. We do not discuss the intermediary condition that the $\Phi_{(d,\mu)}$ are definable but not necessarily \emptyset -definable.

It is clear that clause (i) is preserved by reducts; that is, by the process of restricting the *class* of structures to the reducts in a sublanguage. In fact, by Lemma 3.7 of Elwes [22], it is preserved by parameter-interpretations which are uniform in the sense that the interpreting formulas range through a finite set. However, clause (ii) may be lost under interpretations, though it is preserved under parameter-free bi-interpretations ([22], Lemma 3.7).

In infinite model theory, the definition of o-minimality (see e.g., [20]) places a restriction on the definable sets in one variable: a totally ordered structure is *o-minimal* if every definable set is a finite union of open intervals and points. The Cell Decomposition Theorem in this context yields topological and logical finiteness properties for n -variable definable sets. By the following theorem, which might be viewed as a combinatorial cell decomposition theorem, the story is essentially the same for asymptotic classes, though we have chosen here to have the *definition* focus on the n -variable condition. The proof makes heavy use of clause (ii).

Theorem 1.2.3 (Lemma 2.1.2 of [23]; Theorem 2.1 of [48]) *Suppose that \mathcal{C} is a class of finite structures which satisfies Definition 1.2.2 (clauses (i) and (ii)) for $n = 1$, i.e. for definable sets in one variable. Then \mathcal{C} is an N -dimensional asymptotic class.*

Its proof, an induction on n , is analogous to that of the Cell Decomposition Theorem for o-minimal structures, with asymptotic calculations replacing topological arguments. For a definable subset X of M^{n+1} let $\pi : M^{n+1} \rightarrow M$ denote projection to the first coordinate. Then apply the definition to $\pi(X)$ and the inductive hypothesis to the fibers X_a for $a \in \pi(X)$. Note for every pair (d, μ) that $\{a \in M : h(X_a) = (d, \mu)\}$ is an \emptyset -definable subset of M , by clause (ii), so itself has a specified dimension and measure.

The next section provides many examples of asymptotic classes. As a paradigmatic *non-example*—see Remark 1.5.2 (d)—observe that the collection of all finite total orders is not an asymptotic class: if $\varphi(x, y)$ is the formula $x < y$, then as a ranges through a finite totally ordered structure M , $\varphi(M, a)$ is a subset of M of arbitrary size.

1.3 Examples of asymptotic classes

The most interesting of the examples below are associated with finite fields. We see it as an area of significant interest to find new classes of examples with no connection to finite fields.

Example 1.3.1 The class of all finite fields forms a 1-dimensional asymptotic class, by the main theorem of [13].

Example 1.3.2 A *difference field* is a pair (F, σ) where F is a field and σ is an automorphism of F . Fix a prime p , and positive integers m, n with $m \geq 1, n > 1$, and $(m, n) = 1$. Let $\mathcal{C}_{(m, n, p)}$ be the collection of difference fields $(\mathbb{F}_{p^{kn+m}}, \text{Frob}^k)$ where $k > 0$ and Frob^k denotes the Frobenius automorphism $x \mapsto x^p$; so m, n, p are fixed in the class, but k is varying. Then, by Theorem 3.5.8 of [52], $\mathcal{C}_{(m, n, p)}$ is a 1-dimensional asymptotic class. Note that the fixed field of $\sigma : x \mapsto x^{p^k}$ on $\mathbb{F}_{p^{kn+m}}$ is \mathbb{F}_{p^t} where $t = (k, m)$, so has bounded size. Ryten's result rests upon the main results of Hrushovski [37]: the asymptotic results for difference varieties, and the identification of the ultraproduct theory of $(\bar{\mathbb{F}}_p, x \mapsto x^{p^k})$ with the theory $ACFA_p$, the model companion of the theory of characteristic p difference fields; here $\bar{\mathbb{F}}_p$ denotes the algebraic closure of \mathbb{F}_p .

Example 1.3.3 By further results of Ryten [52, Chapter 5], every family of finite simple groups of *fixed Lie type* is an asymptotic class. The Lie type here in particular determines the Lie rank, that is the number of nodes of the associated Dynkin diagram (or orbits on nodes under the corresponding graph automorphism, in the case of twisted groups). For example, the groups $\text{PSL}_3(q)$, with q varying, form a family of finite simple groups of fixed Lie type and Lie rank 2, corresponding to the Dynkin diagram with two nodes joined by a single edge, so form an asymptotic class. The groups $\text{PSU}_3(q)$, which are subgroups of $\text{PSL}_3(q^2)$, also form an asymptotic class. Here we view the alternating group Alt_n as having Lie rank n , and these do not form an asymptotic class as n grows with the size of the group; indeed, the model theory of the finite alternating groups, like the model theory of groups $\text{PSL}_n(q)$ for fixed q and increasing n , seems to be completely wild. For undecidability of the theory of all finite symmetric groups, or for various families of $n \times n$ matrix groups with unbounded n over a fixed field, see [27], or [12, Section 6.3] for a survey. Also, see [2] for a treatment of non-standard alternating groups as objects in Peano Arithmetic. For separate reasons,

the (simple) cyclic groups of prime order also form an asymptotic class—see Example 1.3.5 below.

Most of the families of finite simple groups are *uniformly parameter bi-interpretable* (even bi-definable), in a natural sense, with finite fields (see Chapter 4 of [52]). Using results of Elwes and Ryten, it follows that the property of being an asymptotic class transfers from the fields to the groups, though care is needed with clause (ii) in Definition 1.2.2, due to the role of parameters in the interpretations.

For example, for any fixed n the groups $\mathrm{PSL}_n(q)$ and $\mathrm{PSU}_n(q)$ are both (uniformly in q) bi-interpretable with the field \mathbb{F}_q ; details can be found in Chapter 5 of [52]. The uniform interpretation of the $\mathrm{PSL}_n(q)$ in \mathbb{F}_q is almost immediate from the definition of $\mathrm{PSL}_n(q)$. More generally, for a given family of Chevalley groups of Lie type \mathbb{L} , such as the family of symplectic groups $\mathrm{PSp}_{2m}(q)$, we can uniformly in \mathbb{F}_q interpret the Lie algebra $\mathbb{L}(\mathbb{F}_q)$, hence the general linear group $\mathrm{GL}(\mathbb{L}(\mathbb{F}_q))$, and inside this the family of root subgroups which generate $\mathrm{Sp}_{2m}(q)$. Since the symplectic group is a product of a bounded number, dependent only on m , of these root subgroups, it is itself uniformly interpretable in \mathbb{F}_q , and hence so is $\mathrm{PSp}_{2m}(q)$. In the other direction, to construct the field \mathbb{F}_q inside $\mathrm{PSL}_n(q)$, the additive structure is given by a root group. The multiplicative structure of the field arises from a torus—which is conjugate to the image in $\mathrm{PSL}_n(q)$ of an appropriate diagonal subgroup of $\mathrm{SL}_n(q)$ —acting on the root group. See also [59] and [42].

For the families of Suzuki and Ree twisted simple groups, the situation is rather more complicated. The construction involves an automorphism of the Dynkin diagram which does not preserve lengths of roots. As a result, these groups are uniformly parameter bi-interpretable not with pure fields, but with difference fields. The class of Suzuki groups ${}^2B_2(2^{2k+1})$ is uniformly parameter bi-interpretable with the class $\mathcal{C}_{(1,2,2)}$, as is the class of Ree groups ${}^2F_4(2^{2k+1})$. The Ree groups ${}^2G_2(3^{2k+1})$ are uniformly parameter bi-interpretable with the members of $\mathcal{C}_{(1,2,3)}$. In these cases we apply Example 1.3.2 above.

Example 1.3.4 The families of simple groups of Lie type all arise as automorphism groups of Tits buildings. The building blocks for these are the so-called ‘rank 2 residues’, which are generalized polygons. Here, a generalized polygon is an incidence structure of points and lines such that the associated bipartite incidence graph—which has the points and lines as vertices with incidence for adjacency—has diameter n and girth $2n$. A generalized n -gon is said to be *thin* if it is an ordinary n -gon, and

is *thick* if every point (respectively line) is incident with at least three lines (respectively points). A thick generalized 3-gon is just a projective plane. The generalized polygons involved in finite simple groups satisfy an additional symmetry condition, the ‘Moufang’ property. Moufang generalized polygons have been classified by Tits and Weiss [61]. In particular, there are seven families of finite Moufang generalized polygons, each such polygon associated with its corresponding ‘little projective group.’ Dello Stritto [19] shows that each of these seven families forms an asymptotic class by proving that the polygons are uniformly parameter bi-interpretable with their corresponding little projective groups, as each corresponding class of groups forms an asymptotic class, by Example 1.3.3 above.

Example 1.3.5 By [48, Theorem 3.14], the collection of all finite cyclic groups is a 1-dimensional asymptotic class. This is hardly surprising, as the multiplicative groups of finite fields are cyclic. In general, the result follows from Szmielew’s Theorem (see for example [31, Theorem A.2.2]), which says that in every abelian group every formula $\varphi(x, \bar{y})$ is equivalent to a boolean combination of formulas of the form $p^m | t(x, \bar{y})$ or $t(x, \bar{y}) = 0$, where p is prime, t is a term, and $p^m | t(x, \bar{y})$ abbreviates $(\exists z) p^m z = t(x, \bar{y})$. By the Compactness Theorem, there is a finite family of such boolean combinations, one of which will be equivalent to φ in each abelian group. The argument then reduces to examining a conjunction of such conditions and their negations. Observe, by Theorem 1.2.3, that it suffices to consider formulas $\varphi(x, \bar{y})$, i.e., families of definable sets in *one* variable x .

Example 1.3.6 Recall that the *random graph* is the unique countably infinite graph that satisfies, for each $n > 0$, the following sentence σ_n , where Rxy denotes that vertices x and y are adjacent:

$$\forall x_1 \dots x_n, \forall y_1, \dots y_n \left[\bigwedge_{1 \leq i, j \leq n} x_i \neq y_j \right. \\ \left. \rightarrow \exists z \left(\bigwedge_{i=1}^n Rzx_i \wedge z \neq x_i \wedge \bigwedge_{i=1}^n \neg Rzy_i \wedge z \neq y_i \right) \right].$$

The *Paley graph* P_q , where q is a prime power with $q \equiv 1 \pmod{4}$, has vertex set \mathbb{F}_q and edge relation given by $x, y \in \mathbb{F}_q$ are adjacent if and only if $x - y$ is a square in \mathbb{F}_q . The collection \mathcal{C} of all Paley graphs forms a 1-dimensional asymptotic class—see [48, Example 3.4].

The essential point here is due to Bollobás and Thomason [9] (see also [10, Ch. XIII.2]). If U and W are disjoint sets of vertices in the Paley graph P_q with $|U \cup W| = m$, and $v(U, W)$ is the number of vertices of P_q not in $U \cup W$ joined to each vertex of U and none of W , then

$$|v(U, W) - 2^{-m}q| \leq \frac{1}{2}(m - 2 + 2^{-m+1})q^{\frac{1}{2}} + m/2.$$

It follows that every non-principal ultraproduct satisfies each σ_n , so is elementarily equivalent to the random graph. As the latter has quantifier elimination, to check that \mathcal{C} is an asymptotic class it suffices to consider quantifier-free formulas, which are handled by the above asymptotic estimates.

Thus the Paley graphs form a class of finite graphs whose theory approximates the random graph and witness that the random graph has the finite model property, that is, every sentence true of the random graph, and in particular the above axioms σ_n for $n \geq 1$, has a finite model and in fact hold in almost all finite graphs.

If one works instead with primes congruent to 3 mod 4, then -1 is a non-square, so the relation R defined above is antisymmetric, and one obtains a 1-dimensional asymptotic class of *Paley tournaments*, whose theory approximates the *random tournament*. Here a *tournament* is a directed graph such that any two vertices are connected by an arc. The analogue of the result of Bollobás and Thomason can be found in [29].

It would be interesting to find other asymptotic classes of graphs approximating the random graph, e.g. not corresponding to edge probability $1/2$. Work of Szönyi [57] may be relevant here.

Example 1.3.7 The random graph has, for each $k > 2$, an arity k analogue, the *countable universal homogeneous k -uniform hypergraph*. Here, a k -uniform hypergraph is just a set equipped with a collection of k -element subsets, the hyperedges. Like the random graph, its theory is axiomatized by ‘extension axioms’, which hold with probability tending to one in finite k -uniform hypergraphs. There is no naïve arity k analogue of the Paley graphs, but Beyarslan [7] has shown that the random k -uniform hypergraph is interpretable in a pseudofinite field, that is, an infinite model of the theory of finite fields. Hence, there is a family of finite k -uniform hypergraphs, uniformly interpretable in finite fields, with an ultraproduct elementarily equivalent to the random k -uniform hypergraph. This certainly yields a weak asymptotic class of finite k -uniform hypergraphs whose theory approximates the random one. So

far as we know, it has not been checked whether it is an asymptotic class, i.e., satisfies Definition 1.2.2(ii).

Example 1.3.8 Recall that a theory T is *strongly minimal* if, in all models of T , every definable subset of the domain is finite or cofinite. Suppose that \mathcal{C} is a class of finite structures such that every non-principal ultraproduct is strongly minimal. It follows rather easily, by Theorem 1.2.3, that \mathcal{C} is a 1-dimensional asymptotic class—see Example 3.9 of [48]. In particular, for every positive integer $d > 2$, the collection of all finite vertex transitive graphs of valency d is a 1-dimensional asymptotic class. Indeed, an ultraproduct is a vertex transitive graph of valency d , and all such are well-known to be strongly minimal; see for example [8], Lemma 2.2.11.

Example 1.3.9 Lastly, let M be a *smoothly approximable structure*—see Section 1.4 directly following this example for a fuller discussion of smooth approximability. Then M is a union of a chain of finite so-called ‘envelopes’. It was shown by Elwes [22] that these envelopes can be chosen to form an asymptotic class. This rests on the asymptotic information on the sizes of definable sets in [18, Proposition 5.2.2]. In particular, suppose that M is *unidimensional*, that is, up to non-orthogonality has a unique family of definable Lie geometries. Then the asymptotic bounds are much tighter than in Definition 1.2.2. Namely, one finds that if the class \mathcal{C} is N -dimensional, $E \in \mathcal{C}$, and D is a d -dimensional definable subset of measure μ of E , then for some fixed constant C depending just on M , it follows that

$$||D| - \mu|E|^{\frac{d}{N}}| < C|E|^{\frac{d}{N} - \frac{1}{N}}.$$

In the non-unidimensional case, the asymptotic behavior is not so clear—consider for example a countably infinite structure consisting of the disjoint union of two \aleph_0 -dimensional vector spaces, one over \mathbb{F}_2 and the other over \mathbb{F}_3 . It is not clear how to approximate these by a class of finite substructures so that a conclusion like that in Theorem 1.2.1 holds. This is the reason for the weaker error term, in the form $o(|M|^{\frac{d}{N}})$, in Definition 1.2.2. Smoothly approximable structures are relevant also to robust classes, and we discuss them further in the next section.

1.4 Smoothly approximable structures

The notion of smooth approximation appears to be due originally to Lachlan; it plays a role already in [15] and [16]. The first systematic investigation of smooth approximation is [39], and a deep theory was developed by Cherlin and Hrushovski in the monograph [18]. There also are two excellent survey accounts, [35] and [17].

A finite substructure N of a structure M is a k -homogeneous substructure of M if all \emptyset -definable relations on M induce \emptyset -definable relations on N , and for every pair \bar{a}, \bar{b} of k -tuples from N , they have the same type in N if and only if they have the same type in M . An \aleph_0 -categorical structure M is *smoothly approximated* if it is the union of a chain $(M_i : i \in \mathbb{N})$ of finite substructures, where for each i , M_i is an $|M_i|$ -homogeneous substructure of M .

As a very basic example, let M be an \aleph_0 -dimensional vector space over the finite field \mathbb{F}_p , and let the M_i for $i \in \mathbb{N}$ form a sequence of finite subspaces with $M_i \leq M_{i+1}$ and with $\bigcup_{i \in \mathbb{N}} M_i = M$. These structures are parsed in the language of \mathbb{F}_p -modules, that is the language for M as a group under addition, and a unary function symbol for multiplication by each element of \mathbb{F}_p . The point here is that if $\text{Dim}(M_i) = n_i$ then we have a natural sequence of embeddings $\text{GL}_{n_0}(p) \leq \text{GL}_{n_1}(p) \leq \text{GL}_{n_2}(p) \leq \dots$, and the union of this sequence of groups has the same orbits on n -tuples from M as $\text{Aut}(M)$, which is $\text{GL}_{\aleph_0}(p)$.

For a slightly more complicated example, suppose that M is endowed with a symplectic form, that is, a non-degenerate bilinear form $\beta : M \times M \rightarrow \mathbb{F}_p$ such that $\beta(v, v) = 0$ for all $v \in M$, and let the M_i form a sequence of finite (even dimensional) subspaces with union M , on each of which β induces a non-degenerate form. Here, β can be given by a family of binary relations, one for each element of \mathbb{F}_p . The fact that each M_i is an $|M_i|$ -homogeneous substructure of M is a consequence of Witt's Lemma – see for example Section 20 of [3].

These two smoothly approximable structures are both examples of *Lie geometries* (see [39] or [18]), about which we say more below. The first is totally categorical, even strongly minimal. The second is not stable (it has the independence property; see, e.g., [31]) but is supersimple of rank 1 (see Remark 1.5.2(c) for more about simple theories). We remark that the random graph is *not* smoothly approximable, even though it is also supersimple of rank 1, \aleph_0 -categorical, and arises in a natural way as a union of a chain of Paley graphs (see Example 1.8.3).

Let M be smoothly approximable. As M is \aleph_0 -categorical, the Ryll-

Nardzewski Theorem yields a function $g : \mathbb{N} \rightarrow \mathbb{N}$ such that for each $k \in \mathbb{N}$, the automorphism group of M , $\text{Aut}(M)$, has at most $g(k)$ orbits on M^k . It follows that if $(M_i : i \in \omega)$ is a sequence of finite substructures witnessing smooth approximability, then $\text{Aut}(M_i)$ has at most $g(k)$ orbits on M_i^k , for each $i, k \in \omega$. This is a very strong condition on a family of arbitrarily large finite permutation groups. In fact, *a posteriori*, the condition holds for all k if one just knows the condition for $k \leq 4$. Using finite permutation group theory—the classification of finite simple groups, the O’Nan-Scott Theorem, and Aschbacher’s structure theory for subgroups of classical groups—the authors in [39] were able to classify all smoothly approximable structures with *primitive* automorphism groups, that is, automorphism groups which preserve no proper non-trivial equivalence relation. In particular, certain building blocks, the *Lie geometries*, were identified. These include pure sets, examples like those described in the preceding paragraphs (possibly with orthogonal or unitary bilinear forms), and their projective and affine versions. There is also the ‘self-dual geometry’, which is really a pair of infinite dimensional vector spaces V, V' over a finite field \mathbb{F}_q equipped with a non-degenerate bilinear map $V \times V' \rightarrow \mathbb{F}_q$, and the ‘quadratic geometry’. The somewhat more mysterious latter is essentially the collection of all quadratic forms associated with a given symplectic form on an \aleph_0 -dimensional vector space over a finite field of characteristic 2.

A very beautiful structure theory of smoothly approximable structures is developed in [18]. We provide a brief overview.

As shown in [15], every \aleph_0 -categorical ω -stable structure—in particular every totally categorical structure, and every stable structure homogeneous over a finite relational language—is smoothly approximable. This is already quite a deep result, and includes rather complicated structures built, by a sequence of finite and affine covers, from pure sets and projective or affine spaces over finite fields. See for example [15], [1] and [32].

Smoothly approximable structures which are not ω -stable are all unstable. Indeed, any stable ω -categorical structure which is not ω -stable interprets a pseudoplane [43], and by [18, Corollary 5.5.5] a pseudoplane cannot be interpreted in any smoothly approximable structure. However, all smoothly approximable structures have a simple theory—see Remark 1.5.2(c) for basic facts about simple theories, and [62] for a general source—and in fact they are supersimple of finite rank (see Remark 1.5.2). Indeed, some crucial ideas in simplicity theory, such as the Independence Theorem, first appeared in [18] (see e.g. Section 5.1

and Proposition 8.4.3 of [18]). They are 1-based (see Section 1.5 for the definition; that they are 1-based may be found in [18]), and thus in particular, no infinite field is definable. Although finite fields play a key role in any family of finite approximating structures, any field involved remains fixed throughout the family.

All groups definable in a smoothly approximable structure are finite-by-abelian-by-finite (see [18]). As noted in Section 1.6 (see Proposition 1.6.2 ff.), extraspecial p -groups of exponent p are smoothly approximable and finite-by-abelian, but not abelian-by-finite.

Cherlin and Hrushovski define the notion of a *Lie coordinatizable structure*. This is a structure bi-interpretable with a *Lie coordinatized structure*, which is, roughly speaking, one coordinatized by a tree of finite height of Lie geometries. A key fact proved in [18] is that the Lie coordinatizable structures are exactly the smoothly approximable structures. Arguments by induction on the height of the coordinatizing tree thus are often used.

Smoothly approximable structures also are *quasi-finitely axiomatizable*. This means that the theory is axiomatized by a single sentence together with a schema of axioms saying that each ‘non-orthogonality class’ of Lie geometries (again, see [18]) is infinite dimensional. In particular, if the smoothly approximable structure is unidimensional, that is, any two interpretable Lie geometries are non-orthogonal, then the theory is axiomatized by a single sentence together with, for each n , a sentence saying that the structure has size at least n . This generalizes an earlier result of Hrushovski [32], itself extending work of Ahlbrandt and Ziegler [1], that all totally categorical structures are quasi-finitely axiomatized in this last sense.

Built into Lie coordinatizability is a theory of the *envelopes*—that is, the finite approximating substructures—of smoothly approximable structures. There are precise results in [18] on the cardinalities of definable sets in envelopes, given by certain polynomials. This is exploited in [22] to show the envelopes can be chosen to form an asymptotic class, so yields Example 1.3.9 above. It also suggests that there may be interesting strengthenings of the notion of asymptotic class, where one considers classes of finite structures in which the sizes of definable sets, with respect to a fixed formula $\varphi(\bar{x}, \bar{y})$, are given not by asymptotic conditions, but by one of finitely many polynomials. So far as we know, this has not been explored. In addition, Lachlan’s theory [44] of ‘shrinking and stretching’, developed for the class of finite structures homogeneous over

a fixed finite relational language, holds in the smoothly approximable context too.

It also follows from the theory that, over a fixed finite language L , if \mathcal{C} is a class of finite structures for which there is some k such that $\text{Aut}(M)$ has at most k orbits on M^4 for all $M \in \mathcal{C}$, then there is $g : \mathbb{N} \rightarrow \mathbb{N}$ such that for all $M \in \mathcal{C}$ and all $k \in \mathbb{N}$, $\text{Aut}(M)$ has at most $g(k)$ orbits on M^k . See for example Theorem 6 of [18].

Smooth approximability is not preserved by reducts, due to problems with the quadratic geometries mentioned above—see the example due to Evans in [18, p. 149]. Yet the class of reducts, namely the *weakly Lie coordinatizable structures*, also is fairly well understood. In particular, the class of weakly Lie coordinatizable structures is characterized by the conjunction of nine model-theoretic properties (Theorem 7 in [18]), among which are \aleph_0 -categoricity, pseudofiniteness (or the finite model property), finiteness of a certain rank, the Independence Theorem, and some more technical conditions. Intriguingly, whereas the bulk of the structure theory of [18] rests ultimately on the classification of finite simple groups, this last result does not, even though it characterizes a class of structures intimately connected to finite simple groups.

1.5 Asymptotic classes and their ultraproducts

We have already seen that investigations of asymptotic classes can be assisted by working with ultraproducts. This is already explicit in [13] on finite fields. Indeed, recall that a *pseudofinite field* is an infinite model of the theory of finite fields. Equivalently, by Ax [4] it is a field F which is perfect, quasifinite (has a unique extension of degree n for each $n > 1$), and is *pseudo-algebraically closed* (PAC) (that is, every absolutely irreducible variety defined over F has an F -rational point). It is shown in [13] that, by Theorem 1.2.1, if F is a pseudofinite field then it is possible to associate with each definable subset D of each power F^t a pair (d, μ) , where d is a nonnegative integer and $\mu \in \mathbb{Q}^{>0}$, such that d is the algebraic-geometric dimension of the Zariski closure of D , and the pairs (d, μ) satisfy certain counting conditions. This led the authors in [48] to introduce the following notion of *measurable structure*. The definition below, taken from [25], is slightly different from Definition 5.1 of [48], but equivalent.

Definition 1.5.1 An infinite L -structure M is *measurable* if there is

a function $h : \text{Def}(M) \rightarrow \mathbb{N} \times \mathbb{R} \cup \{(0, 0)\}$ (we write

$$h(X) = (\text{Dim}(X), \text{Meas}(X)) = (\text{Dim}, \text{Meas})(X))$$

such that the following hold.

1. For each L -formula $\varphi(\bar{x}, \bar{y})$ there is a finite set $D \subset \mathbb{N} \times \mathbb{R}^{>0} \cup \{(0, 0)\}$, so that for all $\bar{a} \in M^m$ we have $h(\varphi(M^n, \bar{a})) \in D$.
2. If $\varphi(M^n, \bar{a})$ is finite then $h(\varphi(M^n, \bar{a})) = (0, |\varphi(M^n, \bar{a})|)$.
3. For every L -formula $\varphi(\bar{x}, \bar{y})$ and all $(d, \mu) \in D_\varphi$, the set $\{\bar{a} \in M^m : h(\varphi(M^n, \bar{a})) = (d, \mu)\}$ is \emptyset -definable.
4. (Fubini) Let $X, Y \in \text{Def}(M)$ and $f : X \rightarrow Y$ be a definable surjection. Then there are $r \in \mathbb{N}$ and $(d_1, \mu_1), \dots, (d_r, \mu_r) \in (\mathbb{N} \times \mathbb{R}^{>0}) \cup \{(0, 0)\}$ so that if $Y_i := \{\bar{y} \in Y : h(f^{-1}(\bar{y})) = (d_i, \mu_i)\}$, then $Y = Y_1 \cup \dots \cup Y_r$ is a partition of Y into non-empty disjoint definable sets. Let $h(Y_i) = (e_i, \nu_i)$ for $i \in \{1, \dots, r\}$. Also let $c := \text{Max}\{d_1 + e_1, \dots, d_r + e_r\}$, and suppose (without loss) that this maximum is attained by $d_1 + e_1, \dots, d_s + e_s$. Then $h(X) = (c, \mu_1\nu_1 + \dots + \mu_s\nu_s)$.

If $X \in \text{Def}(M)$ and $h(X) = (d, \mu)$, we call d the *dimension* of X and μ the *measure* of X , and h the *measuring function*.

We do not emphasize measurable structures in this paper. For more information, see [48] or [25]. We do note the following observations.

Remark 1.5.2 a. If M is measurable and $N \equiv M$ then N is measurable; hence one may speak of a *measurable theory*.

b. If \mathcal{C} is an N -dimensional asymptotic class, then every non-principal ultraproduct of \mathcal{C} is measurable—essentially, the pairs (d, μ) transfer through to the ultraproduct.

c. Shelah introduced in [55] the concept of a *simple* theory (see also [41]). Simplicity is a generalization of stability in which model-theoretic non-forking still provides a good notion of independence. Indeed, it satisfies all the main properties of non-forking in stable theories except stationarity, which controls the number of non-forking extensions of a complete type. Stationarity (i.e., the Finite Equivalence Relation Theorem) is replaced by the ‘Independence Theorem’ in simple theories. The analogue of a superstable theory in stability is, in the context of simplicity, the notion of a *supersimple theory*. In a supersimple theory every definable set has an ordinal-valued ‘ D -rank’; in fact, there are several notions of rank on a definable set in a supersimple theory— D -rank, SU -rank, and S_1 -rank—all of which coincide if any of them is finite. It is shown in [25, Corollary 3.4]—but was noted earlier by Ryten—that if

M is measurable then it is supersimple, and for any definable set X in M , the D -rank of X is at most its dimension and hence is finite.

d. It follows from (b) and (c) that an ultraproduct of an asymptotic class cannot have the *strict order property*: there cannot be a definable partial order (even on a power of the structure) with an infinite chain. This generalizes the observation made at the end of Section 1.2 that the collection of finite linear orders does not form an asymptotic class.

e. There are measurable structures that are not elementarily equivalent to any ultraproduct of an asymptotic class. Vector spaces over an infinite field, in the language of modules over the field, provide one example. A more interesting example of Elwes [22, Section 3.4] consists of a structure with two different pseudofinite field structures (in disjoint languages) of different prime characteristics. It arises by taking the fixed point set of a generic automorphism of the ‘Hrushovski fusion’ [37] of two different algebraically closed fields in distinct positive characteristics. Such a structure cannot be an ultraproduct of finite structures, since no positive integer can be a power of two distinct primes.

We next discuss how some stability-theoretic notions interact with asymptotic classes. The following result enables us to detect in the 1-dimensional case, just from asymptotic information, whether or not every ultraproduct of an asymptotic class is stable.

Proposition 1.5.3 ([48]) *Let \mathcal{C} be a 1-dimensional asymptotic class. Then some ultraproduct of \mathcal{C} is unstable if and only if there is a formula $\varphi(x, \bar{y})$, and for each $k \in \mathbb{N}$ some $M \in \mathcal{C}$ and $\bar{a}_1, \dots, \bar{a}_k \in M^{\ell(\bar{y})}$ with*

- (a) $|\varphi(M, \bar{a}_i)| \geq k$ for each $i = 1, \dots, k$
- (b) $|\varphi(M, \bar{a}_i) \Delta \varphi(M, \bar{a}_j)| \geq k$ for all distinct $i, j \in \{1, \dots, k\}$.

Of the examples of asymptotic classes considered in Section 1.3, only the following have all ultraproducts stable: the class of finite cyclic groups; the asymptotic classes of Example 1.3.8 with all ultraproducts strongly minimal; in Example 1.3.9, if M is a smoothly approximable structure which is ω -categorical and ω -stable (or in particular, which is totally categorical), then every asymptotic class consisting of its envelopes.

We now recall the construction, from a complete theory T , of T^{eq} . For each $n > 0$, and each \emptyset -definable equivalence relation E on n -tuples, one adjoins a new sort interpreted, for $M \models T$, by M^n/E and a function taking each n -tuple to its corresponding E -equivalence class in the new

sort. There is a corresponding language L^{eq} , and all models M of T have corresponding expansions M^{eq} in the language L^{eq} with theory T^{eq} .

Recall also that if A is a subset of a structure M , then $b \in M$ is *algebraic* over A if there is a finite A -definable subset of M containing b . The algebraic closure of A , namely the set of elements algebraic over A , is denoted $\text{acl}(A)$.

A supersimple theory T is said to be *1-based* if, for every $M \models T$ and all subsets A and B of M^{eq} , we have that A and B are independent, in the sense of non-forking, over $\text{acl}^{\text{eq}}(A) \cap \text{acl}^{\text{eq}}(B)$. Using the main theorem of [33] Elwes [22] derived the result below, which *a fortiori* gives structural restrictions on asymptotic classes all of whose ultraproducts are stable. Note that there is an error in [22] stemming from a misunderstanding of [33]—an invalid use of compactness in [22, Lemma 6.4]. A valid argument in its place has been given by Kestner and Pillay (personal communication).

Theorem 1.5.4 *Every measurable stable theory is 1-based.*

As a very special case, note that an algebraically closed field cannot be measurable. To illustrate, for the complex field \mathbb{C} , the map $x \mapsto x^2$ is a surjection $\mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$ which is 2-to-1, contrary to Definition 1.5.1(iv). In fact, there is evidence that measurable fields must be pseudofinite. Scanlon has shown that every infinite measurable field is quasifinite, and easily, every measurable field is perfect—see [48, Theorems 5.18 and 6.1], and also [54]. It is not known if the PAC pseudo-algebraically closed property (see the first paragraph in this section) holds for all measurable fields; this would yield pseudofiniteness.

If M is a 1-dimensional measurable structure—e.g., an ultraproduct of a 1-dimensional asymptotic class—then the algebraic closure operator defines a pregeometry on subsets of M (as it is supersimple of rank 1). In particular it satisfies the *exchange property*: if $b \in \text{acl}(A \cup \{c\}) \setminus \text{acl}(A)$, then $c \in \text{acl}(A \cup \{b\})$. The exchange property can be formalized fairly concretely for 1-dimensional asymptotic classes. The definitions can be finitized, and the formula making c algebraic over $A \cup \{b\}$ can be identified up to finitely many possibilities over the given data. Indeed, suppose that \mathcal{C} is a 1-dimensional asymptotic class, $\varphi(x, \bar{y})$ is a formula, and D is the corresponding subset of $\{0, 1\} \times \mathbb{R}^{>0}$ of dimension-measure pairs as provided in Definition 1.2.2(i). Let $E := \{\mu : (1, \mu) \in D\}$ and for $\mu \in E$ let $\varphi_\mu(\bar{y})$ be a formula defining $\Phi_{(1, \mu)}$ as in Definition 1.2.2(ii). If $M \in \mathcal{C}$, $A \subset M$, and $b \in M$, we say b is in the φ -closure of A , written $b \in \text{cl}_\varphi(A)$, if there is some \bar{a} from A such that $M \models \varphi(b, \bar{a})$ and

$M \not\models \varphi_\mu(\bar{a})$ for each $\mu \in E$. For a set of formulas Σ , we say that b is in the Σ -closure of A , written $b \in \text{cl}_\Sigma(A)$, if $b \in \text{cl}_\varphi(A)$ for some $\varphi \in \Sigma$. We have:

Proposition 1.5.5 [48, Proposition 4.4] *Let \mathcal{C} be a 1-dimensional asymptotic class, $M \in \mathcal{C}$, and $A \subset M$. There are finite sets $\Delta(\varphi)$ and $\Gamma(\varphi)$ of formulas (depending on φ) such that for $a, b \in M$ with $a \in \text{cl}_\varphi(Ab) \setminus \text{cl}_{\Delta(\varphi)}(A)$, if M is ‘large enough’ relative to A and φ , then $b \in \text{cl}_{\Gamma(\varphi)}(Aa)$.*

The notion of a 1-based theory is related to the trichotomy conjecture of Zilber. One version of this conjecture, now known to be false, asserts that if M is a strongly minimal structure, then: either M is *disintegrated*, that is, $\text{acl}(A) = \bigcup(\text{acl}(a) : a \in A)$ for every $A \subset M$; or M is locally modular, equivalently, 1-based; or M interprets an infinite field. Counterexamples to this conjecture were found by Hrushovski [34], but the conjecture has remained extremely influential, and versions hold in key contexts. Furthermore, disintegrated and locally modular strongly minimal sets now are fairly well understood.

It makes sense to investigate the conjecture for 1-dimensional asymptotic classes. The following result, Proposition 4.5 of [48], gives a clear notion of a ‘disintegrated 1-dimensional asymptotic class’, identified by the asymptotic condition (iii).

Proposition 1.5.6 *Let \mathcal{C} be a 1-dimensional asymptotic class. Then the following are equivalent.*

(i) *for every formula $\varphi(x, \bar{y})$, there is a formula $\psi(x, \bar{z})$ and some $K \in \mathbb{N}$ such that if $M \in \mathcal{C}$ with $|M| > K$ and $A \subset M$, then $\text{cl}_\varphi(A) \subset \bigcup_{a \in A} \text{cl}_\psi(\{a\})$;*

(ii) *in every infinite ultraproduct M of members of \mathcal{C} , if $A \subset M$ then $\text{acl}(A) = \bigcup_{a \in A} \text{acl}(\{a\})$;*

(iii) *for every $\varphi(x, \bar{y})$ there is some $K_\varphi \in \mathbb{N}$ such that for all $M \in \mathcal{C}$ with $|M| > K_\varphi$, if $A \subset M$ then $|\text{cl}_\varphi(A)| \leq K_\varphi |A|$.*

It would be of interest to investigate the Zilber trichotomy further for asymptotic classes. Certainly, by one of the main results of [18], all smoothly approximable structures are 1-based. This includes in particular the Lie geometries. We have no idea if, in a 1-dimensional asymptotic class that is not 1-based—that is, ultraproducts of which are not 1-based, and so by Theorem 1.5.4 are unstable—arbitrarily large finite fields must be uniformly interpretable. It would be intriguing to investigate this al-

ready for classes of structures interpretable in finite fields, or even for reducts of finite fields.

1.6 Asymptotic classes of groups

As mentioned in Example 1.3.3, the work of Ryten on difference fields yields the following theorem; we emphasize again that the bound on Lie rank is essential.

Theorem 1.6.1 (Ryten) *If \mathcal{C} is any family of finite simple groups of fixed Lie type, then \mathcal{C} is an asymptotic class.*

The structure of asymptotic classes of groups is an attractive area of study. For groups in which definability is governed by definability in finite fields, or by definable subgroups of cartesian powers (as in one-based groups), one expects good control of definability. It is not clear whether ‘asymptotic class’ is the optimal model theoretic assumption on a family of finite groups; often the same conclusions can be drawn just assuming that all ultraproducts have supersimple finite rank theory. This is developed in [24].

In developing a structure theory one wants, as far as possible, not to assume the classification of finite simple groups (CFSG). It would be wonderful to recover parts of the classification just from model theoretic hypotheses. As one step in this direction, Hrushovski has shown that any family of finite simple groups uniformly definable in finite fields is a family of (possibly twisted) Lie type [36, Theorem 9.2], and the same holds for groups uniformly definable in a family of the difference fields $\mathcal{C}_{(m,n,p)}$ (see [37, Theorem 1.8]—the proof is unpublished). In the same spirit, Theorem 7.5.6 of [18] identifies, by model-theoretic hypotheses, a class of structures closely associated with finite simple groups.

As mentioned in Example 1.3.5, the class of finite cyclic groups is a 1-dimensional asymptotic class. It should be feasible to describe all asymptotic classes of abelian groups. At higher levels of complexity—but within the class of soluble groups—very little is known, though one expects unipotent and Borel subgroups of finite Chevalley groups of fixed Lie type to fall into asymptotic classes. We also mention the following ([48, Proposition 3.11]). A p -group is *extraspecial* if $G' = Z(G) = \Phi(G) \cong C_p$. In particular, extraspecial groups are nilpotent of class 2.

Proposition 1.6.2 *If p is an odd prime, then the class of finite extraspecial groups of exponent p is a 1-dimensional asymptotic class.*

In fact, finite extraspecial groups of odd exponent p are envelopes of a smoothly approximable structure, the unique countably infinite extraspecial group of exponent p . This group has cyclic center, and the quotient by the center is an elementary abelian p -group, equipped with an alternating bilinear form given by the commutator map to the center. Its theory is supersimple but unstable.

The remaining results in this section dealing with families of finite groups are obtained under weaker hypotheses than that of being an ‘asymptotic class’—see our remarks following Theorem 1.6.1. We say that a family \mathcal{C} of finite structures is *supersimple of finite rank* (respectively, *supersimple of rank n*) if all non-principal ultraproducts have these properties. The results below are all analogues of theorems about groups of finite Morley rank. Typically, the proofs use ultraproducts, and facts about measurable groups, or, more generally, groups with a supersimple finite rank theory. In some cases—Propositions 1.6.3(i), 1.6.5, 1.6.6, 1.6.8—the results really belong in some such setting, and there is no use of finiteness or pseudofiniteness.

Proposition 1.6.3 (i) [48, Theorem 3.12] *Let \mathcal{C} be a supersimple rank 1 family of finite groups. Then there is $d \in \mathbb{N}$ such that each group $G \in \mathcal{C}$ has normal subgroups H and N , where $|G : N| \leq d$, $|H| \leq d$, $H \leq Z(N)$, and N/H is abelian.*

(ii) [24] *Let \mathcal{C} be a supersimple rank 2 family of finite groups. Then there is $d \in \mathbb{N}$ such that all groups in \mathcal{C} have a normal subgroup of index at most d which is soluble of derived length at most 4.*

Neither result requires CFSG. Part (ii) was proved by Elwes and Ryten in [26] under the extra assumption that \mathcal{C} is a (2-dimensional) asymptotic class, using CFSG. The bound 4 on the derived length is probably not optimal; it should perhaps be 2, arising from the class of 1-dimensional affine groups $\text{AGL}_1(\mathbb{F}_q)$.

One key ingredient in the proof is the body of results on so-called BFC groups: that is, groups with a finite bound d on the size of all conjugacy classes. If G is such a group with bound d on the size of conjugacy classes, then there is a bound $B(d) \in \mathbb{N}$ such that $|G'| \leq B(d)$; see Wiegold [63], for example, where $B(d) = d^{\frac{1}{2}d^4(\log_2 d)^3}$ is obtained. Another element is a theorem of Schlichting, strengthened by Bergman and Lenstra [6]. It asserts that if G is a group and \mathcal{F} is a family of subgroups of G which is (setwise) invariant under a group K of automorphisms of G such that for some d and every $F_1, F_2 \in \mathcal{F}$ we have $|F_1 : F_1 \cap F_2| \leq d$, then for

some d' dependent only on d , there is a K -invariant subgroup N of G so that $|F : F \cap N| \leq d'$ and $|N : F \cap N| \leq d'$ for all $F \in \mathcal{F}$.

To establish further results, one often needs a version for supersimple theories of the Zilber Indecomposability Theorem, applied to ultraproducts. The version below follows from [62, Theorem 5.5.4].

Theorem 1.6.4 *Let G be a group definable in a supersimple structure of finite rank, and let $\{X_i : i \in I\}$ be a collection of definable subsets of G . Then there exists a definable subgroup H of G such that:*

- (i) $H \leq X_{i_1}^{\pm 1} \dots X_{i_m}^{\pm 1}$ for some $i_1, \dots, i_m \in I$;
- (ii) X_i/H is finite for each $i \in I$.

Moreover, if the collection $\{X_i : i \in I\}$ is invariant under the group K of definable automorphisms of G , then H can be chosen to be K -invariant.

Consequences of Theorem 1.6.4 include the next results. If \mathcal{C} is a family of groups, we say that the subsets X of members G of \mathcal{C} are *uniformly definable* if just finitely many formulas $\varphi(x, \bar{y})$ are required to define the sets X as G ranges through \mathcal{C} .

Proposition 1.6.5 *(from [36, Corollary 7.4]) If \mathcal{C} is a supersimple finite rank family of finite groups with no uniformly definable proper non-trivial normal subgroups, then all but finitely many of the groups in \mathcal{C} are simple.*

Proposition 1.6.6 *(from [36, Corollary 7.1]) If \mathcal{C} is a supersimple finite rank family of finite groups, then the derived subgroups of members of \mathcal{C} are uniformly definable.*

Recall that the *soluble radical* $R(G)$ of a finite group G is its largest soluble normal subgroup.

Proposition 1.6.7 *[24, Theorem 1.1] If \mathcal{C} is a supersimple finite rank family of finite groups, then the soluble radicals $R(G)$ of the groups $G \in \mathcal{C}$ are uniformly definable.*

Proposition 1.6.7 has further structural consequences. For a group G we denote its socle, the direct product of its minimal normal subgroups, by $\text{Soc}(G)$. It follows fairly easily from Proposition 1.6.7 that if \mathcal{C} is a supersimple finite rank class, then the groups $\text{Soc}(G/R(G))$ are uniformly interpretable in G as G ranges through \mathcal{C} , and are a direct product of a *bounded number* of finite simple groups of bounded Lie rank.

There are also the beginnings of a model theory for families of finite *permutation* groups. We view a permutation group model-theoretically

as a pair (X, G) , with a definable group structure on G and a definable action of G on X . By the Orbit-Stabilizer Theorem, if G is transitive on X , we may parse (X, G) as a pair (G, H) , where H is a subgroup of G , the stabilizer of some $x \in X$.

Recall that a permutation group (X, G) is *primitive* if there is no proper non-trivial G -invariant equivalence relation—that is, G -congruence—on X , equivalently, if all point stabilizers G_x for $x \in X$ are maximal subgroups of G . We say that the family \mathcal{C} of finite permutation groups (X, G) is *definably primitive* if, for each non-principal ultraproduct (X^*, G^*) , there is no definable proper non-trivial G^* -congruence on X^* .

Proposition 1.6.8 ([26]) *If \mathcal{C} is a supersimple finite rank class of definably primitive finite permutation groups, then all but finitely many of the permutation groups in \mathcal{C} are primitive.*

Theorem 1.6.9 ([24]) *Let \mathcal{C} be a supersimple finite rank family of finite primitive permutation groups, and suppose that for every ultraproduct (X^*, G^*) , $\text{rk}(X^*) = 1$. Then one of the following holds for ultraproducts (X^*, G^*) .*

(i) $\text{rk}(G^*) = 1$, G^* acts regularly on X^* , and G^* is elementary abelian or torsion-free divisible abelian.

(ii) $\text{rk}(G^*) = 2$, and there is an interpretable pseudofinite field F such that $G \leq \text{AGL}_1(F)$ (the one-dimensional affine group $(F, +).(F, \times)$) in the natural action on F .

(iii) $\text{rk}(G^*) = 3$, there is an interpretable pseudofinite field F , and $\text{PSL}_2(F) \leq G \leq \text{PGL}_2(F)$ in the natural action on the projective line $\text{PG}_1(F)$.

In [47] a structure theory is given for families of finite permutation groups all of whose non-principal ultraproducts are primitive. We do not give details here. One feature concerns families of primitive permutation groups (X, G) such that $\text{Soc}(G)$ is a non-abelian simple group of fixed Lie rank. Except in very specific cases (essentially where the point stabilizers are bounded, or are ‘subfield subgroups’ associated with subfields for which the field extension degree is unbounded) families of this type have primitive ultraproducts. The proof uses much of the above work of Elwes and Ryten, Theorem 1.6.4, and also knowledge of maximal subgroups of finite simple groups. In particular, we have the following result. It generalizes [38, Proposition 8.1], which is over prime fields, but unlike the latter, it makes heavy use of the classification of finite simple groups.

Theorem 1.6.10 ([47]) *Let Chev be a fixed Lie type (possibly twisted) of finite simple groups, and d a positive integer. Let \mathcal{C} be a family of pairs (G, H) where $G = \text{Chev}(q)$ and H is a maximal subgroup of G , and suppose that if $H = \text{Chev}(q_0)$ then the degree $[\mathbb{F}_q : \mathbb{F}_{q_0}] \leq d$. Then the groups H are uniformly definable in the groups G , and \mathcal{C} is an asymptotic class.*

1.7 Robust classes

Robust classes consist of chains of finite structures in which the truth value of every formula, with parameters from some structure in the chain, eventually stabilizes when the formula is interpreted in a sufficiently larger structure. That is, one must “look ahead” in the chain to determine satisfaction of a formula with parameters in some structure in the chain. This framework has provided a setting in which to investigate notions of stability and o-minimality, as well as some provisional versions of complexity of such a chain. In addition to investigating these topics, the interest has mainly been in finding examples.

We begin with the definition of a robust class, and some initial observations, working over an arbitrary—and, unless otherwise specified, finite—first-order language L . Examples are discussed in the next section, and a robust approximation to the ordered additive group of rational numbers is established in Section 1.9.

Definition 1.7.1 A sequence $\mathcal{C} = (M_i : i \in \mathbb{N})$ of finite L -structures forms a *chain* of structures if $M_i \subseteq M_{i+1}$ (as a substructure) for each $i \in \mathbb{N}$. A chain \mathcal{C} is said to be *robust* if for each $n \in \mathbb{N}$ and L -formula $\varphi(x_1, \dots, x_n)$ there is a function $f = f_\varphi : \mathbb{N} \rightarrow \mathbb{N}$ such that for each $i \in \mathbb{N}$ and $a_1, \dots, a_n \in M_i$, and for every $j \geq f(i)$, we have

$$M_{f(i)} \models \varphi(\bar{a}) \text{ if and only if } M_j \models \varphi(\bar{a}).$$

If \mathcal{C} is robust and \bar{a} and M_i are as above, we write $\mathcal{C} \models_{\text{ev}} \varphi(\bar{a})$ if $M_j \models \varphi(\bar{a})$ for all sufficiently large j .

Remark 1.7.2 (a) In what follows, we always assume that f_φ grows as slowly as possible subject to witnessing robustness. Thus, f_φ is uniquely determined by \mathcal{C} .

(b) There is a natural generalization of Definition 1.7.1, where we replace the chain \mathcal{C} by a directed system of finite structures, equipped with specified embeddings. This is our original context, but as nothing

has yet been done with the greater generality, we here use the more concrete version in Definition 1.7.1.

Initial results on robust classes have been obtained over the last few years by the authors (first published here) and in the Ph.D. thesis of the first author's student, Richard Marshall [49]. The study of robust classes is still in its early stages and thus not yet fully systematic. For example, one could envisage a theory of 'asymptotic robust classes', in which the asymptotic estimates described in Definition 1.2.2 are required to hold only when a formula is interpreted in a sufficiently larger structure, but this has not been considered.

Observe that Definition 1.7.1 applies in particular to *sentences*. Thus, if $\mathcal{C} = (M_i : i \in \mathbb{N})$ is a robust class, then for every L -sentence σ , precisely one of σ or $\neg\sigma$ holds in cofinitely many M_i . If \mathcal{C} is an arbitrary chain of finite structures, we define the *asymptotic theory* $T_{\mathcal{C}}^{\text{as}}$ to be the collection of L -sentences true in cofinitely many $M_i \in \mathcal{C}$. If \mathcal{C} is robust then this is a complete theory. Let M be the (countable) direct limit structure of the chain \mathcal{C} . The *limit theory* $T_{\mathcal{C}}^{\text{lim}}$ of \mathcal{C} is defined to be $\text{Th}(M)$. In general, even assuming robustness, we do not expect $T_{\mathcal{C}}^{\text{as}} = T_{\mathcal{C}}^{\text{lim}}$ (see Propositions 1.7.4 and 1.8.2, and the comment following 1.8.2).

Recall that a first-order structure M is *locally finite* if every finite subset of M is contained in a finite substructure of M . Given a countably infinite locally finite structure M , it is natural to ask if it is possible to construct a robust chain with direct limit M . The next result demonstrates that we must refine this initial question to investigate conditions under which we can obtain an *explicitly described* robust chain. This question can take several forms, a theme that we explore throughout the rest of this paper. If a chain of finite structures \mathcal{C}' is a subsequence of a chain \mathcal{C} , we call it a *coarsening* of \mathcal{C} . Obviously, any coarsening of a robust chain is robust.

Proposition 1.7.3 [49, Theorem 2.4.5]

(i) Let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be a chain of finite L -structures. Then there is a coarsening of \mathcal{C} which is robust.

(ii) Every countably infinite locally finite structure is the direct limit of a robust chain.

Proof (i) One systematically, for each structure in the chain and choice of parameters in the structure, replaces the sequence of larger structures by an infinite subsequence in which the formula takes an eventually constant truth value.

(ii) Immediate from (i). \square

Recall that a theory T is *near model complete* if every formula is equivalent modulo T to a boolean combination of existential formulas. Many familiar theories are near model complete; in particular, every model complete theory is near model complete. The theory of pseudofinite fields is near model complete by [40] (see also [14, Section 3]), and so is any complete theory of abelian groups, or, more generally, of modules. Several ‘Hrushovski constructions’ have near model complete but not model complete theories; see for example Baldwin and Shelah [5].

Proposition 1.7.4 *Let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be a chain of finite structures.*

(i) *If $T_{\mathcal{C}}^{\text{as}}$ is near model complete then \mathcal{C} is robust.*

(ii) *If $T_{\mathcal{C}}^{\text{as}}$ is $\forall\exists$ -axiomatized then $T_{\mathcal{C}}^{\text{lim}} \models T_{\mathcal{C}}^{\text{as}}$.*

(iii) *If $T_{\mathcal{C}}^{\text{as}}$ is $\forall\exists$ -axiomatized and complete, then $T_{\mathcal{C}}^{\text{lim}} = T_{\mathcal{C}}^{\text{as}}$.*

Proof (i) Let M be the direct limit of \mathcal{C} . For every tuple \bar{a} from M and existential formula $\varphi(\bar{x})$, we have that $M \models \varphi(\bar{a})$ if and only if $M_i \models \varphi(\bar{a})$ for sufficiently large $i \in \mathbb{N}$. For every formula $\psi(\bar{x})$ there is, modulo $T_{\mathcal{C}}^{\text{as}}$, a formula $\theta(\bar{x})$, which is a boolean combination of existential formulas, and a sentence $\sigma \in T_{\mathcal{C}}^{\text{as}}$, such that $\sigma \models (\forall \bar{x}) \psi(\bar{x}) \leftrightarrow \theta(\bar{x})$. As there is some $N \in \mathbb{N}$ such that $M_i \models \sigma$ for all $i \geq N$, it follows for sufficiently large i that $M_i \models \psi(\bar{x})$ if and only if $M \models \theta(\bar{x})$. Hence \mathcal{C} is robust.

(ii) Each of the $\forall\exists$ -axioms of $T_{\mathcal{C}}^{\text{as}}$ holds in sufficiently large M_i , thus hold in the union M , and hence so do their consequences.

(iii) This is immediate from (ii). \square

Recall that a theory T has the *finite submodel property* if for every $M \models T$ and $\sigma \in T$, there is a finite substructure of M satisfying σ . An extension of the arguments above yields the following.

Proposition 1.7.5 *[49, Theorem 2.4.13] Let L be a finite language and let T be a complete L -theory that is near model complete, $\forall\exists$ -axiomatized, and has the finite submodel property. Then there is a robust chain \mathcal{C} such that $T_{\mathcal{C}}^{\text{as}} = T_{\mathcal{C}}^{\text{lim}} = T$.*

One of the original aims behind the introduction of robust classes is to develop a framework in which a (countable) stable or o-minimal structure might be approximated by a chain of finite structures that reflects these properties. To this end, we propose the following analogues of stability and o-minimality.

Definition 1.7.6 Let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be a robust chain of L -structures with limit M .

(i) The L -formula $\varphi(x_1, \dots, x_m, y_1, \dots, y_n)$ is *unstable in \mathcal{C}* if for all $t \in \mathbb{N}$ there are $\bar{a}_1, \dots, \bar{a}_t \in M^m$ and $\bar{b}_1, \dots, \bar{b}_t \in M^n$ such that for all $i, j \leq t$ we have

$$\mathcal{C} \models_{\text{ev}} \varphi(\bar{a}_i, \bar{b}_j) \Leftrightarrow i \leq j.$$

(ii) The chain \mathcal{C} is *unstable* if and only if some formula is unstable in \mathcal{C} .

(iii) The chain \mathcal{C} is *strongly minimal* if for every formula $\varphi(x, y_1, \dots, y_n)$ there is $n_\varphi \in \mathbb{N}$ such that for all $\bar{a} \in M^n$ either $|\{x \in M : \mathcal{C} \models_{\text{ev}} \varphi(x, \bar{a})\}| \leq n_\varphi$, or $|\{x \in M : \mathcal{C} \models_{\text{ev}} \neg\varphi(x, \bar{a})\}| \leq n_\varphi$.

(iv) Assume that L contains a binary relation $<$ that totally orders M . Then \mathcal{C} is said to be *o-minimal* if for every formula $\varphi(x, y_1, \dots, y_n)$ there is an $n_\varphi \in \mathbb{N}$ such that for all $\bar{a} \in M^n$ we have $\{x \in M : \mathcal{C} \models_{\text{ev}} \varphi(x, \bar{a})\}$ is the union of at most n_φ singletons and open intervals of $(M, <)$.

Proposition 1.7.7 Let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be a robust chain with limit M .

- (i) [49, 4.2.12] For a formula $\varphi(\bar{x}, \bar{y})$, if φ is unstable in \mathcal{C} then φ is unstable in $T_{\mathcal{C}}^{\text{as}}$.
- (ii) If $T_{\mathcal{C}}^{\text{as}}$ is near model complete, then:
 - (a) [49, 4.2.4] a formula φ is stable in \mathcal{C} if φ is stable in $T_{\mathcal{C}}^{\text{lim}}$;
 - (b) \mathcal{C} is strongly minimal if $T_{\mathcal{C}}^{\text{lim}}$ is strongly minimal;
 - (c) if M is totally ordered by $<$, then \mathcal{C} is o-minimal if M is o-minimal.
- (iii) If $T_{\mathcal{C}}^{\text{lim}}$ is near model complete, then the converses to (ii)(a)–(c) hold.

Proof (i) If φ is unstable, then for all t the following sentence holds eventually in \mathcal{C} , and hence belongs to $T_{\mathcal{C}}^{\text{as}}$:

$$\exists \bar{x}_1, \dots, \bar{x}_t \exists \bar{y}_1 \dots \bar{y}_t \left[\bigwedge_{i < j} \varphi(\bar{x}_i, \bar{y}_j) \wedge \bigwedge_{i \geq j} \neg\varphi(\bar{x}_i, \bar{y}_j) \right].$$

(ii) In each case, this follows from the observation that every formula is equivalent, in sufficiently large members of \mathcal{C} , to a boolean combination of existential formulas, and the latter holds eventually in \mathcal{C} if and only if it holds in M . (It is important to note that we *do not claim* that $\mathcal{C} \models_{\text{ev}} \varphi \Leftrightarrow M \models \varphi$ for every formula φ ; indeed, this already may fail for sentences.)

(iii) The argument is similar to that for (ii). \square

The beginnings of a stability theory for robust chains are explored by Marshall in [49, Chapter 4]. For example, a version of Shelah’s φ -rank (for a formula ψ) is defined, and it is shown that for a robust class \mathcal{C} , some formula $\varphi(\bar{x}, \bar{y})$ is unstable in \mathcal{C} if and only if the φ -rank of $\bar{x} = \bar{x}$ is infinite in this sense. An interesting feature here is that the infinitary arguments of Shelah [55, II.2] seem to be unavailable, and the combinatorial result of [30] is used instead. Marshall shows [49, 4.5.19] that if \mathcal{C} is a strongly minimal robust class with $T_{\mathcal{C}}^{\text{as}}$ model complete, then \mathcal{C} is stable, but this has not been proved without the model completeness assumption. An initial theory of Morley rank is also developed. Versions of the independence property and the strict order property for robust classes are defined, and each implies that \mathcal{C} is unstable. Conversely, under the assumption that $T_{\mathcal{C}}^{\text{as}}$ is near model complete and $\forall\exists$ -axiomatized, Marshall proves that \mathcal{C} has either the independence property or the strict order property.

The “look ahead” aspect of satisfaction in a robust chain suggests the introduction of various notions of complexity. Partly adapting [49, Definition 6.2.1], we make the following provisional definitions.

Definition 1.7.8 Let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be a robust chain.

- a. We say \mathcal{C} has *chain complexity* 0 if for every formula $\varphi(\bar{x})$ there is n_φ such that $f_\varphi(i) \leq \text{Max}\{i, n_\varphi\}$. The chain \mathcal{C} has *linear* (respectively, *polynomial*) *chain complexity*, if, for each formula φ , the function f_φ is bounded above by a linear (respectively, polynomial) function.
- b. We say that \mathcal{C} has *linear* (respectively, *polynomial*, *exponential*) *model growth* if the function $i \mapsto |M_i|$ is bounded above by a linear (respectively, polynomial, exponential) function.
- c. The chain \mathcal{C} has *polynomial satisfaction complexity* if for every formula $\varphi(\bar{x})$ there are constants $C > 0$ and $d \in \mathbb{N}$ so that $|M_{f_\varphi(i)}| \leq C|M_i|^d$ for all for all $M_i \in \mathcal{C}$.

Note that chain complexity and satisfaction complexity can be defined for each formula separately. Observe also that Proposition 1.7.3 suggests that chain complexity and model growth play off against each other. Chain complexity and model growth have been explored in [49]; satisfaction complexity is newer and not yet well-explored. Note also that chain complexity 0 implies even linear satisfaction complexity. Satisfaction complexity can be refined by specifying how the constants C

and d depend on φ ; for example one could demand that C be a recursive or even polynomial function of φ or $|\varphi|$.

As a corollary to Proposition 1.7.5, we have

Corollary 1.7.9 [49, 6.2.3] *Let T be a complete theory with quantifier elimination and the finite submodel property. Then there is a robust chain \mathcal{C} with $T_{\mathcal{C}}^{\text{as}} = T_{\mathcal{C}}^{\text{lim}} = T$, and every such chain has chain complexity 0.*

Proof The first assertion is just Proposition 1.7.5. For the second, let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be such a chain. For every formula $\varphi(\bar{x})$ there is a quantifier-free formula $\psi(\bar{x})$ such that T contains the sentence $\sigma := (\forall \bar{x}) \varphi(\bar{x}) \leftrightarrow \psi(\bar{x})$. Choose n_φ least such that $M_i \models \sigma$ for all $i \geq n_\varphi$. Since for all $i > j \geq n_\varphi$ and \bar{a} in M_j we have $M_j \models \psi(\bar{a}) \Leftrightarrow M_i \models \psi(\bar{a})$, it follows that $M_j \models \varphi(\bar{a}) \Leftrightarrow M_i \models \varphi(\bar{a})$. \square

As an extension of 1.7.9, we give a syntactic characterization of robust classes of chain complexity 0. If $L \subset L^+$ are languages and $\mathcal{C} = (M_i : i \in \mathbb{N})$ is a chain of finite L -structures, we say that $\mathcal{C}^+ = (M_i^+ : i \in \mathbb{N})$ is an L^+ -*expansion of \mathcal{C}* if each M_i^+ is an L^+ expansion of M_i and \mathcal{C}^+ is a chain, that is, M_i^+ is an L^+ -substructure of M_{i+1}^+ for each $i \in \mathbb{N}$.

Proposition 1.7.10 *Let $\mathcal{C} = (M_i : i \in \omega)$ be a chain of finite L -structures. The following are equivalent.*

- (i) \mathcal{C} is robust with chain complexity 0.
- (ii) There is a language $L^+ \supset L$ and an expansion \mathcal{C}^+ of \mathcal{C} to L^+ such that $T_{\mathcal{C}^+}^{\text{as}}$ is model complete.

Proof (ii) \Rightarrow (i). Suppose that \mathcal{C} and \mathcal{C}^+ are as in (ii). Let $\varphi(\bar{x})$ be an L -formula. Then there are formulas $\exists \bar{y} \psi(\bar{x}, \bar{y})$ and $\forall \bar{z} \chi(\bar{x}, \bar{z})$, where ψ and χ are quantifier-free L^+ -formulas, such that $T_{\mathcal{C}^+}^{\text{as}}$ contains both $(\forall \bar{x}) \varphi(\bar{x}) \leftrightarrow \exists \bar{y} \psi(\bar{x}, \bar{y})$ and $(\forall \bar{x}) \varphi(\bar{x}) \leftrightarrow \forall \bar{z} \chi(\bar{x}, \bar{z})$. Hence there is $n_\varphi \in \mathbb{N}$ such that each of these sentences holds in M_i^+ for $i \geq n_\varphi$. Let $j > i \geq n_\varphi$ and $\bar{a} \in M_i^{l(\bar{x})}$. If $M_i \models \varphi(\bar{a})$, then there is $\bar{b} \in M_i^{l(\bar{y})}$ such that $M_i \models \psi(\bar{a}, \bar{b})$, and thus, as ψ is quantifier-free, $M_j^+ \models \psi(\bar{a}, \bar{b})$. Hence $M_j \models \varphi(\bar{a})$. If $M_i \models \neg \varphi(\bar{a})$ we argue similarly, using χ .

(i) \Rightarrow (ii). We ‘Morleyize’, uniformly. That is, we expand L to L^+ by introducing, for each L -formula $\varphi(\bar{x})$, a relation symbol $R_\varphi(\bar{x})$, and let

$$T^+ := T_{\mathcal{C}}^{\text{as}} \cup \{(\forall \bar{x}) \varphi(\bar{x}) \leftrightarrow R_\varphi(\bar{x}) : \varphi \text{ an } L\text{-formula}\}.$$

Then it follows from robustness of \mathcal{C} that T^+ is a complete L^+ -theory.

We expand \mathcal{C} to L^+ as follows. For each new relation symbol R_φ , let

n_φ be chosen least so that for all $j > i \geq n_\varphi$ and $\bar{a} \in M_i^{l(\bar{x})}$, we have that $M_i \models \varphi(\bar{a})$ if and only if $M_j \models \varphi(\bar{a})$; such an n_φ exists as \mathcal{C} is robust with chain complexity 0. For $i \geq n_\varphi$, interpret R_φ in M_i by φ and for $i < n_\varphi$ interpret R_φ as the relation induced by the interpretation of R_φ in M_{n_φ} . Then \mathcal{C}^+ is a chain of L^+ -structures, and its asymptotic theory contains T^+ , so is model-complete. \square

1.8 Examples of robust classes

We present several examples of robust classes here and in the next section. The emphasis here is two-fold: to provide examples that illustrate the properties introduced in Section 1.7, and to produce classes with various properties that have prescribed limit structures.

Proposition 1.8.1 *Let M be a smoothly approximable structure, approximated by a chain $\mathcal{C} = (M_i : i \in \omega)$ of finite substructures, where M_i is an $|M_i|$ -homogeneous substructure of M . Then*

- (i) \mathcal{C} is robust with chain complexity 0.
- (ii) The chain \mathcal{C} approximating M can be chosen to have exponential model growth.

Proof (i) We may suppose that $\text{Th}(M)$ admits quantifier elimination by adding a new relation symbol for each formula. Now $\text{Th}(M)$ is model complete, and every sentence in $\text{Th}(M)$ holds in M_i for all sufficiently large i . Since every expansion of M by finitely many constants also is smoothly approximable, the same holds for formulas, that is, \mathcal{C} is robust. By Corollary 1.7.9, \mathcal{C} has chain complexity 0.

(ii) The fact that \mathcal{C} can be chosen to have exponential model growth follows from [18, 5.2.2]. \square

Example 1.8.2 Let L be the language of rings, and in this language let $\mathcal{C} = (M_i : i \in \mathbb{N})$ be any chain of finite fields of characteristic p with union $\bar{\mathbb{F}}_p$, the algebraic closure of \mathbb{F}_p . The theory $T_{\mathcal{C}}^{\text{as}}$ includes the theory of pseudofinite fields, so is near model complete. By Propositions 1.7.4 and 1.7.7 it follows that \mathcal{C} is a strongly minimal robust chain (and is stable).

Note here that $T_{\mathcal{C}}^{\text{as}} \neq T_{\mathcal{C}}^{\text{lim}}$, as $T_{\mathcal{C}}^{\text{as}}$ has the independence property and is thus unstable. The point essentially is that formulas—or sentences—which are boolean combinations of existential formulas hold in the direct limit of \mathcal{C} if and only if they hold eventually in \mathcal{C} , but this is not true for

arbitrary sentences, in particular the axioms of the theory of pseudofinite fields.

Example 1.8.3 Let $p \equiv 1 \pmod{4}$ be prime and M_i be the field $\mathbb{F}_{p^{2^i}}$ for $i \in \mathbb{N}$. Let P_i be the Paley graph on M_i (see Example 1.3.6 for the definition), and put $\mathcal{C} := (P_i : i \in \mathbb{N})$. Then $T_{\mathcal{C}}^{\text{as}}$ is the theory of the random graph, so has quantifier elimination and the finite submodel property. Thus, \mathcal{C} is robust of chain complexity 0, with the random graph as its direct limit.

By probabilistic arguments, it is possible to realize the random graph, or its arity k analogue, as the direct limit of a chain complexity 0 robust chain $(M_i : i \in \mathbb{N})$, with model growth given by the identity function—i.e., $|M_i| = i$ for all i . Indeed, by [60, Theorem 3.2], if Γ_k denotes the universal homogeneous k -uniform hypergraph, then we may write Γ_k as the union of a chain $(M_i : i \in \mathbb{N})$ of finite substructures such that $|M_i| = i$ for each i , and for every sentence σ , if $\Gamma_k \models \sigma$ then $M_i \models \sigma$ for all but finitely many i . This suffices, by quantifier elimination.

Example 1.8.4 [49, Section 3.5] Using the invariants for the elementary theories of Boolean algebras (in the language $(\vee, \wedge, -, 0, 1)$), due to Tarski [58], as described in [11, p.288], it can be shown that the theory of pseudofinite Boolean algebras is complete and near model complete, although not model complete. Hence, any chain $(M_i : i \in \mathbb{N})$ of finite Boolean algebras is robust by 1.7.4. We may obtain the countable atomless Boolean algebra as a direct limit by choosing the embeddings $M_i \rightarrow M_{i+1}$ carefully, for example, putting $M_{i+1} := M_i \times M_i$ for each i , with the diagonal embedding $M_i \rightarrow M_{i+1}$ given by $a \mapsto (a, a)$.

Example 1.8.5 [49, Section 3.4.1] The analysis of theories of abelian groups based on the Szmielw invariants yields that any complete theory of abelian groups is near model complete (see e.g., [31, p.663]). Hence, if \mathcal{C} is a chain of finite abelian groups such that $T_{\mathcal{C}}^{\text{as}}$ is complete, then \mathcal{C} is robust. Since theories of abelian groups are stable, it follows from Proposition 1.7.7 that in this case \mathcal{C} is stable.

Example 1.8.6 [49, Section 3.2] The theory of pseudofinite total orders is complete and near model complete. This can be proved, for example, by an Ehrenfeucht-Fraïssé game argument. It follows that any infinite chain of finite total orders is robust.

It is easy to build such a chain so that it is o-minimal and has direct limit $(\mathbb{Q}, <)$. Let M_n be the natural total order on $\{1, 2, \dots, 2n + 1\}$,

and embed M_n into M_{n+1} via the map $i \mapsto 2i$. The model growth of \mathcal{C} is clearly linear.

Of course, any other countable total order can be realized as the direct limit of a robust chain by Proposition 1.7.3, and $(\mathbb{N}, <)$ and $(\mathbb{Z}, <)$ each is the direct limit of an o-minimal robust chain, by Lemma 1.7.7.

We conclude this section with an example which is closer in spirit to finite model theory, in that it uses Gaifman's Locality Theorem and locality arguments.

Example 1.8.7 Let L be a finite relational language, and let M be a countably infinite L -structure of *finite valency*, in the sense that every $a \in M$ lies in just finitely many tuples satisfying relations of L . Suppose in addition that the automorphism group of M , $\text{Aut}(M)$, is transitive on M . Thus, there is $d \in \mathbb{N}$ such that all elements of M have valency d in the sense above. This assumption is for convenience, and can surely be weakened. At any rate, Cayley graphs of finitely generated groups provide a rich source of examples.

There is a natural notion of distance in M : for $a, b \in M$, we write $d(a, b) = r$ if r is least such that there is a sequence $a = a_0, a_1, \dots, a_r = b$ such that each pair a_i, a_{i+1} lie in a tuple satisfying a relation in M . We further suppose that M is connected, in the sense that $d(a, b)$ is finite for all $a, b \in M$. Lastly, for each $n \in \mathbb{N}$ and $a \in M$, let $S_n(a)$ be $\{x \in M : d(a, x) \leq n\}$ with the L -structure induced from M .

Fix $a \in M$. We recursively construct finite substructures M_n of M as follows. Let $M_1 := S_1(a)$. If M_n has been defined, let M_{n+1} be the union of $M_n \cup S_{n+1}(a)$ and a disjoint isomorphic copy of M_n that is not connected to $M_n \cup S_{n+1}(a)$. Note that each M_n has at least n connected components. Let $\mathcal{C} := (M_n : n > 0)$. Then the direct limit of \mathcal{C} is isomorphic to M .

We claim that $T_{\mathcal{C}}^{\text{as}}$ is complete and near model complete. From this it follows by Proposition 1.7.4 that \mathcal{C} is robust.

We use the presentation of Gaifman's Locality Theorem from [21, Section 2.5]. For an L -structure N and $\bar{a} = (a_1, \dots, a_n)$ from N , let $S_k(\bar{a}) := S_k(a_1) \cup \dots \cup S_k(a_n)$. We first recall that for every L -formula $\varphi(\bar{x})$ and $k \in \mathbb{N}$ one can associate a formula $\varphi^{S_k(\bar{x})}$, called a *local formula*, such that for every L -structure N and \bar{a} in N ,

$$N \models \varphi^{S_k(\bar{x})}(\bar{a}) \text{ if and only if } S_k(\bar{a}) \models \varphi(\bar{a}).$$

A *basic local sentence* has the form

$$\exists x_1 \dots \exists x_n \bigwedge_{1 \leq i < j \leq n} d(x_i, x_j) > 2r \wedge \varphi^{S_r(x_i)}(x_i).$$

Gaifman's Theorem asserts that every first-order L -sentence is logically equivalent to a boolean combination of basic local sentences.

Since M_{n+1} includes some connected components whose union U_n is isomorphic to M_n , it is clear that if σ is a basic local sentence and $M_n \models \sigma$, then $M_{n+1} \models \sigma$; the witnesses for the existential quantifiers of σ in M_n have copies in U_n that witness σ in M_{n+1} . Thus, every basic local sentence is eventually true or eventually false in \mathcal{C} . The completeness of $T_{\mathcal{C}}^{\text{as}}$ follows.

It remains to verify near model completeness. An extension of Gaifman's Theorem (see [28]) asserts that every *formula* $\varphi(\bar{x})$ is logically equivalent to a boolean combination of local formulas and basic local sentences. A local formula $\varphi^{S_r(\bar{x})}$ is a boolean combination of formulas which describe possible atomic diagrams of $S_r(\bar{x})$. Since there is a fixed upper bound on valency in substructures of M , the formula $\varphi^{S_r(\bar{x})}$ can itself be assumed to be a boolean combination of existential formulas. This yields near model completeness.

Note that if M has valency d , then it can be checked that $|M_{n+1}| \leq (d+2)|M_n|$ for all n . It follows that $|M_n| \leq (d+2)^n$ for each n . In particular, \mathcal{C} has exponential model growth. We have not attempted to minimize growth and chain complexity for a chain with limit M .

1.9 A robust approximation of $(\mathbb{Q}, <, +)$.

Our aim is to show that an o-minimal robust class can support some algebraic structure, despite the discreteness of finite total orders, and thus approximate an infinite o-minimal structure on which at least one algebraic operation is defined. The goal here is to construct an o-minimal robust class whose direct limit is $(\mathbb{Q}, <, +)$. The dense ordering can easily be constructed by dovetailing embeddings as in Example 1.8.6; the difficulty lies in incorporating the group operation. Since ordered groups are torsion-free, hence not locally finite, the group operation must be given by a ternary relation symbol.

Theorem 1.9.1 *Let $L = \{<, 0, -, R\}$, where $-$ is a unary function, and R is a ternary relation. Then there is an o-minimal robust class \mathcal{C} such that $T_{\mathcal{C}}^{\text{as}}$ is near model complete, and $T_{\mathcal{C}}^{\text{lim}}$ is the theory of divisible*

ordered abelian groups, with the addition function parsed as a ternary relation R .

Proof For each $n > 0$, let H_n be the L -structure on $\{-n, -n+1, \dots, 0, \dots, n\}$, with $0, <, +$ interpreted naturally, and such that $R(x, y, z)$ holds for $x, y, z \in H_n$ if and only if $x + y = z$. Let $\mathcal{H} = \{H_n : n > 0\}$.

A sequence $(a_k)_{k>0}$ of natural numbers is called *legal* if

- a. $(\forall k > 0) 0 \leq a_k < k$
- b. $(\forall k > 0)(\exists \ell > k)(\forall r \leq k) \ell \equiv a_r \pmod{r}$.

It is not difficult to show that legal sequences exist. For example, one may argue by induction employing the fact (see, e.g., [50, Section 2.3, Exercise 23]) that for arbitrary integers m_1, \dots, m_r the system

$$x \equiv a_1 \pmod{m_1}, \dots, x \equiv a_r \pmod{m_r}$$

has a solution if and only if $a_i \equiv a_j \pmod{(m_i, m_j)}$ for all $1 \leq i < j \leq r$. Note also that the constant sequence $(0)_{k>0}$ is legal. Let \mathcal{S} be the set of all legal sequences.

Let $S = (a_k)_{k>0} \in \mathcal{S}$. From S , we can obtain a sequence $(n_k)_{k>0}$ such that $n_{k+1} \gg n_k^2$ that further satisfies the condition that

$$(\forall r)(\exists \ell)(\forall k \geq \ell)(\forall s \leq r) n_k \equiv a_s \pmod{s}.$$

Let $\mathcal{C}_S := \{H_{n_k} : k > 0\}$. To make \mathcal{C}_S into a chain, we must define how H_{n_k} is embedded into $H_{n_{k+1}}$ for all k . To this end, let $(d_k)_{k>0}$ be a sequence of integers so that $2 \leq d_k \leq n_k$ for all k and in which every $m \geq 2$ appears infinitely often. Then we embed H_{n_k} into $H_{n_{k+1}}$ via the mapping $i \mapsto d_k i$. The intent of this is to ensure that the direct limit structure of \mathcal{C}_S is isomorphic to $(\mathbb{Q}, <, +)$, with $x + y = z$ parsed as the relation $R(x, y, z)$. Finally, let T_S^{as} be the asymptotic theory of \mathcal{C}_S , and T_S^{lim} be the limit theory. Observe that if S, S' are distinct legal sequences, then $T_S^{\text{as}} \neq T_{S'}^{\text{as}}$.

We shall prove that each theory T_S^{as} is near model complete. It follows by Proposition 1.7.4 that \mathcal{C}_S is robust and, as \mathcal{C}_S is totally ordered, that T_S^{as} is complete. Lastly, we prove that \mathcal{C}_S is o-minimal.

For each $M \in \mathcal{C}_S$, we define the *span* of M , $\text{span}(M)$, and the *hull* of M , $\text{hull}(M)$. In fact, $\text{span}(M)$ is isomorphic to $(\mathbb{Z}, <, R, 0)$, and $\text{hull}(M)$ is isomorphic to $(\mathbb{Q}, <, R, 0)$ —that is, the divisible hull of $(\mathbb{Z}, <, R, 0)$ —with R interpreted as the graph of the group operation. In this proof, however, we shall view $\text{span}(M)$ and $\text{hull}(M)$ as many sorted structures, with infinitely many sorts, to ensure that the span and hull of every model of T_S^{as} (and its substructures) is well-defined.

First, $0 \in \text{span}(M)$ and each positive element of $\text{span}(M)$ has, for some $r > 0$, a representative of the form (x_1, \dots, x_r) , where $x_1, \dots, x_r \in M^{>0}$. For intuition, the reader may interpret (x_1, \dots, x_r) as the sum $x_1 + \dots + x_r$. Recursively on $r + s$, we define $(x_1, \dots, x_r)E(y_1, \dots, y_s)$ as the symmetric closure of the relation defined to hold if $x_r \geq y_s$ and $(x_1, \dots, x_{r-1}, x_r - y_s)E(y_1, \dots, y_{s-1})$. For each $x_1, \dots, x_r \in M^{>0}$ as above, $\text{span}(M)$ also contains an element $-(x_1, \dots, x_r)$ with representative $(-x_1, \dots, -x_r)$. Addition is defined on E -classes by putting

$$(x_1, \dots, x_r)/E + (y_1, \dots, y_s)/E = (x_1, \dots, x_r, y_1, \dots, y_s)/E$$

and

$$-(x_1, \dots, x_r)/E + -(y_1, \dots, y_s)/E = -(x_1, \dots, x_r, y_1, \dots, y_s)/E.$$

Lastly, if $x_1, \dots, x_r > 0$ and $y_1, \dots, y_s < 0$, then $(x_1, \dots, x_r)/E + (y_1, \dots, y_s)/E = (z_1, \dots, z_t)/E$, where $z_1, \dots, z_t > 0$, if and only if $(x_1, \dots, x_r)/E = (-y_1, \dots, -y_s)/E + (z_1, \dots, z_t)/E$.

It is convenient to extend the definitions to sequences (x_1, \dots, x_r) where some x_i are positive and some negative. First, for all permutations $\pi \in \text{Sym}(r)$ we define $(x_1, \dots, x_r)E(x_{\pi(1)}, \dots, x_{\pi(r)})$. Then if $x_1, \dots, x_r > 0$ and $x_{r+1}, \dots, x_s < 0$, define $(x_1, \dots, x_s)E(y_1, \dots, y_t)$, where $y_i > 0$ for all i , to hold if $(x_1, \dots, x_r)/E = -(x_{r+1}, \dots, x_s)/E + (y_1, \dots, y_t)/E$.

The group $\text{span}(M)$ is an ordered group, with the ordering defined sortwise by $(x_1, \dots, x_r)/E > 0$ if and only if there are $y_1, \dots, y_s > 0$ with $(x_1, \dots, x_r)E(y_1, \dots, y_s)$. It is now easily checked that $\text{span}(M)$ is a \mathbb{Z} -group. The key point is that for every $n > 1$ and $x_1, \dots, x_r > 0$, there are $i \in \{0, \dots, n-1\}$ and $y_1, \dots, y_{r'} > 0$ such that $n(y_1, \dots, y_{r'}) = (x_1, \dots, x_{r-1}, x_r - i)$. Also M clearly is convex in $\text{span}(M)$.

Now we define $\text{hull}(M)$ to be the divisible hull of $\text{span}(M)$. Since the above definitions are uniform across \mathcal{C}_S , they carry across to models M of T_S^{as} : that is, we may talk of $\text{span}(M)$ and $\text{hull}(M)$, defined as above. As the definitions are quantifier-free, they apply also to substructures.

Model-theoretically, we remark that this construction may be understood as analogous to that of the algebraic closure of a field F : each finite extension of F is interpretable in F , but the full algebraic closure lives on the union of infinitely many sorts of F^{eq} .

The key element in the proof of the theorem is the following assertion.

Claim 1.9.2 *Let $M, N \models T_S^{\text{as}}$, and suppose that M, N have a common*

\mathcal{L} -substructure A . Then $M \equiv_A N$ if and only if $M \cap \text{hull}(A) \cong_A N \cap \text{hull}(A)$.

Proof The left-to-right direction is immediate, since elements of $\text{hull}(A)$ are quantifier-free definable over A .

For the right-to-left direction, we assume that $M \cap \text{hull}(A) = N \cap \text{hull}(A)$ and that M and N are saturated of the same cardinality $> |A|$. We build an isomorphism $\varphi : M \rightarrow N$ that extends $\varphi_0 := \text{id}_A$. We further suppose that φ_0 is extended to $\text{id}_{\text{hull}(A)}$.

The construction of the isomorphism proceeds through a series of steps.

Step 1 Let $1_M := \min\{x \in M : x > 0\}$ and $1_N := \min\{x \in N : x > 0\}$. Define $\varphi(1_M) = 1_N$. Then extend φ to $\text{hull}(A \cup \{1_M\})$.

Step 2 Define $\varphi(\max(M)) = \max(N)$, and extend φ to $\text{hull}(A \cup \{\max(M)\})$.

In Step 1, we must check for all $x \in \text{hull}(A \cup 1_M)$ that $x \in M$ if and only if $\varphi(x) \in N$; likewise, in Step 2, with $\max(M)$ in place of 1_M . This is done carefully in the last paragraph of the argument in Step 3 below, so we omit the details in this case. Note that it is needed here that $\max(M)$ and $\max(N)$ satisfy the same congruence conditions, which holds as $M, N \models T_S^{\text{as}}$.

Once Steps 1 and 2 have been completed, we may suppose that initially $1_M = 1_N \in \text{hull}(A)$, which we denote by 1 in what follows, and that $\max(M) = \max(N) \in \text{hull}(A)$. The remaining task is the following:

Step 3 We must extend φ to $x \in M \setminus \text{dom}(\varphi)$.

Put $C := \{y \in \text{hull}(A) : y < x\}$ and $D := \{y \in \text{hull}(A) : x < y\}$. We first assert that for each $z \in C$ there is some $a_z \in \text{span}(A)$ with $z \leq a_z < x$. Indeed, using the fact that for every n and every $w \in \text{span}(M)$, among the elements $w, w + 1, \dots, w + (n - 1)$ there is an element that is divisible by n , it is easy to see that if the assertion were false then we must have $z < x < z + 1$. With $z = z'/q$, where $z' \in \text{span}(A)$ and $q \in \mathbb{Z} \setminus \{0\}$, it follows that $z' < qx < z' + q$. Then $z', z' + q \in \text{span}(A) \subseteq \text{dom}(\varphi)$, whence, as $\text{span}(M)$ is a \mathbb{Z} -group, $qx \in \text{dom}(\varphi)$. It follows that $x \in \text{dom}(\varphi)$, a contradiction. A similar argument shows that for each $z \in D$ there is some $b_z \in \text{span}(A)$ with $x < b_z \leq z$.

Now let $C' := \{a_z : z \in C\}$ and $D' := \{b_z : z \in D\}$. So $C' < x < D'$, and $C', D' \subset \text{span}(A) \cap M$. The preceding paragraph shows that C' has no greatest element and D' no least element. By saturation of N we see

that there is some $y \in N$ such that $C' < y < D'$ and $x \equiv y \pmod{k}$ for each positive integer k . We put $\varphi(x) = y$, and extend φ to $\text{hull}(A \cup \{x\})$.

A typical element of $\text{hull}(A \cup \{x\})$ has the form $\frac{1}{q}(e_1 + \dots + e_r + nx)$, where $q \in \mathbb{Z}$ and $q \neq 0$, n is a positive integer, and $e_1, \dots, e_r \in A$ (not necessarily distinct). It remains to check that $\frac{1}{q}(e_1 + \dots + e_r + nx) \in M$ if and only if $\frac{1}{q}(e_1 + \dots + e_r + ny) \in N$. Observe first that both $e_1 + \dots + e_r + nx \in \text{span}(M)$ and $e_1 + \dots + e_r + ny \in \text{span}(N)$, and that $e_1 + \dots + e_r + nx > \max(M)$ if and only if $e_1 + \dots + e_r + ny > \max(N)$. Thus, it suffices to see that $e_1 + \dots + e_r + nx \equiv 0 \pmod{q}$ in M if and only if $e_1 + \dots + e_r + ny \equiv 0 \pmod{q}$ in N . This holds since x and y satisfy the same congruences. With this, the proof of the claim is complete. \square

We now finish the proof of the theorem. First, we assert that T_S^{as} is near model complete. By Claim 1.9.2, it follows for all $\bar{a} \in M \models T_S^{\text{as}}$ and substructures A of M that $\text{tp}(\bar{a}/A)$ is determined by the collection of formulas which describe the quantifier-free type of \bar{a} over $\text{hull}(A)$. Such formulas are boolean combinations of existential formulas over A . From this and Proposition 1.7.4, it follows that \mathcal{C}_S is robust. As \mathcal{C}_S is totally ordered, we also see that T_S^{as} is complete.

It remains only to check that \mathcal{C}_S is o-minimal. Let $G := \lim \mathcal{C}_S$, with $x + y = z$ parsed as $R(x, y, z)$. It is evident that G is isomorphic to $(\mathbb{Q}, <, +)$, and hence is o-minimal. The o-minimality of \mathcal{C}_S then follows immediately from the near model completeness of T_S^{as} and Proposition 1.7.7. \square

Remark 1.9.3 By varying the embeddings, we may find, for any $S \in \mathcal{S}$ other robust classes \mathcal{C}' with the same asymptotic theory T_S^{as} but different limits theories. For example, we may realize the infinite cyclic group \mathbb{Z} as $\lim \mathcal{C}'$, by putting $\mathcal{C}' = \mathcal{C}$, equipped with the identity embeddings.

We expect that a variation of this argument should allow us to construct $(\mathbb{Q}, +)$ as the limit of a strongly minimal robust class. It is an open problem to construct an o-minimal robust chain whose limit is a real closed field.

References

- [1] G. Ahlbrandt, M. Ziegler, ‘Quasifinitely axiomatizable totally categorical theories’, *Ann. Pure Appl. Logic* 30 (1986), 63–82.
- [2] J. Allsup, R. W. Kaye, ‘Normal subgroups of non-standard symmetric and alternating groups’, *Arch. Math. Logic* 46 (2007), 107–121.
- [3] M. Aschbacher, *Finite Group Theory*, 2nd Ed., Cambridge University Press, Cambridge, 2000.
- [4] J. Ax, ‘The elementary theory of finite fields’, *Ann. Math.* 88 (1968), 239–271.
- [5] J.T. Baldwin, S. Shelah, ‘Randomness and semigenericity’, *Trans. Amer. Math. Soc.* 349 (1997), 1359–1376.
- [6] G. Bergman, H.W. Lenstra, ‘Subgroups close to normal subgroups’, *J. Algebra* 127 (1989), 80–97.
- [7] O. Beyarslan, ‘Random Hypergraphs in Pseudofinite Fields’, *J. Inst. Math. Jussieu* 9 (2010), 29–47.
- [8] T. Blossier, ‘Ensembles minimaux localement modulaires’, Ph.D. thesis, Université Paris VII, 2001.
- [9] B. Bollobás, A. Thomason, ‘Graphs which contain all small graphs’, *Europ. J. Comb.* 2 (1981), 13–15.
- [10] B. Bollobás, *Random Graphs*, Academic Press, New York, 1985.
- [11] S. Koppelberg, *Handbook of Boolean Algebras* Vol. 1 (Eds. J.D. Monk, R. Bonnet), North-Holland, 1989.
- [12] E.I. Bunina, A.V. Mikhalev, ‘Elementary properties of linear groups and related problems’, *J. Math. Sci.* 123 (2004), 3921–3985.
- [13] Z. Chatzidakis, L. van den Dries, A.J. Macintyre, ‘Definable sets over finite fields’, *J. Reine Angew. Math.* 427 (1992), 107–135.
- [14] Z. Chatzidakis, ‘Model theory of finite and pseudo-finite fields’, *Ann. Pure Appl. Logic* 88 (1997), 95–108.
- [15] G. Cherlin, L. Harrington, and A.H. Lachlan, ‘ \aleph_0 -categorical, \aleph_0 -stable structures’, *Ann. Pure Appl. Logic* 28 (1985), 103–135.
- [16] G. Cherlin, A.H. Lachlan, ‘Stable finitely homogeneous structures’, *Trans. Amer. Math. Soc.* 296 (1986), 815–850.

- [17] G. Cherlin, ‘Large finite structures with few types’, in *Algebraic model theory* (Eds. B. Hart, A.H. Lachlan, M. Valeriote), NATO ASI Series C, vol. 496, Kluwer, Dordrecht, 1997, pp. 53–105.
- [18] G. Cherlin, E. Hrushovski, *Finite structures with few types*, Annals of Mathematics Studies No. 152, Princeton University Press, Princeton, 2003.
- [19] P. Dello Stritto, ‘Asymptotic classes of finite Moufang polygons’, preprint.
- [20] L. van den Dries, *Tame topology and o-minimal structures*, London Math. Soc Lecture Notes 248, Cambridge University Press, 1998.
- [21] H-D. Ebbinghaus, J. Flum, *Finite model theory*, Springer, 1999.
- [22] R. Elwes, ‘Asymptotic classes of finite structures’, *J. Symb. Logic* 72 (2007), 418–438.
- [23] R. Elwes, *Dimension and measure in finite first order structures*, Ph.D. thesis, University of Leeds, 2005.
- [24] R. Elwes, E. Jaligot, H.D. Macpherson, M. Ryten, ‘Groups in supersimple and pseudofinite theories’, preprint.
- [25] R. Elwes, H.D. Macpherson, ‘A survey of asymptotic classes and measurable structures’, in *Model theory with Applications to Algebra and Analysis*, vol. 2 (Eds. Z. Chatzidakis, H.D. Macpherson, A. Pillay, A.J. Wilkie), London Math. Soc. Lecture Notes no. 350, 2008, pp. 125–159.
- [26] R. Elwes, M. Ryten, ‘Measurable groups of low dimension’, *Math. Logic Quarterly* 54 (2008), 374–386.
- [27] Y.L. Ershov, ‘Undecidability of theories of symmetric and finite simple groups’, *Dokl. Akad. Nauk SSSR* 158 (1964), 777–779.
- [28] H. Gaifman, ‘On local and non-local properties’, *Proceedings of the Herbrand Symposium, Logic Colloquium ’81* (Ed. J. Stern), Studies in Logic and the Foundation of Mathematics, vol. 107. Amsterdam: North Holland, 1982, 105–135.
- [29] R. L. Graham, J. H. Spencer, ‘A constructive solution to a tournament problem’, *Canad. Math. Bull.* 14 (1971), 45–48.
- [30] W. Hodges, ‘Encoding orders and trees in binary relations’, *Mathematika* 28 (1981), 67–71
- [31] W. Hodges, *Model Theory*, Cambridge University Press, Cambridge, 1993.
- [32] E. Hrushovski, ‘Totally categorical structures’, *Trans. Amer. Math. Soc.* 313 (1989), 131–159.
- [33] E. Hrushovski, ‘Unimodular minimal structures’, *J. London Math. Soc.* (2) 46 (1992), 385–396.
- [34] E. Hrushovski, ‘A new strongly minimal set’, *Ann. Pure Appl. Logic* 62 (1993), 147–166.
- [35] E. Hrushovski, ‘Finite structures with few types’, in *Finite and infinite combinatorics in sets and logic* (Eds. N.W. Sauer, R.E. Woodrow, B. Sands), NATO ASI Series C vol. 411, Kluwer, Dordrecht, 1993, pp. 175–187.

- [36] E. Hrushovski, ‘Pseudofinite fields and related structures’, in *Model theory and applications* (Eds. L. Bélair, Z. Chatzidakis, P. D’Aquino, D. Marker, M. Otero, F. Point, A. Wilkie), Quaderni di Matematica, vol. 11, Caserta, 2005, 151–212.
- [37] E. Hrushovski, ‘The first order theory of the Frobenius’, preprint, available at <http://front.math.ucdavis.edu/math.LO/0406514>.
- [38] E. Hrushovski, A. Pillay, ‘Definable subgroups of algebraic groups over finite fields’, *J. Reine Angew. Math.* 462 (1995), 69–91.
- [39] W.M. Kantor, M.W. Liebeck, H.D. Macpherson, ‘ \aleph_0 -categorical structures smoothly approximated by finite substructures’, *Proc. London Math. Soc.* (3) 59 (1989), 439–463.
- [40] C. Kiefe, ‘Sets definable over finite fields: their zeta functions’, *Trans. Amer. Math. Soc.* 223 (1976), 45–59.
- [41] B. Kim, A. Pillay, ‘Simple theories’, *Ann. Pure Appl. Logic* 88 (1997), 149–164.
- [42] L. Kramer, G. Röhrle, K. Tent, ‘Defining k in $G(k)$ ’, *J. Alg.* 216 (1999), 77–85.
- [43] A.H. Lachlan, ‘Two conjectures regarding the stability of ω -categorical theories’, *Fund. Math.* 81 (1974), 133–145.
- [44] A.H. Lachlan, ‘On countable stable structures which are homogeneous for a finite relational language’, *Israel J. Math.* 49 (1984), 69–153.
- [45] S. Lang, A. Weil, ‘Number of points of varieties in finite fields’, *Am. J. Math.* 76 (1954), 819–827.
- [46] L. Libkin, *Elements of Finite Model Theory*, Springer, Heidelberg, 2004.
- [47] M.W. Liebeck, H.D. Macpherson, K. Tent, ‘Primitive permutation groups of bounded orbital diameter’, *Proc. London Math. Soc.* (3) 100 (2010), 216–248.
- [48] H.D. Macpherson, C. Steinhorn, ‘One-dimensional asymptotic classes of finite structures’, *Trans. Amer. Math. Soc.* 360 (2008), 411–448.
- [49] R. Marshall, *Robust classes of finite structures*, Ph.D thesis, University of Leeds, 2008.
- [50] I. Niven, H. Zuckerman, H. Montgomery, *An Introduction to the Theory of Numbers* (Fifth Edition), John Wiley & Sons, 1991.
- [51] A. Pillay and C. Steinhorn, ‘On discrete o-minimal structures’, *J. Pure and Applied Logic* 34 (1987), 275–289.
- [52] M. Ryten, *Results around asymptotic and measurable groups*, Ph.D. thesis, University of Leeds, 2007.
- [53] M. Ryten, I. Tomašić, ‘ACFA and measurability’, *Selecta Math.* (New series) 11 (2005), 523–537.
- [54] T. Scanlon, ‘Fields admitting nontrivial strong ordered Euler characteristics are quasifinite’, unpublished manuscript, available at <http://math.berkeley.edu/~scanlon/papers/papers.html>.
- [55] S. Shelah, ‘Simple unstable theories’, *Ann. Pure Appl. Logic* 19 (1980), 177–203.
- [56] W. Szmielew, ‘Elementary properties of abelian groups’, *Fund. Math.* 41 (1955), 203–271.

- [57] T. Szönyi, ‘Some applications of algebraic curves in finite geometry and combinatorics’, in *Surveys in Combinatorics* (Ed. R. Bailey), London Math. Society Lecture Notes 241, Cambridge University Press, Cambridge, 1997, pps. 197–236.
- [58] A. Tarski, ‘Arithmetical classes and types of boolean algebras’, *Bull. of the A.M.S.* 55 (1949), 63.
- [59] S. Thomas, *Classification theory of simple locally finite groups*, Ph.D. thesis, University of London, 1983.
- [60] S. Thomas, ‘Reducts of random hypergraphs’, *Ann. Pure Appl. Logic* 80 (1996), 165–193.
- [61] J. Tits, R. M. Weiss, *Moufang polygons*, Springer, 2002.
- [62] F.O. Wagner, *Simple theories*, Kluwer, Dordrecht, 2000.
- [63] J. Wiegold, ‘Groups with boundedly many finite classes of conjugate elements’, *Proc. Royal Soc. A* 238 (1956), 389–401.