Submitted to the *Journal of Symbolic Logic* on 10th May 2018

MULTIDIMENSIONAL EXACT CLASSES, SMOOTH APPROXIMATION AND BOUNDED 4-TYPES

DANIEL WOLF

Dedicated to the memories of my son Arthur and my mother Valerie

Abstract. In connection with the work of Anscombe, Macpherson, Steinhorn and the present author in [1] we investigate the notion of a multidimensional exact class (R-mec), a special kind of multidimensional asymptotic class (R-mac) with measuring functions that yield the exact sizes of definable sets, not just approximations. We use results about smooth approximation [24] and Lie coordinatisation [14] to prove the following result (Theorem 4.6.4), as conjectured by Macpherson: For any countable language $\mathcal{L}$ and any positive integer d the class $\mathcal{C}(\mathcal{L}, d)$ of all finite $\mathcal{L}$-structures with at most d 4-types is a polynomial exact class in $\mathcal{L}$, where a polynomial exact class is a multidimensional exact class with polynomial measuring functions.

§1. Introduction. The model-theoretic notion of an asymptotic class was introduced by Macpherson and Steinhorn in [36] as a generalisation of the result in [8] of Chatzidakis, van den Dries and Macintyre regarding the size of definable sets in finite fields. This notion has been further generalised by Anscombe, Macpherson, Steinhorn and the present author in [1] and [43] to that of a multidimensional asymptotic class, also known as an R-mac. We will not go over the history of the development of these notions; details can be found in § 1.1 of [43].

In the present work we focus on multidimensional exact classes, also known as R-mecs, which are a special kind of multidimensional asymptotic class where the measuring functions yield the exact sizes of definable sets, not just approximations. We show that multidimensional exact classes and smooth approximation (in the sense of [24]) are intimately related by proving that every smoothly approximable structure gives rise to a muldimensional exact class (Proposition 3.2.1). Using the framework of Lie coordinatisation, as developed by Cherlin and Hrushovski in [14], we then build on Proposition 3.2.1 to prove the main result of this paper, as conjectured by Macpherson:

Main result (Theorem 4.6.4). *For any countable language $\mathcal{L}$ and any positive integer d the class $\mathcal{C}(\mathcal{L}, d)$ of all finite $\mathcal{L}$-structures with at most d 4-types is a polynomial exact class in $\mathcal{L}$, where a polynomial exact class is a multidimensional exact class with polynomial measuring functions.*

2010 *Mathematics Subject Classification.* 03C13, 03C20.

Key words and phrases. Asymptotic class, smooth approximation, Lie coordinatisation.

The author (born Daniel Wood) was funded by the Leeds School of Mathematics through a Graduate Teaching Assistantship. The present work is taken largely from his PhD thesis [43].

We outline the structure of the present work. In §2 we state the definition of an R-mec (and an R-mac), prove some technical lemmas and provide some examples and non-examples. § 3 is about smooth approximation and is where we prove the aforementioned Proposition 3.2.1. In § 4 we move on to Lie coordinatisation, which we use to prove the main result Theorem 4.6.4.

We make extensive use of the Ryll-Nardzewski Theorem throughout this paper. This is well covered in the literature, for example § 1.3 of [18], Theorem 7.3.1 in [21], Theorem 5.1 in [25] and Theorem 4.3.2 in [41]. We refer to [38] and [41] for general model-theoretic notation and terminology.

§2. Multidimensional exact classes. We introduce the central definition of this paper, state and prove some handy lemmas in § 2.2 and then provide some (non-)examples in § 2.3.

2.1. Basic definitions. Let $\mathcal{L}$ be a finitary, first-order language and let $\mathcal{C}$ be a class of finite $\mathcal{L}$-structures. For $m \in \mathbb{N}^+$ define

$$\mathcal{C}(m) := \{(\mathcal{M}, \bar{a}) : \mathcal{M} \in \mathcal{C}, \bar{a} \in M^m\}.$$

The elements of $\mathcal{C}(m)$ are sometimes referred to as *pointed structures*. For completeness we further define $\mathcal{C}(0) := \mathcal{C}$.

Definition 2.1.1 (Definable partition). Let Φ be a partition of $\mathcal{C}(m)$. An element $\pi \in \Phi$ is *definable* if there exists a parameter-free $\mathcal{L}$-formula $\psi(\bar{y})$ with $l(\bar{y}) = m$ such that for every $(\mathcal{M}, \bar{a}) \in \mathcal{C}(m)$ we have $(\mathcal{M}, \bar{a}) \in \pi$ if and only if $\mathcal{M} \models \psi(\bar{a})$. The partition Φ is *definable* if π is definable for every $\pi \in \Phi$.

Definition 2.1.2 (R-mec). Let $\mathcal{C}$ be a class of finite $\mathcal{L}$-structures and let R be a set of functions from $\mathcal{C}$ to $\mathbb{N}$. Then $\mathcal{C}$ is a *multidimensional exact class for R in $\mathcal{L}$*, or *R-mec in $\mathcal{L}$* for short, if for every parameter-free $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$, where $n := l(\bar{x})$ and $m := l(\bar{y})$, there exists a finite definable partition Φ of $\mathcal{C}(m)$ such that for each $\pi \in \Phi$ there exists $h_\pi \in R$ such that

$$|\varphi(\mathcal{M}^n, \bar{a})| = h_\pi(\mathcal{M}) \tag{2.1}$$

for all $(\mathcal{M}, \bar{a}) \in \pi$.

We make some initial observations and remarks on terminology:

Remark 2.1.3.

(i) We call the functions h_π the *measuring functions* and the $\mathcal{L}$-formulas that define the partition Φ the *defining $\mathcal{L}$-formulas*. We often refer to multidimensional exact classes simply as *exact classes*.

(ii) R must be closed under addition and multiplication: If A and B are definable sets, then their disjoint union $A \sqcup B$ is definable and has size $|A| + |B|$ and their cartesian product $A \times B$ is definable and has size $|A| \cdot |B|$. Also note that R must contain a constant function $\mathcal{M} \mapsto k$ for each $k \in \mathbb{N}$, since one can always define a set of any given fixed finite size. We will thus often explcitly state only the generating functions when we describe the elements of a given R. Note that we are assuming here that the structures in $\mathcal{C}$ are arbitrarily large; cf. Corollary 2.2.4.

(iii) If we drop the requirement that the partition Φ be definable, then we call $\mathcal{C}$ a *weak R-mec*. We call (2.1) the *size clause* and the requirement that the partition be definable the *definability clause*. So a weak R-mec need satisfy only the size clause. We sometimes use the term *full R-mec* to emphasise that both the size and definability clauses hold and the term *strictly weak R-mec* to emphasise that only the size clause holds.

(iv) In the $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$ it is important to maintain the distinction between the variables $\bar{x}$ and the variables $\bar{y}$. (Although we use the plural *variables*, either of $\bar{x}$ and $\bar{y}$ could denote a single variable.) The variables $\bar{x}$, which we call *object variables*, are slots for solutions in each $\mathcal{M} \in \mathcal{C}$. The variables $\bar{y}$, which we call *parameter variables*, are slots for parameters from each $\mathcal{M} \in \mathcal{C}$. To aid clarity we sometimes demarcate the two kinds of variables with a semicolon, writing $\varphi(\bar{x}; \bar{y})$.

(v) R-mecs are closed under taking subclasses of $\mathcal{C}$ and supersets of R: If $\mathcal{C}$ is an R-mec in $\mathcal{L}$, then any subclass of $\mathcal{C}$ is also an R'-mec in $\mathcal{L}$ for any superset $R' \supseteq R$. Equivalently: If $\mathcal{C}$ is not an R-mec in $\mathcal{L}$, then no superclass of $\mathcal{C}$ is an R'-mec in $\mathcal{L}$ for any subset $R' \subseteq R$.

Weak R-mecs are closed under taking reducts: Let $\mathcal{C}$ be a weak R-mec in $\mathcal{L}$ and consider some $\mathcal{L}' \subseteq \mathcal{L}$. For $\mathcal{M} \in \mathcal{C}$, let $\mathcal{M}'$ denote the reduct of $\mathcal{M}$ to $\mathcal{L}'$. Then $\{\mathcal{M}' : \mathcal{M} \in \mathcal{C}\}$ is a weak R-mec in $\mathcal{L}'$. Equivalently: Suppose that $\mathcal{C}$ is not a weak R-mec in $\mathcal{L}$ and consider some $\mathcal{L}' \supseteq \mathcal{L}$. For $\mathcal{M} \in \mathcal{C}$, let $\mathcal{M}'$ be an expansion of $\mathcal{M}$ to $\mathcal{L}'$. Then $\{\mathcal{M}' : \mathcal{M} \in \mathcal{C}\}$ is not a weak R-mec in $\mathcal{L}'$. Note that we can't remove the prefix 'weak' here, since taking reducts may affect the definability clause.

We finish this subsection by defining the notion of a multidimensional asymptotic class, which we have already made reference to:

DEFINITION 2.1.4 (R-mac). Let R be a set of functions from $\mathcal{C}$ to $\mathbb{R}^{\geq 0}$ closed under addition and multiplication. Then $\mathcal{C}$ is a *multidimensional asymptotic class for R in* $\mathcal{L}$, or *R-mac in* $\mathcal{L}$ for short, if for every $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$, where $n := l(\bar{x})$ and $m := l(\bar{y})$, there exists a finite definable partition Φ of $\mathcal{C}(m)$ such that for each $\pi \in \Phi$ there exists $h_\pi \in R$ such that

$$\Big| |\varphi(\mathcal{M}^n, \bar{a})| - h_\pi(\mathcal{M}) \Big| = o(h_\pi(\mathcal{M}))$$

for all $(\mathcal{M}, \bar{a}) \in \pi$ as $|M| \to \infty$, where the meaning of the little-o notation is as follows: For every $\varepsilon > 0$ there exists $Q \in \mathbb{N}$ such that for all $(\mathcal{M}, \bar{a}) \in \pi$, if $|M| > Q$, then

$$\Big| |\varphi(\mathcal{M}^n, \bar{a})| - h_\pi(\mathcal{M}) \Big| \leq \varepsilon h_\pi(\mathcal{M}).$$

2.2. Useful lemmas. We state and prove a number of lemmas that we will use later on. We start with the Projection Lemma:

LEMMA 2.2.1 (Projection Lemma). *Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures. Suppose that the definition of an R-mec (Definition 2.1.2) holds for $\mathcal{C}$ and for all $\mathcal{L}$-formulas $\varphi(x, \bar{y})$ with a single object variable x (as opposed to a tuple $\bar{x}$). Then $\mathcal{C}$ is an R-mec in $\mathcal{L}$.*

A proof of the equivalent result for R-macs is given in § 2.4 of [1]. It is adapted from the proof of Theorem 2.1 in [36]. Our proof of Lemma 2.2.1 is a simplified version of the proof in [1].

Proof of Lemma 2.2.1. Consider an arbitrary $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$, where $n := l(\bar{x})$ and $m := l(\bar{y})$. We need to prove that it satisfies both the size and definability clauses. We do this by induction on the length of $\bar{x}$. The base case of the induction is the hypothesis of the lemma.

Let $\bar{x} = (x_1, \dots, x_n)$. By the induction hypothesis we may assume that the size and definability clauses are satisfied by $\varphi(x_1, \dots, x_{n-1}; x_n, \bar{y})$, where the semicolon is used to indicate the division between the object variables and the parameter variables (see Remark 2.1.3(iv)). So we have a finite partition Γ of $\mathcal{C}(1+m) = \{(\mathcal{M}, a, \bar{b}) : \mathcal{M} \in \mathcal{C}, (a, \bar{b}) \in M^{1+m}\}$ with measuring functions $\{f_i : i \in \Gamma\} \subseteq R$ and defining $\mathcal{L}$-formulas $\{\gamma_i(x_n, \bar{y}) : i \in \Gamma\}$.

Consider each $\gamma_i(x_n, \bar{y})$. By the base case of the induction, each $\gamma_i(x_n, \bar{y})$ satisfies the size and definability clauses, so for each $i \in \Gamma$ we have a finite partition $\Phi_i := \{\pi_{i1}, \dots, \pi_{ir_i}\}$ of $\mathcal{C}(m) = \{(\mathcal{M}, \bar{b}) : \mathcal{M} \in \mathcal{C}, \bar{b} \in M^m\}$ with measuring functions $\{g_{ij} : 1 \leq j \leq r_i\} \subseteq R$ and defining $\mathcal{L}$-formulas $\{\psi_{ij}(\bar{y}) : 1 \leq j \leq r_i\}$. We thus have $k := |\Gamma|$ finite partitions of $\mathcal{C}(m)$. We use them to construct a single finite partition Φ of $\mathcal{C}(m)$. Define

$$\pi_{(j_1, \dots, j_k)} := \bigcap_{i \in \Gamma} \pi_{ij_i} \text{ and } J := \{(j_1, \dots, j_k) : 1 \leq j_i \leq r_i, 1 \leq i \leq k\}.$$

Then $\Phi := \{\pi_{(j_1, \dots, j_k)} : (j_1, \dots, j_k) \in J\}$ forms a finite partition of $\mathcal{C}(m)$. We now need to show that this partition works.

We first consider the size clause. For each $\pi_{(j_1, \dots, j_k)}$ we need to find a function $h_{(j_1, \dots, j_k)} \in R$ such that

$$h_{(j_1, \dots, j_k)}(\mathcal{M}) = |\Phi(\mathcal{M}^n, \bar{b})| \tag{2.2}$$

for all $(\mathcal{M}, \bar{b}) \in \pi_{(j_1, \dots, j_k)}$. So fix some arbitrary $(j_1, \dots, j_k)$ and consider an arbitrary pair $(\mathcal{M}, \bar{b}) \in \pi_{(j_1, \dots, j_k)}$. (If $\pi_{(j_1, \dots, j_k)} = \varnothing$, then any function $h \in R$ would be vacuously suitable, so we can ignore this case.) Let $\chi_i(x_1, \dots, x_n, \bar{y})$ denote the $\mathcal{L}$-formula

$$\varphi(x_1, \dots, x_n, \bar{y}) \wedge \gamma_i(x_n, \bar{y}).$$

Then, since the $\mathcal{L}$-formulas $\gamma_i(x_n, \bar{a})$ define the partition Γ, $\varphi(\mathcal{M}^n, \bar{b})$ is partitioned by the $\chi_i(\mathcal{M}^n, \bar{b})$, i.e.

$$\varphi(\mathcal{M}^n, \bar{b}) = \bigcup_{i \in \Gamma} \chi_i(\mathcal{M}^n, \bar{b}), \tag{2.3}$$

where the union is disjoint. Now, for each $i \in \Gamma$ we have

$$\left|\chi_i(\mathcal{M}^n, \bar{b})\right| = \sum_{a \in \gamma_i(\mathcal{M}, \bar{b})} \left|\varphi(\mathcal{M}^{n-1}, a, \bar{b})\right|$$

because $\chi_i(\mathcal{M}^n, \bar{b})$ fibres over $\gamma_i(\mathcal{M}, \bar{b})$. Thus

$$\left|\chi_i(\mathcal{M}^n, \bar{b})\right| = f_i(\mathcal{M}) \cdot \left|\gamma_i(\mathcal{M}, \bar{b})\right|, \tag{2.4}$$

since $\left|\varphi(\mathcal{M}^{n-1}, a, \bar{b})\right| = f_i(\mathcal{M})$ if $\mathcal{M} \models \gamma_i(a, \bar{b})$. But $(\mathcal{M}, \bar{b}) \in \pi_{(j_1,\dots,j_k)} \subseteq \pi_{ij_i}$ and so $\left|\gamma_i(\mathcal{M}, \bar{b})\right| = g_{ij_i}(\mathcal{M})$, which gives

$$\left|\chi_i(\mathcal{M}^n, \bar{b})\right| = f_i(\mathcal{M}) \cdot g_{ij_i}(\mathcal{M})$$

when put into (2.4). Combining this with (2.3) yields

$$\left|\varphi(\mathcal{M}^n, \bar{b})\right| = \sum_{i \in \Gamma} f_i(\mathcal{M}) \cdot g_{ij_i}(\mathcal{M}).$$

So define

$$h_{(j_1,\dots,j_k)}(\mathcal{M}) := \sum_{i=1}^{k} f_i(\mathcal{M}) \cdot g_{ij_i}(\mathcal{M})$$

for all $\mathcal{M} \in \mathcal{C}$ and (2.2) is satisfied as required.

We now come to the definability clause. Let $\psi_{(j_1,\dots,j_k)}(\bar{y})$ denote the formula

$$\bigwedge_{i=1}^{k} \psi_{ij_i}(\bar{y}).$$

Then $(\mathcal{M}, \bar{b}) \in \pi_{(j_1,\dots,j_k)}$ if and only if $\mathcal{M} \models \psi_{(j_1,\dots,j_k)}(\bar{b})$. So the definability clause is also satisfied and so we are done. □

The following lemma shows that R-mecs are closed under adding constant symbols:

Lemma 2.2.2. *Suppose that $\mathcal{C}$ is an R-mec in $\mathcal{L}$. Let $\mathcal{L}'$ be an extension of $\mathcal{L}$ by constant symbols and for $\mathcal{M} \in \mathcal{C}$ let $\mathcal{M}'$ be an $\mathcal{L}'$-expansion of $\mathcal{M}$. Then $\mathcal{C}' := \{\mathcal{M}' : \mathcal{M} \in \mathcal{C}\}$ is an R-mec in $\mathcal{L}'$.*

Proof. This follows straightforwardly from the definition of an R-mec. □

The following lemma shows that if we want to prove that a class $\mathcal{C}$ is an R-mec in $\mathcal{L}$, then it suffices to show that the definition eventually holds for each $\mathcal{L}$-formula:

Lemma 2.2.3. *Suppose that the definition of a multidimensional exact class (Definition 2.1.2) holds for $\varphi(\bar{x}, \bar{y})$, R and the subclass*

$$\mathcal{C}(m)_{>Q} := \{(\mathcal{M}, \bar{a}) : (\mathcal{M}, \bar{a}) \in \mathcal{C}(m) \text{ and } |M| > Q\}$$

of $\mathcal{C}(m)$, where $m := l(\bar{y})$, Q is some positive integer, and R contains the constant function $\mathcal{M} \mapsto k$ for each positive integer $k \leq Q$. Then the definition also holds for $\varphi(\bar{x}, \bar{y})$, R and $\mathcal{C}(m)$.

Proof. By the hypothesis of the lemma there exists a finite partition Φ of $\mathcal{C}(m)_{>Q}$ with measuring functions $\{h_\pi : \pi \in \Phi\}$ and defining $\mathcal{L}$-formulas $\{\psi_\pi(\bar{y}) : \pi \in \Phi\}$. Let

$$\Gamma_i := \{(\mathcal{M}, \bar{a}) : \mathcal{M} \in \mathcal{C}(m) \setminus \mathcal{C}(m)_{>Q} \text{ and } |\varphi(\mathcal{M}^n, \bar{a})| = i\}.$$

Then $\{\Gamma_i : 0 \leq i \leq Q\} \cup \Phi$ is a finite partition of $\mathcal{C}$ with measuring functions $\{g_i : 0 \leq i \leq Q\} \cup \{h_\pi : \pi \in \Phi\}$, where $g_i(\mathcal{M}) := i$ for all $\mathcal{M} \in \mathcal{C}$. So the size clause holds for $\mathcal{C}$.

Let σ_Q be the $\mathcal{L}$-sentence $\exists x_1 \dots \exists x_Q \forall y \bigvee_{1 \leq i \leq Q} y = x_i$, i.e. σ_Q says that there are at most Q elements, and let $\varphi_i(\bar{y})$ be the $\mathcal{L}$-formula $\exists!_i \bar{x}\, \varphi(\bar{x}, \bar{y})$, i.e. $\varphi_i(\bar{a})$

says that $|\varphi(\mathcal{M}^n, \bar{a})| = i$. Then the partition in the previous paragraph is defined by the $\mathcal{L}$-formulas $\{\varphi_i(\bar{y}) \wedge \sigma_Q : 1 \leq i \leq Q\} \cup \{\psi_\pi(\bar{y}) \wedge \neg\sigma_Q : \pi \in \Phi\}$. □

Corollary 2.2.4. *Let $\mathcal{C}$ be a finite class of finite $\mathcal{L}$-structures, where the largest structure in $\mathcal{C}$ has size Q. If R is the set of functions $\{\mathcal{M} \mapsto k : k \in \mathbb{N}, k \leq Q\}$, then $\mathcal{C}$ is an R-mec in $\mathcal{L}$.*

Proof. Use the proof of Lemma 2.2.3. □

Our last useful lemma is a compactness-like result:

Lemma 2.2.5. *Let $\mathcal{C}$ be a class of finite $\mathcal{L}$-structures. For $\mathcal{L}' \subseteq \mathcal{L}$ let $\mathcal{C}_{\mathcal{L}'}$ denote the class of all $\mathcal{L}'$-reducts of structures in $\mathcal{C}$. If $\mathcal{C}_{\mathcal{L}'}$ is an R-mec in $\mathcal{L}'$ for every finite $\mathcal{L}' \subseteq \mathcal{L}$, then $\mathcal{C}$ is an R-mec in $\mathcal{L}$.*

Proof. This follows from Definition 2.1.2, whose first (second-order) quantifier ranges over $\mathcal{L}$-formulas, and the following two facts: Firstly, $\mathcal{L}$-formulas are finite and so any $\mathcal{L}$-formula is an $\mathcal{L}'$-formula for some finite $\mathcal{L}' \subseteq \mathcal{L}$. Secondly, for every $\mathcal{L}'$-formula $\chi(\bar{y})$ (where $m := l(\bar{y})$, for every $\mathcal{L}'$-reduct $\mathcal{M}'$ of an $\mathcal{L}$-structure $\mathcal{M}$ and for every $\bar{a} \in M^m$, $\mathcal{M}' \models \chi(\bar{a})$ if and only if $\mathcal{M} \models \chi(\bar{a})$. □

2.3. Examples and non-examples. We start with an elementary example, which by Remark 2.1.3(v) is necessary for all other examples:

Example 2.3.1. The class $\mathcal{C}$ of finite sets is an R-mec in the language $\mathcal{L}_=$ of pure equality, where R contains the functions $\mathcal{M} \mapsto 0$, $\mathcal{M} \mapsto |M|$ and $\mathcal{M} \mapsto |M| - k$ for each $k \in \mathbb{N}$.

Proof. By the Projection Lemma (Lemma 2.2.1) it suffices to consider an $\mathcal{L}_=$-formula $\varphi(x, \bar{y})$ in one object variable x. Since $\mathcal{L}_=$ has quantifier elimination, $\varphi(x, y_1, \ldots, y_m)$ is equivalent to a formula of the form

$$\bigwedge_{i \in A} x = y_i \wedge \bigwedge_{i \in B} \neg\, x = y_i \;\wedge \bigwedge_{(i,j) \in C} y_i = y_j \;\wedge \bigwedge_{(i,j) \in D} \neg\, y_i = y_j$$

for some $A, B \subseteq \{1, \ldots, m\}$ and $C, D \subseteq \{1, \ldots, m\}^2$. We can now calculate $|\varphi(\mathcal{M}, \bar{a})|$ for different $(\mathcal{M}, \bar{a}) \in \mathcal{C}(m)$:

$$|\varphi(\mathcal{M}, a_1, \ldots, a_m)| = \begin{cases} 0 & \text{if } a_i \neq a_j \text{ for some } (i,j) \in C, \\ 0 & \text{if } a_i = a_j \text{ for some } (i,j) \in D, \\ 0 & \text{if } a_i \neq a_j \text{ for some } (i,j) \in A^2, \\ 0 & \text{if } a_i = a_j \text{ for some } (i,j) \in A \times B, \\ 1 & \text{if } A \neq \varnothing \text{ and } a_i = a_j \text{ for all } (i,j) \in A^2, \text{ or} \\ |M| - b & \text{otherwise, where } b := |\{a_i : i \in B\}|. \end{cases}$$

We can thus find a partition satisfying the size clause. The definability clause follows because the conditions depend only on equality relations and the fixed finite sets A, B, C and D. □

In order to explain the next example we first define the notion of a disjoint union of classes and then prove a lemma:

Definition 2.3.2. Consider $\mathcal{C}_1, \dots, \mathcal{C}_k$, where each $\mathcal{C}_i$ is a class of $\mathcal{L}_i$-structures. Define the *disjoint union* of $\mathcal{C}_1, \dots, \mathcal{C}_k$ to be

$$\mathcal{C}_1 \sqcup \cdots \sqcup \mathcal{C}_k := \{\mathcal{M}_1 \sqcup \cdots \sqcup \mathcal{M}_k : \mathcal{M}_i \in \mathcal{C}_i\},$$

where we define a first-order structure on $\mathcal{M}_1 \sqcup \cdots \sqcup \mathcal{M}_k$ as follows: The domain is $M_1 \cup \cdots \cup M_k$, which we make formally disjoint if necessary. The language is $\mathcal{L}_1 \sqcup \cdots \sqcup \mathcal{L}_k$, which has a sort S_i for each M_i and contains all $\mathcal{L}_i$-symbols for every $i \in \{1, \dots, k\}$, with each $\mathcal{L}_i$-symbol being restricted to the sort S_i.

Lemma 2.3.3. *Let $\mathcal{C}_i$ be an R_i-mec in $\mathcal{L}_i$. Then $\mathcal{C}_1 \sqcup \cdots \sqcup \mathcal{C}_k$ is an R-mec in $\mathcal{L} := \mathcal{L}_1 \sqcup \cdots \sqcup \mathcal{L}_k$, where R is the set generated by $R_1 \cup \cdots \cup R_k$ under addition and multiplication.*

Sketch proof. We restrict our attention to the case $k = 2$, the general case following by induction.

Consider an $\mathcal{L}$-formula $\varphi(\bar{x}_1, \bar{x}_2; \bar{y}_1, \bar{y}_2)$, where $\bar{x}_i$ and $\bar{y}_i$ are of sort S_i. By an induction on the complexity of the formula, one can show that $\varphi(\bar{x}_1, \bar{x}_2; \bar{y}_1, \bar{y}_2)$ is equivalent to a finite disjunction of $\mathcal{L}$-formulas of the form $\chi(\bar{x}_1, \bar{y}_1) \wedge \theta(\bar{x}_2, \bar{y}_2)$, where χ is an $\mathcal{L}_1$-formula, θ is an $\mathcal{L}_2$-formula, and the disjuncts are pairwise inconsistent. Since the domains of $\mathcal{M}_1 \in \mathcal{C}_1$ and $\mathcal{M}_2 \in \mathcal{C}_2$ are disjoint, we have

$$|\chi(\mathcal{M}_1 \sqcup \mathcal{M}_2, \bar{a}_1) \wedge \theta(\mathcal{M}_1 \sqcup \mathcal{M}_2, \bar{a}_2)| = |\chi(\mathcal{M}_1, \bar{a}_1)| \cdot |\theta(\mathcal{M}_2, \bar{a}_2)|.$$

One then proceeds by using the facts that the disjuncts are pairwise inconsistent, thus allowing summation, and that each $\mathcal{C}_i$ is an R-mec. □

Example 2.3.4. Consider the class $\mathcal{C}$ of finite cyclic groups and for arbitrary $k \in \mathbb{N}^+$ define $\mathcal{C}_k := \{C_1 \oplus \cdots \oplus C_k : C_i \in \mathcal{C}\}$. Let $\mathcal{L}$ be the language of groups (with or without a constant symbol for the identity element – recall Lemma 2.2.2). Then $\mathcal{C}_k$ is a multidimensional exact class in $\mathcal{L}'$, where $\mathcal{L}'$ is $\mathcal{L}$ adjoined with a unary predicate P_i for each part of the direct sum:

$$P_i^{C_1 \oplus \cdots \oplus C_k} := \{(0, \dots, 0, \underset{i^{\text{th}} \text{ place}}{\underset{\uparrow}{a}}, 0, \dots, 0) : a \in C_i\}.$$

Sketch proof. Theorem 3.14 in [36] states that $\mathcal{C}$ is a 1-dimensional asymptotic class in $\mathcal{L}$ (see Definition 2.1 in [17] for the definition of an N-dimensional asymptotic class). Inspection of the proof of this theorem shows that $\mathcal{C}$ is in fact an exact class, since the measuring functions yield exact sizes and not just approximaitons. So by Lemma 2.3.3, $\underbrace{\mathcal{C} \sqcup \cdots \sqcup \mathcal{C}}_{k \text{ times}}$ is an exact class in $\underbrace{\mathcal{L} \sqcup \cdots \sqcup \mathcal{L}}_{k \text{ times}}$. We now use the work in § 2.5 of [1] and § 2.4 of [43] regarding interpretability: Since $\mathcal{L}'$ is equipped with the predicates P_i, $\mathcal{C}_k$ and $\mathcal{C} \sqcup \cdots \sqcup \mathcal{C}$ are $\varnothing$-bi-interpretable (see [1] or § 2.4 of [43]) and therefore $\mathcal{C}_k$ is an exact class. □

Remark 2.3.5. We comment on Example 2.3.4. The class $\mathcal{C}$ of finite cyclic groups is both a multidimensional exact class and a 1-dimensional asymptotic class, so one might wonder whether it could be a "1-dimensional exact class". However, the notion of an N-dimensional exact class is inconsistent: Consider two disjoint definable sets $A, B \subseteq M$ with $|A| = \alpha|M|^{a/N}$ and $|B| = \beta|M|^{b/N}$, where $a > b$. Then their union $A \cup B$, which is definable, has size $\alpha|M|^{a/N} +$

$\beta|M|^{b/N}$, which cannot be expressed in the form $\mu|M|^{d/N}$ for a dimension–measure pair (d, μ). This is not an issue for an N-dimensional *asymptotic* class, since $|M|^{a/N}$ swamps $|M|^{b/N}$ as $|M| \to \infty$. It is also not an issue for a multidimensional exact class, where one is not bound to dimension–measure pairs.

Example 2.3.6 (Proposition 4.4.2 in [19]). Consider the class of homocyclic groups

$$\mathcal{C} := \{(\mathbb{Z}/p^n\mathbb{Z})^m : p \text{ is prime and } n, m \in \mathbb{N}^+\}$$

in the language $\mathcal{L} := \{+\}$. This class is an R-mec, where R consists of functions of the form

$$\sum_{i=0}^{r} \sum_{j=-rd}^{rd} c_{ij} p^{m(in+j)},$$

where r is the length of the object-variable tuple of the given $\mathcal{L}$-formula (see Remark 2.1.3(iv)); d is a positive integer that is constructively determined by the $\mathcal{L}$-formula; and the c_{ij} are integers that depend on the $\mathcal{L}$-formula, with $c_{ij} := 0$ whenever $in + j < 0$. (Each group $(\mathbb{Z}/p^n\mathbb{Z})^m \in \mathcal{C}$ is determined by a triple (p, n, m), so by defining a function on such triples we also define a function on $\mathcal{C}$.)

Further examples will arise as this paper progresses. We now turn our attention to non-examples, which are often just as interesting.

Non-Example 2.3.7 (Example 3.1 in [36]). The class $\mathcal{C}$ of all finite linear orders in (any extension of) the language $\mathcal{L} = \{<\}$ does not form a weak R-mec for any R.

Proof. Let $\varphi(x, y)$ be the formula $x < y$ and consider the finite linear order $\mathcal{M}_k := \{a_0 < \cdots < a_k\}$. Then $|\varphi(\mathcal{M}_k, a_i)| = i$. As we let k increase and let i vary we define arbitrarily many subsets of distinct sizes. Thus no finite number of functions can yield $|\varphi(\mathcal{M}_k, a_i)|$ for all $k, i \in \mathbb{N}$. Let's make that argument a little more rigorous.

By way of contradiction, suppose that there exists R such that $\mathcal{C}$ forms a weak R-mec. So for the formula $\varphi(x, y)$ there exists a finite partition Φ of $\mathcal{C}(1)$ with measuring functions $\{h_\pi : \pi \in \Phi\} \subseteq R$. Let $t := |\Phi|$ and consider the finite linear order finite linear order $\mathcal{M}_t$. Then t measuring functions are not enough for this structure, since there are $t + 1$ different sizes of the definable subsets, namely $|\varphi(\mathcal{M}_t, a_0)| = 0, \ldots, |\varphi(\mathcal{M}_t, a_t)| = t$. A contradiction. □

The following non-example is informative, as it shows that the choice of language in Example 2.3.6 is important:

Non-Example 2.3.8. Let p be prime. Then the class $\{\mathbb{Z}/p^n\mathbb{Z} : n \in \mathbb{N}^+\}$ of multiplicative monoids in (any extension of) the language $\mathcal{L} = \{\times\}$ does not form a weak R-mec for any R.

Proof. Let R be any set of functions from $\mathcal{C}$ to $\mathbb{N}$ and let $\varphi(x, y)$ be the formula $\exists z\, (x = z \times y)$. Then $|\varphi(\mathbb{Z}/p^n\mathbb{Z}, p^i)| = p^{n-i}$. So as we let n increase and let i vary we define arbitrarily many subsets of distinct sizes. Thus, by the same argument given in the proof of Non-Example 2.3.7, no finite number of measuring functions can suffice for $|\varphi(\mathbb{Z}/p^n\mathbb{Z}, p^i)|$ for all $n, i \in \mathbb{N}$. □

REMARK 2.3.9.

(i) Non-Examples 2.3.7 and 2.3.8 are special cases of the general fact that an ultraproduct of a weak multidimensional asymptotic class cannot have the strict order property; see § 3.4 of [1]. (See Definition 2.14 in [7] or Exercise 8.2.4 in [41] for a definition of the strict order property.)

(ii) The issue preventing Non-Example 2.3.8 from being an R-mec is the unbounded exponent n. If the exponent is bounded, then one can have an R-mec, as shown by the work of Bello Aguirre in [4] and [5].

The following two non-examples concern ultraproducts, the random graph and the random tournament,[1] which are covered extensively in the literature, for instance [3], Exercise 2.5.19 in [38] and Exercise 1.2.4 in [41] (ultraproducts), p. 232 of [10], p. 17 of [12], §§ 1–2 of [18], p. 435 of [30], pp. 50–52 of [38] and Exercise 3.3.1 in [41] (the random graph and the random tournament).

NON-EXAMPLE 2.3.10. The random graph is not elementarily equivalent to an ultraproduct of a multidimensional exact class.

PROOF. See [1] or Non-Example 2.3.12 in [43]. □

NON-EXAMPLE 2.3.11. The random tournament is not elementarily equivalent to an ultraproduct of a multidimensional exact class.

PROOF. See [1] or Non-Example 2.3.14 in [43]. Note that in the latter the tournament relation $a \twoheadrightarrow b$ is shown incorrectly as $a \,\dot{}\, b$. □

REMARK 2.3.12. The situation is quite different for asymptotic classes: The random graph is elementarily equivalent to any infinite ultraproduct of the class of Paley graphs, which is a 1-dimensional asymptotic class (Example 3.4 in [36]), and the random tournament is elementarily equivalent to any infinite ultraproduct of the class of Paley tournaments, which is also a 1-dimensional asymptotic class (Example 3.5 in [36]). This is an interesting phenomenon, especially in light of Theorem 7.5.6 in [14] and Theorem 4.6.4. We will discuss it further in Question 5.3.

§3. Smooth approximation and exact classes.

The goal of this section is to prove Proposition 3.2.1, which states that finite structures smoothly approximating an $\aleph_0$-categorical structure form a multidimensional exact class. In § 3.1 we define the notion of smooth approximation and then provide some examples. In § 3.2 we state and prove the result.

3.1. Smooth approximation.

The notion of smooth approximation was introduced by Lachlan in the 1980s, arising as a generalisation of $\aleph_0$-categorical, $\aleph_0$-stable structures [13], in particular Corollary 7.4 of that paper. [9], [27], [31], [32] and [33] are also relevant, but the key texts on smooth approximation itself are [24] by Kantor, Liebeck and Macpherson and [14] by Cherlin and Hrushovski. A history of the development of the notion is to be found in § 1.1 of [14] and there is a survey article [34], which also contains improvements and errata to [24].

[1] Due to its different guises, the random graph goes by various names, including the 'Rado graph' and 'the generic (countable homogeneous) graph'. The random tournament has similar aliases.

Smooth approximation also arises in the context of asymptotic classes in [16], [17], [36] and [37].

For $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ we use the notation $\mathcal{N} \leq \mathcal{M}$ to mean that $\mathcal{N}$ is an $\mathcal{L}$-substructure of $\mathcal{M}$.

Definition 3.1.1 (Homogenous substructure). Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. $\mathcal{N}$ is a *homogeneous substructure*[2] of $\mathcal{M}$, notationally $\mathcal{N} \leq_{\text{hom}} \mathcal{M}$, if $\mathcal{N} \leq \mathcal{M}$ and for every $k \in \mathbb{N}^+$ and every pair $\bar{a}, \bar{b} \in N^k$, $\bar{a}$ and $\bar{b}$ lie in the same $\text{Aut}(\mathcal{M})$-orbit if and only if $\bar{a}$ and $\bar{b}$ lie in the same $\text{Aut}_{\{N\}}(\mathcal{M})$-orbit, where

$$\text{Aut}_{\{N\}}(\mathcal{M}) := \{\sigma \in \text{Aut}(\mathcal{M}) : \sigma(N) = N\}.$$

Definition 3.1.2 (Smooth approximation). An $\mathcal{L}$-structure $\mathcal{M}$ is *smoothly approximable* if $\mathcal{M}$ is $\aleph_0$-categorical and there exists a sequence $(\mathcal{M}_i)_{i<\omega}$ of finite homogeneous substructures of $\mathcal{M}$ such that $M_i \subset M_{i+1}$ for all $i < \omega$ and $\bigcup_{i<\omega} M_i = M$. We say that $\mathcal{M}$ is *smoothly approximated* by the $\mathcal{M}_i$.

We provide some examples of smoothly approximable structures, starting with a trivial example:

Example 3.1.3. Let $\mathcal{M}$ be a countably infinite set in the language of equality. Enumerate $\mathcal{M}$ as $(a_i : i < \omega)$ and let $\mathcal{M}_i = \{a_0, \ldots, a_i\}$. Then each $\mathcal{M}_i$ is a finite homogeneous substructure of $\mathcal{M}$ and $\mathcal{M} = \bigcup_{i<\omega} \mathcal{M}_i$.

Example 3.1.4. Consider a language $\mathcal{L} := \{I_1, I_2\}$, where I_1 and I_2 are binary relation symbols. Let $\mathcal{M}$ be a countable $\mathcal{L}$-structure where $I_1^{\mathcal{M}}$ and $I_2^{\mathcal{M}}$ are equivalence relations such that $I_1^{\mathcal{M}}$ has infinitely many classes, $I_2^{\mathcal{M}}$ refines $I_1^{\mathcal{M}}$, every I_1-equivalence class contains infinitely many I_2-equivalence classes, and every I_2-equivalence class is infinite; that is, $\mathcal{M}$ is partitioned into infinitely many I_1-equivalence classes, each of which is then partitioned into infinitely many I_2-equivalence classes, each of which is infinite. Note that $\mathcal{M}$ is unique up to isomorphism and hence $\aleph_0$-categorical, since the structure is first-order expressible in $\mathcal{L}$.

Enumerate the I_1-equivalence classes as $(a_i : i < \omega)$ and the I_2-equivalence classes within each a_i as $(a_{ij} : j < \omega)$. Finally, enumerate the elements of each a_{ij} as $(a_{ijk} : k < \omega)$. Let $\mathcal{M}_{(r,s,t)} := \{a_{ijk} : i \leq r, j \leq s, k \leq t\}$. Then each $\mathcal{M}_{(r,s,t)}$ is a finite homogeneous substructure of $\mathcal{M}$ and $\mathcal{M} = \bigcup_{r<\omega} \mathcal{M}_{(r,r,r)}$.

Note that this example straightforwardly generalises to the case of n nested equivalence relations for any $n < \omega$.

Example 3.1.5. Let $\mathcal{M}$ be the direct sum of ω-many copies of the additive group $\mathbb{Z}/p^2\mathbb{Z}$, where p is some fixed prime. Note that $\mathcal{M}$ is $\aleph_0$-categorical, which can be seen via Szmielew invariants (see Appendix A.2 in [21]). Let $\mathcal{M}_i$ consist of the first i copies of $\mathbb{Z}/p^2\mathbb{Z}$. Then each $\mathcal{M}_i$ is a finite homogeneous substructure of $\mathcal{M}$ and $\mathcal{M} = \bigcup_{i<\omega} \mathcal{M}_i$.

3.2. Smooth approximation is exact. We now come to Proposition 3.2.1, the central result of this section. We first give the main proof, leaving the necessary technical lemmas until afterwards.

[2] We define 'homogeneous substructure' as one term, not as the conjunction of two words; that is, 'homogeneous substructure' does not mean a substructure that is homogeneous.

PROPOSITION 3.2.1. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure smoothly approximated by finite homogeneous substructures $(\mathcal{M}_i)_{i<\omega}$. Then there exists R such that $\mathcal{C} := \{\mathcal{M}_i : i < \omega\}$ is an R-mec in $\mathcal{L}$.*

PROOF. Let $\varphi(\bar{x}, \bar{y})$ be an $\mathcal{L}$-formula with $n := l(\bar{x})$ and $m := l(\bar{y})$.

We first cover the size clause. We use the Ryll-Nardzewski Theorem: Since $\mathcal{M}$ is $\aleph_0$-categorical, $\mathrm{Aut}(\mathcal{M})$ acts oligomorphically on $\mathcal{M}$ and thus $\mathcal{M}^m$ has only finitely many $\mathrm{Aut}(\mathcal{M})$-orbits, say $\Theta_1, \dots, \Theta_k$. We use these orbits to define a finite partition $\pi_1, \dots, \pi_k$ of $\mathcal{C}(m) = \{(\mathcal{M}_i, \bar{a}) : i < \omega, \bar{a} \in {M_i}^m\}$:

$$(\mathcal{M}_i, \bar{a}) \in \pi_j :\Longleftrightarrow \; \bar{a} \in \Theta_j.$$

Define $\pi_j^{\mathcal{M}_i} := \{\bar{a} \in {M_i}^m : (\mathcal{M}_i, \bar{a}) \in \pi_j\}$ and let $\bar{a}, \bar{b} \in {M_i}^m$. Then

$$\begin{aligned} \bar{a}, \bar{b} \in \pi_j^{\mathcal{M}_i} \iff & \bar{a}, \bar{b} \in \Theta_j \\ & \Longrightarrow \bar{a} \text{ and } \bar{b} \text{ lie in the same } \mathrm{Aut}_{\{M_i\}}(\mathcal{M})\text{-orbit} \\ & \qquad (\text{since } \mathcal{M}_i \leq_{\mathrm{hom}} \mathcal{M}) \\ & \Longrightarrow |\varphi({\mathcal{M}_i}^n, \bar{a})| = |\varphi({\mathcal{M}_i}^n, \bar{b})|. \end{aligned} \tag{3.5}$$

We justify the last implication: Since $\bar{a}$ and $\bar{b}$ lie in the same $\mathrm{Aut}_{\{M_i\}}(\mathcal{M})$-orbit, there is some $\sigma \in \mathrm{Aut}_{\{M_i\}}(\mathcal{M})$ such that $\sigma(\bar{a}) = \bar{b}$. But $\sigma \upharpoonright M_i$ is an automorphism of $\mathcal{M}_i$ and thus $\mathcal{M}_i \models \varphi(\bar{c}, \bar{a})$ if and only if $\mathcal{M}_i \models \varphi(\sigma(\bar{c}), \sigma(\bar{a}))$. Therefore $\sigma \colon \varphi({\mathcal{M}_i}^n, \bar{a}) \to \varphi({\mathcal{M}_i}^n, \bar{b})$ is a bijection and hence $|\varphi({\mathcal{M}_i}^n, \bar{a})| = |\varphi({\mathcal{M}_i}^n, \bar{b})|$.

Define $h_j(\mathcal{M}_i) := |\varphi({\mathcal{M}_i}^n, \bar{a})|$, where $\bar{a}$ is some arbitrary element of $\pi_j^{\mathcal{M}_i}$ (if no such $\bar{a}$ exists, then the value of h_j at $\mathcal{M}_i$ can be chosen to be anything, say 0); this function is well-defined by (3.5). Then $\pi_1, \dots, \pi_k$ and $h_1, \dots, h_k$ satisfy the size clause.

We now come to the definability clause. We use the Ryll-Nardzewski Theorem again: Each orbit Θ_j is the solution set of an isolated m-type and so the $\mathcal{L}$-formula isolating this type defines Θ_j in $\mathcal{M}$; let $\psi_j(\bar{y})$ be the isolating formula for Θ_j. So $\mathcal{M} \models \psi_j(\bar{a})$ if and only f $\bar{a} \in \Theta_j$. We claim that the following is eventually true, i.e. there exists $Q \in \mathbb{N}$ such that for each ψ_j, if $i > Q$, then

$$\mathcal{M}_i \models \psi_j(\bar{a}) \iff \bar{a} \in \pi_j^{\mathcal{M}_i} \tag{3.6}$$

for every $\bar{a} \in {M_i}^m$. By Lemma 2.2.3 this suffices to prove the definability clause.

We prove this claim: Apply Lemma 3.2.6 to ψ_j to obtain $Q_j \in \mathbb{N}$ such that if $i > Q_j$ and $\bar{a} \in {M_i}^m$, then

$$\mathcal{M} \models \psi_j(\bar{a}) \iff \mathcal{M}_i \models \psi_j(\bar{a}). \tag{3.7}$$

Let $Q := \max\{Q_j : 1 \leq j \leq k\}$. Consider $\bar{a} \in {M_i}^m$ with $i > Q$. Then

$$\mathcal{M}_i \models \psi_j(\bar{a}) \overset{(3.7)}{\iff} \mathcal{M} \models \psi_j(\bar{a}) \iff \bar{a} \in \Theta_j \iff \bar{a} \in \pi_j^{\mathcal{M}_i}$$

and so (3.6) holds. □

Definition 3.2.2 (Canonical language). We define the *canonical language* [3] of $\mathcal{M}$ to be

$$\mathcal{L}^* := \mathcal{L} \cup \{P_\Theta : \Theta \text{ is a } \mathrm{Aut}(\mathcal{M})\text{-orbit of } \mathcal{M}\},$$

where each P_Θ is a new unary predicate symbol. We expand $\mathcal{M}$ to an $\mathcal{L}^*$-structure $\mathcal{M}^*$ by defining the assignment of each P_Θ in $\mathcal{M}^*$ to be Θ. We expand each $\mathcal{M}_i$ to an $\mathcal{L}^*$-structure ${\mathcal{M}_i}^*$ by defining the assignment of each P_Θ to be $\Theta \cap M_i$.

Lemmas 3.2.3 and 3.2.4 are standard and we state them without proof:

Lemma 3.2.3. $\mathrm{Aut}(\mathcal{M}) = \mathrm{Aut}(\mathcal{M}^*)$.

Lemma 3.2.4. $\mathrm{Th}(\mathcal{M}^*)$ *has quantifier elimination; in particular, any* $\mathcal{L}^*$*-formula is equivalent in* $\mathrm{Th}(\mathcal{M}^*)$ *to a quantifier-free* $(\mathcal{L}^* \setminus \mathcal{L})$*-formula.*

Lemma 3.2.5. $\mathcal{M}^*$ *is smoothly approximated by* $({\mathcal{M}_i}^*)_{i<\omega}$.

Proof. Since $\mathcal{M}$ is $\aleph_0$-categorical, by Lemma 3.2.3 and the Ryll-Nardzewski Theorem, $\mathcal{M}^*$ is also $\aleph_0$-categorical. Also note that each ${\mathcal{M}_i}^*$ is a finite $\mathcal{L}^*$-substructure of $\mathcal{M}^*$. It remains to show that ${\mathcal{M}_i}^* \leq_{\mathrm{hom}} \mathcal{M}^*$. If $\bar{a}, \bar{b} \in {\mathcal{M}_i}^*$ lie in the same $\mathrm{Aut}(\mathcal{M}^*)_{\{M_i\}}$-orbit, then $\bar{a}$ and $\bar{b}$ lie in the same $\mathrm{Aut}(\mathcal{M}^*)$-orbit, since $\mathrm{Aut}(\mathcal{M}^*)_{\{M_i\}} \subseteq \mathrm{Aut}(\mathcal{M}^*)$. Now suppose that $\bar{a}, \bar{b} \in {\mathcal{M}_i}^*$ lie in the same $\mathrm{Aut}(\mathcal{M}^*)$-orbit. By Lemma 3.2.3, $\bar{a}$ and $\bar{b}$ lie in the same $\mathrm{Aut}(\mathcal{M})$-orbit. Thus, since $\mathcal{M}_i \leq_{\mathrm{hom}} \mathcal{M}$, there exists $\sigma \in \mathrm{Aut}(\mathcal{M})_{\{M_i\}}$ such that $\sigma(\bar{a}) = \bar{b}$. But $\sigma \in \mathrm{Aut}(\mathcal{M}^*)_{\{M_i\}}$, again by Lemma 3.2.3, and so $\bar{a}$ and $\bar{b}$ lie in the same $\mathrm{Aut}(\mathcal{M}^*)_{\{M_i\}}$-orbit. □

Lemma 3.2.6. *Let* $\chi(\bar{y})$ *be an* $\mathcal{L}$*-formula with* $m := l(\bar{y})$*. Then there exists* $Q \in \mathbb{N}$ *such that if* $i > Q$ *and* $\bar{c} \in {M_i}^m$*, then*

$$\mathcal{M} \models \chi(\bar{c}) \iff \mathcal{M}_i \models \chi(\bar{c}).$$

Proof. Consider $\mathcal{M}^*$. By Lemma 3.2.4, $T := \mathrm{Th}(\mathcal{M}^*)$ has quantifier elimination and thus there is a quantifier-free $\mathcal{L}^*$-formula $\delta(\bar{y})$ such that $\forall \bar{y}\, (\chi(\bar{y}) \leftrightarrow \delta(\bar{y})) \in T$. Thus by compactness there is an $\mathcal{L}^*$-sentence $\tau \in T$ such that

$$\tau \models \forall \bar{y}\, (\chi(\bar{y}) \leftrightarrow \delta(\bar{y})). \tag{3.8}$$

By Lemma 3.2.5 and the $\forall\exists$-axiomatisation of T (see the proof of Proposition 5.4 in [24]), there exists $Q \in \mathbb{N}$ such that ${\mathcal{M}_i}^* \models \tau$ for all $i > Q$. Now, consider some arbitrary $\bar{c} \in {M_i}^m$ with $i > Q$. Since δ is quantifier-free and ${\mathcal{M}_i}^* \leq \mathcal{M}^*$,

$$\mathcal{M}^* \models \delta(\bar{c}) \iff {\mathcal{M}_i}^* \models \delta(\bar{c}).$$

Hence by (3.8) we have

$$\mathcal{M}^* \models \chi(\bar{c}) \iff {\mathcal{M}_i}^* \models \chi(\bar{c})$$

because $\mathcal{M}^* \models \tau$ and ${\mathcal{M}_i}^* \models \tau$. But χ is an $\mathcal{L}$-formula and thus

$$\mathcal{M} \models \chi(\bar{c}) \iff \mathcal{M}_i \models \chi(\bar{c}),$$

as required. □

[3] Note that the term *canonical language* is sometimes used to refer to the smaller language $\mathcal{L}^* \setminus \mathcal{L}$. We avoid this usage.

§4. Lie coordinatisation. The goal of this section is to use Lie coordinatisation to prove the main result Theorem 4.6.4, as conjectured by Macpherson. As such, our account of Lie coordinatisation is streamlined for this purpose and we leave some important notions from [14] by the wayside, most notably orientation and orthogonality. That being said, we make explicit a number of details that are only implicit in [14], especially in our proofs of Theorem 4.4.1 and Proposition 4.5.2. Our presentation is based primarily on [14], with input from [15].

The history of Lie coordinatisation is complex and we give only a very brief summary; see § 1 of [11] and §§ 1.1–1.2 of [14] for a more detailed picture. The notion was developed by Cherlin and Hrushovski as (inter alia) an attempt to find a structure theory for smoothly approximable structures, building on the work of Kantor, Liebeck and Macpherson in [24]. Deep links between other model-theoretic notions were discovered through their investigation (§ 1.2 of [14]). In particular, it was shown that Lie coordinatisability and smooth approximation are equivalent (Theorem 2 in [14]). Note that the classification of finite simple groups plays a fundamental role, albeit in the background.

In contrast to its mathematical depth, Lie coordinatisation has made only a shallow footprint in the literature, in part due to the development of simple theories (see pp. 8–10 of [14]). The first publication on the topic was the paper [22] by Hrushovski, in joint work with Cherlin. Some technical issues were found in this paper (see p. 7 of [14]) and corrected results were published in [11], which is essentially an abridgement of the main text [14]. The paper [15] by Chowdhury, Hart and Sokolović makes significant contributions and Hrushovski has published some further work on quasifiniteness in [23]. There are also some unpublished notes [20] by Hill and Smart. Lie coordinatisation arises in the context of asymptotic classes in [16], [17], [36] and [37].

We now outline the structure of this section. In § 4.1 we go over the basic concepts of Lie coordinatisation and in § 4.2 we provide two examples of Lie coordinatisable structures. § 4.3 develops the notion of an envelope, which is fundamental to the rest of the section. We then move on to § 4.4, where we state and sketch a proof of a result (Theorem 4.4.1) that allows us to apply Proposition 3.2.1 to obtain a short version of Macpherson's conjecture (Corollary 4.4.2). § 4.5 then provides us with the extra information needed to prove the full version of the conjecture in § 4.6.

4.1. Lie geometries and Lie coordinatisation. We state the definition of Lie coordinatisation. We need to go over a number of preliminaries first, starting with Lie geometries. We refer the reader to chapter 7 of [2] for the terminology and theory of vector spaces with forms.

Definition 4.1.1 (Linear Lie geometry). Let K be a finite field. A *linear Lie geometry* over K is one of the following six kinds [4] of structures:

1. *A degenerate space.* An infinite set in the language of equality.
2. *A pure vector space.* An infinite-dimensional vector space V over K.
3. *A polar space.* Two infinite-dimensional vector spaces V and W over K with a non-degenerate bilinear form $V \times W \to K$.

[4] We use the word 'kind' in order to avoid overuse of the word 'type'.

4. *A symplectic space.* An infinite-dimensional vector space V over K with a symplectic bilinear form $V \times V \to K$.
5. *A unitary space.* An infinite-dimensional vector space V over K with a unitary sesquilinear form $V \times V \to K$.
6. *An orthogonal space.* An infinite-dimensional vector space V over K with a quadratic form $V \to K$ whose associated bilinear form is non-degenerate.

Remark 4.1.2. We comment on Definition 4.1.1.

(i) We consider linear Lie geometries as two-sorted structures (V, K), with a sort V in the language of groups with an abelian group structure, a sort K in the language of rings with a field structure, and a function $K \times V \to V$ for scalar multiplication. We call V the *vector sort* and K the *field sort*. (See pp. 5 and 12 of [41] for a summary of multi-sorted structures and languages.) The elements of K are named by constant symbols.[5] In the polar case, the vector sort is $V \cup W$ in the language of groups equipped with an equivalence relation with precisely two classes V and W, each with an abelian group structure.

(ii) We have ignored quadratic Lie geometries (Definition 2.1.4 in [14]), as we do not need to consider them, save only to rule them out in the proof of Proposition 4.5.2. They arise from the fact that in characteristic 2 every symplectic bilinear form has many associated quadratic forms.

Fact 4.1.3. *Every linear Lie geometry has quantifier elimination and is $\aleph_0$-categorical.*

Proof. Lemmas 2.2.8 and 2.3.19 in [14]. □

Definition 4.1.4 (Projective Lie geometry). Let L be a linear Lie geometry. Define an equivalence relation $\sim$ on $L \setminus \operatorname{acl}(\varnothing)$ by $a \sim b :\Longleftrightarrow \operatorname{acl}(a) = \operatorname{acl}(b)$, where acl denotes the usual model-theoretic algebraic closure in L. The *projectivisation* of L is then the quotient structure $(L \setminus \operatorname{acl}(\varnothing))/{\sim}$. A *projective Lie geometry* is a structure that is the projectivisation of some linear Lie geometry.

Remark 4.1.5 (comment after Definition 2.1.7 in [14]). By quantifier elimination (Fact 4.1.3), algebraic closure is just linear span and so a projective Lie geometry is a projective geometry in the usual sense.

Definition 4.1.6 (Affine Lie geometry). An *affine Lie geometry* is a structure of the form $(V, A, \oplus, -)$, where V is the vector sort of a linear Lie geometry (but not a degenerate space), A is a set, $\oplus\colon V \times A \to A$ is a regular group action and $-\colon A \times A \to V$ is such that $a = v \oplus b$ implies $a - b = v$. Here 'regular' means that for every $a, b \in A$ there exists a unique $v \in V$ such that $a = v \oplus b$. In the polar case the structure is $(V, W, A, \oplus, -)$, where $\oplus\colon V \times A \to A$ is a regular group action and $-\colon A \times A \to V$ is such that $a = v \oplus b$ implies $a - b = v$.

The notions of canonical and stable embeddedness are fundamental to Lie coordinatisation:

[5] Note that this is what the prefix 'basic' refers to in Definition 2.1.6 in [14]. Since we always name the field elements by constant symbols, we suppress this prefix.

Definition 4.1.7. Consider an $\mathcal{L}$-structure $\mathcal{N}$ and an $\mathcal{L}'$-structure $\mathcal{M}$ such that the underlying set M is an $\mathcal{L}_N$-definable subset of N. Let $c \in \mathcal{N}^{\mathrm{eq}}$ be a canonical parameter for M. (See § 8.2 of [38] or § 8.4 of [41] for an introduction to canonical parameters.)

(i) $\mathcal{M}$ is *canonically embedded* in $\mathcal{N}$ if the $\mathcal{L}'_\varnothing$-definable relations of $\mathcal{M}$ are precisely the $\mathcal{L}_c$-definable relations on $\mathcal{M}$; that is, for every $n \in \mathbb{N}^+$, a subset $D \subseteq M^n$ is $\mathcal{L}'_\varnothing$-definable in the structure $\mathcal{M}$ if and only if it is $\mathcal{L}_c$-definable in the structure $\mathcal{N}$. (The notation $\mathcal{L}'_\varnothing$ isn't strictly necessary, since $\mathcal{L}' = \mathcal{L}'_\varnothing$, but the subscript $\varnothing$ is added to emphasise $\varnothing$-definability.)
(ii) $\mathcal{M}$ is *stably embedded* in $\mathcal{N}$ if every $\mathcal{L}_N$-definable relation on $\mathcal{M}$ is $\mathcal{L}_M$-definable in a uniform way; that is, for every $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$, where $n := l(\bar{x})$ and $m := l(\bar{y})$, if $\varphi(\mathcal{N}^n, \bar{a}) \subseteq M^n$ for every $\bar{a} \in N^m$, then there exists an $\mathcal{L}$-formula $\varphi'(\bar{x}, \bar{z})$, where $r := l(\bar{z})$, such that for every $\bar{a} \in N^m$ there exists $\bar{a}' \in M^r$ such that $\varphi(\mathcal{N}^n, \bar{a}) = \varphi'(\mathcal{N}^n, \bar{a}')$. (Note that we need not have $m = r$.)
(iii) $\mathcal{M}$ is *fully embedded* in $\mathcal{N}$ if $\mathcal{M}$ is both canonically and stably embedded in $\mathcal{N}$.

Intuitively, $\mathcal{M}$ is fully embedded in $\mathcal{N}$ if $\mathcal{N}$ cannot place any additional structure on $\mathcal{M}$.

We won't need the following definition until § 4.5, but it follows on from the previous definitions.

Definition 4.1.8 (Localisation). Let P be a projective Lie geometry, arising from a linear Lie geometry L. Suppose that P is fully embedded in an $\mathcal{L}$-structure $\mathcal{M}$. The *localisation* P/A of P over a finite set $A \subset M$ is defined as follows: Let f be the bilinear/sesquilinear form on L, where for a degenerate or pure vector space we define $f(v, w) := 0$ for all $v, w \in L$ and for an orthogonal space f is the bilinear form associated to the quadratic form on L. Define

$$L_A^\perp := \{v \in L : f(v, w) = 0 \text{ for all } w \in \operatorname{acl}(A) \cap L\}$$

or, in the polar case,

$$\begin{aligned} L_A^\perp := \{v \in V : f(v, w) = 0 \text{ for all } w \in \operatorname{acl}(A) \cap W\} \\ \cup \{v \in W : f(v, w) = 0 \text{ for all } w \in \operatorname{acl}(A) \cap V\}. \end{aligned}$$

Let $L_A^\perp/(L_A^\perp \cap \operatorname{acl}(A))$ be the quotient space, in the usual sense of a quotient of abelian groups. (This makes sense by Remark 4.1.5.) Then P/A is the projectivisation of $L_A^\perp/(L_A^\perp \cap \operatorname{acl}(A))$; that is, let $\sim$ be as in Definition 4.1.4 and then quotient $L_A^\perp/(L_A^\perp \cap \operatorname{acl}(A))$ by $\sim$.

We are now ready to state the definition of Lie coordinatisation itself:

Definition 4.1.9 (Lie coordinatisation). Let $\mathcal{M}$ be an $\mathcal{L}$-structure. A *Lie coordinatisation* of $\mathcal{M}$ is an $\mathcal{L}_\varnothing$-definable partial order $<$ of M that forms a tree of finite height with an $\mathcal{L}_\varnothing$-definable root w such that the following condition holds: For every $a \in M \setminus \{w\}$ either the immediate predecessor u of a has only finitely many immediate successors (which implies $a \in \operatorname{acl}(u)$) or, if $a \notin \operatorname{acl}(u)$, then there exist $b < a$ and an $\mathcal{L}_b$-definable projective Lie geometry J fully embedded in $\mathcal{M}$ such that either

(i) $a \in J$ or, if $a \notin J$, then
(ii) there exist $c \in M$ with $b < c < a$ and an $\mathcal{L}_c$-definable affine Lie geometry (V, A) fully embedded in $\mathcal{M}$ such that $a \in A$, the projectivisation of V is J, and $J < V < A$,

where for subsets $X, Y \subset M$ the notation $X < Y$ means that every element of X lies in a lower level of the tree than every element of Y. We call the Lie geometries J and (V, A) *coordinatising geometries*. By a *Lie coordinatised structure* we mean a structure equipped with a Lie coordinatisation.

Definition 4.1.10 (Lie coordinatisability). An $\mathcal{L}$-structure $\mathcal{M}$ is *Lie coordinatisable* if it is $\varnothing$-bi-interpretable (see § 2.5 of [1] or § 2.4 of [43]) with a Lie coordinatised structure that has finitely many 1-types.

Remark 4.1.11.

(i) We have actually defined so-called 'weak Lie coordinatisability' (p. 17 of [14]), since in Definition 4.1.9 we did not stipulate the orientation condition relating to quadratic coordinatising geometries (Definition 2.1.10 in [14]). This condition is important and cannot be ignored in general, but we can ignore it because we do not need to consider quadratic Lie geometries (Remark 4.1.2(ii)). For brevity we thus suppress the prefix 'weak', the sketch proof of Theorem 4.4.1 being an exception. Note that the orientation condition is also ignored in [15] for the same reason (p. 517 of [15]).

Fact 4.1.12. *If $\mathcal{M}$ is Lie coordinatisable, then $\mathcal{M}$ is $\aleph_0$-categorical.*

Proof. Lemma 2.3.19 in [14]. □

Remark 4.1.13. The distinction between Lie coordinatisation and Lie coordinatisability is important to maintain in general, but we freely move from the latter to the former by adding finitely many sorts from $\mathcal{M}^{\mathrm{eq}}$ to $\mathcal{M}$.

The following is one of the deep results of [14]:

Fact 4.1.14. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Then $\mathcal{M}$ is Lie coordinatisable if and only if $\mathcal{M}$ is smoothly approximable.*

Proof. Theorem 2 in [14]. □

4.2. Examples. We give two examples of Lie coordinatisable structures, returning to Examples 3.1.4 and 3.1.5. This is no coincidence, as shown by Fact 4.1.14.

Example 4.2.1. Consider a language $\mathcal{L} := \{I_1, I_2\}$, where I_1 and I_2 are binary relation symbols. Let $\mathcal{M}$ be a countable $\mathcal{L}$-structure where $I_1^{\mathcal{M}}$ and $I_2^{\mathcal{M}}$ are equivalence relations such that $I_1^{\mathcal{M}}$ has infinitely many classes, $I_2^{\mathcal{M}}$ refines $I_1^{\mathcal{M}}$, every I_1-equivalence class contains infinitely many I_2-equivalence classes, and every I_2-equivalence class is infinite; that is, $\mathcal{M}$ is partitioned into infinitely many I_1-equivalence classes, each of which is then partitioned into infinitely many I_2-equivalence classes, each of which is infinite. We claim that $\mathcal{M}$ is Lie coordinatisable.

We first outline the tree structure. At the root we place $\ulcorner \mathcal{M} \urcorner$ (the canonical parameter of $\mathcal{M}$ in $\mathcal{M}^{\mathrm{eq}}$, which is $\varnothing$-definable), above which we place the I_1-classes, as imaginary elements of $\mathcal{M}^{\mathrm{eq}}$. Above each I_1-class we then place the

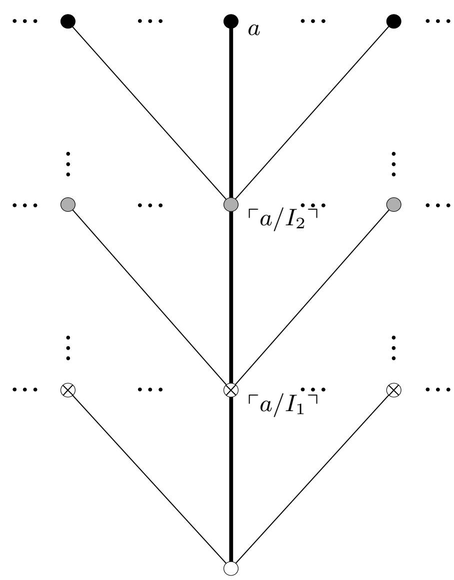

FIGURE 1. A finite fragment of the tree from Example 4.2.1, with the branch leading to the element a in bold. The nodes are shaded according to membership: The white node is $\ulcorner \mathcal{M} \urcorner$, the crossed nodes are elements of $\mathcal{M}/I_1$, the grey nodes are elements of $(a/I_1)/I_2$, and the black nodes are elements of a/I_2. The small dots represent the rest of the tree.

I_2-classes, again as imaginary elements of $\mathcal{M}^{\text{eq}}$, with every I_2-class above the I_1-class in which the I_2-class is contained. Finally, above each I_2-class we place the elements of $\mathcal{M}$ contained in that I_2-class. So this tree has height 3 and infinite width at each level.

Let's explain the notation used in Figure 1. So consider some arbitrary $a \in M$. For $j = 1$ or 2, let a/I_j denote the I_j-class that contains a and let $\ulcorner a/I_j \urcorner$ denote the same I_j-class but as a member of $\mathcal{M}^{\text{eq}}$; so $\ulcorner a/I_j \urcorner \in \mathcal{M}^{\text{eq}}$ is a canonical parameter for the a-definable subset $a/I_j \subset M$. We define $(a/I_1)/I_2$ and $\ulcorner (a/I_1)/I_2 \urcorner$ similarly.

We now use this notation to check that Definition 4.1.9 holds for the tree. The imaginary element $\ulcorner a/I_1 \urcorner$ lies in the $\ulcorner \mathcal{M} \urcorner$-definable degenerate projective geometry $\mathcal{M}/I_1$ and $\ulcorner \mathcal{M} \urcorner < \ulcorner a/I_1 \urcorner$. The imaginary element $\ulcorner a/I_2 \urcorner$ lies in the $\ulcorner a/I_1 \urcorner$-definable degenerate projective geometry $(a/I_1)/I_2$ and $\ulcorner a/I_1 \urcorner < \ulcorner a/I_2 \urcorner$. Finally, the real element a lies in the $\ulcorner a/I_2 \urcorner$-definable degenerate projective geometry a/I_2 and $\ulcorner a/I_2 \urcorner < a$. Adjoining a finite number of sorts from $\mathcal{M}^{\text{eq}}$ (recall Remark 4.1.13), each of these geometries is fully embedded in $\mathcal{M}$. (Note that $\mathcal{M}/I_2$ is *not* fully embedded, since I_1 defines extra structure on $\mathcal{M}/I_2$ that is not definable within $\mathcal{M}/I_2$ using equality alone.) So $\mathcal{M}$ is indeed Lie coordinatisable.

Remark 4.2.2. This example generalises to the case where we have n equivalence relations $I_1, \dots, I_n$ such that there are infinitely many I_1-classes, I_{j+1} refines I_j and every I_j-class contains infinitely many I_{j+1}-classes (for $1 \leq j \leq n-1$), and every I_n-class is infinite. At the base of the tree (the 0^{th} level) we place $\ulcorner \mathcal{M} \urcorner$. At the j^{th} level (for $1 \leq j \leq n-1$) we place the I_j-classes, as imaginary elements of $\mathcal{M}^{\text{eq}}$, with every I_j-class above the I_{j-1}-class in which the I_j-class is contained. Finally, at the top of the tree (the n^{th} level) we place the elements of M, with each $a \in M$ placed above $\ulcorner a/I_n \urcorner$.

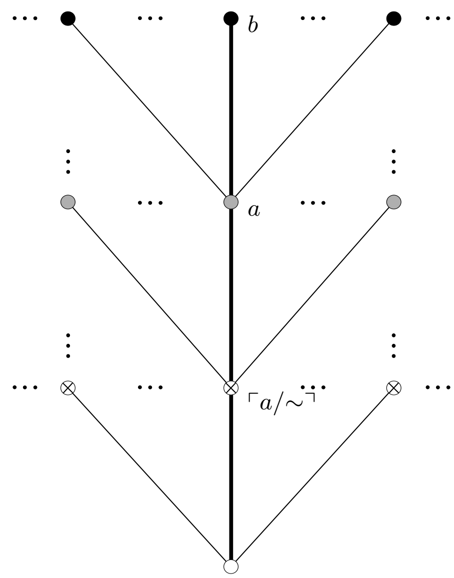

Figure 2. A finite fragment of the tree from Example 4.2.3, with the branch leading to the element $b \in \mathcal{M}_a$ in bold. The nodes are shaded according to membership: The white node is the zero vector, the crossed nodes are elements of $P(\mathcal{M}_0)$, the grey nodes are elements of $a/\sim$, and the black nodes are elements of $\mathcal{M}_a$. Note that there are only finitely many (in fact $p-1$) nodes immediately above each crossed node. The small dots represent the rest of the tree.

Example 4.2.3 (Example 2.1.11 in [14]). Let $\mathcal{L} := \{0, +\}$ and let p be a fixed prime number. (The case $p = 2$ *is* allowed.) We define M to be the direct sum of ω-many copies of $\mathbb{Z}/p^2\mathbb{Z}$, i.e.

$$M := \{(a_i)_{i<\omega} : a_i \in \mathbb{Z}/p^2\mathbb{Z} \text{ and } a_i = 0 \text{ for all but finitely many } i\}.$$

(We specify the direct sum because it is countable, unlike the direct product.) The set M naturally forms an $\mathcal{L}$-structure $\mathcal{M}$, the $\mathcal{L}$-structure arising componentwise from the $\mathcal{L}$-structure of the group $\mathbb{Z}/p^2\mathbb{Z}$. Explicitly: $0^{\mathcal{M}} := (0)_{i<\omega}$ and

$(a_i)_{i<\omega} + (b_i)_{i<\omega} := (a_i + b_i)_{i<\omega}$. For brevity we write 0 for $0^{\mathcal{M}}$. We claim that $\mathcal{M}$ is Lie coordinatisable.

We first introduce some notation: For $v \in M$ let $\mathcal{M}_v := \{a \in M : pa = v\}$, where $pa := \underbrace{a + a + \cdots + a}_{p \text{ times}}$. Observe that $\mathcal{M}_0$ has a vector space structure over $\mathbb{F}_p$ and thus is a linear Lie geometry over $\mathbb{F}_p$. Let $P(\mathcal{M}_0)$ be the projectivisation of $\mathcal{M}_0$ (Definition 4.1.4). Then $P(\mathcal{M}_0) = \mathcal{M}_0 \setminus \{0\}/{\sim}$, where $a \sim b$ if and only if $a = rb$ for some $r \in \mathbb{F}_p$ (recall Remark 4.1.5). So $|a/{\sim}| = p - 1$ for all $a \in \mathcal{M}_0$. Adjoining a sort for $P(\mathcal{M}_0)$ (recall Remark 4.1.13), we also have that $P(\mathcal{M}_0)$ is fully embedded in $\mathcal{M}$.

We now outline the tree structure. At the root we place 0, above which we place the elements of $P(\mathcal{M}_0)$, considered as imaginary elements of $\mathcal{M}^{\mathrm{eq}}$. On the next level we place the elements of $\mathcal{M}_0 \setminus \{0\}$, with each a placed above $\ulcorner a/{\sim} \urcorner$. Finally, the top level contains the elements of $\mathcal{M} \setminus \mathcal{M}_0$, with each $b \in \mathcal{M}_a$ placed above a. So we have a tree of height 3 and infinite width at each level, although the second level comprises an infinite amount of finite branching. Note that we're using the fact here that if $b \in \mathcal{M} \setminus \mathcal{M}_0$, then $b \in \mathcal{M}_a$ for some $a \in \mathcal{M}_0$. Proof: Suppose $pb \neq 0$. So $pb = a$ for some $a \in M$. Then $pa = p(pb) = p^2 b = 0$, since $p^2 c = 0$ for all $c \in M$. So $b \in \mathcal{M}_a$ and $a \in \mathcal{M}_0$, as required.

Let's check that Definition 4.1.9 holds for this tree. So consider some arbitrary non-zero $a \in \mathcal{M}_0$ and $b \in \mathcal{M}_a$. See Figure 2 for an illustration. The imaginary element $\ulcorner a/{\sim} \urcorner$ lies in the 0-definable projective geometry $P(\mathcal{M}_0)$, which is fully embedded, as noted in the previous paragraph, and $0 < \ulcorner a/{\sim} \urcorner$. The real element a is algebraic over $\ulcorner a/{\sim} \urcorner$, since $a/{\sim}$ is $\ulcorner a/{\sim} \urcorner$-definable and finite, again as noted in the previous paragraph, and $\ulcorner a/{\sim} \urcorner < a$. This leaves us with the top level of the tree, which we deal with in the next paragraph.

Firstly, observe that $0 < a < b$. The real element a defines an affine geometry $(\mathcal{M}_0, \mathcal{M}_a)$, where $\mathcal{M}_0$ is the $\mathbb{F}_p$-vector space, $\mathcal{M}_a$ is the $\mathcal{M}_0$-affine space, and the action $\mathcal{M}_0 \times \mathcal{M}_a \to \mathcal{M}_a$ is given by $(u, v) \mapsto u + v$. (This action is well-defined, since $p(u + v) = pu + pv = 0 + a = a$ and so $u + v \in \mathcal{M}_a$.) As we have already noted, the projectivisation of $\mathcal{M}_0$ is $P(\mathcal{M}_0)$, which is a fully embedded, 0-definable projective geometry, and we have $b \in \mathcal{M}_a$ by assumption. So the tree structure does indeed satisfy the definition of Lie coordinatisation.

Remark 4.2.4. This example generalises to the direct sum of ω-many copies of $\mathbb{Z}/p^n\mathbb{Z}$, for any $n \in \mathbb{N}^+$. When $n = 1$, the tree structure is the same as in the case $n = 2$, except that $\mathcal{M}_0 \setminus \{0\}$ forms the top level, since $\mathcal{M} \setminus \mathcal{M}_0 = \varnothing$. When $n \geq 3$, the first three levels (0, $P(\mathcal{M}_0)$ and $\mathcal{M}_0$) are the same, but at the third level one places the elements of $\{b \in M : b \in \mathcal{M}_a \text{ for some } a \in \mathcal{M}_0\}$, instead of simply $\mathcal{M} \setminus \mathcal{M}_0$, and at the $(j+1)^{\text{th}}$ level (for $1 \leq j \leq n$) one places $\{c \in M : c \in \mathcal{M}_b \text{ for some } b \text{ in the } j^{\text{th}} \text{ level}\}$. The $(n+1)^{\text{th}}$ level is the upper-most level.

Remark 4.2.5. In both Example 4.2.1 and Example 4.2.3 the tree is nicely stratified, namely root–degenerate–degenerate–degenerate in the former and root–projective–algebraic–affine in the latter. This need not be the case, however: There are Lie coordinatising trees containing maximal chains of different lengths.

For example, one could take the disjoint union (in a suitable language, with a common root) of two Lie coordinatising trees of different heights.

4.3. Standard systems of geometries and envelopes. We develop the key notion of an envelope of a Lie coordinatised structure. Our presentation is a simplified version of that given in [14], streamlined for the purpose of stating and proving Proposition 4.5.2. We begin with the notion of a standard system of geometries:

Definition 4.3.1 (Standard system of geometries). Let $\mathcal{M}$ be a Lie coordinatised $\mathcal{L}$-structure. A *standard system of geometries* on $\mathcal{M}$ is a $\varnothing$-definable function $J\colon A \to \mathcal{P}(M)$ whose domain A is the set of realisations of a 1-type over $\varnothing$ in $\mathcal{M}$ and whose image is a set of projective coordinatising Lie geometries of the same kind, i.e. $J(a)$ and $J(b)$ are isomorphic for every $a, b \in A$. By '$\varnothing$-definable' we mean that there is an $\mathcal{L}$-formula $\varphi(x, y)$ such that $\varphi(\mathcal{M}, a) = J(a)$ for every $a \in A$. We call A the *domain* of J, which we denote by $\operatorname{dom}(J)$.

Definition 4.3.2 (Dimension function).

(i) Let J be a Lie geometry over a field K. An *approximation* of J is a finite-dimensional geometry over K of the same kind as J. For example, if J is the projectivisation of a pure vector space over a finite field K, then an approximation of J is a finite-dimensional projective space over K, or if J is a degenerate space, then an approximation of J is a finite set in the language of equality.

(ii) Let $\mathcal{M}$ be a Lie coordinatised structure. A *dimension function* is a function μ on a finite set of S of standard systems of geometries on $\mathcal{M}$ that assigns an approximation to each $J \in S$, i.e. $\mu(J)$ is an approximation of $J(a)$ for some $a \in \operatorname{dom}(J)$. (This is independent of the choice of a, since $J(a)$ is the same kind of projective Lie geometry for every $a \in \operatorname{dom}(J)$.) We call S the *domain* of μ, which we denote by $\operatorname{dom}(\mu)$.

Definition 4.3.3 (μ-Envelope). Let $\mathcal{M}$ be a Lie coordinatised structure. Then a *μ-envelope* is a pair (E, μ) consisting of a finite subset $E \subset M$ and a dimension function μ for which the following three conditions holds:

(i) E is algebraically closed in $\mathcal{M}$. (Note that this implies that E is a substructure of $\mathcal{M}$.)

(ii) For every $a \in M \setminus E$ there exist $J \in \operatorname{dom}(\mu)$ and $b \in \operatorname{dom}(J) \cap E$ such that $\operatorname{acl}(E) \cap J(b)$ is a proper subset of $\operatorname{acl}(E, a) \cap J(b)$.

(iii) For every $J \in \operatorname{dom}(\mu)$ and for any $b \in \operatorname{dom}(J) \cap E$, $J(b) \cap E$ and $\mu(J)$ are isomorphic.

Remark 4.3.4.

(i) We often denote a μ-envelope by E, rather than (E, μ), leaving the dimension function as implicit. We similarly often use the term 'envelope', rather than 'μ-envelope'.

(ii) It may help the reader's intuition to know that envelopes form homogeneous substructures of $\mathcal{M}$ (Lemma in 3.2.4 [14]). Indeed, this is how the left-to-right direction of Fact 4.1.14 is proved (pp. 61–62 of [14]).

(iii) In general one can have countably infinite approximations and envelopes, but we do not need to consider them.

The following definition is fundamental to the work in § 4.5 and arises from Definitions 3.1.1.4(5) and 5.2.1 in [14]:

Definition 4.3.5. Consider a μ-envelope (E, μ), where $\mathrm{dom}(\mu) = \{J_1, \ldots, J_s\}$. For each J_i we define $d_E(J) := \dim \mu(J)$, where $\dim \mu(J) := |\mu(J)|$ if $\mu(J)$ is a pure set. We further define $d^*_E(J) := (-\sqrt{q})^{d_E(J)}$, where q is the size of the base finite field of $\mu(J)$, or, if $\mu(J)$ is a pure set, then we define $d^*_E(P) := d_E(P)$. (Taking $-\sqrt{q}$, rather than just q, does initially look strange. It is done purely for unitary spaces: see the end of the proof of Proposition 4.5.2.) Finally, we define $\bar{d}^*(E) := (d^*_E(J_1), \ldots, d^*_E(J_s))$.

We illustrate the preceding definitions by returning to Examples 4.2.1 and 4.2.3:

Example 4.3.6 (continuation of Example 4.2.1). Put simply, an example of an envelope in this case is a subset $E \subseteq \mathcal{M}$ that intersects a fixed number (n_1) of I_1-classes, a fixed number (n_2) of I_2-classes within each of these I_1-classes, and a fixed number (n_3) of elements within each of these I_2-classes. So, up to isomorphism, an envelope is given by a triple (n_1, n_2, n_3). Using the enumerations from Example 3.1.4, two examples of envelopes are

$$E_1 := \{a_{ijk} : 1 \leq i \leq 3, 1 \leq j \leq 6, 1 \leq k \leq 1\}$$

$$\text{and } E_2 := \{a_{ijk} : 19 \leq i \leq 21, 3 \leq j \leq 8, 2015 \leq k \leq 2015\}.$$

The triple for both E_1 and E_2 is $(n_1, n_2, n_3) = (3, 6, 1)$. Let's explain this in terms of standard systems of geometries and dimension functions.

Consider the following three standard systems of geometries:

- $J_\alpha \colon \mathcal{M} \to \mathcal{M}^{\mathrm{eq}}$, $J_\alpha(a) := \ulcorner \mathcal{M}/I_1 \urcorner$;
- $J_\beta \colon \mathcal{M} \to \mathcal{M}^{\mathrm{eq}}$, $J_\beta(a) := \ulcorner (a/I_1)/I_2 \urcorner$; and
- $J_\gamma \colon \mathcal{M} \to \mathcal{M}^{\mathrm{eq}}$, $J_\gamma(a) := \ulcorner a/I_2 \urcorner$.

A dimension function μ on $\{J_\alpha, J_\beta, J_\gamma\}$ assigns an approximation to each of $J_\alpha(a)$, $J_\beta(a)$ and $J_\gamma(a)$, where $a \in M$ is arbitrary. An approximation of a given geometry is determined by the dimension of the approximation, which in this case is equal to the size of the approximation, since all the projective geometries are degenerate. Thus μ is determined by a choice of triple (n_1, n_2, n_3), as mentioned in the previous paragraph. So, if $\mu = (n_1, n_2, n_3)$, then a μ-envelope is an envelope with associated triple (n_1, n_2, n_3). Furthermore, again because all the projective geometries in this example are degenerate, we have $\bar{d}^*(E) = (n_1, n_2, n_3)$.

Example 4.3.7 (continuation of Example 4.2.3). Consider the standard system of geometries $J \colon \{0\} \to \mathcal{M}^{\mathrm{eq}}$, where $J(0) := P(\mathcal{M}_0)$. A dimension function μ on J assigns an approximation to $P(\mathcal{M}_0)$, i.e. a finite-dimensional subspace of $P(\mathcal{M}_0)$. So, as in Example 4.2.1, μ is determined by an integer. A μ-envelope is then a finite power of $\mathbb{Z}/p^2\mathbb{Z}$; that is, a subset $E := \{(a_i)_{i<\omega} \in M : a_i \neq 0 \text{ only if } i = t_j \text{ for some } j\}$ given by n integers $t_1, \ldots, t_n \in \mathbb{N}^+$, where $n := \dim \mu(J)$. Thus E is determined, up to isomorphism, by n. Since the base field is $\mathbb{F}_p$, which has size p, we have $\bar{d}^*(E) = (-\sqrt{p})^n$. This is a 1-tuple because there is only one standard system in the domain of μ.

4.4. Macpherson's conjecture, short version. We now take a big step towards proving Theorem 4.6.4 by proving a shorter version, namely Corollary 4.4.2, where the existence of a multidimensional exact class is asserted but the nature of the measuring functions is not specified. We first provide a sketch proof of part 2 of Theorem 6 in [14], as this result is crucial to our proof of Corollary 4.4.2. The key ingredients needed to prove the result are contained in [14], namely Propositions 4.4.3, 4.5.1 and 8.3.2 and their proofs, but the (non-trivial) argument putting them together is not made completely explicit. We state the result in a way that is convenient for our present purposes, but it is essentially the same as the original statement in [14], the only significant difference being the use of the equivalence of Lie coordinatisation and smooth approximation (Fact 4.1.14).

Theorem 4.4.1. *Let $\mathcal{L}$ be a finite language and let $d \in \mathbb{N}^+$. Define $\mathcal{C}(\mathcal{L}, d)$ to be the class of all finite $\mathcal{L}$-structures with at most d 4-types. Then there is a finite partition $\mathcal{F}_1, \dots, \mathcal{F}_k$ of $\mathcal{C}(\mathcal{L}, d)$ such that the $\mathcal{L}$-structures in each $\mathcal{F}_i$ smoothly approximate an $\mathcal{L}$-structure $\mathcal{F}_i^*$. Moreover, the $\mathcal{F}_i$ are definably distinguishable: For each $\mathcal{F}_i$ there exists an $\mathcal{L}$-sentence χ_i such that for all $\mathcal{M} \in \mathcal{C}(\mathcal{L}, d)$ above some minimum size, $\mathcal{M} \models \chi_i$ if and only if $\mathcal{M} \in \mathcal{F}_i$.*

Sketch proof.[6] We first show that there cannot exist infinitely many pairwise elementarily inequivalent Lie coordinatisable $\mathcal{L}$-structures with the same skeletal type, where a skeletal type is, roughly speaking, a full description of the Lie coordinatising tree structure in an extended language $\mathcal{L}_{sk}$; see § 4.2 of [14] for the full definition. So, for a contradiction, suppose that there are in fact infinitely many such $\mathcal{L}$-structures $\{\mathcal{N}_i : i < \omega\}$ with the same skeletal type S. Working in $\mathcal{L}_{sk}$, by a judicious choice of ultrafilter we can take a non-principal ultraproduct $\mathcal{N}^*$ of the $\mathcal{N}_i$ such that $\mathcal{N}^* \not\equiv \mathcal{N}_i$ for all $i < \omega$. We may assume that $\mathcal{N}^*$ is countable by moving to a countable elementary substructure. Since the skeletal type S is expressible in $\mathcal{L}_{sk}$ (this is a general fact of skeletal types, not just S) and true in each $\mathcal{N}_i$, by Łos's theorem $\mathcal{N}^*$ is Lie coordinatised and has skeletal type S. Work in chapter 4 of [14], especially Proposition 4.4.3 and its proof, shows that every Lie coordinatised structure is quasifinitely axiomatised, and thus in particular $\mathcal{N}^*$ is quasifinitely axiomatised. Put roughly, this means that $\mathrm{Th}(\mathcal{N}^*)$ is axiomatised by a sentence σ and an axiom schema of infinity, where we consider $\mathrm{Th}(\mathcal{N}^*)$ as an $\mathcal{L}'$-theory in a finite language $\mathcal{L}'$ containing $\mathcal{L}_{sk}$. This axiom schema of infinity holds for all the $\mathcal{N}_i$ because they each have the same skeletal type as $\mathcal{N}^*$. Furthermore, again by Łos's theorem, there exists $j < \omega$ such that $\mathcal{N}_j \models \sigma$. Thus $\mathcal{N}^* \equiv \mathcal{N}_j$, a contradiction.

We now return to the original class $\mathcal{C} := \mathcal{C}(\mathcal{L}, d)$. We take an infinite ultraproduct $\mathcal{U}^*$ of the structures in $\mathcal{C}$. We take this ultraproduct in a non-standard model of set theory, working with some suitable Gödel coding of formulas, which

[6] The main argument was given by Hrushovski in email correspondence and Macpherson provided essential input by working out key details. The contribution of the present author lay in working through further details and writing up the proof.

allows us to consider $\mathcal{U}^*$ as an $\mathcal{L}^*$-structure, where $\mathcal{L}^*$ is the ultrapower of the language $\mathcal{L}$; that is, $\mathcal{L}^*$ extends $\mathcal{L}$ by including infinitary formulas with non-standard Gödel numbers, although the number of free variables in any given formula remains finite. We may again assume that $\mathcal{U}^*$ is countable by moving to a countable elementary substructure. $\mathcal{U}^*$ is 4-quasifinite (Definition 2.1.1 in [14]) and thus by Theorem 3 in [14] is weakly Lie coordinatisable. So by Proposition 7.5.4 in [14] the $\mathcal{L}$-reduct $\mathcal{U}$ of $\mathcal{U}^*$ is also weakly Lie coordinatisable. The $\mathcal{L}$-structure $\mathcal{U}$ thus has a skeletal type. By the first part of the proof there can be only finitely many pairwise elementarily inequivalent Lie coordinatisable $\mathcal{L}$-structures with this skeletal type, say $\mathcal{F}_1^*, \dots, \mathcal{F}_k^*$. By Proposition 4.4.3 each $\mathcal{F}_i^*$ has a characteristic sentence, say χ_i. The χ_i yield a partition $\mathcal{C} = \mathcal{F}_1 \cup \ldots \cup \mathcal{F}_k$, where each χ_i is true in all $\mathcal{M} \in \mathcal{F}_i$ and false in all $\mathcal{M} \in \mathcal{F}_j$ for $j \neq i$, potentially with the exception of some small structures. Moreover, again by Proposition 4.4.3, this partition is such that each $\mathcal{M} \in \mathcal{F}_i$ is an envelope of $\mathcal{F}_i^*$ and so by work in chapter 3 of [14] the structures in $\mathcal{F}_i$ smoothly approximate $\mathcal{F}_i^*$.

Note that the work cited from chapter 4 of [14] is written in terms of Lie coordinatisability, but inspection of the proofs shows that weak Lie coordinatisability suffices. □

Corollary 4.4.2 (Macpherson's conjecture, short version). *For any countable language $\mathcal{L}$ and any $d \in \mathbb{N}^+$ there exists R such that the class $\mathcal{C}(\mathcal{L}, d)$ of all finite $\mathcal{L}$-structures with at most d 4-types is an R-mec in $\mathcal{L}$.*

Proof. Let $\mathcal{C} := \mathcal{C}(\mathcal{L}, d)$. The reader should recall Remark 2.1.3(v), as we will use it at various points in this proof.

First suppose that $\mathcal{L}$ is finite. By Theorem 4.4.1, $\mathcal{C}$ can be finitely partitioned into subclasses $\mathcal{F}_1, \dots, \mathcal{F}_k$ such that the structures in each $\mathcal{F}_i$ smoothly approximate an $\mathcal{L}$-structure $\mathcal{F}_i^*$. Thus by Proposition 3.2.1 each $\mathcal{F}_i$ is an R_i-mec in $\mathcal{L}$ for some R_i. Let $R_{\mathcal{L}} := R_1 \cup \dots \cup R_k$. We claim that $\mathcal{C}$ is an $R_{\mathcal{L}}$-mec in $\mathcal{L}$.

We prove this claim: Let $\varphi(\bar{x}, \bar{y})$ be an $\mathcal{L}$-formula with $n := l(\bar{x})$ and $m := l(\bar{y})$. Since each $\mathcal{F}_i$ is an R_i-mec, we have a suitable finite partition Φ_i of each $\mathcal{F}_i(m)$. Then $\Phi_1 \cup \dots \cup \Phi_k$ is a finite partition of $\mathcal{C}(m)$ and so $\mathcal{C}$ is a weak $R_{\mathcal{L}}$-mec in $\mathcal{L}$. It remains to show that the definability clause holds. We again use Theorem 4.4.1: For each $\mathcal{F}_i$ there is an $\mathcal{L}$-sentence χ_i such that $\mathcal{M} \models \chi_i$ if and only if $\mathcal{M} \in \mathcal{F}_i$, for sufficiently large $\mathcal{M}$. So, by conjoining χ_i to the defining $\mathcal{L}$-formulas of each Φ_i, we satisfy the definability clause, using Lemma 2.2.3 to deal with the finite number of potential exceptions. So the claim is proved.

Now suppose that $\mathcal{L}$ is infinite. Consider some arbitrary finite $\mathcal{L}' \subset \mathcal{L}$ and let $\mathcal{C}_{\mathcal{L}'}$ denote the class of all $\mathcal{L}'$-reducts of structures in $\mathcal{C}$. Each structure in $\mathcal{C}_{\mathcal{L}'}$ has at most d 4-types, since a reduct cannot have more types than the original structure. Thus, by the first part of the proof, $\mathcal{C}_{\mathcal{L}'}$ is an $R_{\mathcal{L}'}$-mec in $\mathcal{L}'$. (It could be the case that $\mathcal{C}_{\mathcal{L}'}$ is a proper subclass of the class of all finite $\mathcal{L}'$-structures with at most d 4-types, but that wouldn't matter, since a subclass of an R-mec is also an R-mec.) Let $\mathbb{L}$ be the set of all finite subsets of $\mathcal{L}$ and define

$$R := \bigcup_{\mathcal{L}' \in \mathbb{L}} R_{\mathcal{L}'}.$$

Then each $\mathcal{C}_{\mathcal{L}'}$ is an R-mec in $\mathcal{L}'$ by Remark 2.1.3(v). Therefore $\mathcal{C}$ is an R-mec in $\mathcal{L}$ by Lemma 2.2.5. □

Remark 4.4.3. The reader may well be wondering what's so special about 4-types. Well, firstly, if there is a bound on the number of n-types, then there is a bound on the number of k-types for all $k \leq n$. So in the statement of Theorem 4.6.4 we could replace 4-types with n-types for any $n > 4$ and the result would still go through. As for 4 itself, the explanation goes deeper and we will not go into detail. However, put *very* roughly, the number 4 arises because the projective linear group preserves the cross-ratio, which is a projective invariant on 4-tuples of colinear points. The classification of finite simple groups also plays a role. Details can be found in § 6 of [1], [24] and [34]. Note that in [34] the original bound on 5-types, as given in [24], is improved to one on 4-types.

4.5. Definable sets in envelopes. Corollary 4.4.2 provides no information about the structure of R, only its existence. In this section we use Lie geometries to ascertain information about the nature of R. We first need to define a rank:

Definition 4.5.1 (Definition 2.2.1 in [14]). Let $\mathcal{M}$ be an $\mathcal{L}$-structure and let $D \subseteq M$ be a parameter-definable set. We define the *CH-rank* of D as follows:

(i) $\text{rk}(D) > 0$ if and only if D is infinite.
(ii) For $n \in \mathbb{N}$, $\text{rk}(D) \geq n+1$ if and only if there exist parameter-definable subsets $D_1, D_2 \subset M$ and functions $\pi\colon D_1 \to D$ and $f\colon D_1 \to D_2$ such that
  - $\text{rk}(\pi^{-1}(d)) = 0$ for all $d \in D$;
  - $\text{rk}(D_2) > 0$; and
  - $\text{rk}(f^{-1}(d)) \geq n$ for all $d \in D_2$.

If $\text{rk}(D) > n$ for all $n \in \mathbb{N}$, then we define $\text{rk}(D) = \infty$.

The following result provides us with information about the sizes of definable sets in envelopes, which we will then use in §4.6 to shed light on the structure of R in Corollary 4.4.2. It uses Definition 4.3.5 and is a generalisation of Proposition 5.2.2 in [14]. Proposition 5.2.2 is essentially about the formula $x = x$, since it concerns the sizes of envelopes, rather than the sizes of definable subsets of envelopes, and arbitrary formulas with parameters arise only as part of the proof. In contrast, Proposition 4.5.2 concerns arbitrary formulas with parameters from the outset and so more complexity arises. We also go into considerably more detail on certain points than in the proof given in [14].

Proposition 4.5.2. *Let $\mathcal{E}$ be an ordered family of envelopes of a Lie coordinatised $\mathcal{L}$-structure $\mathcal{M}$ such that $\text{dom}(\mu) = \text{dom}(\mu')$ for all $(E, \mu), (E', \mu') \in \mathcal{E}$ and such that the parity and signature of orthogonal spaces are constant on the family, where by 'ordered family' we mean that for all $(E, \mu), (E', \mu') \in \mathcal{E}$ either $E \subseteq E'$ or $E' \subseteq E$. Let $\bar{a} \in M^m$ (where m is arbitrary), let $D_{\bar{a}} \subseteq M$ be an $\mathcal{L}_{\bar{a}}$-definable set and let s be the size of the common domian of the dimension functions. Then there exists a polynomial $\rho \in \mathbb{Q}[\mathbf{X}_1, \ldots, \mathbf{X}_s]$ and an integer $Q \in \mathbb{N}$ such that $|D_{\bar{a}} \cap E| = \rho(\bar{d}^*(E))$ for all $(E, \mu) \in \mathcal{E}$ with $|E| > Q$ and $\bar{a} \in E^m$.*

Remark 4.5.3. We offer a brief explanation of the parity/signature assumption; full details can be found in § 21 of [2]. The parity of a finite-dimensional orthogonal space V refers to $\dim(V)$, distinguishing between odd and even dimension. The signature refers to the quadratic form on V, there being only two possibilities (up to equivalence); in the even-dimensional case this is determined by the Witt index and in the odd-dimensional case by the hyperbolic hyperplane. The assumption is important, but its use is restricted to the calculations at the end of the proof, and there only in the orthogonal case.

Proof of Proposition 4.5.2. Let $\varphi(x, \bar{a})$ be the $\mathcal{L}_{\bar{a}}$-formula that defines $D_{\bar{a}}$. So $D_{\bar{a}} = \varphi(\mathcal{M}, \bar{a})$. By Fact 4.1.12 and the Ryll-Nardzewski Theorem we may assume without loss of generality that $\varphi(x, \bar{a})$ defines the set of realisations of a 1-type $r(x)$ over $\bar{a}$ in $\mathcal{M}$. So $D_{\bar{a}} = r(\mathcal{M})$. Also note that since $\mathcal{E}$ is ordered by $\subseteq$, either $|D_{\bar{a}} \cap E| = \varnothing$ for all $(E, \mu) \in \mathcal{E}$ or there exists $Q \in \mathbb{N}$ such that $|D_{\bar{a}} \cap E| \neq \varnothing$ for all $(E, \mu) \in \mathcal{E}$ with $|E| > Q$. In the former case we can set $Q := 0$ and $\rho := 0$. So we henceforce assume that we are in the latter case. With these two assumptions in hand, we are now in a position to start the main line of argument. We proceed by induction on CH-rank.

First suppose that $\mathrm{rk}(D_{\bar{a}}) = 0$. Then $D_{\bar{a}}$ is finite. Let $k := |D_{\bar{a}}|$. Since $D_{\bar{a}}$ is both finite and $\bar{a}$-definable, $D_{\bar{a}} \subseteq \mathrm{acl}(\bar{a})$. Thus, since envelopes are algebraically closed (Definition 4.3.3), $D_{\bar{a}} \subseteq E$ for all $E \in \mathcal{E}$ with $\bar{a} \in E^m$. So $|D_{\bar{a}} \cap E| = |D_{\bar{a}}| = k$ for all $E \in \mathcal{E}$ with $\bar{a} \in E^m$. Hence the constant polynomial $\rho := k$ suffices.

Now consider the case $\mathrm{rk}(D_{\bar{a}}) > 0$. Then $D_{\bar{a}}$ is infinite. Assume as the induction hypothesis that the result holds for any parameter-definable subset of M with CH-rank strictly less than $\mathrm{rk}(D_{\bar{a}})$. Let $d \in D_{\bar{a}}$. For a contradiction, suppose that every step in the tree below d is algebraic; that is, if $c_0 < c_1 < \cdots < c_t = d$ is the chain leading to d, where c_0 is the root of the tree, then each c_{i+1} is algebraic over its immediate predecessor c_i. We claim that $d \in \mathrm{acl}(\varnothing)$.

We prove this claim. We proceed by induction on i to show that $c_i \in \mathrm{acl}(\varnothing)$ for every i, and so in particular $d = c_t \in \mathrm{acl}(\varnothing)$. Since the root is $\varnothing$-definable (Definition 4.1.9), $c_0 \in \mathrm{dcl}(\varnothing) \subseteq \mathrm{acl}(\varnothing)$. Now suppose that $c_i \in \mathrm{acl}(\varnothing)$. Then $c_{i+1} \in \mathrm{acl}(\mathrm{acl}(\varnothing))$, since $c_{i+1} \in \mathrm{acl}(c_i)$ by our supposition. But $\mathrm{acl}(\mathrm{acl}(\varnothing)) = \mathrm{acl}(\varnothing)$, since algebraic closure is idempotent, and hence $c_{i+1} \in \mathrm{acl}(\varnothing)$. So the claim is proved.

We now use the claim to derive a contradiction. Since $d \in \mathrm{acl}(\varnothing)$, there exists some $\mathcal{L}$-formula $\chi(x)$ such that $\mathcal{M} \models \chi(d)$ and $\chi(\mathcal{M})$ is finite. So $\chi(x) \in \mathrm{tp}(d/\varnothing) \subseteq \mathrm{tp}(d/\bar{a}) = r(x)$ and hence $D_{\bar{a}} = r(\mathcal{M}) \subseteq \chi(\mathcal{M})$ is finite, a contradiction.

So by the contradiction there exists $c \leq d$ such that c is not algebraic over its immediate predecessor. Take c to be minimal, i.e. lowest in the tree. By Definition 4.1.9 the non-algebraicity of c implies that c lies in a coordinatising geometry J, where J is b-definable for some $b < c$. The minimality of c implies that J is a projective Lie geometry, since the vector and affine parts of a coordinatising affine Lie geometry lie above the projectivisation of the vector part. Recalling Remark 4.1.2(ii), the same argument applies to quadratic geometries: The affine part Q of a coordinatising quadratic geometry, namely the set of quadratic forms on which the vector part V acts by translation, lies above V in the tree, V being a symplectic space. So the minimality of c implies that J is the projectivisation of V.

Case 1: The element b is the root. Then $b \in \mathrm{dcl}(\varnothing)$ and so J is $\varnothing$-definable. We define a set that is central to our argument:

$$S := \{(c', d') \in M^2 : \mathrm{tp}((c', d')/\bar{a}) = \mathrm{tp}((c, d)/\bar{a})\}.$$

Let S_i be the projection of S to the i^{th} coordinate. Then S_1 is the set of realisations of $\mathrm{tp}(c/\bar{a})$ and S_2 is the set of realisations of $\mathrm{tp}(d/\bar{a})$, as proved in the next paragraph. Then $S_1 \subseteq J$, since $c \in J$ and J is $\varnothing$-definable, and $S_2 = D_{\bar{a}}$, since $\mathrm{tp}(d/\bar{a}) = r(x)$.

We prove the claim that S_1 is the set of realisations of $\mathrm{tp}(c/\bar{a})$: If $c' \in S_1$, then it is immediate from the definition of S that $c' \models \mathrm{tp}(c/\bar{a})$. Now suppose that $c' \models \mathrm{tp}(c/\bar{a})$. By the Ryll-Nardzewski Theorem, $\mathcal{M}$ is saturated and thus there exists $\sigma \in \mathrm{Aut}(\mathcal{M}/\bar{a})$ such that $\sigma(c) = c'$; we'll use this trick several more times and henceforth won't cite it explicitly. Thus $(c', \sigma(d)) = (\sigma(c), \sigma(d)) \in S$ and hence $c' \in S_1$. So the claim is proved. The proof of the claim that S_2 is the set of realisations of $\mathrm{tp}(d/\bar{a})$ proceeds symmetrically.

Let's now consider the intersection of $D_{\bar{a}}$ with an envelope. So take some arbitrary $(E, \mu) \in \mathcal{E}$ with $|E| > Q$ and $\bar{a} \in E$. Since $D_{\bar{a}} \cap E \neq \varnothing$, we may assume without loss of generality that $d \in E$, for if $d \notin E$, then we may take some $d' \in D_{\bar{a}} \cap E$ and repeat the previous arguments for this new element d'.

Define

$$S_E := \{(c', d') \in S : d' \in E\}.$$

We will use this set to calculate the size of $D_{\bar{a}} \cap E$, but we first need to go over some preliminaries. Let S_{Ei} be the projection of S_E to the i^{th} coordinate. Then $S_{E2} = S_2 \cap E = D_{\bar{a}} \cap E$. We claim that $S_{E1} = S_1 \cap E$.

We prove this claim. Let $c' \in S_{E1}$. Then $(c', d') \in S_E$ for some $d' \in E$. Now, $c' \leq d'$ and so $c' \in \mathrm{dcl}(d')$. Thus, since envelopes are algebraically closed (by definition), $c' \in E$. So $c' \in S_1 \cap E$ (since $S_{E1} \subseteq S_1$), as required. Now let $c' \in S_1 \cap E$. Let $d'' \in D \cap E$. Since $\mathrm{tp}(d''/\bar{a}) = \mathrm{tp}(d/\bar{a})$, there exists $\sigma \in \mathrm{Aut}(\mathcal{M}/\bar{a})$ such that $\sigma(d) = d''$. Let $c'' := \sigma(c)$. Then $(c'', d'') \in S_E$. By the same argument used earlier in this paragraph, $c'' \in E$. Now, $\mathrm{tp}(c''/\bar{a}) = \mathrm{tp}(c'/\bar{a})$ and so there exists $\sigma' \in \mathrm{Aut}(\mathcal{M}/\bar{a})$ such that $\sigma'(c'') = c'$. Now, since envelopes are homogeneous substructures (Lemma 3.2.4 in [14] and Definition 3.1.2) and $c', c'' \in E$, we may assume that $\sigma(E) = E$. Let $d' := \sigma'(d'')$. Then $d' \in E$, since $d'' \in E$. Hence $(c', d') \in S_E$ and so $c' \in S_{E1}$, as required. So the claim is proved.

We introduce some further definitions: For $c' \in S_1$ let $c'/S_2 := \{d' : (c', d') \in S\}$ and $c'/S_{E2} := \{d' : (c', d') \in S_E\}$, and for $d' \in S_2$ let $d'/S_1 := \{c' : (c', d') \in S\}$ and $d'/S_{E1} := \{c' : (c', d') \in S_E\}$. The sizes of the c'/S_{E2} and the d'/S_{E1} are in fact independent of c' and d', as we now show.

First consider some arbitrary $c' \in S_{E1}$. Let $D_{\bar{a}c}$ be the set of realisations of $\mathrm{tp}(d/\bar{a}c)$. Then, by the definition of S, $D_{\bar{a}c} = c/S_2$. Let $d' \in c'/S_{E2}$. Then, since $\mathrm{tp}((c', d')/\bar{a}) = \mathrm{tp}((c, d)/\bar{a})$, there exists $\sigma \in \mathrm{Aut}(\mathcal{M}/\bar{a})$ such that $\sigma(c', d') = (c, d)$. We claim that $\sigma\colon c'/S_2 \to c/S_2$ is a bijection. Injectivity is immediate. It is well-defined, since if $d'' \in c'/S_2$, then $\sigma(c', d'') = (c, \sigma(d'')) \in S$ and so $\sigma(d') \in c/S_2$. It is surjective, since if $d'' \in c/S_2$, then $\sigma^{-1}(c, d'') = (c', \sigma^{-1}(d'')) \in S$

and so $\sigma^{-1}(d'') \in c'/S_2$. So the claim is proved. Now, as mentioned previously, envelopes are homogeneous substructures. So, since $d, d' \in E$, we may assume that $\sigma(E) = E$. Thus

$$\begin{aligned} |c'/S_{E2}| &= |c'/S_2 \cap E| \\ &= |c/S_2 \cap E| \\ &= |D_{\bar{a}c} \cap E| \end{aligned} \tag{4.9}$$

for all $c' \in S_{E1}$.

Now consider some arbitrary $d' \in S_{E2}$. Since $c \leq d$, $c \in \operatorname{dcl}(d)$. Thus, since $\operatorname{tp}(d'/\bar{a}) = \operatorname{tp}(d/\bar{a})$, there exists a unique $c' \in M$ such that $(c', d') \in S$. But $d' \in E$ and so $(c', d') \in S_E$. Hence

$$|d'/S_{E1}| = 1 \tag{4.10}$$

for all $d' \in S_{E2}$.

We are now in a position to calculate the size of S_E and thereby also that of $D_{\bar{a}} \cap E$. Let's first calculate $|S_E|$ in terms of $|S_{E1}|$:

$$\begin{aligned} |S_E| &= \sum_{c' \in S_{E1}} |c'/S_{E2}| \\ &= |S_{E1}| \cdot |D_{\bar{a}c} \cap E| \quad \text{(by (4.9))}. \end{aligned} \tag{4.11}$$

And now in terms of $|S_{E2}|$:

$$\begin{aligned} |S_E| &= \sum_{d' \in S_{E2}} |d'/S_{E1}| \\ &= |S_{E2}| \quad \text{(by (4.10))}. \end{aligned} \tag{4.12}$$

So, since $S_{E2} = D_{\bar{a}} \cap E$, (4.11) and (4.12) yield

$$|D_{\bar{a}} \cap E| = |S_{E1}| \cdot |D_{\bar{a}c} \cap E|. \tag{4.13}$$

First consider S_{E1}. We previously proved that $S_{E1} = S_1 \cap E$. We also showed that S_1 is the set of realisations of $\operatorname{tp}(c/\bar{a})$ and that S_1 is a subset of J. By the Ryll-Nardzewski Theorem, $\operatorname{tp}(c/\bar{a})$ is isolated and so S_1 is $\bar{a}$-definable. So S_1 is an $\bar{a}$-definable subset of a projective geometry. Thus, as we will show later (after Case 2), there exists a polynomial $\rho_1 \in \mathbb{Q}[\mathbf{X}_1, \ldots, \mathbf{X}_s]$ such that $\rho_1(\bar{d}^*(E)) = |S_1 \cap E|$.

Now consider $D_{\bar{a}c}$, which is a parameter-definable subset of M, again by the Ryll-Nardzewski Theorem. We have $\operatorname{rk}(D_{\bar{a}c}) < \operatorname{rk}(D_{\bar{a}})$, as proved in the following paragraph, and thus by the induction hypothesis there exists a polynomial $\rho_2 \in \mathbb{Q}[\mathbf{X}_1, \ldots, \mathbf{X}_s]$ such that $|D_{\bar{a}c} \cap E| = \rho_2(\bar{d}^*(E))$.

We prove the claim that $\operatorname{rk}(D_{\bar{a}c}) < \operatorname{rk}(D_{\bar{a}})$. Let $n := \operatorname{rk}(D_{\bar{a}c})$. We previously showed that $D_{\bar{a}c} = c/S_2$. We also showed that for every $c' \in S_1$ there exists $\sigma \in \operatorname{Aut}(\mathcal{M}/\bar{a})$ such that $\sigma(c/S_2) = c'/S_2$, which thus means $\operatorname{rk}(c'/S_2) = n$ for every $c' \in S_1$. Define $f \colon D_{\bar{a}} \to S_1$ by $f(d') := c'$, where c' is such that $(c', d') \in S$. As we showed earlier, for every $d \in S_2$ there is precisely one c' such that $(c', d') \in S$, so f is well-defined. Then, since $f^{-1}(c') = c'/S_2$, we have $\operatorname{rk}(f^{-1}(c')) = n$ for every $c' \in S_1$. Also note that $\operatorname{rk}(S_1) > 0$, since S_1 is infinite (because c is not algebraic over its immediate predecessor). Thus, taking

$D := D_1 := D_{\bar{a}}$, $D_2 := S_1$ and π to be the identity map in Definition 4.5.1, we see that $\mathrm{rk}(D_{\bar{a}}) \geq n+1 > \mathrm{rk}(D_{\bar{a}c})$. So the claim is proved.

Define $\rho := \rho_1 \cdot \rho_2$. Then (4.13) gives us the desired result:

$$\begin{aligned} |D_{\bar{a}} \cap E| &= |S_{E1}| \cdot |D_{\bar{a}c} \cap E| \\ &= \rho_1(\bar{d^*}(E)) \cdot \rho_2(\bar{d^*}(E)) \\ &= \rho(\bar{d^*}(E)). \end{aligned}$$

End of Case 1.

Case 2: The element b is not the root. Since c is minimal, b and each element below b (except the root) is algebraic over its immediate predecessor. Thus, by the same induction used earlier in the proof, $b \in \mathrm{acl}(\varnothing)$. Thus, by inspection of Definition 4.1.9, we see that we may add to $\mathcal{L}$ a constant symbol for b without affecting the Lie coordinatising tree. Adding the new constant symbol preserves the inequality $\mathrm{rk}(D_{\bar{a}c}) < \mathrm{rk}(D_{\bar{a}})$, again since $b \in \mathrm{acl}(\varnothing)$, but it makes J $\varnothing$-definable. We may thus simply repeat the argument given in Case 1 in the extended language $\mathcal{L}_b$. *End of Case 2.*

We now prove our earlier claim that there exists a polynomial $\rho_1 \in \mathbb{Q}[\mathbf{X}_1, \ldots, \mathbf{X}_s]$ such that $\rho_1(\bar{d^*}(E)) = |S_1 \cap E|$. The set S_1 is an $\mathcal{L}_{\bar{a}}$-definable subset of J and thus, since J is fully embedded in $\mathcal{M}$, S_1 is $\bar{a}$-definable in the language of J; we may assume that $\bar{a}$ lies in J by stable embeddedness. We now consider the localisation $J/\bar{a}$ of J at $\bar{a}$ (Definition 4.1.8). J fibres over $J/\bar{a}$, where two elements lie in the same fibre if and only if they have the same algebraic closure over $\bar{a}$. These fibres all have the same finite size, where this size is determined by $\mathrm{tp}(\bar{a})$. Now, S_1 might not respect these fibres; that is, the intersection of S_1 with each fibre might vary in size. However, since the fibres are finite, there are only finitely many possible sizes for these intersections and so we can $\bar{a}$-definably partition the set of fibres according to size. We then consider the intersection of each part of the partition with E: We calculate the size of the base of the fibres, which is a $\varnothing$-definable subset of $J/\bar{a}$, and then multiply this result by the size of the fibre. We then sum these results to obtain $|S_1 \cap E|$. So, in short, by localising J at $\bar{a}$, it suffices to consider $\varnothing$-definable subsets of projective Lie geometries. It remains to do the explicit calculations in each kind of projective Lie geometry. We use quantifier elimination (Fact 4.1.3).

A projectivisation of a degenerate space. Projectivisation in this case is trivial. The only $\varnothing$-definable set is the whole space itself. (We can rule out $\varnothing$ because $D_{\bar{a}} \cap E \neq \varnothing$.) So $S_1 = J$. Thus, since $J \cap E = \mu(J)$ (Definition 4.3.3), where μ is the dimension function of E, we have $|S_1 \cap E| = d_E(J)$, as required.

A projectivisation of a pure vector space. The only $\varnothing$-definable set is again the whole space itself. So $S_1 = J$. Thus, going via the approximation of the linear space, which has dimension $\dim \mu(J) + 1$, we have

$$|S_1 \cap E| = \frac{q^{\dim \mu(J)+1} - 1}{q-1} = q^{\dim \mu(J)} + 1 = (-\sqrt{q})^{2 \dim \mu(J)} + 1 = d_E(J)^2 + 1,$$

as required.

A projectivisation of a polar space. This is the same as the vector space case, except that we can define either half of the space or the whole space. If the former, then the answer is the same as that in the vector space case. If the latter, then we multiply this answer by 2.

A projectivisation of a symplectic space. Since there is only one 1-type, this case is the same as the pure vector space case.

A projectivisation of a unitary space. The calculations can be found in the proof of Proposition 5.2.2 in [14]. Note that it is this case that forces us to consider $(-\sqrt{q})^{\dim \mu(J)}$, rather than just $q^{\dim \mu(J)}$.

An projectivisation of an orthogonal space. The calculations can again be found in the proof of Proposition 5.2.2 in [14]. Note that this is where the assumption regarding constant signature and parity is used (Remark 4.5.3). Also note that there is a small typographical error in the calculations: On p. 91 of [14] it should state $n(2i+j,\alpha) = q^i n(j,\alpha) + q^{j-1}(q^{2i} - q^i)$, the original term $q^i n(i,\alpha)$ being incorrect.

One final note: The calculations for unitary and orthogonal spaces in [14] are actually done in the linear Lie geometry, rather than in the projectivisation. However, by a similar fibering argument to the one used earlier with the localisation, this is sufficient. □

4.6. Macpherson's conjecture, full version. We introduce the notion of a polynomial exact class, enabling us to state and prove Theorem 4.6.4, the main result of the present work.

Definition 4.6.1 (Polynomial exact class). Let $\mathcal{L}$ be a language and $\mathcal{C}$ a class of finite $\mathcal{L}$-structures. Then $\mathcal{C}$ is a *polynomial exact class* in $\mathcal{L}$ if there exist

- $R \subseteq \mathbb{Q}[\mathbf{X}_1, \dots, \mathbf{X}_k]$ for some $k \in \mathbb{N}^+$,
- $\mathcal{L}$-formulas $\delta_1(\bar{x}_1, \bar{y}_1), \dots, \delta_k(\bar{x}_k, \bar{y}_k)$ and
- $\bar{a}_1 \in M^{l(\bar{y}_1)}, \dots, \bar{a}_k \in M^{l(\bar{y}_k)}$ for each $\mathcal{M} \in \mathcal{C}$

such that $\mathcal{C}$ is an R-mec in $\mathcal{L}$ where

$$h(\mathcal{M}) = h\left(|\delta_1(\mathcal{M}^{l(\bar{x}_1)}, \bar{a}_1)|, \dots, |\delta_k(\mathcal{M}^{l(\bar{x}_1)}, \bar{a}_k)|\right)$$

for every $h \in R$ and for every $\mathcal{M} \in \mathcal{C}$.

Remark 4.6.2. If we replace 'R-mec' with 'R-mac' in Definition 4.6.1, then we define a *polynomial asymptotic class*. In this case we allow polynomials with irrational coefficients.

Note that any 1-dimensional asymptotic class is a polynomial asymptotic class, since we may take δ to be the $\mathcal{L}$-formula $x = x$ and h to be the polynomial $\mu \mathbf{X}^d$, where (d, μ) is the dimension–measure pair.

Example 4.6.3 (Theorem 4.3.2 in [19]). The class of finite vector spaces is a polynomial asymptotic class.

Theorem 4.6.4 (Macpherson's conjecture, full version). *For any countable language $\mathcal{L}$ and for any $d \in \mathbb{N}^+$ the class $\mathcal{C}(\mathcal{L}, d)$ of all finite $\mathcal{L}$-structures with at most d 4-types is a polynomial exact class in $\mathcal{L}$.*

Proof. By Corollary 4.4.2 we know that $\mathcal{C} := \mathcal{C}(\mathcal{L}, d)$ is a multidimensional exact class. It remains to show that the measuring functions are polynomial in the sense of Definition 4.6.1.

Recall our use of Theorem 4.4.1 in the proof of Corollary 4.4.2: We partitioned $\mathcal{C}$ into subclasses $\mathcal{F}_1, \ldots, \mathcal{F}_k$ such that the $\mathcal{L}$-structures in each $\mathcal{F}_i$ smoothly approximate an $\mathcal{L}$-structure $\mathcal{F}_i^*$. By the work in [14] each $\mathcal{F}_i$ is a class of envelopes for $\mathcal{F}_i^*$, which is Lie coordinatisable. So Proposition 4.5.2 implies that $\mathcal{C}$ is a polynomial exact class, since each coordinatising Lie geometry is fully embedded in and thus (by definition) also definable in $\mathcal{F}_i^*$.

We address some details arising from this proof. Firstly, by the Projection Lemma (Lemma 2.2.1) it suffices to consider $\mathcal{L}$-formulas in one object variable, as we do in Proposition 4.5.2. Secondly, by Lemma 3.2.6 the intersection $\varphi(\mathcal{F}_i^*, \bar{a}) \cap \mathcal{M}$ is equal to the relativisation $\varphi(\mathcal{M}, \bar{a})$ for all $\mathcal{M} \in \mathcal{F}_i$ above some minimum size, so by Lemma 2.2.3 it suffices to consider the intersection. Thirdly, since $\mathcal{C}$ is an exact class, rather than just an asymptotic class, the measuring functions are determined by the formula and thus it is not necessary to show that the polynomials given by Proposition 4.5.2 are uniform in the parameter $\bar{a}$; this point is important because the measuring functions cannot depend on the parameters. Lastly, the hypothesis of constant parity and signature in the statement of Proposition 4.5.2 can be satisfied by partitioning each $\mathcal{F}_i$ into (up to) four subclasses, each with constant parity and signature. □

REMARK 4.6.5. Theorem 4.6.4 generalises Theorem 3.8 in [36] and Proposition 4.1 in [17].

§5. Open questions.

We pose a number of questions arising from the present work. In doing so we refer to the important model-theoretic notions of stability and (super)simplicity, which we have so far only mentioned in passing. We do not define these notions, but instead direct the reader to the vast literature on them, [7], [26] and [41] being good introductions. We also consider the notion of homogeneity, which is easier to define:

DEFINITION 5.1. An $\mathcal{L}$-structure $\mathcal{M}$ is *homogeneous* if $\mathcal{M}$ is countable and every isomorphism between substructures of $\mathcal{M}$ extends to an automorphism of $\mathcal{M}$.

Note that the word 'homogeneous' is overused in mathematics, especially in model theory. What we call 'homogeneous' might be called 'ultrahomogeneous' by other authors. See the comment after Definition 2.1.1 in [35].

FACT 5.2. *Let $\mathcal{L}$ be a finite relational language.*

(i) If $\mathcal{M}$ is a homogeneous $\mathcal{L}$-structure, then $\mathcal{M}$ is $\aleph_0$-categorical.
(ii) If $\mathcal{M}$ is an $\aleph_0$-categorical $\mathcal{L}$-structure, then $\mathrm{Th}(\mathcal{M})$ *has quantifier elimination if and only if $\mathcal{M}$ is homogeneous.*
(iii) If $\mathcal{M}$ is a stable homogeneous $\mathcal{L}$-structure, then $\mathcal{M}$ is $\aleph_0$-stable.

QUESTION 5.3. By Fact 5.2 and Corollary 7.4 in [13], if $\mathcal{L}$ is a finite relational language and $\mathcal{M}$ is a stable homogeneous $\mathcal{L}$-structure, then $\mathcal{M}$ is smoothly approximable and thus by Proposition 3.2.1 is elementarily equivalent to an ultraproduct of a multidimensional exact class. Does the converse hold? That is, if $\mathcal{L}$ is a finite relational language and $\mathcal{M}$ is a homogeneous $\mathcal{L}$-structure that is elementarily equivalent to an ultraproduct of a multidimensional exact class, then is $\mathcal{M}$ necessarily stable?

Recalling Remark 2.3.12, answering this question might shed some light on the role in Theorem 7.5.6 in [14] of the generic bipartite graph, which is neither stable nor smoothly approximable.

The following two questions were suggested to the present author by Ivan Tomašić:

QUESTION 5.4. What is the relationship between the work of Krajíček, Scanlon and others on Euler characteristics and R-macs and R-mecs? [29], [40], [28], [39], [42]

The notion of a generalised measurable structure, as developed in [1], also appears to be related, but a thorough investigation has yet to be carried out.

QUESTION 5.5. What are the interactions between polynomial exact classes and varieties with a polynomial number of points over finite fields?

The work of Brion and Peyre in [6] would be a good starting point for research into this question, as it suggests that algebraic varieties homogeneous under a linear algebraic group may provide a generic example of a polynomial exact class.

Acknowledgements. The concepts explored in this paper were introduced to me by Dugald Macpherson, my supportive and patient PhD supervisor, who also provided invaluable commentary on earlier versions of this paper. I gained further insights from Sylvy Anscombe, Gregory Cherlin, Immanuel Halupczok, Ehud Hrushovski and Ivan Tomašić. The TEX Stack Exchange helped me overcome some LATEX issues that arose whilst writing this paper, while my wife Mandy helped me overcome some non-LATEX issues that arose whilst writing this paper.

REFERENCES

[1] S. Anscombe, H.D. Macpherson, C. Steinhorn and D. Wolf, 'Multidimensional asymptotic classes', 2017. In preparation.

[2] M. Aschbacher, *Finite group theory*, Cambridge, UK: Cambridge University Press, 2^{nd} ed., 2000. Cambridge studies in advanced mathematics vol. 10.

[3] J.L. Bell and A. Slomson, *Models and ultraproducts: an introduction*, Mineola, NY, USA: Dover Publications, 2006. This is an unabridged republication of the 1974 third revised printing of the work originally published by the North-Holland Publishing Company in 1969.

[4] R.I. Bello Aguirre, 'Generalised Stability of ultraproducts of finite residue rings', *Archive for Mathematical Logic*, 2016. To appear. Preprint available online at https://arxiv.org/abs/1503.04454.

[5] R.I. Bello Aguirre, 'Model theory of finite and pseudofinite rings', PhD thesis, University of Leeds, 2016. Available online at http://etheses.whiterose.ac.uk/15771.

[6] M. Brion and E. Peyre, 'Counting points of homogeneous varieties over finite fields', *Journal für die reine und angewandte Mathematik*, vol. 645: pp. 105–124, 2010.

[7] E. Casanovas, *Simple theories and hyperimaginaries*, New York, USA: Cambridge University Press, 2011. ASL Lecture Notes in Logic vol. 39.

[8] Z. Chatzidakis, L. van den Dries and A. Macintyre, 'Definable sets over finite fields', *Journal für die reine und angewandte Mathematik*, vol. 427: pp. 107–135, 1992.

[9] G. Cherlin, and A.H. Lachlan, 'Stable finitely homogeneous structures', *Transactions of the American Mathematical Society*, vol. 296: pp. 815–850, 1986.

[10] G. Cherlin, 'Homogeneous tournaments revisited', *Geometriae Dedicata*, vol. 26: pp. 231–240, 1988.

[11] G. Cherlin, 'Large finite structures with few types', *Algebraic model theory*, edited by B.T. Hart, A.H. Lachlan and M.A. Valeriote, pp. 53–105, Dordrecht, the Netherlands: Kluwer Academic Publishers, 1997. NATO ASI Series C: Mathematical and Physical Sciences, vol. 496.

[12] G. Cherlin, 'The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments', *Memoirs of the American Mathematical Society*, vol. 131, no. 621, 1998.

[13] G. Cherlin, L. Harrington and A.H. Lachlan, '$\aleph_0$-categorical, $\aleph_0$-stable structures', *Annals of Pure and Applied Logic*, vol. 28: pp. 103–135, 1985. (This is sometimes incorrectly referenced as pp. 227–270 of vol. 18 (1980), which in fact refers to an entirely different paper by Cherlin and Shelah in the *Annals of Mathematical Logic*, which was renamed the *Annals of Pure and Applied Logic* in 1983.).

[14] G. Cherlin and E. Hrushovski, *Finite structures with few types*, Princeton, USA: Princeton University Press, 2003. Annals of Mathematics Studies vol. 152. Additional notes are available online at https://www.math.rutgers.edu/~cherlin/Paper/notes.pdf.

[15] A. Chowdhury, B. Hart and Ž. Sokolović, 'Affine covers of Lie geometries and the amalgamation property', *Proceedings of the London Mathematical Society*, vol. 85: pp. 513–563, 2002.

[16] R. Elwes, 'Dimension and measure in finite first order structures', PhD thesis, University of Leeds, 2005.

[17] R. Elwes, 'Asymptotic classes of finite structures', *Journal of Symbolic Logic*, vol. 72: pp. 418–438, 2007.

[18] D.M. Evans, 'Examples of $\aleph_0$-categorical structures', *Automorphisms of first-order structures*, edited by R. Kaye and H.D. Macpherson, pp. 33–72, New York, USA: Oxford University Press, 1994.

[19] D. García, D. Macpherson and C. Steinhorn, 'Pseudofinite structures and simplicity', *Journal of Mathematical Logic*, vol. 15: pp. 1–41, 2015.

[20] C. Hill and C. Smart, 'A re-iteration of *Finite structures with few types*, Chapters 1–3', 2006. Unpublished manuscript.

[21] W. Hodges, *Model theory*, New York, USA: Cambridge University Press, 1993. Encyclopedia of Mathematics and its Applications vol. 42.

[22] E. Hrushovski, 'Finite structures with few types', *Finite and infinite combinatorics and set theory*, edited by N.W. Sauer, R.E. Woodrow and B. Sands, pp. 175–187, Dordrecht, the Netherlands: Kluwer Academic Publishers, 1993. NATO ASI Series C: Mathematical and Physical Sciences, vol. 411.

[23] E. Hrushovski, 'On pseudo-finite dimensions', *Notre Dame Journal of Formal Logic*, vol. 54: pp. 463–495, 2013.

[24] W.M. Kantor, M.W. Liebeck and H.D. Macpherson, '$\aleph_0$-categorical structures smoothly approximated by finite substructures', *Proceedings of the London Mathematical Society*, vol. 59: pp. 439–463, 1989.

[25] R. Kaye and H.D. Macpherson, 'Models and groups', *Automorphisms of first-order structures*, edited by R. Kaye and H.D. Macpherson, pp. 3–31, New York, USA: Oxford University Press, 1994.

[26] B. Kim, *Simplicity theory*, New York, USA: Oxford University Press, 2014. Oxford Logic Guides vol. 53.

[27] J. Knight and A.H. Lachlan, 'Shrinking, stretching, and codes for homogeneous structures', *Classification theory*, edited by J.T. Baldwin, pp. 192–229, Berlin and Heidelberg, Germany: Springer-Verlag, 1987. Lecture Notes in Mathematics vol. 1292.

[28] J. Krajíček, 'Approximate Euler characteristic, dimension, and weak pigeonhole principles', *Journal of Symbolic Logic*, vol. 69: pp. 201–214, 2004.

[29] J. Krajíček and T. Scanlon, 'Combinatorics with definable sets: Euler characteristics and Grothendieck rings', *The Bulletin of Symbolic Logic*, vol. 6: pp. 311–330, 2000.

[30] A.H. Lachlan, 'Countable homogeneous tournaments', *Transactions of the American Mathematical Society*, vol. 284: pp. 431–461, 1984.

[31] A.H. Lachlan, 'On countable stable structures which are homogeneous for a finite relational language', *Israel Journal of Mathematics*, vol. 49: pp. 69–153, 1984.

[32] A.H. Lachlan, 'Structures coordinatized by indiscernible sets', *Annals of Pure and Applied Logic*, vol. 34: pp. 245–273, 1987.

[33] A.H. Lachlan, 'Stable finitely homogeneous structures: a survey', *Algebraic model theory*, edited by B.T. Hart, A.H. Lachlan and M.A. Valeriote, pp. 145–159, Dordrecht, the Netherlands: Kluwer Academic Publishers, 1997. NATO ASI Series C: Mathematical and Physical Sciences, vol. 496.

[34] H.D. Macpherson, 'Homogeneous and smoothly approximated structures', *Algebraic model theory*, edited by B.T. Hart, A.H. Lachlan and M.A. Valeriote, pp. 161–179, Dordrecht, the Netherlands: Kluwer Academic Publishers, 1997. NATO ASI Series C: Mathematical and Physical Sciences, vol. 496.

[35] H.D. Macpherson, 'A survey of homogeneous structures', *Discrete Mathematics*, vol. 311: pp. 1599–1634, 2011.

[36] H.D. Macpherson and C. Steinhorn, 'One-dimensional asymptotic classes of finite structures', *Transactions of the American Mathematical Society*, vol. 360: pp. 411–448, 2008.

[37] H.D. Macpherson and C. Steinhorn, 'Definability in classes of finite structures', *Finite and algorithmic model theory*, edited by J. Esparza, C. Michaux and C. Steinhorn, pp. 140–176, New York, USA: Cambridge University Press, 2011. LMS Lecture Notes Series vol. 379.

[38] D. Marker, *Model theory: an introduction*, New York, USA: Springer-Verlag, 2002. Graduate Texts in Mathematics vol. 217.

[39] M.J. Ryten and I. Tomašić, 'ACFA and measurability', *Selecta Mathematica*, vol. 11: pp. 523–537, 2005/2006.

[40] T. Scanlon, 'Fields admitting ordered nontrivial strong Euler characteristics are quasifinite', 2000. Unpublished note. Available online at https://math.berkeley.edu/~scanlon/papers/quasi.pdf.

[41] K. Tent and M. Ziegler, *A course in model theory*, New York, USA: Cambridge University Press, 2012. ASL Lecture Notes in Logic vol. 40.

[42] I. Tomašić, 'Fields with measure and automorphism', *Bulletin of the London Mathematical Society*, vol. 41: pp. 429–440, 2009.

[43] D. Wolf, 'Model theory of multidimensional asymptotic classes', PhD thesis, University of Leeds, 2016. Available online at http://etheses.whiterose.ac.uk/15316.

Formerly of the
SCHOOL OF MATHEMATICS
UNIVERSITY OF LEEDS
LEEDS LS2 9JT
UNITED KINGDOM
Email: dwolfeu@gmail.com