

A Crèche Course in Model Theory

Preface	4
1 Preliminaries and notation	5
1 Structures	5
2 Tuples	6
3 Terms	7
4 Substructures	8
5 Formulas	9
6 Yet more notation	11
2 Theories and elementarity	13
1 Logical consequences	13
2 Elementary equivalence	15
3 Embeddings and isomorphisms	17
4 Completeness	19
5 The Tarski-Vaught test	20
6 The downward Löwenheim-Skolem theorem	20
7 Elementary chains	22
3 Ultraproducts and compactness	23
1 Filters and ultrafilters	23
2 Direct products	24
3 Quotient structures	25
4 Ultraproducts	26
5 The compactness theorem	28
6 Finite axiomatizability	29
7 Realization of types and upward Löwenheim-Skolem	30
4 Types and morphisms	32
1 Semilattices and filters	32
2 Distributive lattices and prime filters	34
3 Types as filters	36
4 Morphisms	38
5 Some relational structures	41
1 Dense linear orders	41
2 Random graphs	43
3 Notes	46
6 Fraïssé limits	47
1 Rich models.	47
2 Weaker notions of universality and homogeneity	52

3	The amalgamation property	53
4	Notes	53
7	Some algebraic structures	54
1	Abelian groups	54
2	Torsion-free abelian groups	55
3	Divisible abelian groups	56
4	Commutative rings	58
5	Integral domains	59
6	Algebraically closed fields	60
7	Hilbert's Nullstellensatz	61
8	Saturation and homogeneity	64
1	Saturated structures	64
2	Homogeneous structures	66
3	The monster model	67
9	Preservation theorems	72
1	Lyndon-Robinson Lemma	72
2	Quantifier elimination by back-and-forth	74
10	Geometry and dimension	76
1	Algebraic and definable elements	76
2	Strongly minimal theories	78
3	Independence and dimension	79
11	Countable models	82
1	The omitting types theorem	82
2	Prime and atomic models	83
3	Countable categoricity	85
4	Small theories	86
5	A toy version of a theorem of Zil'ber	88
6	Notes	89
12	Definability and automorphisms	90
1	Many-sorted structures	90
2	The eq-expansion	91
3	The definable closure in the eq-expansion	93
4	The algebraic closure in the eq-expansion	94
5	Elimination of imaginaries	95
6	Imaginaries, the true story	97
7	Uniform elimination of imaginaries	98
13	Invariant sets	101
1	Invariant sets and types	101
2	Morley sequences and indiscernibles	102
3	From coheirs to Ramsey to indiscernibles	104
4	Expansions	105

14 Lascar invariant sets	107
1 Lascar strong types	107
2 The Lascar graph and Newelski's theorem	109
3 Kim-Pillay types	112
4 Notes	114
 15 Stable formulas (tentative, a scribbling block)	 115
1 Externally definable sets	115
2 The ladder property	117
3	117
4 Notes e references	118
 16 Some combinatorics	 119
1 A general setting	119
2 The combinatorial bare bones of stability	120
3 Vapnik-Chervonenkis dimension	122
4 Epsilon-approximations	123
5 A piercing problem	128
6 Appendix: Farkas' lemma	129

Preface

These are the notes of a course that I have been giving for many years (first in Amsterdam and then) in Turin. A few chapters are missing and will be added soon. I keep the most recent version in <https://github.com/domenicozambella/creche>

Below I list a few peculiarities of the exposition.

- ▷ [Fraïssé limits](#) are presented in a slightly general setting which accommodates more examples (and I'll add a few more). This is used e.g. to discuss saturation.
- ▷ [Quantifier elimination for ACF](#) and Hilbert's Nullstellensatz are presented with more details than usual (which may annoy some experts but hopefully help some students).
- ▷ The proof of the [Omitting Types Theorem](#) uses a model theoretic construction (different from the standard syntactic proof).
- ▷ [Imaginaries and the eq-expansion](#) are introduced from the (equivalent) dual perspective: the canonical name of a definable set is the set itself.
- ▷ [Lascar and Kim-Pillay strong types](#) are introduced in a slightly unconventional way. I am grateful to Krzysztof Krupiński for helping me squaring the circle with KP-types.
- ▷ [Ramsey Theorem](#) is derived from the existence of coheir sequences. Admittedly, this not the shortest proof, but it may be interesting. I plan to add a proof of Hindman's theorem along the same lines.

Chapter 1

Preliminaries and notation

This chapter introduces the syntax and semantic of first order logic. We assume the reader is not completely unfamiliar with first order logic.

1 Structures

A **(first order) language L** (also called **signature**) is a triple that consists of

- ▷ a set L_{fun} whose elements are called **function symbols**.
- ▷ a set L_{rel} whose elements are called **relation symbols**.
- ▷ a function that assign to every $f \in L_{\text{fun}}$ and $r \in L_{\text{rel}}$ non negative integers n_f and n_r that we call **arity** of the function, respectively relation, symbol. In words we say that f is a **n_f -ary** function symbol, similarly for r . We may say **constant** for 0-ary function symbol.

Warning: it is custom to use the symbol L to denote both the language and the set of formulas (to be defined below) associated to it. With $|L|$ we denote the cardinality of $L_{\text{fun}} \cup L_{\text{rel}} \cup \omega$.

A **(first order) structure M** of signature L (for short **L -structure**) consists of

- ▷ A set that we call the **domain** or **support** and denote with the same symbol M used for the whole structure
- ▷ A function that assign to every $f \in L_{\text{fun}}$ a total map $f^M : M^{n_f} \rightarrow M$ and to every $r \in L_{\text{rel}}$ a relation $r^M \subseteq M^{n_r}$. We call f^M and r^M the **interpretation** of f , respectively r in M .

Recall that by definition $M^0 = \{\emptyset\}$. Therefore the interpretation of a constant c is a function that maps the unique element of M^0 into an element of M . We will identify c^M with $c^M(\emptyset)$.

We may use the word **model** as a synonym of structure. But beware that in some context the word is used to denote some particular sort of structures.

If M is an L -structure and $A \subseteq M$ is any subset we write $L(A)$ the language obtained adding to L_{fun} the elements of A as constants. In this context the elements of A are called **parameters**. There is a canonical expansion of M to an $L(A)$ -structure by setting $a^M = a$ for every $a \in A$.

1.1 Example The **language of additive groups** consists of the following function symbols

- ▷ a constant (that is, a function symbol of arity 0): 0
- ▷ a unary function symbol (that is, of arity 1): $-$
- ▷ a binary function symbol (that is, of arity 2): $+$

In the **language of multiplicative groups** the three symbols above are replaced with 1 , $^{-1}$, and \cdot respectively. Any group is a structure in either of these two signatures with the obvious interpretation. Needless to say, not all structures with these signature are groups.

The **language of (unitary) rings** contains all function symbols above with exception of $^{-1}$. The **language of ordered rings** also contains the binary relation symbol $<$. \square

The following is a less straightforward example. The reason behind the choice of the language of vector spaces will be clear in Section 4 below.

- 1.2 Example** Let K be a field. The **language of vector spaces over K** , that here we denote by L_K , extends that of additive groups with a unary function symbols k for every $k \in K$.

Recall that a vector space over K is an abelian group M together with an action of K over M , that is, a function $\mu : K \times M \rightarrow M$. To view a vector space over K as an L_K -structure we interpret the group symbols in the obvious way and each $k \in K$ as the function $\mu(k, -)$. \square

The languages in the examples above, with the exception of that of ordered rings, are **functional languages**, i.e. $L_{\text{rel}} = \emptyset$. In the following we consider two important examples of **relational languages** where $L_{\text{fun}} = \emptyset$.

- 1.3 Example** The **language of strict orders** contains only the binary relation symbol $<$. The **language of graphs** also contains only a binary relation symbol (there is no canonical choice). In combinatorics a graph is a set of unordered pairs. To associate to a graph a first order structure we identify identify sets of unordered pairs with symmetric irreflexive relations. \square

2 Tuples

Let A be any set. A **tuple of elements of A** is a function $a : \alpha \rightarrow A$ for some ordinal α . We call α the **length** of a and denote it by $|a|$. If a is surjective, we call a an **enumeration** of A . Occasionally, we may say **sequence** for tuple – a word which we use also when the domain is not necessarily an ordinal (though, typically, an ordered set).

We write a_i for the i -th element of a , that is, the element $a(i)$, where $i < \alpha$. We may present a tuple with the notation $a = \langle a_i : i < \alpha \rangle$; when α is finite, we may also write $a = a_0, \dots, a_{\alpha-1}$.

The set of tuples of elements of A length α is denoted by A^α . The set of tuples of length $< \alpha$ is denoted by $A^{<\alpha}$. For instance, $A^{<\omega}$ is the set of all finite tuples of elements of A . When α is finite we confuse A^α with α -th Cartesian power of A . In particular A^1 is confused with A .

If $a, b \in A^{|a|}$ and h is a function defined on A , we write $h(a) = b$ for $h(a_i) = b_i$. We often confuse the pair $\langle a, b \rangle$ with the tuple of pairs $\langle a_i, b_i \rangle$. The context will resolve the ambiguity.

Note that there is a unique tuple of length 0, we call it the **empty tuple**. In fact, for

all A , even when A is empty, \emptyset is the unique element of A^0 .

We will often concatenate tuples. If a and b are tuples we will write ab or a,b indifferently.

3 Terms

Fix a infinite set V whose elements we call **variables**. We use the letters x, y, z , ecc. to denote variables or tuple of variables. We rarely refer to V explicitly and we always assume it is large enough for our needs.

We fix a signature L for the whole section.

1.4 Definition A **term** is a finite sequence of elements of $L_{\text{fun}} \cup V$ that are obtained with the following inductive procedure

- o. every variable, intended as a tuple of length 1, is a term;
- i. If $f \in L_{\text{fun}}$ and t is a tuple obtained concatenating n_f terms, then ft is a term. By ft we understand the tuple obtained prefixing t by f .

We will say **L-term** when we wish to specify the language L . □

Note that any constant f , intended as a tuple of length 1, is a term (by i, the term f is obtained concatenating $n_f = 0$ terms and prefixing by f). Terms that do not contain variables are called **closed terms**.

The intended meaning of, for instance, the term $++xyz$ is $(x+y)+z$. The first expression uses an **prefix notation** the second the **infix notation**. When convenient we (informally) use the infix notation and add parenthesis to improve legibility and avoid ambiguity.

The following lemma shows that the prefix notation allows to write unambiguous terms without the use of parenthesis.

1.5 Lemma (unique legibility of terms) Let a be a sequence of terms. Suppose a can be obtained both by the concatenation of the terms t_1, \dots, t_n or by the concatenation of the terms s_1, \dots, s_m . Then $n = m$ and $s_i = t_i$.

Proof By induction on $|a|$. If $|a| = 0$ then $n = m = 0$ and there is nothing to prove. Suppose the claim holds for tuples of length k and let $a = a_1, \dots, a_{k+1}$. Then a_1 is the first element of both t_1 and s_1 . If a_1 is a variable, say x , then t_1 and s_1 are the term x and $n = m = 1$. Otherwise a_1 is a function symbol, say f , then $t_1 = f\bar{t}$ and $s_1 = f\bar{s}$ where \bar{t} and \bar{s} are obtained concatenating the terms t'_1, \dots, t'_p and s'_1, \dots, s'_p . Now apply the induction hypothesis to a_2, \dots, a_{k+1} and to the terms $t'_1, \dots, t'_p, t_2, \dots, t_n$ and $s'_1, \dots, s'_p, s_2, \dots, s_m$. □

If $x = x_1, \dots, x_n$ is a tuple of distinct variables and $s = s_1, \dots, s_n$ is a tuple of terms, we write $t[x/s]$ for the sequence obtained replacing x by s coordinatewise. Proving that $t[x/s]$ is indeed a term is tedious and can safely skip.

If t is a term and x_1, \dots, x_n are (tuples of) variables we write $t(x_1, \dots, x_n)$ to declare that the variables occurring in t are among those that occur in x_1, \dots, x_n . When a term has been presented as $t(x, y)$, we write $t(s, y)$ for $t[x/s]$.

Finally we define the interpretation of a term in a structure M . We begin with closed terms. These are interpreted in 0-ary functions, i.e. an elements of the structure.

1.6 Definition Let t be a closed $L(M)$ term. We write t^M for the *interpretation of t* . This is defined by induction of the syntax of t

- i. if $t = f \bar{t}$, where $f \in L_{\text{fun}}$ and \bar{t} is a tuple obtained concatenation the terms t_1, \dots, t_{n_f} , then $t^M = f^M(t_1^M, \dots, t_{n_f}^M)$.

Note that in i we have used Lemma 1.5 in an essential way. In fact this ensures that the sequence \bar{t} uniquely determines the terms t_1, \dots, t_{n_f} . \square

The inductive definition above is based on the case $n_f = 0$, i.e. for f a constant (or a parameter). When $t = c$, a constant, \bar{t} is the empty tuple, then $t^M = c^M(\emptyset)$ which we abbreviate with c^M . In particular, if $t = a$, a parameter, then $t^M = a^M = a$.

Now we generalize the interpretation to all (non necessarily closed) terms. If $t(x)$ is a term then we define $t^M(x) : M^{|x|} \rightarrow M$ to be the function that maps $a \mapsto t(a)^M$.

4 Substructures

1.7 Definition Fix a signature L and let M and N be two L -structures. We say that M is a *substructure* of N , and write $M \subseteq N$, if

1. the domain of M is a subset of the domain of N
2. $f^M = f^N \upharpoonright M^{n_f}$ for every $f \in L_{\text{fun}}$
3. $r^M = r^N \cap M^{n_r}$ for every $f \in L_{\text{rel}}$.

Note that when f is a constant 2 becomes $f^M = f^N$, in particular the substructures of N contains at least all constants of N .

If a set $A \subseteq N$ is such that

1. $f^N[A^{n_f}] \subseteq A$ for every $f \in L_{\text{fun}}$

then there is a unique substructure $M \subseteq N$ with domain A , namely, the structure with the following interpretation

2. $f^M = f^N \upharpoonright M^{n_f}$ (which is a good definition by the assumption on A);
3. $r^M = r^N \cap M^{n_r}$.

It is usual to confuse subsets of N that satisfy 1 with the unique substructure they support.

It is immediate to verify that the intersection of an arbitrary family of substructures of N is a substructure of N . Therefore, for any given $A \subseteq N$ we may define the *substructure of N generated A* as the intersection of all substructures of N that contain A . We write $\langle A \rangle_N$. The following easy proposition $\langle A \rangle_N$ gives more concrete representation of $\langle A \rangle_N$

1.8 Lemma The following hold for every $A \subseteq N$

$$1 \quad \langle A \rangle_N = \left\{ t^N : t \text{ a closed } L(A)\text{-term} \right\}$$

$$\begin{aligned}
2 \quad \langle A \rangle_N &= \left\{ t^N(a) : t(x) \text{ an } L\text{-term and } a \in A^{|x|} \right\} \\
3 \quad \langle A \rangle_N &= \bigcup_{n \in \omega} A_n, \text{ where } A_0 = A \\
&\quad A_{n+1} = A_n \cup \left\{ f^N(a) : f \in L_{\text{fun}}, a \in A_n^{n_f} \right\}. \quad \square
\end{aligned}$$

1.9 Example Let L be the multiplicative (or additive) language of groups. Let N be a group, which we consider as an L -structure in the natural way. Then the substructures of N are exactly the subgroups of N and $\langle A \rangle_N$ is the group generated by $A \subseteq N$. A similar claim is true when L is the signature of vector spaces and N is a vector space. The choice of the language is more or less fixed if we want that the algebraic and the model theoretic notion of substructure coincide. \square

5 Formulas

Fix a language L and a set of variables V . A **formula** is a finite sequence of symbols in $L_{\text{fun}} \cup L_{\text{rel}} \cup V \cup \{ \doteq, \perp, \neg, \vee, \exists \}$. The last set contains the logical symbols that are called respectively

\doteq **equality**; \perp **contradiction**; \neg **negation**;
 \vee **disjunction**; \exists **existential quantifier**.

Syntactically, \doteq behaves like a binary relation symbol. So, for convenience set $n_{\doteq} = 2$. However \doteq is considered as a logic symbol because its semantic is fixed (it is always interpreted in the diagonal).

The definition below uses the prefix notation which simplifies the proof of the unique legibility lemma. However, in practice we always use the infix notation: $t \doteq s$, $\varphi \vee \psi$, etc.

1.10 Definition A **formula** is any finite sequence is obtained with the following inductive procedure

- o. if $r \in L_{\text{rel}} \cup \{=\}$ and t is a tuple obtained concatenating n_r terms then $r t$ is a formula. Formulas of this form are called **atomic**;
- i. if φ e ψ are formulas then the following are formulas: \perp , $\neg \varphi$, $\vee \varphi \psi$, and $\exists x \varphi$, for any $x \in V$. \square

We use L to denote both the language and the set of formulas. We write L_{at} for the set of atomic formulas and L_{qf} for the set of **quantifier-free formulas** i.e., formulas where \exists does not occur.

The proof of the following is similar to the analogous lemma for terms.

1.11 Lemma (unique legibility of formulas) Let a be a sequence of formulas. Suppose a can be obtained both by the concatenation of the formulas $\varphi_1, \dots, \varphi_n$ or by the concatenation of the formulas ψ_1, \dots, ψ_m . Then $n = m$ and $\varphi_i = \psi_i$. \square

A formula is **closed** if all its variables occur under the scope of a quantifier. Closed formulas are also called **sentences**. We will do without a formal definition of *occurs*

under the scope of a quantifier which is too lengthy. An example suffices: all occurrences of x are under the scope a quantifiers in the formula $\exists x \varphi$. These occurrences are called **bonded**. The formula $x \doteq y \wedge \exists x \varphi$ has **free** (i.e., not bond) occurrences of x and y .

Let x is a tuple of variables and t is a tuple of terms such that $|x| = |t|$. We write $\varphi[x/t]$ for the formula obtained substituting t for all free occurrences of x , coordinatewise.

We write $\varphi(x)$ to declare that the free variables in the formula φ are all among those of the tuple x . In this case we write $\varphi(t)$ for $\varphi[x/t]$.

We will often use without explicit mention the following useful syntactic decomposition of formulas with parameters.

1.12 Lemma *For every formula $\varphi(x) \in L(A)$ there is a formula $\psi(x; z) \in L$ and a tuple of parameters $a \in A^{|z|}$ such that $\varphi(x) = \psi(x; a)$. \square*

Just as a term $t(x)$ is a name for a function $t(x)^M : M^{|x|} \rightarrow M$, a formula $\varphi(x)$ is a name for a subset $\varphi(x)^M \subseteq M^{|x|}$ which we call **the subset of M defined by $\varphi(x)$** . It is also very common to write $\varphi(M)$ for the set defined by $\varphi(x)$. In general sets of the form $\varphi(M)$ for some $\varphi(x) \in M$ are called **definable**.

1.13 Definition of truth *For every formula φ with variables among those of the tuple x we define $\varphi(x)^M$ by induction as follows*

$$o1. \quad (\doteq t s)(x)^M = \{a \in M^{|x|} : t^M(a) = s^M(a)\}$$

$$o2. \quad (r t_1 \dots t_n)(x)^M = \{a \in M^{|x|} : \langle t_1^M(a), \dots, t_n^M(a) \rangle \in r^M\}$$

$$i0. \quad \perp(x)^M = \emptyset$$

$$i1. \quad (\neg \xi)(x)^M = M^{|x|} \setminus \xi(x)^M$$

$$i2. \quad (\vee \xi \psi)(x)^M = \xi(x)^M \cup \psi(x)^M$$

$$i3. \quad (\exists y \varphi)(x)^M = \bigcup_{a \in M} (\varphi[y/a])(x)^M$$

Condition i2 assumes that ξ and ψ are uniquely determined by $\vee \xi \psi$. This is guaranteed by the unique legibility of formulas, Lemma 1.11. Analogously, o1 e o2 assume Lemma 1.5.

The case when x is the empty tuple is far from trivial. Note that $\varphi(\emptyset)^M$ is a subset of $M^0 = \{\emptyset\}$. Then there are two possibilities either $\{\emptyset\}$ or \emptyset . We will read them as two **truth values**: **True** and **False**, respectively. If $\varphi^M = \{\emptyset\}$ we say that **φ is true in M** , if $\varphi^M = \emptyset$, we say that **φ is false in M** . We write $M \models \varphi$, respectively $M \not\models \varphi$. Or we may say that **M models φ** , respectively **M does not model φ** . It is immediate to verify at

$$\varphi(M) = \{a \in M^{|x|} : M \models \varphi(a)\}.$$

6 Yet more notation

The following logical connectives are defined as abbreviations:

\top	stands for	$\neg \perp$	tautology
$\varphi \wedge \psi$	stands for	$\neg[\neg\varphi \vee \neg\psi]$	conjunction
$\varphi \rightarrow \psi$	stands for	$\neg\varphi \vee \psi$	implication
$\varphi \leftrightarrow \psi$	stands for	$[\varphi \rightarrow \psi] \wedge [\psi \rightarrow \varphi]$	bi-implication
$\varphi \nleftrightarrow \psi$	stands for	$\neg[\varphi \leftrightarrow \psi]$	exclusive disjunction
$\forall x \varphi$	stands for	$\neg \exists x \neg \varphi$	universal quantifier

We agree that \rightarrow e \leftrightarrow bind less than \wedge e \vee . Unary connectives (quantifiers and negation) bind stronger than binary connectives. For example

$$\exists x \varphi \wedge \psi \rightarrow \neg \zeta \vee \vartheta \quad \text{reads as} \quad [(\exists x \varphi) \wedge \psi] \rightarrow [(\neg \zeta) \vee \vartheta]$$

We say that $\forall x \varphi(x)$ is the **universal closure** of $\varphi(x)$ and that $\exists x \varphi(x)$ is the **existential closure**. We say that $\varphi(x)$ **holds in M** when its universal closure is true in M . We say that $\varphi(x)$ is **consistent in M** when its existential closure is true in M .

The semantic of conjunction and disjunction is associative. Then for any finite set of formulas $\{\varphi_i : i \in I\}$ we can write without ambiguities

$$\bigwedge_{i \in I} \varphi_i \qquad \bigvee_{i \in I} \varphi_i$$

When $x = x_1, \dots, x_n$ is a tuple of variables we write $\exists x \varphi$ or $\exists x_1, \dots, x_n \varphi$ for $\exists x_1 \dots \exists x_n \varphi$. With first order sentences we are able to say that $\varphi(M)$ has at least n elements (also, no more than, or exactly n). It is convenient to use the following abbreviations.

$$\begin{aligned} \exists^{\geq n} x \varphi(x) & \text{ stands for } \exists x_1, \dots, x_n \left[\bigwedge_{1 \leq i \leq n} \varphi(x_i) \wedge \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j \right]. \\ \exists^{\leq n} x \varphi(x) & \text{ stands for } \neg \exists^{\geq n+1} x \varphi(x) \\ \exists^= n x \varphi(x) & \text{ stands for } \exists^{\geq n} x \varphi(x) \wedge \exists^{\leq n} x \varphi(x) \end{aligned}$$

1.14 Exercise Let M be an L -structure and let $\psi(x), \varphi(x, y) \in L$. Write a sentence true in M exactly when

- a. $\psi(M) \in \{\varphi(a, M) : a \in M\}$;
- b. $\{\varphi(a, M) : a \in M\}$ contains at least two sets;
- c. $\{\varphi(a, M) : a \in M\}$ contains only sets that are pairwise disjoint. □

1.15 Exercise Let M be a structure in the signature of graphs as defined in Example 1.3 (but not necessarily a graph). Write a sentence φ such that,

- a. $M \models \varphi$ if and only if there is an $A \subseteq M$ such that $r^M \subseteq A \times \neg A$.

Remark: φ assert an asymmetric version of the property below

- b. $M \models \psi$ if and only if there is an $A \subseteq M$ such that $r^M \subseteq (A \times \neg A) \cup (\neg A \times A)$.

Assume M is a graph, what required in b is equivalent to saying that M is a *bipartite graph*, or equivalently that it has *chromatic number 2* i.e., we can color the vertices with 2 colors so that no two adjacent vertices share the same color. \square

Chapter 2

Theories and elementarity

1 Logical consequences

A **theory** is a set $T \subseteq L$ of sentences. We write $M \models T$ if $M \models \varphi$ for every $\varphi \in T$. If $\varphi \in L$ is a sentence we write $T \vdash \varphi$ when

$$M \models T \Rightarrow M \models \varphi \quad \text{for every } M.$$

In words, we say that φ is a **logical consequence** of T or that φ **follows from** T . If S is a theory $T \vdash S$ has a similar meaning. If $T \vdash S$ and $S \vdash T$ we say that T and S are **logically equivalent**. We may say that T **axiomatizes** S (or vice versa).

We say that a theory is **consistent** if it has a model. With the notation above, T is consistent if and only if $T \not\vdash \perp$.

The **closure of T under logical consequence** is the set $\text{ccl}(T)$ which is defined as follows:

$$\text{ccl}(T) = \left\{ \varphi \in L : \text{sentence such that } T \vdash \varphi \right\}$$

If T is a finite set, say $T = \{\varphi_1, \dots, \varphi_n\}$ we write $\text{ccl}(\varphi_1, \dots, \varphi_n)$ for $\text{ccl}(T)$. If $T = \text{ccl}(T)$ we say that T is **closed under logical consequences**.

The **theory of M** is the set of sentences that hold in M and is denoted by $\text{Th}(M)$. More generally, if \mathcal{K} is a class of structures, $\text{Th}(\mathcal{K})$ is the set of sentences that hold in every model in \mathcal{K} . That is

$$\text{Th}(\mathcal{K}) = \bigcap_{M \in \mathcal{K}} \text{Th}(M)$$

The class of all models of T is denoted by $\text{Mod}(T)$. We say that \mathcal{K} is **axiomatizable** if $\text{Mod}(T) = \mathcal{K}$ for some theory T . If T is finite we say that \mathcal{K} is **finitely axiomatizable**. To sum up

$$\begin{aligned} \text{Th}(M) &= \left\{ \varphi : M \models \varphi \right\} \\ \text{Th}(\mathcal{K}) &= \left\{ \varphi : M \models \varphi \text{ for all } M \in \mathcal{K} \right\} \\ \text{Mod}(T) &= \left\{ M : M \models T \right\} \end{aligned}$$

2.1 Example Let L be the language of multiplicative groups. Let T_g be the set containing the universal closure of following three formulas

$$\text{ass.} \quad (x \cdot y) \cdot z = x \cdot (y \cdot z);$$

$$\text{inv.} \quad x \cdot x^{-1} = x^{-1} \cdot x = 1;$$

$$\text{neu.} \quad x \cdot 1 = 1 \cdot x = x.$$

Then T_g axiomatizes the theory of groups, i.e. $\text{Th}(\mathcal{K})$ for \mathcal{K} the class of all groups. Let φ be the universal closure of the following formula

can. $z \cdot x = z \cdot y \rightarrow x = y$.

As φ formalizes the cancellation property then $T_g \vdash \varphi$, that is, φ is a logical consequence of T_g . Now consider the sentence ψ which is the universal closure of

com. $x \cdot y = y \cdot x$.

So, commutative groups model ψ and non commutative groups model $\neg\psi$. Hence neither $T_g \vdash \psi$ nor $T_g \vdash \neg\psi$. We say that T_g **does not decide** ψ . \square

Note that even when T is a very concrete set, $\text{ccl}(T)$ may be more difficult to grasp. In the example above T_g contains three sentences but $\text{ccl}(T_g)$ is an infinite set containing sentences that code theorems of group theory yet to be proved.

2.2 Remark The following properties say that ccl is a finitary closure operator.

1. $T \subseteq \text{ccl}(T)$ (extensive)
2. $\text{ccl}(T) = \text{ccl}(\text{ccl}(T))$ (increasing)
3. $T \subseteq S \Rightarrow \text{ccl}(T) \subseteq \text{ccl}(S)$ (idempotent)
4. $\text{ccl}(T) = \bigcup \{ \text{ccl}(S) : S \text{ finite subset of } T \}$. (finitary)

Properties 1-3 are easy to verify while 4 requires the compactness theorem. \square

In the next example we list a few algebraic theories with straightforward axiomatization.

2.3 Example We write T_{ag} for the theory of abelian groups which contains the universal closure of following

- a1. $(x + y) + z = y + (x + z)$;
- a2. $x + (-x) = 0$;
- a3. $x + 0 = x$;
- a4. $x + y = y + x$.

The theory T_r of (unitary) rings extends T_{ag} with

- a5. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$;
- a6. $1 \cdot x = x \cdot 1 = x$;
- a7. $(x + y) \cdot z = x \cdot z + y \cdot z$;
- a8. $z \cdot (x + y) = z \cdot x + z \cdot y$.

The theory of commutative rings T_{cg} contains also com of examples 2.1. The theory of ordered rings T_{or} extends T_{cr} with

- o1. $x < z \rightarrow x + y < z + y$;
- o2. $0 < x \wedge 0 < z \rightarrow 0 < x \cdot z$.

\square

The axiomatization of vector spaces is less straightforward.

2.4 Example Fix a field K . The language L_K extends the language of additive groups with a unary function for every element of K . The theory of vector fields over K extends T_{ag} with the following axioms (for all $h, k, l \in K$)

$$\text{m1. } h(x + y) = hx + hy$$

$$\text{m2. } lx = hx + kx, \quad \text{where } l = h +_K k$$

$$\text{m3. } lx = h(kx), \quad \text{where } l = h \cdot_K k$$

$$\text{m4. } 0_K x = 0$$

$$\text{m5. } 1_K x = x$$

The symbols 0_K and 1_K denote the zero and the unit of K . The symbols $+_K$ and \cdot_K denote the sum and the product in K . These are not part of L_K , they are symbols we use in the metalanguage. \square

2.5 Example Recall from Example 1.3 that we represent a graph with a symmetric irreflexive relation. Therefore **theory of graphs** contains the following two axioms

$$1. \neg r(x, x);$$

$$2. r(x, y) \rightarrow r(y, x).$$

\square

Our last example is a trivial one.

2.6 Example Let L be the empty language The **theory of infinite sets** is axiomatized by the sentences $\exists^{\geq n} x (x = x)$ for all positive integer n . \square

2.7 Exercise Prove that $\text{ccl}(\varphi \vee \psi) = \text{ccl}(\varphi) \cap \text{ccl}(\psi)$. \square

2.8 Exercise Prove that $T \cup \{\varphi\} \vdash \psi$ then $T \vdash \varphi \rightarrow \psi$. \square

2.9 Exercise Prove that $\text{Th}(\text{Mod}(T)) = \text{ccl}(T)$. \square

2 Elementary equivalence

The following is a fundamental notion in model theory.

2.10 Definition We say that M and N are **elementarily equivalent** if

$$\text{ee. } N \models \varphi \Leftrightarrow M \models \varphi, \quad \text{for every } \varphi \in L.$$

In this case we write $M \equiv N$. More generally, we write $M \equiv_A N$ and say that M and N are **elementarily equivalent over A** if the following hold

$$a. A \subseteq M \cap N$$

$$\text{ee'}. \text{ equivalence ee above holds for every sentence } \varphi \in L(A).$$

\square

The case when A is the whole domain of M is particularly important.

2.11 Definition When $M \equiv_M N$ we write $M \preceq N$ and say that M is an **elementary substructure** of N . \square

In the definition above the use of the term *substructure* is appropriate by the following lemma.

2.12 Lemma If M and N are such that $M \equiv_A N$ and A is the domain of a substructure of M then A is also the domain a substructure of N and the two substructures coincide.

Proof Let f be a function symbol and let r be a relation symbol. It suffices to prove that $f^M(a) = f^N(a)$ for every $a \in A^{n_f}$ and that $r^M \cap A^{n_r} = r^N \cap A^{n_r}$.

If $b \in A$ is such that $b = f^M a$ then $M \models fa = b$. So, from $M \equiv_A N$, we obtain $N \models fa = b$, hence $f^N a = b$. This proves $f^M(a) = f^N(a)$.

Now let $a \in A^{n_r}$ and suppose $a \in r^M$. Then $M \models ra$ and, by elementarity, $N \models ra$, hence $a \in r^N$. By symmetry $r^M \cap A^{n_r} = r^N \cap A^{n_r}$ follows. \square

It is not easy to prove that two structures are elementary equivalent. A direct verification is unfeasible even for the most simple structures. It will take a few chapters before we are able to discuss concrete examples.

We generalize the definition of $\text{Th}(M)$ to include parameters

$$\text{Th}(M/A) = \left\{ \varphi : \text{sentence in } L(A) \text{ such that } M \models \varphi \right\}.$$

The following proposition is immediate

2.13 Proposition For every pair of structures M and N and every $A \subseteq M \cap N$ the following are equivalent

- a. $M \equiv_A N$;
- b. $\text{Th}(M/A) = \text{Th}(N/A)$;
- c. $N \models \varphi(a) \Leftrightarrow M \models \varphi(a)$ for every $\varphi(x) \in L$ and every $a \in A^{|x|}$.
- d. $\varphi(M) \cap A^{|x|} = \varphi(N) \cap A^{|x|}$ for every $\varphi(x) \in L$.

\square

If we restate a and c of the proposition above when $A = M$ we obtain that the following are equivalent

- a'. $M \preceq N$;
- c'. $\varphi(M) = \varphi(N) \cap M^{|x|}$ for every $\varphi(x) \in L$.

Note that c' extends to all definable sets what Definition 1.7 requires for a few basic definable sets.

2.14 Example Let G be a group which we consider as a structure in the multiplicative language of groups. We show that if G is simple and $H \preceq G$ then also H is simple. Recall that G is simple if all its normal subgroups are trivial, equivalently, if for every $a \in G \setminus \{1\}$ the set $\{gag^{-1} : g \in G\}$ generates the whole group G .

Assume H is not simple then there $a, b \in H$ be such that b is not product of elements of $\{hah^{-1} : h \in H\}$. hence for every n

$$H \models \neg \exists x_1, \dots, x_n (b = x_1 a x_1^{-1} \cdots x_n a x_n^{-1})$$

By elementarity the same hold in G . Hence G is not simple. \square

2.15 Exercise Let $A \subseteq M \cap N$. Prove that $M \equiv_A N$ if and only if $M \equiv_B N$ for every finite $B \subseteq A$. \square

2.16 Exercise Let $M \preceq N$ and let $\varphi(x) \in L(M)$. Prove that $\varphi(M)$ is finite if and only if

$\varphi(N)$ is finite and in this case $\varphi(N) = \varphi(M)$. \square

2.17 Exercise Let $M \preceq N$ and let $\varphi(x, y) \in L$. Suppose there are finitely many sets of the form $\varphi(a, N)$ for some $a \in N^{|x|}$. Prove that all these sets are definable over M . \square

2.18 Exercise Consider \mathbb{Z}^n as a structure in the additive language of groups with the natural interpretation. Prove that $\mathbb{Z}^n \not\cong \mathbb{Z}^m$ for every positive integers $n \neq m$. Hint: in \mathbb{Z}^n there are at most $2^n - 1$ elements that are not congruent modulo 2. \square

3 Embeddings and isomorphisms

Here we prove that isomorphic structures are elementary equivalent and a few related results.

2.19 Definition An **embedding** of M into N is an injective map $h : M \rightarrow N$ such that

1. $a \in r^M \Leftrightarrow ha \in r^N$ for every $r \in L_{\text{rel}}$ and $a \in M^{n_r}$;
2. $f^M(a) = f^N(ha)$ for every $f \in L_{\text{fun}}$ and $a \in M^{n_f}$.

Note that when $c \in L_{\text{fun}}$ is a constant then 2 reads $hc^M = c^N$. Therefore that $M \subseteq N$ if and only if $\text{id}_M : M \rightarrow N$ is an embedding.

An surjective embedding is an **isomorphism** or, when domain and codomain coincide, an **automorphism**. \square

Condition 1 above and the assumption that h is injective can be summarized in the following

- 1'. $M \models r(a) \Leftrightarrow N \models r(ha)$ for every $r \in L_{\text{rel}} \cup \{=\}$ and every $a \in M^{n_r}$.

Note also that, by straightforward induction on syntax, from 2 we obtain

- 2' $ht^M(a) = t^N(ha)$ for every term $t(x)$ and every $a \in M^{|x|}$.

Combining these two properties, and by straightforward induction on the syntax, we obtain

3. $M \models \varphi(a) \Leftrightarrow N \models \varphi(ha)$ for every $\varphi(x) \in L_{\text{qf}}$ and every $a \in M^{|x|}$.

Recall that L_{qf} for the set of quantifier-free formulas. It is worth noting that when $M \subseteq N$ and $h = \text{id}_M$ then 3 becomes

- 3' $M \models \varphi(a) \Leftrightarrow N \models \varphi(a)$ for every $\varphi(x) \in L_{\text{qf}}$ and for every $a \in M^{|x|}$.

Finally we prove that first order truth is preserved under isomorphism. We say that a map $h : M \rightarrow N$ **fixes** $A \subseteq M$ (pointwise) if $\text{id}_A \subseteq h$. An isomorphism that fixes A is also called an **A-isomorphism**.

2.20 Theorem If $h : M \rightarrow N$ is an isomorphism then for every $\varphi(x) \in L$

$$\# \quad M \models \varphi(a) \Leftrightarrow N \models \varphi(ha) \text{ for every } a \in M^{|x|}$$

In particular, if h is an A -isomorphism then $M \equiv_A N$.

Proof We proceed by induction of the syntax of $\varphi(x)$. When $\varphi(x)$ is atomic $\#$ holds by 3 above. Induction for the Boolean connectives is straightforward so we only

need to consider the existential quantifier. Assume as induction hypothesis that

$$M \models \varphi(a, b) \Leftrightarrow N \models \varphi(ha, hb) \quad \text{for every tuple } a \in M^{|x|} \text{ and } b \in M.$$

We prove that # holds for the formula $\exists y \varphi(x, y)$.

$$\begin{aligned} M \models \exists y \varphi(a, y) &\Leftrightarrow M \models \varphi(a, b) \quad \text{for some } b \in M \\ &\Leftrightarrow N \models \varphi(ha, hb) \text{ for some } b \in M \quad (\text{by induction hypothesis}) \\ &\Leftrightarrow N \models \varphi(ha, c) \quad \text{for some } c \in N \quad (\Leftarrow \text{by surjectivity}) \\ &\Leftrightarrow N \models \exists y \varphi(ha, y). \end{aligned} \quad \square$$

2.21 Corollary *If $h : M \rightarrow N$ is an isomorphism then $h[\varphi(M)] = \varphi(N)$ for every $\varphi(x) \in L$.* \square

We can now give a few very simple examples of elementarily equivalent structures.

2.22 Example Let L be the language of strict orders. Consider intervals of \mathbb{R} (or in \mathbb{Q}) as structures in the natural way. The intervals $[0, 1]$ and $[0, 2]$ are isomorphic, hence $[0, 1] \equiv [0, 2]$ follows from Theorem 2.20. Clearly, $[0, 1]$ is a substructure of $[0, 2]$. However $[0, 1] \not\preceq [0, 2]$, in fact the formula $\forall x (x \leq 1)$ holds in $[0, 1]$ but is false in $[0, 2]$.

The example above shows that $M \subseteq N$ and $M \equiv N$ does not imply $M \preceq N$.

Now we prove that $(0, 1) \preceq (0, 2)$. By Exercise 2.15 above, it suffices to verify that $(0, 1) \equiv_B (0, 2)$ for every finite $B \subseteq (0, 1)$. This follows again by Theorem 2.20 as $(0, 1)$ and $(0, 2)$ are B -isomorphic for every finite $B \subseteq (0, 1)$. \square

For the sake of completeness we also give the definition of homomorphism.

2.23 Definition A **homomorphism** is a total map $h : M \rightarrow N$ such that

1. $a \in r^M \Rightarrow ha \in r^N$ for every $r \in L_{\text{rel}}$ and $a \in M^{n_r}$;
2. $h f^M(a) = f^N(ha)$ for every $f \in L_{\text{fun}}$ and $a \in M^{n_f}$. \square

Note that only one implication is required in 1.

2.24 Exercise Prove that if $h : N \rightarrow N$ is an automorphism and $M \preceq N$ then $h[M] \preceq N$. \square

2.25 Exercise Let L be the empty language. Let $A, D \subseteq M$. Prove that the following are equivalent

1. D is definable over M ;
2. either D is finite and $A \subseteq D$ or $\neg D$ is finite and $\neg D \subseteq A$.

Hint: as L -structures are plain sets, every bijection $f : M \rightarrow M$ is an automorphism. \square

2.26 Exercise Prove that if $\varphi(x)$ is an existential formula (i.e. of the form $\exists y \psi(x, y)$ for $\psi(x, y) \in L_{\text{qt}}$) and $h : M \rightarrow M$ is an embedding then

$$M \models \varphi(a) \Rightarrow N \models \varphi(ha) \quad \text{for every } a \in M^{|x|}. \quad \square$$

2.27 Exercise Assume that the language L contains only the symbol $+$. Consider \mathbb{Z} as

a structure of signature L with the usual interpretation of $+$. Prove that there is no existential formula that defines the set of even integers. Hint: use exercise 2.26. \square

4 Completeness

A theory T is **maximal consistent** if it is consistent and there is no consistent theory $S \supset T$. Equivalently, if T contains every sentence φ **consistent with** T , that is, such that $T \cup \{\varphi\}$ is consistent. Clearly a maximal consistent theory is closed under logical consequences.

A theory T is **complete** if $\text{ccl}T$ is maximal consistent. It is not easy to prove that a theory is complete. Explicit examples will be given in the next chapters.

2.28 Proposition *The following are equivalent*

- a. T is maximal consistent;
- b. $T = \text{Th}(M)$ for some structure M ;
- c. T is consistent and $\varphi \in T$ or $\neg\varphi \in T$ for every sentence φ .

Proof To prove $\text{a} \Rightarrow \text{b}$, assume that T is consistent. Then there is $M \models T$. Therefore $T \subseteq \text{Th}(M)$. As T is maximal consistent $T = \text{Th}(M)$. Implication $\text{b} \Rightarrow \text{c}$ is immediate. As for $\text{c} \Rightarrow \text{a}$ note that if $T \cup \{\varphi\}$ is consistent then $\neg\varphi \notin T$ therefore $\varphi \in T$ follows from c. \square

The proof of the proposition below is left as an exercise for the reader.

2.29 Proposition *The following are equivalent*

- a. T is complete;
- b. there is a unique maximal consistent theory S such that $T \subseteq S$;
- c. T is consistent and $T \vdash \text{Th}(M)$ for every $M \models T$;
- d. T is consistent and $T \vdash \varphi$ or $T \vdash \neg\varphi$ for every sentence φ ;
- e. T is consistent and $M \equiv N$ for every pair of models of T . \square

2.30 Exercise Prove that the following are equivalent

- a. T is complete;
- b. for every sentence φ , or $T \vdash \varphi$ or $T \vdash \neg\varphi$ but not both.

By contrast prove that the following are *not* equivalent

- a. T is maximal consistent;
- b. for every sentence φ , or $\varphi \in T$ or $\neg\varphi \in T$ but not both.

Hint: consider the theory containing all sentences where the symbol \neg occurs an even number of times. This theory is not consistent as it contains \perp . \square

2.31 Exercise Prove that if T has exactly 2 maximal consistent extension T_1 and T_2 then there is a sentence φ such that $T, \varphi \vdash T_1$ and $T, \neg\varphi \vdash T_2$. \square

5 The Tarski-Vaught test

There is no natural notion of *smallest* elementary substructure containing a set of parameters A . The downward Löwenheim-Skolem, which we prove in the next section, is the best result that holds in full generality. Given an arbitrary $A \subseteq N$ we shall construct a model $M \preceq N$ containing A that is small in the sense of cardinality. The construction selects one by one the elements of M that are required to realise the condition $M \preceq N$. Unfortunately, Definition 2.11 supposes full knowledge of the truth in M and it may not be applied during the construction. The following lemma comes to our rescue with a property equivalent to $M \preceq N$ that only mention the truth in N .

2.32 Lemma(Tarski-Vaught test) *For every $A \subseteq N$ the following are equivalent*

1. A is the domain of a structure $M \preceq N$;
2. for every formula $\varphi(x) \in L(A)$, with $|x| = 1$,
 $N \models \exists x \varphi(x) \Rightarrow N \models \varphi(b)$ for some $b \in A$.

Proof $1 \Rightarrow 2$

$$\begin{aligned} N \models \exists x \varphi(x) &\Rightarrow M \models \exists x \varphi(x) \\ &\Rightarrow M \models \varphi(b) && \text{for some } b \in M \\ &\Rightarrow N \models \varphi(b) && \text{for some } b \in M \end{aligned}$$

$2 \Rightarrow 1$ we first check that A is the domain of a substructure of N , that is, $f^N a \in A$ for every function symbol f and every $a \in A^{n_f}$. This follows from 2 with $\varphi(x) = f a = x$ for $\varphi(x)$.

Write M for the substructure of N with domain A . By induction on the syntax we prove

$$M \models \varphi(a) \Leftrightarrow N \models \varphi(a) \quad \text{for every } a \in M^{|x|}.$$

If $\varphi(x)$ is atomic the claim follows from $M \subseteq N$ and the remarks underneath Definition 2.19. The case of Boolean connectives is straightforward, so only the existential quantifier requires a proof

$$\begin{aligned} M \models \exists x \varphi(a, x) &\Leftrightarrow M \models \varphi(a, b) && \text{for some } b \in M \\ &\Leftrightarrow N \models \varphi(a, b) && \text{for some } b \in M \text{ by induction hypothesis} \\ &\Leftrightarrow N \models \exists x \varphi(a, x) && \Leftarrow \text{uses 2.} \quad \square \end{aligned}$$

2.33 Exercise Consider \mathbb{R} in the language of strict orders. Prove that $\mathbb{R} \setminus \{0\} \preceq \mathbb{R}$. \square

6 The downward Löwenheim-Skolem theorem

The main theorem of this section was proved by Löwenheim at the beginning of the last century. Skolem gave a simpler proof immediately afterwards. At the time, the result was perceived as paradoxical.

A few years earlier, the works of Zermelo and Fraenkel had convinced the mathematical community that the whole of set theory could be formalised in a first order language. The downward Löwenheim-Skolem theorem implies the existence of an

infinite countable model M of set theory: this is the so-called **Skolem paradox**. The existence of M seems paradoxical because, in particular, a sentence that formalises the subset axiom holds in M . Therefore M contains an element b which, in M , is the set of subsets of the natural numbers. But the set of elements of b is a subset of M , and therefore it is countable. In fact, this is not a contradiction, because the expression ‘all subsets of the natural numbers’ does not have the same meaning in M as it has in the real world. The notion of cardinality, too, acquires a different meaning. In the language of set theory, there is a first order sentence that formalises the fact that b is uncountable: the sentence says that there is no bijection between b and the natural numbers. Therefore the bijection between the elements of b and the natural numbers (which exists in the real world) does not belong to M . The notion of equinumerosity has a different meaning in M and in the real world, but those who live in M cannot realise this.

2.34 Downward Löwenheim-Skolem Theorem *Let N be an infinite L -structure and let $A \subseteq N$. Then there is a structure M of cardinality $\leq |L(A)|$ such that $A \subseteq M \preceq N$.*

Proof Set $\lambda = |L(A)|$. Below we construct a chain $\langle A_i : i < \lambda \rangle$ of subsets of N . The chain begins at $A_0 = A$ and is continuous at limit ordinals. Finally we set $M = \bigcup_{i < \lambda} A_i$. All A_i will have cardinality $\leq \lambda$ so $|M| \leq \lambda$ follows.

Now we construct A_{i+1} given A_i . Assume as induction hypothesis that $|A_i| \leq \lambda$. Then $|L(A_i)| \leq \lambda$. For some fixed variable x let $\langle \varphi_k(x) : k < \lambda \rangle$ be an enumeration of the formulas in $L(A_i)$ that are consistent in N . For every k pick $a_k \in N$ such that $N \models \varphi_k(a_k)$. Define $A_{i+1} = A_i \cup \{a_k : k < \lambda\}$. Then $|A_{i+1}| \leq \lambda$ is clear.

Now use Tarski-Vaught test to prove $M \preceq N$. Suppose $\varphi(x) \in L(M)$ is consistent in N . As finitely many parameters occur in formulas, $\varphi(x) \in L(A_i)$ for some i . Then $\varphi(x)$ is among the formulas enumerated above and $A_{i+1} \subseteq M$ contains a solution of $\varphi(x)$. □

We will need to adapt the construction above to meet more requirements on the model M . To better control the elements that enters M it is convenient to add one element at the time (above we add λ elements at each stage). We need to enumerate formulas with care if we want to complete the construction by stage λ .

2.35 Second proof of the downward Löwenheim-Skolem Theorem From set theory we know there is a bijection $\pi : \lambda \rightarrow \lambda^2$ such that $\pi_1(i)$ and $\pi_2(i)$, the two components of $\pi(i)$, are both $\leq i$. Suppose we have defined the set A_j for every $j \leq i$. Let $\langle \varphi_k(x) : k < \lambda \rangle$ be an enumeration of the formulas in $L(A_{\pi_1(i)})$. Let b be a solution of the $\varphi_k(x)$ for $k = \pi_2(i)$ and define $A_{i+1} = A_i \cup \{b\}$.

We use Tarski-Vaught test to prove $M \preceq N$. Let $\varphi(x) \in L(M)$ be consistent in N . As π is a bijection, there is an i such that $\varphi(x)$ is the $\pi_2(i)$ -th formula enumerated in $L(A_{\pi_1(i)})$. Hence a witness of $\varphi(x)$ is enumerated in M at stage i . □

2.36 Exercise Assume L is countable and let $M \preceq N$ have arbitrary (large) cardinality. Let $A \subseteq N$ be countable. Prove that there is a countable model K such that $A \subseteq K \preceq N$ and $K \cap M \preceq M$ (in particular, $K \cap M$ is a model). Hint: adapt the construction used to prove the downward Löwenheim-Skolem Theorem. □

7 Elementary chains

An **elementary chain** is a chain $\langle M_i : i < \lambda \rangle$ of structures such that $M_i \preceq M_j$ for every $i < j < \lambda$. The **union** (or **limit**) of the chain is the structure with as domain the set $\bigcup_{i < \lambda} M_i$ and as relations and functions the union of the relations and functions of M_i . It is plain that all structures in the chain are substructures of the limit.

2.37 Lemma *Let $\langle M_i : i \in \lambda \rangle$ be an elementary chain of structures. Let N be the union of the chain. Then $M_i \preceq N$ for every i .*

Proof By induction on the syntax of $\varphi(x) \in L$ we prove

$$M_i \models \varphi(a) \Leftrightarrow N \models \varphi(a) \quad \text{for every } i < \lambda \text{ and every } a \in M_i^{|x|}$$

As remarked in 3' of Section 3, the claim holds for quantifier-free formulas. Induction for Boolean connectives is straightforward so we only need to consider the existential quantifier

$$\begin{aligned} M_i \models \exists y \varphi(a, y) &\Rightarrow M_i \models \varphi(a, b) && \text{for some } b \in M_i. \\ &\Rightarrow N \models \varphi(a, b) && \text{for some } b \in M_i \subseteq N \end{aligned}$$

where the second implication follows from the induction hypothesis. Vice versa

$$N \models \exists y \varphi(a, y) \Rightarrow N \models \varphi(a, b) \quad \text{for some } b \in N$$

As a is a finite tuple, $a \in M_j^{|x|}$ for some j . We can also assume that $b \in M_j$. Then we can apply the induction hypothesis to $\varphi(x, y)$

$$\begin{aligned} &\Rightarrow M_j \models \varphi(a, b) && \text{for some } b \in M_j \\ &\Rightarrow M_j \models \exists y \varphi(a, y) \\ &\Rightarrow M_i \models \exists y \varphi(a, y) \end{aligned}$$

The last implication holds because $M_i \preceq M_j$ or, in case $j < i$, because $M_j \preceq M_i$. □

2.38 Exercise Let $\langle M_i : i \in \lambda \rangle$ be an elementary chain of substructures of N . Let M be the union of the chain. Prove that $M \preceq N$. □

2.39 Exercise Give an alternative answer to Exercise 2.36 using the downward Löwenheim-Skolem Theorem (instead of its proof). Hint: construct two countable chains of countable models such that $K_i \cap M \subseteq M_i \preceq N$ and $A \cup M_i \subseteq K_{i+1} \preceq N$. The required model is $K = \bigcup_{i \in \omega} K_i$. In fact it is easy to check that $K \cap M = \bigcup_{i \in \omega} M_i$. □

Chapter 3

Ultraproducts and compactness

1 Filters and ultrafilters

The matter in this section will be reconsidered in a more general setting in Chapter 4.

Let I be any set. A **filter on I** (or a **filter of $\mathcal{P}(I)$**), is a non empty set $F \subseteq \mathcal{P}(I)$ such that for every $a, b \in \mathcal{P}(I)$:

$$f1 \quad a \in F \text{ and } a \subseteq b \Rightarrow b \in F$$

$$f2 \quad a \in F \text{ and } b \in F \Rightarrow a \cap b \in F$$

A filter F is **proper** if $F \neq \mathcal{P}(I)$, equivalently if $\emptyset \notin F$. Otherwise it is **improper**. By f1 above, F is proper if and only if $\emptyset \notin F$. A filter F is **principal** if $F = \{a \subseteq I : b \subseteq a\}$ for some set $b \subseteq I$. We say that $\{b\}$ **generates** F . When I is finite every filter F is principal generated by $\{\cap F\}$. Non principal filters on I exist as soon as I is infinite. In fact, if I is infinite it is easy to check that the following is filter

$$F = \{a \subseteq I : a \text{ is cofinite in } I\}$$

where **cofinite (in I)** means that $I \setminus a$ is finite. This is called the **Fréchet's filter on I** . The Fréchet's filter is the minimal non principal filter.

3.1 Exercise Let I be infinite. Prove that the Fréchet's filter on I contains all non principal filters on I . □

A proper filter F is **maximal** if there is no proper filter H such that $F \subset H$. A proper filter F is an **ultrafilter** if for every $a \subseteq I$:

$$a \notin F \Rightarrow \neg a \in F \quad \text{where } \neg a \stackrel{\text{def}}{=} I \setminus a$$

Below we prove that ultrafilters are exactly the maximal filters.

Let $B \subseteq \mathcal{P}(I)$ the **filter generated by B** is the intersection of all filters containing B . It is easy to check that the intersection of a family of filters is a filter, so the notion is well-defined. The following easy proposition gives a workable characterization of the filter generated by a set.

3.2 Proposition The filter generated by B is $\{a \subseteq I : \cap C \subseteq a \text{ for some finite } C \subseteq B\}$. □

We say that B has the **finite intersection property** if $\cap C \neq \emptyset$ for every finite $C \subseteq B$. The following proposition is immediate.

3.3 Proposition The following are equivalent

1. the filter generated by B is non principal;
 2. B has the finite intersection property.
-

A proper filter F is **prime** if for every $a, b \subseteq I$.

$$a \cup b \in F \Rightarrow a \in F \text{ or } b \in F$$

Prime filters coincide with maximal filter. (However, in Chapter 4 we introduce more a general context where primeness is distinct from maximality.)

3.4 Proposition For every filter F on I , the following are equivalent:

1. F is a maximal filter;
2. F is a prime filter;
3. F is an ultrafilter.

Proof $1 \Rightarrow 2$. Suppose $a, b \notin F$, where F is maximal. We claim that $a \cup b \notin F$. By maximality, there is a $c \in F$ such that $a \cap c = b \cap c = \emptyset$. Therefore $(a \cup b) \cap c = \emptyset$. Hence $a \cup b \notin F$.

$2 \Rightarrow 3$. It suffices to note that $I = a \cup \neg a \in F$.

$3 \Rightarrow 1$. If $a \notin F$, where F is an ultrafilter, then $\neg a \in F$ and no proper filter contains $F \cup \{a\}$. \square

3.5 Proposition Let $B \subseteq \mathcal{P}(I)$ have the finite intersection property, then B is contained in a maximal filter.

Proof First prove that the union of a chain of subsets of $\mathcal{P}(I)$ with the finite intersection property has the finite intersection property. Let \mathcal{B} be such a chain and suppose for a contradiction that $\bigcup \mathcal{B}$ does not have the finite intersection property. Fix a finite $C \subseteq \bigcap \mathcal{B}$ such that $\bigcap C = \emptyset$. As C is finite, $C \subseteq B$ for some $B \in \mathcal{B}$. Hence B does not have the finite intersection property. A contradiction.

Now apply Zorn lemma to obtain a $B \subseteq \mathcal{P}(I)$ which is maximal among the sets with the finite intersection property. It is immediate that B is a filter. \square

3.6 Exercise Prove that all principal ultrafilters are generated by a singleton. \square

2 Direct products

In this and the next section $\langle M_i : i \in I \rangle$ is a sequence of L -structures. (We are abusing of the word sequence, I is only a set.) The **direct product** of this sequence is a structure denoted by

$$\prod_{i \in I} M_i \quad \text{below this is abbreviated by } N$$

and is defined by conditions 1-3 below. If $M_i = M$ for all $i \in I$ we say that N is a **direct power** of M and denote it by M^I .

The domain of N is the set of the functions

$$\begin{aligned} 1. \quad \hat{a} &: I \rightarrow \bigcup_{i \in I} M_i \\ \hat{a} &: i \mapsto \hat{a}i \in M_i \end{aligned}$$

We confuse tuples of elements of N with tuple-valued functions, for instance, the tuple $\hat{a} = \langle \hat{a}_1 \dots \hat{a}_n \rangle$ is confused with the function $\hat{a} : i \mapsto \hat{a}i = \langle \hat{a}_1 i, \dots, \hat{a}_n i \rangle$. But at

a first reading it might be convenient to pretend that all functions and relations are unary.

The interpretation of $f \in L_{\text{fun}}$ is defined as follows

$$2. \quad (f^N \hat{a})i = f^{M_i}(\hat{a}i) \quad \text{for all } i \in I,$$

The interpretation of $r \in L_{\text{rel}}$ is the product of the relations r^{M_i} , that is, we define

$$3. \quad \hat{a} \in r^N \Leftrightarrow \hat{a}i \in r^{M_i} \quad \text{for all } i \in I.$$

The following proposition is immediate.

3.7 Proposition *If only the connectives \wedge, \forall and \exists occur in $\varphi(x) \in L$ then for every $\hat{a} \in N^{|x|}$*

$$\# \quad N \models \varphi(\hat{a}) \Leftrightarrow M_i \models \varphi(\hat{a}i) \quad \text{for all } i \in I.$$

Proof By induction on syntax. First note that we can extend 2 to all terms $t(x)$

$$2'. \quad (t^N \hat{a})i = t^{M_i}(\hat{a}i) \quad \text{for all } i \in I.$$

Combining 3 and 2' we obtain that for every $r \in L_{\text{rel}} \cup \{=\}$ and every L -term $t(x)$

$$3'. \quad N \models rt\hat{a} \Leftrightarrow M_i \models rt\hat{a}i \quad \text{for all } i \in I$$

This shows that $\#$ holds for $\varphi(x)$ atomic. Induction for the connectives \wedge, \forall and \exists is immediate. \square

A consequence of Proposition 3.7 is that the direct product of groups, rings and vector spaces are structures of the same sort. Notably, the product of fields is not a field.

3 Quotient structures

If E is an equivalence relation on N we write $[c]_E$ for the equivalence class of $c \in N$. We use the same symbol for the equivalence relation on M^n defined as follows: if $a = a_1, \dots, a_n$ and $b = b_1, \dots, b_n$ are n -tuples of elements of N then $a E b$ means that $a_i E b_i$ holds for all i . By $[a]_E$ we mean both the equivalence class of $a \in M^n$ or the tuple of equivalence classes $[a_1]_E, \dots, [a_n]_E$. The context will disambiguate.

An equivalence relation E on a structure N is a **congruence** if for every $f \in L_{\text{fun}}$

$$c1. \quad a E b \Rightarrow f^N a E f^N b;$$

When E is a congruence on N we write N/E for the structure that has as domain the set of E -equivalence classes in N and the following interpretation of $f \in L_{\text{fun}}$ and $r \in L_{\text{rel}}$:

$$c2. \quad f^{N/E}[a]_E = [f^N a]_E;$$

$$c3. \quad [a]_E \in r^{N/E} \Leftrightarrow [a]_E \cap r^N \neq \emptyset.$$

By c1 this interpretation is well-defined. We call N/E the **quotient structure**. The reader will recognize it as a familiar notion by the following proposition (which is not required in the following and requires the notion of homomorphism, see Definition 2.23. Recall that the **kernel** of a total map $h : N \rightarrow M$ is the equivalence relation E such that

$$a E b \Leftrightarrow ha = hb$$

for every $a, b \in N$.

3.8 Proposition Let $h : N \rightarrow M$ be a surjective homomorphism and let E be the kernel of h . Then there is an isomorphism k that makes the following diagram commute

$$\begin{array}{ccc} N & \xrightarrow{h} & M \\ \downarrow \pi & \nearrow k & \\ N/E & & \end{array}$$

where $\pi : a \mapsto [a]_E$ is the projection map. □

! Quotients clutter the notation with brackets. To avoid the mess, we prefer to reason in N and tweak the satisfaction relation. We replace equality with E and adapt the interpretation of relation symbols to c3 above. The following definition makes this precise (an inductive definition is required for instance in the proof of Łoś Theorem). Warning: it is not standard.

3.9 Definition For t_1, t_2 closed terms of $L(N)$ define

$$1^* \quad N/E \models^* t_1 = t_2 \Leftrightarrow [t_1^N]_E = [t_2^N]_E$$

For t a tuple of closed terms of $L(N)$ and $r \in L_{\text{rel}}$ a relation symbol

$$2^* \quad N/E \models^* rt \Leftrightarrow [t^N]_E \in r^{N/E}$$

Finally the definition is extended to all sentences $\varphi \in L(N)$ by induction in the usual way

$$3^* \quad N/E \models^* \neg \varphi \Leftrightarrow \text{not } N/E \models^* \varphi$$

$$4^* \quad N/E \models^* \varphi \wedge \psi \Leftrightarrow N/E \models^* \varphi \text{ and } N/E \models^* \psi$$

$$5^* \quad N/E \models^* \exists x \varphi(x) \Leftrightarrow N/E \models^* \varphi(\hat{a}) \text{ for some } a \in N. \quad \square$$

Now, by induction on the syntax of formulas one can prove \models^* does what required. In particular, $N/E \models^* \varphi(a) \Leftrightarrow \varphi(b)$ for every $a E b$.

3.10 Proposition Let E be a congruence relation of N . Then the following are equivalent for every $\varphi(x) \in L$

$$1. \quad N/E \models^* \varphi(a);$$

$$2. \quad N/E \models \varphi([a]_E). \quad \square$$

4 Ultraproducts

Assume the notation of the previous sections. In particular, $\langle M_i : i \in I \rangle$ is a sequence of structures and N is the direct product of this sequence.

Let F be a filter on I . We define the following congruence relation on N

$$\hat{a} \sim_F \hat{c} \Leftrightarrow \{ i \in I : \hat{a}i = \hat{c}i \} \in F.$$

It is easy to verify that \sim_F is indeed an equivalence relation: reflexivity and symmetry are evident, transitivity follows from f2 in Section 1. It is also clear that \sim_F a congruence.

We now introduce an ad hoc notion quotient structure N/F by tweaking the definition in Section 3. Namely, we replace c3 by the following

$$\text{c3'}. \quad [\hat{a}]_{\sim_F} \in r^{N/F} \Leftrightarrow \{i : \hat{a}i \in r^{M_i}\} \in F$$

We will not elaborate on the difference between c3 and c3'. We refer the interested reader to Exercise 3.14 and 3.15.

The structure N/F is called the **reduced product** of the structures $\langle M_i : i \in I \rangle$ or, when $M_i = M$ for all $i \in I$, the **reduced power** of M . When F is an ultrafilter we say **ultraproduct**, respectively **ultrapower**.

We the notation introduced in Definition 3.9 (see also Proposition 3.10) we are ready to state a fundamental theorem about ultraproducts.

3.11 Łoś Theorem *Let $\varphi(x) \in L$ and let F be an ultrafilter on I . Then for every $\hat{a} \in N^{|x|}$, the following are equivalent*

1. $N/F \models^* \varphi(\hat{a})$;
2. $\{i : M_i \models \varphi(\hat{a}i)\} \in F$.

Proof We proceed by induction of the syntax of $\varphi(x)$. Suppose $\varphi(x)$ is of the form $rt(x)$ for some tuple of terms $t(x)$ and $r \in L_{\text{rel}} \cup \{=\}$. Then $1 \Leftrightarrow 2$ follows easily from c3'.

We prove the inductive step for the connectives \neg , \wedge , and \exists . We begin with \neg . This is the only place in the proof where the assumption that F is an *ultrafilter* is required. From the inductive hypothesis we obtain

$$N/F \models^* \neg\varphi(\hat{a}) \Leftrightarrow \{i : M_i \models \varphi(\hat{a}i)\} \notin F$$

So, as F is an *ultrafilter*

$$\Leftrightarrow \{i : M_i \models \neg\varphi(\hat{a}i)\} \in F.$$

Now consider \wedge . Assume as inductive hypothesis that the equivalence $1 \Leftrightarrow 2$ holds for $\varphi(x)$ and $\psi(x)$, then

$$N/F \models^* \varphi(\hat{a}) \wedge \psi(\hat{a}) \Leftrightarrow \{i : M_i \models \varphi(\hat{a}i)\} \in F \text{ and } \{i : M_i \models \psi(\hat{a}i)\} \in F$$

As filters are closed under intersection we obtain

$$\Leftrightarrow \{i : M_i \models \varphi(\hat{a}i) \wedge \psi(\hat{a}i)\} \in F.$$

Finally consider $\exists y$. Assume as inductive hypothesis that the equivalence $1 \Leftrightarrow 2$ holds for $\varphi(x, y)$, then

$$\begin{aligned} N/F \models^* \exists y \varphi(\hat{a}, y) &\Leftrightarrow N/F \models^* \varphi(\hat{a}, \hat{b}) && \text{for some } \hat{b} \in N \\ &\Leftrightarrow \{i : M_i \models \varphi(\hat{a}i, \hat{b}i)\} \in F && \text{for some } \hat{b} \in N. \end{aligned}$$

We claim this is equivalent to

$$\Leftrightarrow \{i : M_i \models \exists y \varphi(\hat{a}i, y)\} \in F.$$

In fact, direction \Rightarrow is trivial and, as for \Leftarrow , we choose as \hat{b} a sequence that picks a witness of $M_i \models \exists y \varphi(\hat{a}i, y)$ if it exists, some arbitrary element of M_i otherwise. \square

Let a^I denote the element of M^I that has constant value a . The following is an immediate consequence of Łoś theorem.

3.12 Corollary *For every $a \in M$*

$$M^I/F \models^* \varphi(a^I) \Leftrightarrow M \models \varphi(a). \quad \square$$

Often one identifies M with its image under the map $h : a \mapsto [a^I]_F$ and say that M^I/F is an elementary extension of M .

3.13 Corollary *Every infinite structure has a proper elementary extension.*

Proof Let M be an infinite structure and let F be a *non principal* ultrafilter on ω . By the above remark, it suffices to show that $h[M]$ is a *proper* substructure of M^ω/F . As M is infinite, there is an injective function $\hat{d} \in M^\omega$. Then for every $a \in M$ the set $\{i : \hat{d}i = a\}$ is either empty or a singleton and, as F is non principal, it does not belong to F . So, by Łoś Theorem $M^\omega/F \models^* \hat{d} \neq a^I$ for every $a \in M$, that is, $[\hat{d}]_F \notin h[M]$. \square

3.14 Exercise Prove that c3 differs from 3c' only if $\{i : r^{M_i} = \emptyset\} \in F$.

3.15 Exercise Let o be some fresh element. Let M'_i is a superstructures of M_i with domain $M_i \cup \{o\}$, relations $r^{M'_i} = r^{M_i} \cup \{o\}^{n_r}$, and functions $f^{M'_i}$ just any extension of f^{M_i} . Denote by N' the direct product of the sequence $\langle M'_i : i \in I \rangle$. The element of N' that is identically o is denoted by o^I .

It is easy to verify that the \sim_F -equivalence classes that do not contain o^I form substructure of

Let N'/\sim_F be the quotient structure as defined in Section 3. Prove that N/F is isomorphic to the substructure of N'/\sim_F that contains all equivalence classes but for the one of o^I . \square

5 The compactness theorem

A theory is **finitely consistent** if all its finite subsets are consistent. The following theorem is the *fiat lux* of model theory.

3.16 Compactness Theorem *Every finitely consistent theory is consistent.*

Proof Let T be a finitely consistent theory. We claim that the structure N/F which we define below is a model of T . Let I be the set of consistent sentences I in the language L . For every $\zeta \in I$ pick some $M_\zeta \models \zeta$. For any sentence $\varphi \in L$ we write X_φ for the following subset of I

$$X_\varphi = \{\zeta \in I : \zeta \vdash \varphi\}$$

Clearly φ is consistent if and only if $X_\varphi \neq \emptyset$. Moreover $X_{\varphi \wedge \psi} = X_\varphi \cap X_\psi$. hence, as T is finitely consistent, the set $B = \{X_\varphi : \varphi \in T\}$ has the finite intersection property. Therefore B extends to an ultrafilter F on I . Define

$$N = \prod_{\zeta \in I} M_\zeta.$$

We claim that $N/F \models T$. By Łoś Theorem, for every sentence $\varphi \in L$

$$N/F \models \varphi \Leftrightarrow \{\zeta : M_\zeta \models \varphi\} \in F.$$

By the definition of F , for every $\varphi \in T$, the set $X_\varphi \subseteq \{\zeta : M_\zeta \models \varphi\}$ belongs to F .

Therefore $N/F \models T$, et lux fuit. \square

The compactness theorem can be formulated in the following apparently stronger way.

3.17 Corollary *If $T \vdash \varphi$ then there is some finite $S \subseteq T$ such that $S \vdash \varphi$.*

Proof Suppose $S \not\vdash \varphi$ for every finite $S \subseteq T$. Then for every finite $S \subseteq T$ there is a model $M \models S \cup \{\neg\varphi\}$. In other words, $T \cup \{\neg\varphi\}$ is finitely consistent. By compactness $T \cup \{\neg\varphi\}$ hence $T \not\vdash \varphi$. \square

3.18 Exercise Let $\Phi \subseteq L$ be a set of sentences and suppose that $\vdash \psi \leftrightarrow \bigvee \Phi$ for some sentence ψ . Prove that there is a finite $\Phi_0 \subseteq \Phi$ such that $\vdash \psi \leftrightarrow \bigvee \Phi_0$. \square

6 Finite axiomatizability

A theory T is **finitely axiomatizable** if $\text{ccl}(S) = \text{ccl}(T)$ for some finite S . The following theorem shows that we can restrict the search of S to the subsets of T .

3.19 Proposition *For every theory T the following are equivalent*

1. T is finitely axiomatizable;
2. there a finite $S \subseteq T$ such that $S \vdash T$.

Proof Only $1 \Rightarrow 2$ requires a proof. If T is finitely axiomatizable, there is a sentence φ such that $\text{ccl}(\varphi) = \text{ccl}(T)$. Then $T \vdash \varphi \vdash T$. By Proposition 3.17 there is a finite $S \subseteq T$ such that $S \vdash \varphi$. Then also $S \vdash T$. \square

If L is empty, then every structure is a model. The **theory of infinite sets** is the set of sentences that hold in every infinite structure.

3.20 Example The theory of infinite sets is not finitely axiomatizable. Define

$$T_\infty = \{ \exists^{\geq n} x (x = x) : n \in \omega \}$$

Every infinite set is a model of T_∞ and, vice versa, every model of T_∞ is an infinite set. Then $\text{ccl}(T_\infty)$ is the theory of infinite sets. Suppose for a contradiction that T_∞ is finitely axiomatizable. By Proposition 3.19, $\exists^{\geq n} x (x = x) \vdash T_\infty$ for some n . Any set of cardinality $n + 1$ proves that this is not the case. \square

The following is a less trivial example. It proves a claim we made in Exercise 1.15.

3.21 Example Write T_{gph} for the theory of graphs, see Example 2.5. Let \mathcal{K} be the following class of structures

$$\mathcal{K} = \left\{ M \models T_{\text{gph}} : r^M \subseteq (A \times \neg A) \cup (\neg A \times A) \text{ for some } A \subseteq M \right\}$$

We prove that \mathcal{K} is axiomatizable but not finitely axiomatizable, i.e. $\mathcal{K} = \text{Mod}(T)$ for some theory T , but T cannot be chosen finite.

A **path of length n** in M is a sequence $c_0, \dots, c_n \in M$ such that $M \models r(c_i, c_{i+1})$ for every $0 \leq i < n$. A path is **closed** if $c_0 = c_n$. We claim that the following theory axiomatizes \mathcal{K} .

$$T = \left\{ \neg \exists x_0, \dots, x_{2n+1} \left[\bigwedge_{i=0}^{2n} r(x_i, x_{i+1}) \wedge x_0 = x_{2n+1} \right] : n \in \omega \right\}$$

In words, T says that all closed paths have even length. Inclusion $\mathcal{K} \subseteq \text{Mod}(T)$ is clear, we prove $\text{Mod}(T) \subseteq \mathcal{K}$. Let $M \models T$ and let $A_o \subseteq M$ contain exactly one point for every connected component of M . Define

$$A = \left\{ b : M \models \exists x_0, \dots, x_{2n} \left[\bigwedge_{i=1}^{2n-1} r(x_i, x_{i+1}) \wedge a = x_0 \wedge x_{2n} = b \right], a \in A_o, n \in \omega \right\}$$

We claim that $r^M \subseteq (A \times \neg A) \cup (\neg A \times A)$, hence $M \in \mathcal{K}$. We need to verify that if $r(b, c)$ then neither $b, c \in A$ nor $b, c \in \neg A$. Suppose for a contradiction that $r(b, c)$ and $b, c \in A$ (the case $b, c \in \neg A$ is similar). As b and c belong to the same connected component, there are two paths b_0, \dots, b_{2n} and c_0, \dots, c_{2m} that connect $a = b_0 = c_0 \in A_o$ to $b = b_{2n}$ and $c = c_{2m}$. Then $a, b_1, \dots, b_{2n}, c_{2m}, \dots, c_1, a$ is a closed path of odd length. A contradiction.

We now prove that \mathcal{K} is not finitely axiomatizable. By Proposition 3.19 it suffices to note that no finite $S \subseteq T$ axiomatizes \mathcal{K} . \square

7 Realization of types and upword Löwenheim-Skolem

Recall that a **type** is a set of formulas. When we present types we usually declare the variables that may occur in it – we write $p(x), q(x)$, etc. where x is a tuple of variables. Clearly, when x is the empty tuple, $p(x)$ is just a theory. We identify a finite types with the conjunction of the formulas contained in it.

We write $M \models p(a)$ if $M \models \varphi(a)$ for every $\varphi(x) \in p$. We say that a is a **solution** or a **realization** of $p(x)$. An equivalent notation is $M, a \models p(x)$ or, when M is clear from the context, $a \models p(x)$. We say that $p(x)$ is **consistent in M** it has a solution in M . In this case we may write $M \models \exists x p(x)$. We say that $p(x)$ is **consistent** tout court if it is consistent in some model.

We say that a type $p(x)$ is **finitely consistent** if all its finite subsets are consistent. If they are all consistent in the same model M , we say that $p(x)$ is **finitely consistent in M** . The following theorem shows that the latter notion, which is trivial for theories, is very interesting for types.

3.22 Compactness Theorem for types *Every finitely consistent type $p(x) \subseteq L$ is consistent. Moreover, if $p(x) \subseteq L(M)$ is finitely consistent in M then it is realised in some elementary extension of M .*

Proof Let L' be an expansion of L with a fresh constant symbols c , a tuple of constants of the same length as x . Then $p(c)$ is a finitely consistent theory in the language L' . By the compactness theorem there is an L' -structure $N' \models p(c)$. Let N be the reduced of N' to L . That is the L -structure with the same domain and the same interpretation of N' on the symbols of L . Then N realises $p(x)$.

As for the second claim, let a be an enumeration of M . We can assume that $p(x)$ has the form $p'(x, a)$ for some $p'(x, z) \in L$. Define

$$q(z) = \{ \varphi(z) : M \models \varphi(a) \}$$

Clearly, $q(z) \cup p'(x, z)$ is finitely consistent (in M). By the first part of the proof

there is a model N' such that $N' \models p'(a', b')$ for some a', b' . Let h be the function $\{\langle a, a' \rangle\}$. Then $h : M \rightarrow N'$ is an embedding and we can identify M with $h[M]$. The proposition is proved if we can show that $h[M] \preceq N'$. For any $\varphi(z) \in L$ we have

$$h[M] \models \varphi(ha) \Leftrightarrow M \models \varphi(a) \Leftrightarrow \varphi(z) \in q(z) \Rightarrow N' \models \varphi(a').$$

Then $h[M] \preceq N'$ follows because all sentences in $L(M)$ have the form $\varphi(a)$ for some $\varphi(z) \in L$. □

The following corollary is historically important.

3.23 Upward Löwenheim-Skolem Theorem *Every infinite structure has arbitrarily large elementary extensions.*

Proof Let $x = \langle x_i : i < \lambda \rangle$ be a tuple distinct variables, where λ is an arbitrary cardinal. The type $p(x) = \{x_i \neq x_j : i < j < \lambda\}$ is finitely consistent in every infinite structure and every structure that realises p has cardinality $\geq \lambda$. Hence the claim follows from Theorem 3.22. □

Chapter 4

Types and morphisms

In Section 1 and 2 we introduce distributive lattices and prime filters and prove Stone's representation theorem for distributive lattices. In Section 3 we discuss lattices that arise from sets of formulas and their prime filters (prime types). These sections are mainly required for the discussion of Hilbert's Nullstellensatz in Section 7.7 below.

There is a lot of notation in this chapter. The reader may skim through and return to it when we refer to it.

1 Semilattices and filters

A **preorder** is a set \mathbb{P} with a transitive and symmetric relation which we usually denote by \leq . If $A, B \subseteq \mathbb{P}$ we write $A \leq B$ if $a \leq b$ for every $a \in A$ and $b \in B$. We write $a \leq B$ and $A \leq b$ for $\{a\} \leq B$ and $A \leq \{b\}$ respectively.

Quotienting a preorder by the equivalence relation

$$a \sim b \Leftrightarrow a \leq b \text{ and } b \leq a$$

gives a **(partial) order**. We often do not distinguish between a preorder and the partial order associated to it. Preorders are very common; here we are interested in the one induced by the relation of logical consequence

$$\varphi \leq \psi \Leftrightarrow \varphi \vdash \psi,$$

or, more generally, by the relation of **logical consequence over a theory T** , that is

$$\varphi \leq \psi \Leftrightarrow T \cup \{\varphi\} \vdash \psi.$$

A partial order \mathbb{P} is a **lower semilattice** if for each pair $a, b \in \mathbb{P}$ there is a maximal element c such that $c \leq \{a, b\}$. We call c the **meet** of a and b . The meet is unique and is denoted by $a \wedge b$. Dually, a partial order is an **upper semilattice** if for each pair of elements a and b there is a minimal element c such that $\{a, b\} \leq c$. This c is called the **join** of a and b . The join is unique and is denoted by $a \vee b$. A **lattice** is simultaneously a lower and an upper semilattice.

An element c such that $c \leq \mathbb{P}$ is called a **lower bound** or a **bottom**. An element such that $\mathbb{P} \leq c$ is called an **upper bound** or a **top**. Lower and upper bounds are unique and will be denoted by **0**, respectively **1**. Other symbols common in the literature are \perp , respectively \top . A semilattice is **bounded** if it has both an upper and a lower bound.

For the rest of this section we assume that \mathbb{P} is a bounded lower semilattice.

The meet is associative and commutative

$$\begin{aligned} (a \wedge b) \wedge c &= a \wedge (b \wedge c) \\ a \wedge b &= b \wedge a. \end{aligned}$$

Hence we may unambiguously write $a_1 \wedge \cdots \wedge a_n$. When $C \subseteq \mathbb{P}$ is finite, we write $\wedge C$ for the meet of all the elements of C . We agree that $\wedge \emptyset = 1$.

In an upper semilattice, the dual properties hold for the join. We write $\vee C$ for the join of all elements in C and we agree that $\vee \emptyset = 0$.

A **filter** of \mathbb{P} is a non-empty set $F \subseteq \mathbb{P}$ that satisfies the following for all $a, b \in \mathbb{P}$

f1. $a \in F$ and $a \leq b \Rightarrow b \in F$

f2. $a, b \in F \Rightarrow a \wedge b \in F$.

We say that F is a **proper filter** if $F \neq \mathbb{P}$, equivalently if $0 \notin F$. We say that F is **principal** if $F = \{b : a \leq b\}$ for some $a \in \mathbb{P}$. More precisely, we say that F is the **principal filter generated by a** . A proper filter F is **maximal** if there is no filter H such that $F \subset H \subset \mathbb{P}$.

We say that a filter F is **maximal relative to c** if $c \notin F$ and $c \in H$ for every $H \supset F$. So, a filter F is maximal if it is maximal relative to 0. We say that F is **relatively maximal** when it is maximal relative to some c .

For $B \subseteq \mathbb{P}$ we define the **filter generated by B** to be the intersection of all the filters containing B . It is easy to verify that this is indeed a filter. When B is a finite set, the filter generated by B is the principal filter generated by $\wedge B$. In general we have the following.

4.1 Proposition For every $B \subseteq \mathbb{P}$, the filter generated by B is the set

$$\{a : \wedge C \leq a \text{ for some finite non-empty } C \subseteq B\}.$$

□

This has the following important consequence.

4.2 Proposition Let $B \subseteq \mathbb{P}$ and let $c \in \mathbb{P}$. If $\wedge C \not\leq c$ for every finite non-empty $C \subseteq B$, then B is contained in a maximal filter relative to c .

Proof Let \mathcal{F} be the set of filters F such that $B \subseteq F$ and $c \notin F$. By Proposition 4.1, \mathcal{F} is non-empty. It is immediate that \mathcal{F} is closed under unions of arbitrary chains. Then, by Zorn's lemma, \mathcal{F} has a maximal element. □

4.3 Exercise Prove that the following are equivalent

1. F is maximal relative to c ;
2. for every $a \notin F$ there is $d \in F$ such that $d \wedge a \leq c$.

□

4.4 Exercise (A generalization of the exercise above.) Let $B \subseteq \mathbb{P}$ and let $c \in \mathbb{P}$ be such that $\wedge C \not\leq c$ for every finite non-empty $C \subseteq B$. Prove that the following are equivalent

1. B is a maximal filter relative to c ;
2. $a \notin B \Rightarrow b \wedge a \leq c$ for some $b \in B$.

□

4.5 Exercise Let $F \subseteq \mathbb{P}$ be a non-principal filter. Is F always contained in a maximal non-principal filter? □

2 Distributive lattices and prime filters

Let \mathbb{P} be a lattice. We say that \mathbb{P} is **distributive** if for every $a, b, c \in \mathbb{P}$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

Throughout this section we assume that \mathbb{P} is a bounded distributive lattice.

A proper filter F is **prime** if for every $a, b \in \mathbb{P}$

$$a \vee b \in F \Rightarrow a \in F \text{ or } b \in F.$$

4.6 Proposition *Every relatively maximal filter of \mathbb{P} is prime.*

Proof Let F be maximal relative to c and assume that $a \notin F$ and $b \notin F$. Then, by Exercise 4.3, there are $d_1, d_2 \in F$ such that $d_1 \wedge a \leq c$ and $d_2 \wedge b \leq c$. Let $d = d_1 \wedge d_2$. Then $d \wedge a \leq c$ and $d \wedge b \leq c$ and therefore $(d \wedge a) \vee (d \wedge b) \leq c$. Hence, by distributivity, $d \wedge (a \vee b) \leq c$. Then $a \vee b \notin F$. \square

The **Stone space** of \mathbb{P} is a topological space that we denote by $S(\mathbb{P})$. The points of $S(\mathbb{P})$ are the prime filters of \mathbb{P} . The closed sets of the **Stone topology** are arbitrary intersections of sets of the form

$$[a]_{\mathbb{P}} = \{ F : \text{prime filter such that } a \in F \}.$$

for $a \in \mathbb{P}$. In other words, the sets above form a base of closed sets of the Stone topology. Using 1 and 3 in the following proposition the reader can easily check that this is indeed a base for a topology topology.

4.7 Proposition *For every $a, b \in \mathbb{P}$ we have*

1. $[0]_{\mathbb{P}} = \emptyset$;
2. $[1]_{\mathbb{P}} = S(\mathbb{P})$;
3. $[a]_{\mathbb{P}} \cup [b]_{\mathbb{P}} = [a \vee b]_{\mathbb{P}}$;
4. $[a]_{\mathbb{P}} \cap [b]_{\mathbb{P}} = [a \wedge b]_{\mathbb{P}}$.

Proof The verification is immediate. Only 3 requires that the filters in $S(\mathbb{P})$ are prime. \square

The closed subsets of $S(\mathbb{P})$ ordered by inclusion form a distributive lattice. The following is a representation theorem for distributive lattices.

4.8 Theorem *The map $a \mapsto [a]_{\mathbb{P}}$ is an embedding of \mathbb{P} in the lattice of the closed subsets of $S(\mathbb{P})$. In particular*

1. $0 \mapsto \emptyset$;
2. $1 \mapsto S(\mathbb{P})$;
3. $a \vee b \mapsto [a]_{\mathbb{P}} \cup [b]_{\mathbb{P}}$;
4. $a \wedge b \mapsto [a]_{\mathbb{P}} \cap [b]_{\mathbb{P}}$.

Proof It is immediate that the map above preserves the order and Proposition 4.7 shows that it preserves the lattice operations. We prove that the map is injective. Let $a \neq b$, say $a \not\leq b$. We claim that $[a]_{\mathbb{P}} \not\subseteq [b]_{\mathbb{P}}$. There is a filter F that contains a and is

maximal relative to b . By Proposition 4.6 such an F is prime. Then $F \in [a]_{\mathbb{P}} \setminus [b]_{\mathbb{P}}$. \square

4.9 Theorem *With the Stone topology, $S(\mathbb{P})$ is a compact space.*

Proof Let $\langle [a_i]_{\mathbb{P}} : i \in I \rangle$ be basic closed sets such that for every finite $J \subseteq I$

$$\bigcap_{i \in J} [a_i]_{\mathbb{P}} \neq \emptyset.$$

We claim that

$$\bigcap_{i \in I} [a_i]_{\mathbb{P}} \neq \emptyset.$$

By 4 of Proposition 4.7 and a we obtain that $\bigwedge C \not\leq 0$ for every finite $C \subseteq \{a_i : i \in I\}$. By Proposition 4.2, there is a maximal (relative to 0) filter containing $\{a_i : i \in I\}$. By Proposition 4.6, such filter is prime and it belongs to the intersection in b. \square

Let $a, b \in \mathbb{P}$. If $a \wedge b = 0$ and $a \vee b = 1$, we say that b is the **complement** of a (and vice versa). The complement of an element need not exist. If the complement exists it is unique, the complement of a is denoted by $\neg a$.

4.10 Lemma *Let $U \subseteq S(\mathbb{P})$ be a clopen set. Then $U = [a]_{\mathbb{P}}$ for some $a \in \mathbb{P}$ (and therefore a has a complement).*

Proof As both U and $S(\mathbb{P}) \setminus U$ are closed, for some sets $A, B \subseteq \mathbb{P}$

$$\begin{aligned} \bigcap_{x \in A} [x]_{\mathbb{P}} &= U \\ \bigcap_{y \in B} [y]_{\mathbb{P}} &= S(\mathbb{P}) \setminus U. \end{aligned}$$

By compactness, that is Theorem 4.9, there are some finite $A_0 \subseteq A$ and $B_0 \subseteq B$ such that

$$\bigcap_{x \in A_0} [x]_{\mathbb{P}} \cap \bigcap_{y \in B_0} [y]_{\mathbb{P}} = \emptyset$$

Let $a = \bigwedge A_0$ and $b = \bigwedge B_0$. From claim 4 of Proposition 4.7 we obtain $[a]_{\mathbb{P}} \cap [b]_{\mathbb{P}} = \emptyset$. Therefore $U \subseteq [a]_{\mathbb{P}} \subseteq S(\mathbb{P}) \setminus [b]_{\mathbb{P}} \subseteq U$. Hence $U = [a]_{\mathbb{P}}$ and $b = \neg a$. \square

A **Boolean algebra** is a bounded distributive lattice where every element has a complement. In a Boolean algebra, the sets $[a]_{\mathbb{P}}$ are clopen and they form also a base of open set of the topology of $S(\mathbb{P})$. A topology that has a base of clopen sets is called **zero-dimensional**. By the following proposition the Stone topology is Hausdorff.

A proper filter of a Boolean algebra is an **ultrafilter** if either $a \in F$ or $\neg a \in F$ for every a .

4.11 Proposition *Let \mathbb{P} be a Boolean algebra. Then the following are equivalent*

1. F is maximal;
2. F is prime;
3. F is an ultrafilter.

Proof Implication $2 \Rightarrow 3$ is obtained observing that $a \vee \neg a \in F$. The rest is immediate. \square

4.12 Exercise Suppose we defined $S(\mathbb{P})$ as the set of relatively maximal filters. Which essential property would not hold? \square

3 Types as filters

This section we work with a fixed set of formulas Δ all with free variables among those of some fixed tuple \mathbf{x} . In this section we do not display \mathbf{x} in the notation. Subsets of Δ are called **Δ -types**.

We associate to Δ a bounded lattice $\mathbb{P}(\Delta)$. This is the closure under conjunction and disjunction of the formulas in $\Delta \cup \{\perp, \top\}$. The (pre)order relation in $\mathbb{P}(\Delta)$ is given by

$$\psi \leq \varphi \Leftrightarrow T \cup \{\psi\} \vdash \varphi,$$

for some fixed theory T . In this section, to lighten notation, we absorb T in the symbol \vdash . If p is a Δ -type we denote by $\langle p \rangle_{\mathbb{P}}$ the filter in $\mathbb{P}(\Delta)$ generated by p .

4.13 Lemma For every Δ -type p

$$\langle p \rangle_{\mathbb{P}} = \left\{ \varphi \in \mathbb{P}(\Delta) : p \vdash \varphi \right\}.$$

In particular p is consistent if and only if $\langle p \rangle_{\mathbb{P}}$ is a proper filter.

Proof Inclusion \subseteq is clear. Inclusion \supseteq is a consequence of the Compactness Theorem. In fact $p \vdash \varphi$ implies that $\psi \vdash \varphi$ for some formula ψ that is conjunction of formulas in p . Then $\varphi \in \langle p \rangle_{\mathbb{P}}$ follows from $\psi \in \langle p \rangle_{\mathbb{P}}$ and $\psi \leq \varphi$. \square

We say that $p \subseteq \Delta$ is a **principal Δ -type** if $\langle p \rangle_{\mathbb{P}}$ is a principal. The following lemma is an immediate consequence of the Compactness Theorem. Note that in 3 the formula φ is arbitrary, possibly not even in $\mathbb{P}(\Delta)$.

4.14 Lemma For every Δ -type p the following are equivalent

1. p is principal;
2. $\psi \vdash p \vdash \psi$ for some formula ψ (here ψ is any formula, it need not be in Δ);
3. $\varphi \vdash p$ where φ is conjunction of formulas in p .

\square

Proof Implications $1 \Rightarrow 2$ is immediate by Lemma 4.13. To prove $2 \Rightarrow 3$ suppose $\psi \vdash p \vdash \psi$. Apply compactness to obtain a formula φ , conjunction of formulas in p , such that $\varphi \vdash \psi$. Implications $3 \Rightarrow 1$ is trivial. \square

4.15 Definition We say that $p \subseteq \Delta$ is a **prime Δ -type** if $\langle p \rangle_{\mathbb{P}}$ is a prime filter. We say that p is a **complete Δ -type** if $\langle p \rangle_{\mathbb{P}}$ is a maximal filter. \square

Though in general neither Δ nor $\mathbb{P}(\Delta)$ are closed under negation, Lemma 4.13 has the following consequence.

4.16 Proposition For every consistent Δ -type p the following are equivalent

1. p is complete;
2. p is consistent and either $p \vdash \varphi$ or $p \vdash \neg\varphi$ for every formula $\varphi \in \Delta$.

Proof $2 \Rightarrow 1$. Assume 2. As p is consistent, $\langle p \rangle_{\mathbb{P}}$ is a proper filter. To prove that it is maximal suppose $\varphi \notin \langle p \rangle_{\mathbb{P}}$. Then $p \not\vdash \varphi$ and from 2 it follows that $p \cup \{\neg\varphi\}$ is inconsistent. Hence no proper filter contains $p \cup \{\varphi\}$.

$1 \Rightarrow 2$. Suppose that p is complete and $p \not\vdash \varphi$. Then $\langle p \rangle_{\mathbb{P}}$ is a maximal filter. Then $\varphi \notin \langle p \rangle_{\mathbb{P}}$ and, as $\langle p \rangle_{\mathbb{P}}$ is maximal, $p \cup \{\varphi\}$ generates the improper filter. Then $p \cup \{\varphi\}$ is inconsistent. Hence $p \vdash \neg\varphi$. \square

Given a model M and a tuple $c \in M^{|x|}$ the **Δ -type of c in M** is the sets

$$\Delta\text{-tp}_M(c) = \{ \varphi(x) \in \Delta : M \models \varphi(c) \}$$

When the model M is clear from the context we omit the subscript. When x and c are the empty tuple, we write **$\text{Th}_\Delta(M)$** for $\Delta\text{-tp}_M(c)$.

4.17 Lemma *For every Δ -type p the following are equivalent*

1. p is prime;
2. $p \cup \{ \neg\varphi : \varphi \in \Delta \text{ such that } p \not\vdash \varphi \}$ is consistent;
3. $\{ \varphi \in \Delta : p \vdash \varphi \} = \Delta\text{-tp}_M(c)$ for some model M and some tuple $c \in M^{|x|}$.

Proof Implications $2 \Rightarrow 3 \Rightarrow 1$ are clear, we prove $1 \Rightarrow 2$. By compactness if the type in 2 is inconsistent then there are finitely many formulas $\varphi_1, \dots, \varphi_n \in \Delta$ such that $p \not\vdash \varphi_i$ and

$$p \vdash \bigvee_{i=1}^n \varphi_i.$$

hence p is not prime. \square

The following corollary is immediate. When it comes to verifying that a given Δ -type is prime, it simplifies the proof.

4.18 Corollary *For every Δ -type p the following are equivalent*

1. p is prime;
2. $p \vdash \bigvee_{i=1}^n \varphi_i \Rightarrow p \vdash \varphi_i$ for some $i \leq n$, for every n and every $\varphi_1, \dots, \varphi_n \in \Delta$. \square

The set Δ in this section contains only formulas with variables among those of the tuple x . In the following it is convenient to consider sets Δ that are closed under substitution of variables with any other variable. Then we may write **$\Delta_{|x}$** for the set of formulas in Δ with variables among those of the tuple x . But when x is clear from the context Δ may be used for $\Delta_{|x}$. The set of prime Δ -types is denoted by **$S(\Delta)$** , and we write **$S_x(\Delta)$** for $S(\Delta_{|x})$.

The most common Δ used in the sequel is the set of all formulas in $L(A)$ and the underlying theory is $\text{Th}(M/A)$ for some given model M containing A . In this case we write **$\text{tp}_M(c/A)$** for $\Delta\text{-tp}_M(c)$ or, when A is empty, **$\text{tp}_M(c)$** . The set $S_x(\Delta)$ is denoted by $S_x(A)$. The topology on $S_x(A)$ is generated by the clopen

$$[\varphi(x)] = \{ p \in S_x(A) : \varphi(x) \in p \}.$$

Sometimes Δ is the set of all formulas of a given syntactic form. Then we use some

more suggestive notation that we summarize below.

4.19 Notation The following are some of the most common Δ -types and Δ -theories

1. $\mathbf{at}\text{-}\mathbf{tp}(c), \quad \mathbf{Th}_{\mathbf{at}}(M) \quad \text{when } \Delta = L_{\mathbf{at}}$
2. $\mathbf{at}^{\pm}\text{-}\mathbf{tp}(c), \quad \mathbf{Th}_{\mathbf{at}^{\pm}}(M) \quad \text{when } \Delta = L_{\mathbf{at}^{\pm}}$
3. $\mathbf{qf}\text{-}\mathbf{tp}(c), \quad \mathbf{Th}_{\mathbf{qf}}(M) \quad \text{when } \Delta = L_{\mathbf{qf}}.$

Clearly, the types/theories in 2 and 3 are equivalent. Most used is the theory $\mathbf{Th}_{\mathbf{at}^{\pm}}(M/M)$ which is called the **diagram of M** and has a dedicated symbol: **$\mathbf{Diag}(M)$** . \square

4.20 Remark Let $A \subseteq M \cap N$. The following are equivalent

1. $N \models \mathbf{Diag}\langle A \rangle_M$;
2. $\langle A \rangle_M$ is a substructure of N . \square

4 Morphisms

First we set the meaning that the word **map** has in these notes.

4.21 Definition A **map** consists a triple $f : M \rightarrow N$ where

1. M is a set (usually a structure) called the **domain of the map**;
2. N is a set (usually a structure) called the **codomain the map**;
3. f is a function with **domain of definition** $\text{dom } f \subseteq M$ and **image** $\text{img } f \subseteq N$.

By **cardinality of $f : M \rightarrow N$** we understand the cardinality of the function f . \square

If $\text{dom } f = M$ we say that the map is **total**; if $\text{img } f = N$ we say that it is **surjective**. The **composition** of two maps and the **inverse** of a map are defined in the obvious way.

4.22 Definition Let Δ be a set of formulas with free variables in the tuple $x = \langle x_i : i < \lambda \rangle$. The map $h : M \rightarrow N$ is a **Δ -morphism** if it **preserves the truth** of all formulas in Δ . By this we mean that

$$p. \quad M \models \varphi(a) \Rightarrow N \models \varphi(ha) \quad \text{for every } \varphi(x) \in \Delta \text{ and every } a \in (\text{dom } h)^{|x|}.$$

Notation: if a is the tuple $\langle a_i : i < \lambda \rangle$ then ha is the tuple $\langle ha_i : i < \lambda \rangle$. \square

When $\Delta = L$ we say **elementary map** for Δ -morphism. When $\Delta = L_{\mathbf{at}}$ we say **partial homomorphism** and when $\Delta = L_{\mathbf{at}^{\pm}}$ we say **partial embedding** or **partial isomorphism**. The reason for the latter name is explained in Remark 4.23.

It is immediate to verify that a partial embedding which is total is an embedding (so, there is not conflict with Definition 2.19 above). Similarly, a partial embedding which is a total and surjective is an isomorphism. The precise connection between partial and total homo/iso-morphisms is discussed in following remark.

4.23 Remark For every map $h : M \rightarrow N$ the following are equivalent

1. $h : M \rightarrow N$ is a partial isomorphism;

2. there is an isomorphism $k : \langle \text{dom } h \rangle_M \rightarrow \langle \text{img } h \rangle_N$ that extends h .

Moreover, k unique. The equivalence holds replacing isomorphism by homomorphism (in which case we in 2 we obtain an surjection). The extension k is obtained defining $k(t(a)) = t(ha)$.

A similar fact holds for partial homomorphisms with epimorphism (surjective homomorphism) for isomorphism.

Unfortunately, there is no hope to generalize this fact to Δ -morphisms as, in general, there is no notion of generated Δ -elementary substructure. \square

We use Δ -morphisms to compare, locally, two structures (or different loci of the same structure). There are different ways to do this, in the proposition below we list a few synonymous expressions. But first some more notation. When x is a fixed tuple of variables, $a \in M^{|x|}$ and $b \in N^{|x|}$ we write

$$\begin{aligned} M, a &\Rightarrow_{\Delta} N, b && \text{if } M \models \varphi(a) \Rightarrow N \models \varphi(b) \text{ for every } \varphi(x) \in \Delta. \\ M, a &\equiv_{\Delta} N, b && \text{if } M \models \varphi(a) \Leftrightarrow N \models \varphi(b) \text{ for every } \varphi(x) \in \Delta. \end{aligned}$$

The following equivalences are immediate and will be used without explicit reference.

4.24 Proposition *For every given set of formulas Δ and every map $h : M \rightarrow N$ the following are equivalent*

1. $h : M \rightarrow N$ is a Δ -morphism;
2. $M, a \Rightarrow_{\Delta} N, ha$ for every $a \in (\text{dom } h)^{|x|}$;
3. $\Delta\text{-tp}_M(a) \subseteq \Delta\text{-tp}_N(ha)$ for every $a \in (\text{dom } h)^{|x|}$;
4. $N, ha \models p(x)$ for every $a \in (\text{dom } h)^{|x|}$ and $p(x) = \Delta\text{-tp}_M(a)$. \square

4.25 Remark Condition p in Definition 4.22 apply to tuples x of any length, in particular to the empty tuple. In this case $\varphi(x)$ is a sentence, $a \in (\text{dom } h)^0 = \{\emptyset\}$ is the empty tuple, and p asserts that $\text{Th}_{\Delta}(M) \subseteq \text{Th}_{\Delta}(N)$. When $h = \emptyset$ this is actually all what p says. In fact $\emptyset^{|x|} = \emptyset$ unless $|x| = 0$. Still, $\text{Th}_{\Delta}(M) \subseteq \text{Th}_{\Delta}(N)$ may be a non trivial requirement. \square

4.26 Definition *We call $\text{Th}_{\Delta}(M)$ the **Δ -characteristic** of M . We say that T decides the Δ -characteristic if $\text{Th}_{\Delta}(M) = \text{Th}_{\Delta}(N)$ for all $M, N \models T$. In other words, for any pair of models of T , the empty map $\emptyset : M \rightarrow N$ is a Δ -morphism. When $\Delta = L_{\text{at}^{\pm}}$ we say simply **characteristic**, then T decides the characteristic $\langle \emptyset \rangle_M \simeq \langle \emptyset \rangle_N$ for all $M, N \models T$. \square*

We conclude this section with a couple of propositions that break Theorem 2.20 into parts. More interestingly, in Chapter 9 we shall prove a sort of converse of Propositions 4.28 and 4.29.

This is not immediately required in the following. There is a relation between some properties of Δ -morphisms closure of Δ logical connectives. When $C \subseteq \{\forall, \exists, \neg, \vee, \wedge\}$ is a set of connectives, we write **$C\Delta$** for the closure of Δ with respect to all connectives in C . We write $\neg\Delta$ for the set containing the negation of the formulas in Δ . Warning: do not confuse $\neg\Delta$ with $\{\neg\}\Delta$.

It is clear that Δ -morphisms are $\{\wedge, \vee\}$ - Δ -morphisms.

4.27 Proposition *For every given set of formulas Δ and every injective map $h : M \rightarrow N$ the following are equivalent*

- a. $h : M \rightarrow N$ is a $\neg\Delta$ -morphism;
- b. $h^{-1} : M \rightarrow N$ is a Δ -morphism. □

4.28 Proposition *For every set of formulas Δ , every Δ -embedding $h : M \rightarrow N$ is a $\{\exists\}$ - Δ -morphism.*

Proof Formulas in $\{\exists\}\Delta$ have the form $\exists y \varphi(x, y)$ where y is a finite tuples of variables and $\varphi(x, y) \in \Delta$. For every tuple $a \in (\text{dom } h)^{|x|}$ we have:

$$\begin{aligned}
 M \models \exists y \varphi(a, y) &\Rightarrow M \models \varphi(a, b) \text{ per una tupla } b \in M^{|y|} \\
 &\Rightarrow N \models \varphi(ha, hb) \\
 &\Rightarrow N \models \varphi(ha, c) \text{ per una tupla } c \in N^{|y|} \\
 &\Rightarrow N \models \exists y \varphi(ha, y).
 \end{aligned}$$

Note that the second implication requires the totality of $h : M \rightarrow N$ which guarantees that $\varphi(a, y)$ has a solution in $\text{dom } h$. □

When $h : M \rightarrow N$ is injective the following is a corollary of Propositions 4.27 and 4.28. The general proof is the ‘dual’ of that of Propositions 4.28.

4.29 Proposition *For every set of formulas Δ , every surjective Δ -morphism $h : M \rightarrow N$ is a $\{\forall\}$ - Δ -morphism. □*

Chapter 5

Some relational structures

In the first section we prove that theory of *dense linear orders without endpoints* is ω -categorical. That is, any two such countable orders are isomorphic. This is an easy classical result of Cantor. In this chapter we examine Cantor's construction (a so-called *back-and-forth* construction) in great detail. In the second section we apply the same technique to prove that the theory of the *random graph* is ω -categorical.

1 Dense linear orders

The **language of strict orders**, which in this section we denote by L , contains only a binary relation symbol $<$. A structure M of signature L is a **strict order** if it models (the universal closure of) the following formulas

1. $x \not< x$ irreflexive;
2. $x < z < y \rightarrow x < y$ transitive.

Note that the following is an immediate consequence of 1 and 2.

$$x < y \rightarrow y \not< x \quad \text{antisymmetric.}$$

We say that the order is **total** or **linear** if

- li. $x < y \vee y < x \vee x = y$ linear or total.

An order is **dense** if

- nt. $\exists x, y (x < y)$ non trivial;
- d. $x < y \rightarrow \exists z (x < z < y)$ dense.

We need to require the existence of two comparable elements, then d implies that dense orders are in fact infinite. We say that the ordering has no **endpoints** if

- e. $\exists y (x < y) \wedge \exists y (y < x)$ without endpoints.

We denote by T_{lo} the theory strict linear orders and by T_{dlo} the theory of dense linear orders without endpoints. Clearly, these are consistent theories: \mathbb{Q} with the usual ordering is a model of T_{dlo} .

We introduce some notation to improve readability of the proof of the following theorem. Let A and B be subsets of an ordered set. We write $A < B$ if $a < b$ for every $a \in A$ and $b \in B$. We write $a < B$ e $A < b$ for $\{a\} < B$, respectively $A < \{b\}$. Let $M \models T_{lo}$. Then $M \models T_{dlo}$ if and only if for every finite $A, B \subseteq M$ such that $A < B$ there is a c such that $A < c < B$. In fact axiom d is evident and axioms nt and e are obtained taking replacing A and/or B by the empty set.

Now we prove the first of a series of lemmas that we call **extension lemmas**. Recall that in the language of strict orders a map $k : M \rightarrow N$ is a partial isomorphism if

$$M \models a < b \Leftrightarrow N \models ka < kb \quad \text{for every } a, b \in \text{dom } k.$$

(When $M, N \models T_{\text{lo}}$ direction \Rightarrow suffices.)

5.1 Lemma Fix $M \models T_{\text{lo}}$ and $N \models T_{\text{dlo}}$. Let $k : M \rightarrow N$ be a finite partial isomorphism and let $b \in M$. Then there is a partial isomorphism $h : M \rightarrow N$ that extends k and is defined in b .

Proof Given a finite partial isomorphism $k : M \rightarrow N$ define

$$\begin{aligned} X &= \{x \in \text{dom}(k) : x < b\}; \\ Y &= \{y \in \text{dom}(k) : b < y\}. \end{aligned}$$

The sets X and Y are finite and partition $\text{dom } k$, and $X < Y$. As $k : M \rightarrow N$ is a partial isomorphism, $k[X] < k[Y]$. Then in N there is an element c such that $k[X] < c < k[Y]$. It is easy to check that setting $h = k \cup \{b, c\}$ we obtain the required extension. \square

The following is an equivalent version of Lemma 5.1.

5.2 Corollary Let $M \models T_{\text{lo}}$ be countable and let $N \models T_{\text{dlo}}$. Let $k : M \rightarrow N$ be a finite partial isomorphism. Then there is a (total) embedding $h : M \rightarrow N$ that extends k .

Proof Let $\langle a_i : i < \omega \rangle$ be an enumeration of M . Define by induction a chain of finite partial isomorphisms $h_i : M \rightarrow N$ such that $a_i \in \text{dom } h_{i+1}$. The construction starts with $h_0 = k$. At stage $i + 1$ we chose any finite partial isomorphism $h_{i+1} : M \rightarrow N$ that extends h_i and is defined in a_i . This is possible by Lemma 5.1. In the end we set

$$h = \bigcup_{i \in \omega} h_i.$$

It is immediate to verify that $h : M \rightarrow N$ is the required embedding. \square

5.3 Exercise Prove that the extension Lemma 5.1 characterizes models of T_{dlo} among models of T_{lo} . That is, if N is a model of T_{lo} such that the conclusion of Lemma 5.1 holds, then $M \models T_{\text{dlo}}$. \square

We are now ready to prove that any two countable models of T_{dlo} are isomorphic which is a classical result of Cantor's. Actually what we prove is slightly more general than that. In fact Cantor's theorem is obtained from the theorem below by setting $k = \emptyset$, which we are allowed to, because all models have the same empty characteristic (cfr. Remark 4.25).

5.4 Theorem Every finite partial isomorphism $k : M \rightarrow N$ between models of T_{dlo} extends to an isomorphism $g : M \rightarrow N$.

The following is the archetypal **back-and-forth** construction. It is important to note that it does not mention linear orders at all. It only uses the extension Lemma 5.1. In many other contexts, as soon as we have an extension lemma, we can apply verbatim the same construction (cfr. Theorem 6.6).

Proof Fix some enumerations of M and N , say $\langle a_i : i < \omega \rangle$, respectively $\langle b_i : i < \omega \rangle$. We define by induction a chain of finite partial isomorphisms $g_i : M \rightarrow N$ such that $a_i \in \text{dom } g_{i+1}$ and $b_i \in \text{img } g_{i+1}$. In the end we set

$$g = \bigcup_{i \in \omega} g_i$$

We begin by setting $g_0 = k$. The inductive step consists of two half-steps that we may call the *forth-step* and *back-step*. In the forth-step we define $g_{i+1/2}$ such that $a_i \in \text{dom } g_{i+1/2}$. In the back-step to define g_{i+1} such that $b_i \in \text{img } g_{i+1}$.

By the extension lemma 5.1 there is a finite partial isomorphism $g_{i+1/2} : M \rightarrow N$ that extends g_i and is defined in a_i . Now apply the same lemma to extend $(g_{i+1/2})^{-1} : N \rightarrow M$ to a finite partial isomorphism $(g_{i+1})^{-1} : N \rightarrow M$ defined in b_i . \square

Let λ be an infinite cardinal. We say that a theory is **λ -categorical** if any two models of T of cardinality λ are isomorphic. From Theorem 5.4, taking $k = \emptyset$, we obtain the following.

5.5 Corollary *The theory T_{dlo} is ω -categorical.* \square

We also obtain that T_{dlo} is a complete theory. This is consequence of the following general fact.

5.6 Proposition *If T is λ -categorical, for some $\lambda \geq |L|$, then T is complete.*

Proof Let M and N be any two models of T_{dlo} . Applying the upward and/or downward Löwenheim-Skolem theorem, we may assume they both have cardinality λ . (Here we use that $\lambda \geq |L|$.) Hence $M \simeq N$ and in particular $M \equiv N$. \square

5.7 Exercise Prove that T_{dlo} is not λ -categorical for any uncountable λ . \square

5.8 Exercise Prove that \mathbb{R} , in the language of strict orders, is not isomorphic to $\mathbb{R} \setminus \{0\}$. \square

5.9 Exercise Let L be the language of strict orders augmented with countably many constants $\{c_i : i \in \omega\}$. Let T be the theory that extends T_{dlo} with the axioms $c_i < c_{i+1}$ for all i . Prove that T is complete. Find three non isomorphic models of this theory. For N a rightly chosen model of T , prove the statement in Lemma 5.1, with M any model of T . \square

5.10 Exercise Show that in Theorem 5.6 the assumption $\lambda \geq |L|$ is necessary. (Hint. Let ν be an uncountable cardinal. The language contains only the ordinals $i < \nu$ as constants. The theory T says that there are infinitely many elements and either $i = 0$ for every $i < \nu$, or $i \neq j$ for every $i < j < \nu$. Prove that T is ω -categorical but incomplete.) \square

2 Random graphs

Recall that the **language of graphs**, which in this section we denote by L , contains only a binary relation r . A **graph** structure of signature L such that

1. $\neg r(x, x)$ irreflexive;
2. $r(x, y) \rightarrow r(y, x)$ symmetric.

An element of a graph M is called a **vertex** or a **node**. An **arc** is an unordered pair of vertices $\{a, b\} \subseteq M$ such that $M \models r(a, b)$. In words we may say that a is adjacent to b .

A **random graph** is a graph that also satisfies the following axioms for every n

nt. $\exists x, y, z \ (x \neq y \neq z \neq x)$ non trivial;

$r_n. \bigwedge_{i,j=1}^n x_i \neq y_j \rightarrow \exists z \bigwedge_{i=1}^n [r(x_i, z) \wedge \neg r(z, y_i)]$ for every $n \in \mathbb{Z}^+$.

The theory of graphs is denoted by T_{gph} and the theory of random graphs is denoted by T_{rg} . The scheme of axioms r_n plays the same role of density in the previous section. It says that given two non empty disjoint sets X and Y of cardinality $\leq n$ there is a vertex z that is adjacent to all vertices in X and to no vertex in Y . Note that it might well be that $z \in Y$, however $z \in X$ would contradict 1.

Some readers may doubt if T_{rg} is consistent.

5.11 Proposition *There exists a random graph.*

Proof The domain of N is the set of natural numbers. Let $N \models r(n, m)$ hold if the n -th prime number divides m or vice versa m -th prime number divides n . \square

5.12 Proposition *Random graphs are infinite.*

Proof Suppose for a contradiction that M is a finite random graph. As M contains at least 3 elements, we can fix two distinct elements $a, b \in M$ and require a z adjacent to all elements of $M \setminus \{a, b\}$ and no element of $\{a, b\}$. As remarked above, $z \in \{a, b\}$, then a and b are not adjacent. Now let w be adjacent to all vertices in $M \setminus \{a\}$ but not to a . The only option is $w = b$, but this implies that a is adjacent to b . Contradiction. \square

The following is the analogous of Lemma 5.1 for random graphs. Recall that in the language of graphs a map $k : M \rightarrow N$ is a partial isomorphism if

$$M \models r(a, b) \Leftrightarrow N \models r(ka, kb) \quad \text{for every } a, b \in \text{dom } k.$$

5.13 Lemma *Fix $M \models T_{\text{gph}}$ and $N \models T_{\text{rg}}$. Let $k : M \rightarrow N$ be a finite partial isomorphism and let $b \in M$. Then there is a partial isomorphism $h : M \rightarrow N$ that extends k and is defined in b .*

Proof The schema of the proof is the same as in Lemma 5.1, so we use the same notation. Assume $b \notin \text{dom } k$ and define

$$X = \{x \in \text{dom } k : M \models r(x, b)\} \quad \text{e} \quad Y = \{y \in \text{dom } k : M \models \neg r(y, b)\}.$$

These two sets are finite and disjoint, then so are $k[A]$ and $k[B]$. Then there is a c such that

$$\bigwedge_{a \in X} r(ka, c) \quad \wedge \quad \bigwedge_{a \in Y} \neg r(ka, c).$$

As $k[A] \cup k[B] = \text{img } k$, it is immediate to verify that $h = k \cup \{(b, c)\}$ is the require extension. \square

With the same proof of Corollary 5.2 obtain the following.

5.14 Corollary *Let $M \models T_{\text{gph}}$ be countable and let $N \models T_{\text{rg}}$. Let $k : M \rightarrow N$ be a finite partial isomorphism. Then there is a (total) embedding $h : M \rightarrow N$ that extends k .* \square

The proof of Theorem 5.4 gives the following theorem and its corollary.

- 5.15 Theorem** Every finite partial isomorphism $k : M \rightarrow N$ between models of T_{rg} extends to an isomorphism $g : M \rightarrow N$. □
- 5.16 Corollary** The theory T_{rg} is ω -categorical (and therefore complete). □
- 5.17 Exercise** Let $a, b, c \in N \models T_{\text{rg}}$. Prove that $r(a, N) = r(b, N) \cap r(c, N)$ occurs only in the trivial case $a = b = c$. □
- 5.18 Exercise** Prove that for every $b \in N \models T_{\text{rg}}$ the set $r(b, N)$ is a random graph. Assume N is countable. Is every random graph $M \subseteq N$ of the form $r(b, N)$ for some $b \in N$? □
- 5.19 Exercise** Let N be free union of two random graphs. That is, $N = N_1 \sqcup N_2$ and $r^N = r^{N_1} \sqcup r^{N_2}$, with N_1 and N_2 some countable random graphs. By \sqcup we denote the disjoint union. Prove that N is not a random graph. Show that N_1 is not a \emptyset -definable subset of N . Write a first order sentence $\psi(x, y)$ true if x and y belong to the same connected component of N . Axiomatize the class \mathcal{K} of graphs that are free union of two random graphs. □
- 5.20 Exercise** Prove that T_{rg} is not λ -categorical for any uncountable λ . Hint. Prove that there is a random graph N of cardinality λ where every vertex is adjacent to $< \lambda$ vertices. Compare it with its complement graph (the graph that has arcs between pairs that are non adjacent in N). □
- 5.21 Exercise** Prove that T_{rg} is not finitely axiomatizable. Hint. Given a random graph N and a set P of cardinality $n + 1$ show that you can add arcs to $M \sqcup P$ and make it satisfy axiom r_n but not r_{n+1} . □
- 5.22 Exercise** Let $A \subseteq N \models T_{\text{rg}}$ and let $\varphi(x) \in L(A)$, where $|x| = 1$. Prove that if $\varphi(N)$ is finite then $\varphi(N) \subseteq A$. □
- 5.23 Exercise (Peter J. Cameron)** Prove that for every countable graph M the following are equivalent

1. M is either the infinite complete graph, the infinite empty graph, or the random graph;
2. if $M_1, M_2 \subseteq M$ are such that $M_1 \sqcup M_2 = M$, then $M_1 \simeq M$ or $M_2 \simeq M$.

Hint. For $2 \Rightarrow 1$ first show that if $r(a, M) = \emptyset$ for some $a \in M$ then the graph is null and if $\{b\} \cup r(b, M) = M$ for some b then the graph is complete. Also note that 2 implies that a similar property holds for any finite partition of M , not just bi-partitions. It follows that if 2 and there is a finite A such that $\bigwedge_{a \in A} r(a, M) = \emptyset$ then the graph is null and if there is a finite B such that $B \cup \bigvee_{b \in B} r(b, M) = M$ then the graph is complete.

Suppose M is not a random graph. Fix some finite, disjoint A and B such that no c satisfies both $r(A, c) = A$ and $r(B, c) = \emptyset$. Let $M_1 = \{c : r(A, c) \neq A\} \cup B$

and $M_2 = B \cup \{c : r(B, c) \neq \emptyset\} \setminus M_1$. This is a partition of M . Now note that $\bigwedge_{a \in A} r(a, M_1) = \emptyset$ and $\bigvee_{b \in B} r(b, M_2) = M_2$. \square

3 Notes

The following is a well-written accessible survey on the amazing model theoretic properties of the random graph.

[1] Peter J. Cameron, *The random graph* (2013), available at <https://arxiv.org/abs/1301.7544>.

Chapter 6

Fraïssé limits

We introduce *Fraïssé limits*, also homogeneous-universal or generic structures, which here we call *rich models*, after Poizat. Rich models generalize the examples in Chapters 5 and the many more to come. We also prove *elimination of quantifiers* for the theories introduced in Chapters 5.

1 Rich models.

We now define *categories of models and partial morphisms*. These are examples of concrete categories as intended in category theory. However, apart from the name, in what follows we dispense with all notions of category theory as they would make the exposition less basic than intended (and not provide much more technical instruments). We refer the interested reader to the references at the end of the chapter.



Warning: the terminology introduced in this section is not standard. In the literature many details are left implicit. This is generally safe when the category used is fixed. Here we prefer to be more explicit because, when discussing saturation and quantifier elimination, it helps to compare different categories.

A **category (of models and partial morphisms)** is a class \mathcal{M} which is disjoint union of two classes: \mathcal{M}_{ob} and \mathcal{M}_{ar} . The first is the class of **objects** and contains structures with a common signature L which we call **models**. The second is the class of **arrows** and contains (partial) maps between models which we call **morphisms**. We require that the identity maps are morphisms and that composition of two morphisms is again a morphism. This makes \mathcal{M} a well-defined category.

For example, \mathcal{M}_{ob} could consist of all models of some theory T_0 and \mathcal{M}_{ar} of all partial isomorphisms between these. Alternatively, as morphisms we could take elementary maps between models. At a first reading the reader may assume \mathcal{M} is as in one of these two examples. In the general case we need to make some assumptions on \mathcal{M} .

6.1 Definition For ease of reference we list together all properties required below

- c0. between any two models M and N there is some morphism;
- c1. the identity map $\text{id}_A : M \rightarrow M$ is a morphism, for any $A \subseteq M$;
- c2. morphisms are invertible maps and the inverse of a morphism is a morphism;
- c3. morphisms preserve the truth of L_{at} -formulas;
- c4. if M is a model, every elementary map $k : M \rightarrow N$ is a morphism (and, N is a model);
- c5. if $k_i : M \rightarrow N$ is a chain of morphisms, then $\bigcup_{i < \lambda} k_i : M \rightarrow N$ is a morphism.
- c6. if $k' : M \rightarrow N$ is a morphism for every finite $k' \subseteq k$, then $k : M \rightarrow N$ is a morphism. \square

Note that c0 says that the category is **connected** and that c1 implies that any restriction of a morphism is a morphism. Given c1, we can rephrase c0 by saying that $\varnothing : M \rightarrow N$ is a morphism for any pair of models M and N .

A consequence of c4 is that the class of models is closed under elementary equivalence. In fact, if $N \equiv M$ then $\varnothing : M \rightarrow N$ is elementary, hence it is a morphism and, in particular, and N is a model.

We call c6 the **finite character of morphisms**. Note that it implies 5. We shall never use 6 in these notes, but most natural categories of models and partial morphisms have this property.

The following two definitions assume c2. The generalization to non injective morphisms is not straightforward (in fact, there are two generalizations: *projective* and *inductive*). These generalizations are not very common and will not be considered here.

6.2 Definition Assume that \mathcal{M} satisfies c0-c2 of Definition 6.1. We say that a model N is **λ -rich** if for every model M of cardinality $\leq \lambda$ and every morphism $k : M \rightarrow N$ of cardinality $< \lambda$ there is a total morphism $h : M \rightarrow N$ that extends k . We say that N is **rich** tout court if it is λ -rich for $\lambda = |N|$. When $\mathcal{M}_{\text{ob}} = \text{Mod}(T_0)$ for some theory T_0 and \mathcal{M}_{ar} is clear from the context, we say **rich model of T_0** . □

Rich models are also called *Fraïssé limits* or *homogeneous-universal* for a reason that will soon be clear; they are also called *generic*. Unfortunately these names are either too long or too generic, so we opt for the less common term *rich* proposed by Poizat.

The following two notions are deeply connected with richness.

6.3 Definition Assume that \mathcal{M} satisfies c0-c2 of Definition 6.1. We say that a model N is **λ -universal** if for every model M of cardinality $\leq \lambda$ there is a total morphism $k : M \rightarrow N$. We say that a model N is **λ -homogeneous** if every $k : N \rightarrow N$ of cardinality $< \lambda$ extends to a bijective morphism $h : N \rightarrow N$ (an automorphism when c3 below holds).

Note that the larger \mathcal{M}_{ar} , the stronger notion of homogeneity. When \mathcal{M}_{ar} contains all partial isomorphisms between models (the largest class of morphisms considered here), it is common to say **λ -ultrahomogeneous** for λ -homogeneous.

As above, when $\lambda = |N|$ we say **universal**, respectively **homogeneous** and **ultrahomogeneous**. □

In Section 5.1 we implicitly used $\mathcal{M}_{\text{ob}} = \text{Mod}(T_{\text{lo}})$ and partial isomorphisms as \mathcal{M}_{ar} . In Section 5.2 we used $\mathcal{M}_{\text{ob}} = \text{Mod}(T_{\text{gph}})$ and again partial isomorphisms as \mathcal{M}_{ar} . Corollary 5.2 proves that every model of T_{dlo} is ω -rich. Corollary 5.14 claims the analogous fact for T_{rg} .

In the following we fequently work under the following assumption (even when not all properties are strictly necessary).

6.4 Assumption Assume $|L| \leq \lambda$ and suppose that \mathcal{M} satisfies c0-c5 of Definition 6.1

The assumption on the cardinality L is only necessary to apply the downward Löwenheim-Skolem Theorem when required.

6.5 Proposition (Assume 6.4) *The following are equivalent*

1. N is a λ -rich model;
2. for every morphism $k : M \rightarrow N$ of cardinality $< \lambda$ and every $b \in M$ there is a morphism $h : M \rightarrow N$ that extends k and is defined in b .

Proof Closure under union of chains of morphisms, which is ensured by c5, immediately yields $2 \Rightarrow 1$. As for implication $1 \Rightarrow 2$ we only need to consider the case $\lambda < |M|$. By the downward Löwenheim-Skolem theorem there is an $M' \preceq M$ of cardinality λ containing b . Let $h : M' \rightarrow N$ be the total morphism obtained from 1. By c4, the map $h : M \rightarrow N$ is a composition of morphisms, hence a morphism. \square

The following theorem subsumes both Theorem 5.4 and Theorem 5.15.

6.6 Theorem (Assume 6.4) *Let M and N be two rich models of the same cardinality λ . Then every morphism $k : M \rightarrow N$ of cardinality $< \lambda$ extends to an isomorphism.*

Proof When $\lambda = \omega$, we can take the proof of Theorem 5.4 and replace *partial isomorphism* by *morphism* and the references to Lemma 5.1 by references to Proposition 6.5. As for uncountable λ , we only need to extend the construction through limit stages. By c5 we can simply take the union. \square

6.7 Corollary (Assume 6.4) *Then all rich models of cardinality λ are isomorphic.* \square

It is obvious that λ -rich models are λ -universals and, by Theorem 6.6, they are λ -homogeneous. These two notions are weaker than richness. For instance, when M is as in Section 5.2, the countable graph that has no arc is trivially ultrahomogeneous but it is not universal and a fortiori not rich. On the other hand if we add to a countable random graph an isolated point we obtain an universal graph which is not ultrahomogeneous. However, when taken together, these two properties are equivalent to richness.

6.8 Theorem (Assume 6.4) *The following are equivalent:*

1. N is rich;
2. N is homogeneous and universal.

Proof Implication $1 \Rightarrow 2$ is clear as noted above, so we prove $2 \Rightarrow 1$. Let $k : M \rightarrow N$ be a morphism of cardinality $< |N|$ and let $b \in M$. As N is universal, there is a morphism $f : M \rightarrow N$ defined on $\text{dom } k \cup \{b\}$. By c2 the map $k \circ f^{-1} : N \rightarrow N$ is a morphism of cardinality $< |N|$. By homogeneity it has an extension to a bijective morphism $h : N \rightarrow N$. It is immediate that $h \circ f : M \rightarrow N$ is the required extension of k . \square

6.9 Exercise Let N be the structure obtained by adding to a countable random graph an isolated point. Show that this graph is homogeneous if \mathcal{M}_{ar} is the class of elementary maps but not if \mathcal{M}_{ar} is the class of partial isomorphisms. \square

Next theorem is one of the most important facts about λ -rich models

6.10 Theorem (Assume 6.4) *Then every morphism between λ -rich models is elementary. In particular, λ -rich models are elementary equivalent.*

Proof Let $k : M \rightarrow N$ be a morphism between rich models. It suffices to prove that every finite restriction of k is elementary. By c1, we may as well assume that k itself is finite. It suffices to construct $M' \preceq M$ and $N' \preceq N$ together with a morphism $h : M \rightarrow N$ that extends k and maps M' bijectively to N' . Then by c4 the map $h : M' \rightarrow N'$ is also a morphism because it is the composition of morphisms. Finally, by c3, it is an isomorphism, in particular an elementary map.

We construct simultaneously a chain $\langle A_i : i < \lambda \rangle$ of subsets of M , a chain $\langle B_i : i < \lambda \rangle$ of subsets of N , and a chain functions $\langle h_i : i < \lambda \rangle$ such that $h_i : M \rightarrow N$ are morphisms, $\text{dom } h_i \subseteq A_i$ and $\text{img } h_i \subseteq B_i$. In the end we set

$$M' = \bigcup_{i < \lambda} A_i, \quad N' = \bigcup_{i < \lambda} B_i, \quad h = \bigcup_{i < \lambda} h_i.$$

The chains start with $A_0 = \text{dom } k$, $B_0 = \text{img } k$ and $h_0 = k$. As usual at limit stages we take the union. Now we consider successor stages. Fix some enumerations of M and N . At stage $i + 1$ we also fix some enumerations of A_i , B_i , $L_x(A_i)$, and $L_x(B_i)$, where $|x| = 1$. Let $\langle i_1, i_2 \rangle$ be the i -th pair of ordinals $< \lambda$. If the i_2 -th formula in $L_x(A_{i_1})$ is consistent in M , pick any solution a . Also, if the i_2 -th formula in $L_x(B_{i_1})$ is consistent in N , pick any solution b . Let a' be the i_2 -th element of A_{i_1} and let b' be the i_2 -th element of B_{i_1} .

Let $h_{i+1} : M \rightarrow N$ be an extension of h_i such that $\text{dom } h_{i+1} = \text{dom } h_i \cup \{a'\}$ and $\text{img } h_{i+1} = \text{img } h_i \cup \{b'\}$. Note that the latter requirement is met applying Proposition 6.5 to $h_i^{-1} : N \rightarrow M$. Define $A_{i+1} = A_i \cup \{a, h_{i+1}^{-1}b'\}$ and $B_{i+1} = B_i \cup \{b, h_{i+1}a'\}$.

The elements a and b , added to M' and N' , ensure that $M' \preceq M$ and $N' \preceq N$, by the Tarski-Vaught test. In fact every formula in $L(M')$ belongs to $L(A_{i_1})$ for some i_1 and it occurs as i_2 -formula in the enumeration that has been fixed along with the construction. Therefore at stage $i = \langle i_1, i_2 \rangle$ a witness is added to M' . The same happens for formulas in $L(N')$.

A similar argument shows that the elements a' and b' , added to the domain and range of h , ensure that the map $h : M' \rightarrow N'$ is both total and surjective. \square

By the theorem below we can unambiguously speak of **the theory of rich models**. In fact it immediately follows that if M is μ -rich and N is λ -rich then $M \equiv N$.

6.11 Exercise Let L be the language of strict orders. Prove that $\mathbb{Q} \preceq \mathbb{R}$. \square

Assume for simplicity that L is countable. Then all ω -rich models belong to $\text{Mod}(T)$ for some complete theory T . It is interesting to ask if the converse is true: if T is the theory of the ω -rich models of \mathcal{M} , do all models in $\mathcal{M} \cap \text{Mod}(T)$ are rich? The answer is affirmative when T is either T_{dlo} or T_{rg} , but this is not always the case, see Example 6.14 for an easy counterexample. An affirmative answer implies that the theory of rich models is ω -categorical. An interesting variant of this question is considered in Exercise 8.8.

For convenience of the reader, we instantiate Theorem 6.10 with the two categories used in Chapter 5.

6.12 Corollary *Partial isomorphisms between models of T_{dlo} and between models of T_{rg} are elementary maps.* \square

When partial isomorphism between models of a given theory T coincide with el-

elementary maps, it is always by a fundamental reason. Let us introduce some terminology. Let T be a consistent theory. We say that **T has (or admits) elimination of quantifiers** if for every formula $\varphi(x) \in L$ there is a quantifier-free formula $\psi(x) \in L_{\text{qf}}$ such that

$$T \vdash \psi(x) \leftrightarrow \varphi(x).$$

We will discuss general criteria for elimination of quantifiers in Chapter 9, here we report without proof the following theorem.

6.13 Theorem *Let T be a theory that decides the characteristic of its models. The following are equivalent*

1. T has elimination of quantifiers;
2. every partial isomorphism between models T is an elementary map. □

For the time being when saying that T has *quantifier elimination* we intend 2 of the theorem above.

In the next chapter we introduce important examples of ω -rich models that do not have an ω -categorical theory. These are algebraic structures (groups, fields etc.) hence more complex than pure relational structures. So, we conclude this section with an example of this phenomenon in almost trivial context.

6.14 Example Let L contain a unary predicate r_n for every positive integer n . The theory T_0 contains the axioms $\neg \exists x [r_n(x) \wedge r_m(x)]$ for $n \neq m$ and $\exists^{\leq n} x r_n(x)$ for every n . Work in the category of models of T_0 and partial isomorphisms. Let T be the theory that extends T_0 with the axioms $\exists^=n x r_n(x)$ for every n . Let $q(x)$ be the type $\{\neg r_n(x) : n \in \omega\}$. There are models of T that do not realize $q(x)$, hence T is not ω -categorical. It is easy to verify that the following are equivalent.

1. N is ω -rich model;
2. $N \models T$ and $q(N)$ is infinite.

The reader may use Theorem 6.10 to prove that T is complete and has elimination of quantifiers. It is also easy to verify that every uncountable model of T is rich and consequently that T is uncountably categorical. □

6.15 Exercise Let T_0 and \mathcal{M} be as in Example 6.14 except that we restrict the language to the relations r_0, \dots, r_n for a fixed n . Do ω -rich models of T_0 exist? Is their theory ω -categorical? What if we add to T_0 the axiom $r_0(x) \vee \dots \vee r_n(x)$? □

6.16 Exercise The language contains only the binary relations $<$ and e . The theory T_0 says that $<$ is a strict linear order and that e is an equivalence relation. Let $\mathcal{M}_{\text{ob}} = \text{Mod}(T_0)$ and let \mathcal{M}_{ar} be the class of partial isomorphisms between models. Do rich models exist? Can we axiomatize their theory? Is it ω -categorical? □

6.17 Exercise The language contains only two binary relations. The theory T_0 says that they are equivalence relations. Let $\mathcal{M}_{\text{ob}} = \text{Mod}(T_0)$ and let \mathcal{M}_{ar} be the class of partial isomorphisms between models. Do rich models exist? Can we axiomatize their theory? Is it ω -categorical? □

- 6.18 Exercise** In the language of graphs let T_0 say that there are no cycles (equivalently, there is at most one path between any two nodes). In combinatorics these graphs are called *forests*, their connected components, *trees*. Let $\mathcal{M}_{\text{ob}} = \text{Mod}(T_0)$ and let \mathcal{M}_{ar} be the class of partial isomorphisms between models. Do rich models exist? Can we axiomatize their theory? Is it ω -categorical? \square

2 Weaker notions of universality and homogeneity

We want to extend the equivalence in Theorem 6.8 to λ -rich models. For that we need to weaken the notions of λ -universality and λ -homogeneity.

- 6.19 Definition** We say that a structure N is *weakly λ -homogeneous* if for every $b \in N$ every morphism $k : N \rightarrow N$ of cardinality $< \lambda$ extends to one defined in b . The term *back-and-forth λ -homogeneous* is also used. \square

The following easy exercise on back-and-forth is required in the sequel.

- 6.20 Exercise** Assume that \mathcal{M} satisfies c0-c5 of Definition 6.1. Prove that any weakly λ -homogeneous structure of cardinality λ is homogeneous. \square
- 6.21 Definition** We say that a structure N is *weakly λ -universal* if for every model M and every $A \subseteq M$ of cardinality $< \lambda$ there is a morphism $k : M \rightarrow N$ such that $A \subseteq \text{dom } k$. \square

- 6.22 Lemma** (Assume 6.4) Let N be a weakly λ -homogeneous model. Let $A \subseteq N$ have cardinality $\leq \lambda$ and let $k : N \rightarrow N$ be a morphism of cardinality $< \lambda$. Then there is a model $M \preceq N$ containing A and an automorphism $h : M \rightarrow M$ that extends k .

Proof Similar to the proof of Theorem 6.10. We shall construct simultaneously a chain $\langle A_i : i < \lambda \rangle$ of subsets of N and a chain functions $\langle h_i : i < \lambda \rangle$, such that $h_i : N \rightarrow N$ are morphisms. In the end we will set

$$M = \bigcup_{i < \lambda} A_i \quad \text{and} \quad h = \bigcup_{i < \lambda} h_i$$

The chains start with $A_0 = A \cup \text{dom } k \cup \text{img } k$ and $h_0 = k$. As usual, at limit stages we take the union. Now we consider successor stages. At stage i we fix some enumerations of A_i and of $L_x(A_i)$, where $|x| = 1$. Let $\langle i_1, i_2 \rangle$ be the i -th pair of ordinals $< \lambda$. If the i_2 -th formula in $L_x(A_{i_1})$ is consistent in N , let a be any of its solutions. Also let b be the i_2 -th element of A_{i_1} . Let $h_{i+1} : N \rightarrow N$ be a minimal morphism that extends h_i and is such that $b \in \text{dom } h_{i+1} \cap \text{img } h_{i+1}$. Define $A_{i+1} = A_i \cup \{a, h_{i+1}b, h_{i+1}^{-1}b\}$. \square

- 6.23 Theorem** (Assume 6.4) For every model N the following are equivalent

1. N is λ -rich;
2. N is weakly λ -universal and weakly λ -homogeneous.

Proof Implication $1 \Rightarrow 2$ is clear. To prove $2 \Rightarrow 1$ we generalize the proof of Theorem 6.8. We assume 2 fix some morphism $k : M \rightarrow N$ of cardinality $< \lambda$ and let $b \in M$. By weak λ -universality there is a morphism $f : M \rightarrow N$ with domain of definition $\text{dom } k \cup \{b\}$. The map $f \circ k^{-1} : N \rightarrow N$ has cardinality $< \lambda$ and, by

Lemma 6.22, it has an extension to an automorphism $h : N' \rightarrow N'$ for some $N' \preceq N$ containing $\text{img } k \cup \text{img } f$. Then $h \circ f : M \rightarrow N$ extends k and is defined on b . \square

3 The amalgamation property

We say that \mathcal{M} has the **amalgamation property** if for every pair of morphisms $f_1 : M \rightarrow M_1$ and $f_2 : M \rightarrow M_2$ there are some total morphism $g_1 : M_1 \rightarrow N$ and $g_2 : M_2 \rightarrow N$ such that $g_1 \circ f_1(a) = g_2 \circ f_2(a)$ for every $a \in \text{dom}(g_1 \circ f_1) \cap \text{dom}(g_2 \circ f_2)$.

We write $M \leq N$ for $M \subseteq N$ and $\text{id}_M : M \rightarrow N$ is a morphism. We say that $\langle M_i : i < \lambda \rangle$ is a **\leq -chain** if $M_i \leq M_j$ for all $i < j < \lambda$.

We say that $h : M' \rightarrow N'$ **extends** $k : M \rightarrow N$ if $k \subseteq h$, $M \leq M'$, and $N \leq N'$. As we assume that morphisms are invertible, we may express the amalgamation property in the following convenient form: every morphism extends to a total one.

For the following theorem to hold we need the following property:

6.24 Definition We say that \mathcal{M} is **closed under union of \leq -chains** if

c7. if $\langle M_i : i < \lambda \rangle$ is a \leq -chain, then $M_i \leq \bigcup_{j < \lambda} M_j$ for all $i < \lambda$. \square

The following is a general existence theorem for rich models.

6.25 Theorem Let λ be a cardinal such that $\lambda = \lambda^{<\lambda}$. (Assume 6.4) Assume further that \mathcal{M} has the amalgamation property and is closed under union of \leq chains. Then there is a rich model N of cardinality λ .

Proof We construct N as union of a \leq -chain of models $\langle N_i : i < \lambda \rangle$ such that $|N_i| = \lambda$ for all $i < \lambda$. Let N_0 be any model of cardinality λ . At stage $i + 1$, let $f : M \rightarrow N_i$ be the least morphism (in a well-ordering that we specify below) such that $|f| < |M| \leq \lambda$ and f has no extension to an embedding $f' : M \rightarrow N_i$. Apply the amalgamation property to obtain a total morphism $f' : M \rightarrow N'$ that extends $f : M \rightarrow N_i$. By the downward Löwenheim-Skolem Theorem we may assume $|N'| = \lambda$. Let $N_{i+1} = N'$. At limit stages take the union.

The well-ordering mentioned needs to be chosen so that in the end we forget nobody. So, first at each stage we well-order the isomorphism-type of the morphisms $f : M \rightarrow N_i$ such that $|f| < |M| \leq \lambda$. Then the required well-ordering is obtained by dovetailing all these well-orderings. The length of this enumeration is at most $2^\lambda \cdot \lambda^{<\lambda}$, which is λ by hypothesis.

We check that N is rich. Let $f : M \rightarrow N$ be a morphism and $|f| < |M| \leq \lambda$. The cofinality of λ is larger than $|f|$, hence $\text{img } f \subseteq N_i$ for some $i < \lambda$. So $f : M \rightarrow N_i$ is a morphism and at some stage j we have ensured the existence of a total morphism of $f' : M \rightarrow N_{j+1}$ that extends f . \square

4 Notes

[1] Michael Lieberman and Jirí Rosický, *Classification theory for accessible categories* (2014), available at <https://arxiv.org/abs/1404.2528>.

Chapter 7

Some algebraic structures

The main result in this chapter is Corollary 7.24, the elimination of quantifiers in algebraically closed fields, which in algebra is called Chevalley's Theorem on constructible sets. From it we derive Hilbert's Nullstellensatz 7.24. Finally, we isolate the model theoretic properties of those types that correspond to the algebraic notions of prime and radical ideal of polynomials.

The first sections of this chapter are not a pre-requisite for Sections 4–7, at the cost of a few repetitions in the latter.

7.1 Notation Recall that when $A \subseteq M$ we denote by $\langle A \rangle_M$ the substructure of M generated by A . Then $\langle A \rangle_M \subseteq N$ is equivalent to $N \models \text{Diag} \langle A \rangle_M$. The diagram of a structure has been defined at the end of Section 4.3.

In this chapter, whenever some $A \subseteq M \models T$ are fixed, by *model* we mean superstructures of $\langle A \rangle_M$ that models T . The notions of *logical consequence*, *consistency*, *completeness*, etc. are modified accordingly, and we write \vdash for $T \cup \text{Diag} \langle A \rangle_M \vdash$.

We say that a type $p(x)$ is **trivial** if $\vdash p(x)$. □

1 Abelian groups

In this section the language L is that of additive groups. The theory T_{ag} of **abelian groups** is axiomatized by the universal closure of the following axioms

$$\text{a1 } (x + y) + z = y + (x + z);$$

$$\text{a2 } x + (-x) = (-x) + x = 0;$$

$$\text{a3 } x + 0 = 0 + x = x;$$

$$\text{a4 } x + y = y + x.$$

Let x be a tuple of variables of length α , an ordinal. We write $L_{\text{ter},x}$ for the set of terms $t(x)$ with free variables among x . On this set we define the equivalence relation

$$t(x) \sim s(x) = T_{\text{ag}} \vdash t(x) = s(x).$$

We define the group operations on $L_{\text{ter},x}/\sim$ in the obvious way. We denote by $\mathbb{Z}^{\oplus \alpha}$ the set of tuples of integers of length α that are almost always 0. The group operations on $\mathbb{Z}^{\oplus \alpha}$ are defined coordinate-wise. The following immediate proposition implies in particular that $L_{\text{ter},x}/\sim$ is isomorphic to $\mathbb{Z}^{\oplus \alpha}$.

7.2 Proposition *Let $A \subseteq M \models T_{\text{ag}}$. Then for every formula $\varphi(x) \in L_{\text{at}}(A)$ there are $n \in \mathbb{Z}^{\oplus \alpha}$ and $c \in \langle A \rangle_M$ such that*

$$\vdash \varphi(x) \leftrightarrow \sum_{i < \alpha} n_i x_i = c.$$

where $n = \langle n_i : i < \alpha \rangle$ and $x = \langle x_i : i < \alpha \rangle$. \square

Proof Up to equivalence over T_{ag} the formula $\varphi(x)$ has the form $s(x) = t(a)$ for some parameter-free terms $s(x)$ and $t(z)$. Over $\text{Diag } \langle A \rangle_M$, we can replace $t(a)$ with a single element of $\langle A \rangle_M$ and write $s(x)$ as the linear combination shown above. \square

7.3 Definition Let $M \models T_{\text{ag}}$. For $A \subseteq M$ and $c \in M$. We say that c is **independent from A** if $\langle A \rangle_M \cap \langle c \rangle_M = \{0\}$. Otherwise we say that c is **dependent from A** . The **rank** of M is the least cardinality of a subset $A \subseteq M$ such that all elements in M are dependent from A . We denote it by **rank M** . \square

Note that when M is a vector space the condition $\langle A \rangle_M \cap \langle c \rangle_M = \{0\}$ is equivalent to saying that c is not a linear combination of vectors in A and $\text{rank } M$ coincides with the dimension of M . In fact, what we do here for abelian groups could be easily generalized to D -modules, where D is any integral domain, and in particular to vector spaces.

The following proposition gives a convenient syntactic characterization of independence.

7.4 Proposition Let $A \subseteq M \models T_{\text{ag}}$. Then for every $c \in M$ the following are equivalent.

1. $\langle A \rangle_M \cap \langle c \rangle_M = \{0\}$;
2. $p(x) = \text{at-tp}_M(c/A)$ is trivial (see Notation 4.19 and 7.1).

Proof $1 \Rightarrow 2$ By Proposition 7.2, a non trivial formula in $p(x)$ may be assumed to have the form $nx = a$ for some $n \in \mathbb{Z} \setminus \{0\}$ and some $a \in \langle A \rangle_M \setminus \{0\}$. If such a formula is satisfied by c then $0 \neq a \in \langle A \rangle_M \cap \langle c \rangle_M$.

$2 \Rightarrow 1$ If $\langle A \rangle_M \cap \langle c \rangle_M \neq \{0\}$ then $nc = a$ for some $a \in \langle A \rangle_M \setminus \{0\}$. Then c satisfies the non trivial equation $nx = a$. \square

7.5 Remark Let $k : M \rightarrow N$ be a partial embedding and let a be an enumeration of $\text{dom } k$. We claim that $k \cup \{ \langle b, c \rangle \} : M \rightarrow N$ is a partial embedding for every $b \in M$ and $c \in N$ that are independent from a , respectively ka . In fact, it suffices to check that $M, a, b \equiv_{\text{at}} N, ka, c$. Suppose $\varphi(x, z) \in L_{\text{at}}$ is such that $M \models \varphi(a, b)$. Then by independence $\varphi(a, x)$ is trivial, i.e.

$$T_{\text{ag}} \cup \text{Diag } \langle a \rangle_M \vdash \varphi(a, x).$$

As $\langle a \rangle_M$ and $\langle ka \rangle_N$ are isomorphic structures

$$T_{\text{ag}} \cup \text{Diag } \langle ka \rangle_N \vdash \varphi(ka, x).$$

Therefore $N \models \varphi(ka, c)$. The proof that $N \models \varphi(ka, c)$ implies $M \models \varphi(a, b)$ is similar. \square

2 Torsion-free abelian groups

The theory of **torsion-free abelian groups** extends T_{ag} with the following axioms for all positive integers n

st $nx = 0 \rightarrow x = 0$.

We denote this theory by T_{tfag} . It is not difficult to see that in a torsion-free abelian group every equation of the form $nx = a$ has at most one solution.

7.6 Proposition *Let $M \models T_{\text{tfag}}$ be uncountable. Then $\text{rank } M = |M|$.*

Proof Let $A \subseteq M$ have cardinality $< |M|$. We claim that M contains some element that is independent from A . It suffices to show that the number of elements that are dependent from A is $< |M|$. If $c \in M$ is dependent from A then, by Proposition 7.7, it is a solution of some formula $L_{\text{at}}(A)$. As there is no torsion, such a formula has at most one solution. Therefore the number of elements that are dependent from A is at most $|L_{\text{at}}(A)|$, that is $\max\{|A|, \omega\}$. If M is uncountable the claim follows. \square

7.7 Proposition *Let $A \subseteq M \models T_{\text{tfag}}$. Let $p(x) = \text{at}^\pm\text{-tp}(b/A)$, where $b \in M$. Then one of the following holds:*

1. b is independent from A ;
2. $M \models \varphi(b)$ for some $\varphi(x) \in L_{\text{at}}(A)$ such that $\vdash \varphi(x) \rightarrow p(x)$.

Note the similarity with Example 6.14, where the independent type is $q(x)$ and the isolating formulas are the $r_i(x)$.

It is important to observe that the set A above may be infinite. This is essential to obtain Corollary 7.11, and it is one of the main differences between this example and the examples encountered in Chapter 5.

Proof If b is dependent from A , then b satisfies a non trivial atomic formula $\varphi(x)$ which we claim is the formula required in 2. It suffices to show that $\varphi(x)$ implies a complete $L_{\text{at}^\pm}(A)$ -type. Clearly this type must be $p(x)$. Let $\varphi(x)$ have the form $nx = a$ for some $n \in \mathbb{Z} \setminus \{0\}$ and $a \in \langle A \rangle_M \setminus \{0\}$. We show that for every $m \in \mathbb{Z}$ and every $c \in \langle A \rangle_M$ one of the following holds

- a. $\vdash nx = a \rightarrow mx = c$;
- b. $\vdash nx = a \rightarrow mx \neq c$.

Suppose not for a contradiction that neither a nor b holds and fix models N_1, N_2 and some $b_i \in N_i$ such that

- a'. $N_1 \models nb_1 = a$ and $N_1 \models mb_1 \neq c$;
- b'. $N_2 \models nb_2 = a$ and $N_2 \models mb_2 = c$.

From b' we obtain $N_2 \models ma = nc$. As N_1 is torsion-free, from a' we obtain $N_1 \models ma \neq nc$. But $ma = nc$ is a formula with parameters in $\langle A \rangle_M$, so it should have the same truth value in all superstructures of $\langle A \rangle_M$, a contradiction. \square

3 Divisible abelian groups

The theory of **divisible abelian groups** extends T_{tfag} with the following axioms for all integers $n \neq 0$

$$\text{div}_n \exists x nx = y.$$

We denote this theory by T_{dag} .

7.8 Proposition Let $M \models T_{\text{tfag}}$ and let $\varphi(x) \in L_{\text{at}}(A)$, where $|x| = 1$, be consistent. Then $N \models \exists x \varphi(x)$ for every model $N \models T_{\text{dag}} \cup \text{Diag}(A)_M$.

Note that the proposition mentions T_{tfag} first and then T_{dag} .

The claim in the proposition holds more generally for all $\varphi(x) \in L_{\text{qf}}$ and also when x is a tuple of variables. This follows from Lemma 7.10, whose proof uses the proposition.

Proof We can assume that $\varphi(x)$ has the form $nx = a$ for some $n \in \mathbb{Z}$ and some $a \in \langle A \rangle_N$. If $n = 0$, then $a = 0$ since $\varphi(x)$ is consistent, and the claim is trivial. Otherwise a solution exist in N by axiom div_n . \square

7.9 Exercise Prove a converse of Proposition 7.8. Let $A \subseteq N \models T_{\text{ag}}$ and let x be a single variable. Prove that if $N \models \exists x \varphi(x)$ for every consistent $\varphi(x) \in L_{\text{at}}(A)$, then $N \models T_{\text{dag}}$. \square

We are ready to prove that divisible abelian groups of infinite rank are ω -rich. We shall be slightly more general than that.

7.10 Lemma Let $k : M \rightarrow N$ be a partial isomorphism of cardinality $< \lambda$, where $M \models T_{\text{tfag}}$ and $N \models T_{\text{dag}}$ has rank $\geq \lambda$. Then for every $b \in M$ there is $c \in N$ such that $k \cup \{ \langle b, c \rangle \} : M \rightarrow N$ is a partial isomorphism.

Proof Let a be an enumeration of $\text{dom } k$ and let $p(x, z) = \text{at}^\pm\text{-tp}(b, a)$. The required c has to realize $p(x, ka)$. We consider two cases. If b is dependent from a , then Proposition 7.7 yields a formula $\varphi(x, z) \in L_{\text{at}}$ such that

- i. $\vdash \varphi(x, a) \rightarrow p(x, a)$
- ii. $\varphi(x, a)$ is consistent.

By isomorphism, i and ii hold with a replaced by ka . Then by Proposition 7.8 the formula $\varphi(x, ka)$ has a solution $c \in N$.

The second case, which has no analogue in Lemma 5.1, is when b is independent from a . Then by Remark 7.5 we may choose c to be any element of N independent from ka . Such an element exists because N has rank at least λ . \square

7.11 Corollary Every uncountable model of T_{dag} is rich in the category of models of T_{tfag} and partial isomorphisms. In particular T_{dag} is uncountably categorical, it is complete, and has quantifier elimination.

Proof Any uncountable $N \models T_{\text{dag}}$ has rank $|N|$, therefore it is rich by Lemma 7.10; categoricity and completeness follow. As for quantifier elimination, suppose $k : M \rightarrow N$ is a partial isomorphism between models of T_{dag} . If $M \preceq M'$ and $M \preceq N'$ are elementary superstructures of uncountable cardinality then $k : M' \rightarrow N'$ is elementary by Theorem 6.10 and this suffices to conclude that $k : M \rightarrow N$ is elementary. \square

7.12 Exercise Prove that every model of T_{dag} is ω -ultraomogeneous (independency of cardinality and rank). \square

4 Commutative rings

In this section L is the **language of (unital) rings**. It contains two constants 0 and 1 the unary operation $-$ and two binary operations $+$ and \cdot . The theory of rings contains the following axioms

a1-a4 as for abelian groups

$$r1 \quad (x \cdot y) \cdot z = y \cdot (x \cdot z),$$

$$r2 \quad 1 \cdot x = x \cdot 1 = x,$$

$$r3 \quad (x + y) \cdot z = x \cdot z + y \cdot z,$$

$$r4 \quad z \cdot (x + y) = z \cdot x + z \cdot y.$$

All the rings we consider are **commutative**

$$c \quad x \cdot y = y \cdot x.$$

We denote the theory of commutative rings by T_{cr} .

In what follows the theory $T_{cr} \cup \text{Diag}\langle A \rangle_M$ is implicit in the sense of Notation 7.1, so it is important to remember that $\text{Diag}\langle A \rangle_M$ is not trivial even when $A = \emptyset$. In fact, $\text{Diag}\langle \emptyset \rangle_M$ determines the characteristic of the models.

Let $A \subseteq M \models T_{cr}$ and let x be a tuple of variables of length α , an ordinal. We write $L_{\text{ter},x}(A)$ for the set of terms $t(x)$ with free variables among x and parameters in A . On this set we define the equivalence relation

$$t(x) \sim s(x) \quad = \quad \vdash \quad t(x) = s(x).$$

On $L_{\text{ter},x}(A)/\sim$ we define the ring operations in the obvious way so that $L_{\text{ter},x}(A)/\sim$ is a commutative ring. We denote by $A[x]$ the set of polynomials with variables among x and parameters in $\langle A \rangle_M$. The ring operations on $A[x]$ are defined as usual. The following proposition (which is clear, but tedious to prove) implies in particular that $L_{\text{ter},x}(A)/\sim$ is isomorphic to $A[x]$. For simplicity we state it only for $|x| = 1$.

7.13 Proposition *Let $A \subseteq M \models T_{cr}$ and let x be a single variable. Then for every formula $\varphi(x) \in L_{\text{at}}(A)$ there is a unique $n < \omega$ and a unique tuple $a = \langle a_i : i \leq n \rangle$ of elements of $\langle A \rangle_M$ such that $a_n \neq 0$ and*

$$\vdash \quad \varphi(x) \leftrightarrow \sum_{i \leq n} a_i x^i = 0. \quad \square$$

The integer n in the proposition above is called the **degree of $\varphi(x)$** .

7.14 Definition *Let $A \subseteq M \models T_{cr}$. We say that an element $b \in M$ is **transcendent over A** if the type $p(x) = \text{at-tp}_M(b/A)$ is trivial (see Notation 4.19 and 7.1). Otherwise we say that b is **algebraic over A** . The **transcendence degree** of M is the least cardinality of a subset $A \subseteq M$ such that no element of M is transcendent over A .* \square

7.15 Remark Remark 7.5 holds here with ‘independent’ replaced by ‘transcendent’ and T_{ag} replaced by T_{cr} . \square

5 Integral domains

Let $a \in M \models T$. We say that a is a **zero divisor** if $ab = 0$ for some $b \in M \setminus \{0\}$. An **integral domain** is a commutative ring without zero divisors. The theory of integral domains contains the axioms of commutative rings and the following

nt. $0 \neq 1$

id. $x \cdot y = 0 \rightarrow x = 0 \vee y = 0$.

We denote the theory of integral domains by T_{id} .

7.16 Proposition *Let $M \models T_{\text{id}}$ be uncountable. Then M has transcendence degree $|M|$.*

Proof Every polynomial has finitely many solutions in an integral domain and there are $|L(A)|$ polynomials over A . \square

7.17 Proposition *Let $A \subseteq M \models T_{\text{id}}$. For $b \in M$ let $p(x) = \text{at}^\pm\text{-tp}(b/A)$. Then one of the following holds*

1. b is transcendental over A ;
2. $M \models \varphi(b)$ for some $\varphi(x) \in L_{\text{at}}(A)$ such that $\vdash \varphi(x) \rightarrow p(x)$.

Note the similarity with Example 6.14, where the transcendental type is $q(x)$ and the isolating formulas are the $r_i(x)$.

As in Proposition 7.7, the set A may be infinite. This is essential to obtain Corollary 7.20.

Proof Suppose b is not transcendental, i.e. it satisfies a non trivial atomic formula. Let $\varphi(x) \in L_{\text{at}}(A)$ be a non trivial formula with minimal degree such that $\varphi(b)$. We prove that $\varphi(x)$ implies a complete $L_{\text{at}^\pm}(A)$ -type. Clearly this type must be $p(x)$. We prove that for any $\xi(x) \in L_{\text{at}}(A)$ one of the following holds

1. $\vdash \varphi(x) \rightarrow \xi(x)$
2. $\vdash \varphi(x) \rightarrow \neg \xi(x)$.

Let us write $a(x) = 0$ and $a'(x) = 0$ for the formulas $\varphi(x)$ and $\xi(x)$, respectively. Let n be the degree of the polynomial $a(x)$. If $\langle A \rangle_M$ is a field, choose a polynomial $d(x)$ of maximal degree such that

- a. $d(x)t(x) = a(x)$
- a'. $d(x)t'(x) = a'(x)$,

for some polynomials $t(x)$ and $t'(x)$.

If $\langle A \rangle_M$ is not a field, polynomials $d(x)$, $t(x)$ and $t'(x)$ as above exist with coefficients in the field of fractions of $\langle A \rangle_M$. Then a and a' hold up to a factor in $\langle A \rangle_M$ which we absorb in $a(x)$ and $a'(x)$.

From a we get $d(b) = 0$ or $t(b) = 0$. In the first case, as $a(x)$ has minimal degree, we conclude that $t(x)$ is constant. This implies that any zero of $a(x)$ is also a zero of $a'(x)$, that is, it implies 1.

Now suppose $t(b) = 0$. Then the minimality of the degree of $a(x)$ implies that $d(x) = d$, where d is a nonzero constant. If $\langle A \rangle_M$ is a field, apply Bézout's identity

to obtain two polynomials $c(x)$ and $c'(x)$ such that $d = a(x)c(x) + a'(x)c'(x)$. Then $a(x)$ and $a'(x)$ have no common zeros, and 2 follows. If $\langle A \rangle_M$ is not a field, we use Bézout's identity in the field of fractions of $\langle A \rangle_M$ and obtain $d'd = a(x)c(x) + a'(x)c'(x)$ for some $d' \in \langle A \rangle_M \setminus \{0\}$. Then we reach the same conclusion. \square

6 Algebraically closed fields

Let $a, b \in M \models T_{\text{cr}}$. We say that b is the **inverse** of a if $a \cdot b = 1$. A field is a commutative ring where every non-zero element has an inverse. The **theory of fields** contains T_{id} and the axiom

$$\text{f. } \exists y [x \neq 0 \rightarrow x \cdot y = 1].$$

Fields are structures in the signature of rings: the language contains no symbol for the multiplicative inverse. So, substructures of fields are merely integral domains.

The theory of **algebraically closed field**, which we denote by T_{acf} , also contains the following axioms for every positive integer n

$$\text{ac}_n. \exists x (x^n + z_{n-1}x^{n-1} + \dots + z_1x + z_0 = 0)$$

This not a complete theory because it does not decide the characteristic of the field. For a prime p , we define the theory T_{acf}^p , which contains T_{acf} and the axiom

$$\text{ch}_p. 1 + \dots (p \text{ times}) \dots + 1 = 0.$$

The theory T_{acf}^0 contains the negation of ch_p for all p . Note that all models of T_{acf}^p have the same characteristic in the model theoretic sense defined in 4.26. In fact, any two fields M and N with the same (algebraic) characteristic have isomorphic prime subfield, that is, $\langle \emptyset \rangle_M \simeq \langle \emptyset \rangle_N$.

7.18 Proposition *Let $A \subseteq N \models T_{\text{acf}}$ and let x be a single variable. Then $N \models \exists x \varphi(x)$ for every $\varphi(x) \in L_{\text{at}}(A)$ that is consistent with $T_{\text{id}} \cup \text{Diag}\langle A \rangle_M$.*

Note that the proposition mentions T_{acf} first and then T_{id} .

The claim in the proposition holds more generally for all $\varphi(x) \in L_{\text{qf}}$ when x is a tuple of variables. This follows from Lemma 7.19 whose proof uses the proposition.

Proof Up to equivalence $\varphi(x)$ has the form $a_nx^n + \dots + a_1x + a_0 = 0$ for some $a_i \in \langle A \rangle_N$. Choose n minimal. If $n = 0$, then $a_0 = 0$ since $\varphi(x)$ is consistent, and the claim is trivial. Otherwise $a_n \neq 0$ and the claim follows from f and ac_n . \square

7.19 Lemma *Let $k : M \rightarrow N$ be a partial isomorphism of cardinality $< \lambda$, where $M \models T_{\text{id}}$ and $N \models T_{\text{acf}}$ has transcendence degree $\geq \lambda$. Then for every $b \in M$ there is $c \in N$ such that $k \cup \{ \langle b, c \rangle \} : M \rightarrow N$ is a partial isomorphism.*

The following is the proof of Lemma 7.10 which we have pasted here for the sake of the lazy reader.

Proof Let a be an enumeration of $\text{dom } k$ and let $p(x, z) = \text{at}^\pm\text{-tp}_M(c, a)$. The required c has to realize $p(x, ka)$. We consider two cases. If b is algebraic over a , then Proposition 7.17 yields a formula $\varphi(x, z) \in L_{\text{at}}$ such that

$$\text{i. } \varphi(x, a) \rightarrow p(x, a)$$

ii. $\varphi(x, a)$ is consistent.

By isomorphism i and ii hold with a replaced by ka . Then by Proposition 7.18 the formula $\varphi(x, ka)$ has a solution in $c \in N$.

The second case, which has no analogue in Lemma 5.1, is when b is transcendent over a . Then by Remark 7.15 we may choose c to be any element of N transcendent over ka . This exists because N has transcendence degree $\geq \lambda$. \square

7.20 Corollary *Every uncountable model of T_{acf}^p is rich in the category of integral domains of characteristic p and partial isomorphisms. In particular T_{acf}^p is uncountably categorical, it is complete, and has quantifier elimination.*

The proof requires the same argument as Corollary 7.11. However here we have to fix the characteristic of the integral domain, otherwise the resulting category is not connected.

Proof Every uncountable $N \models T_{\text{acf}}^p$ has transcendence degree $|N|$, therefore it is rich by Lemma 7.19. Categoricity and completeness follow.

For quantifier elimination, let $k : M \rightarrow N$ be a partial embedding between models of T_{acf}^p . Let M' and N' be elementary superstructures of M and N respectively of sufficiently large cardinality. As M' and N' are rich, $k : M' \rightarrow N'$ is elementary by Theorem 6.10. Hence $k : M \rightarrow N$ is also elementary. \square

7.21 Exercise Prove that every model of T_{acf} is ω -ultrahomogeneous (independently of cardinality and transcendence degree). \square

7 Hilbert's Nullstellensatz

In this section we fix a tuple of variables x and a subset A of an integral domain. We denote by $\Delta(A)$ the set of formulas of the form $t(x) = 0$ where $t(x)$ is a term with parameters in A . So $\Delta(A)$ -types are (possibly infinite) systems of polynomial equations with coefficients in $\langle A \rangle_M$.

For convenience we define the closure of $p(x)$ under logical consequences as follows

$$\text{ccl } p(x) = \{t(x) = 0 : p(x) \vdash t(x)=0\}$$

Remember that we always work under the assumptions made in Notation 7.1. In particular, in this section we work over the theory $T_{\text{id}} \cup \text{Diag}\langle A \rangle_M$. In general, closure under logical consequences is an elusive notion. Hence Propositions 7.22 and 7.23 are useful because they give a model theoretical, respectively algebraic, characterisation of $\text{ccl } p(x)$.

7.22 Proposition *Let $A \subseteq M \models T_{\text{id}}$ and let $p(x)$ be a $\Delta(A)$ -type. Fix some N of sufficiently large cardinality such that $\langle A \rangle_M \subseteq N \models T_{\text{acf}}$. Then*

$$\text{ccl } p(x) = \{t(x) = 0 : N \models \forall x [p(x) \rightarrow t(x)=0]\}.$$

It suffices to choose N such that $|A| < |N|$ and $|x| \leq |N|$; note that x has possibly infinite length. In Corollary 7.24 below we will considerably strengthen this proposition for finite x .

Proof Only the inclusion \supseteq requires a proof. Fix some polynomial $t(x)=0 \notin \text{ccl } p(x)$. Then $p(x) \wedge t(x) \neq 0$ is consistent, so there is a model M' such that $M' \models p(a) \wedge t(a) \neq 0$ for some $a \in M'^{|x|}$. Then there is a partial isomorphism $h : M' \rightarrow N$ that extends id_A and is defined on a , provided N is large enough to accommodate a . This implies that $p(x) \wedge t(x) \neq 0$ has a solution in N . Hence $t(x)$ does not belong to the set on the r.h.s. \square

Recall that $A[x]$ denotes the ring of polynomials with variables in x and coefficients in $\langle A \rangle_M$. We identify $\Delta(A)$ and $A[x]$ in the obvious way. Consequently, a $\Delta(A)$ -type $p(x)$ is identified with a set of polynomials $p \subseteq A[x]$. For $p \subseteq A[x]$ we write **rad p** for the **radical ideal generated by p** , that is, the intersection of all prime ideals containing p . When $p = \text{rad } p$, we say that **p is a radical ideal**. Recall from algebra that if $p \subseteq A[x]$ is an ideal then

$$\# \quad \text{rad } p = \left\{ t(x) : t^n(x) \in p \text{ for some positive integer } n \right\},$$

which explains the name.

7.23 Proposition *Let $A \subseteq M \models T_{\text{id}}$ and let $p(x)$ be a $\Delta(A)$ -type. Then $\text{ccl } p(x) = \text{rad } p$.*

The proposition holds with a similar proof for the broader class of rings without nilpotent elements. (Which does not come as a surprise.)

Proof (\supseteq) We claim that $\text{ccl } p(x)$ is an ideal. In fact, for every pair of $L(A)$ -terms $t(x)$ and $s(x)$

$$\begin{aligned} t(x) = 0 &\vdash s(x)t(x) = 0 \\ s(x) = t(x) = 0 &\vdash s(x) + t(x) = 0 \end{aligned}$$

Moreover, as integral domains do not have nilpotent elements

$$t^n(x) = 0 \vdash t(x) = 0$$

By # above, $\text{ccl } p(x)$ is a radical ideal which proves $\text{ccl } p(x) \supseteq \text{rad } p$.

(\subseteq) We fix some $t(x) \notin \text{rad } p$ and prove that $p(x) \wedge t(x) \neq 0$ is consistent. Let q be some prime ideal containing p such that $t(x) \notin q$. As q is prime, the ring $A[x]/q$ is an integral domain. The polynomials that vanish in $A[x]/q$ at $x + q$ are exactly those in q . Hence $A[x]/q$ witnesses the consistency of $q(x) \wedge t(x) \neq 0$. \square

When x is a finite tuple of variables, we can extend the validity of Proposition 7.22 to the case $A = M = N$.

7.24 Corollary (Hilbert's Nullstellensatz) *Let $N \models T_{\text{acf}}$ and let $p(x)$, where $|x| < \omega$, be a $\Delta(N)$ -type. Then*

$$\text{rad } p = \text{ccl } p(x) = \left\{ t(x) : N \models \forall x [p(x) \rightarrow t(x) = 0] \right\}.$$

Proof Let N' be a large elementary extension of N . By Proposition 7.22, the claim holds for N' . By Hilbert's Basis Theorem, the ideal generated by p is finitely generated hence $p(x)$ is equivalent to a formula (cfr. Exercise 7.27). Therefore, by elementarity, the claim holds for N . \square

Hilbert's Nullstellensatz comes in two variants. The one in Corollary 7.24 is sometimes referred to as the *strong* Nullstellensatz. The weaker variant is stated in Exercise 7.26.

We conclude this section by showing that the notions of primeness for types and ideals coincide (if we restrict to types closed under logical consequences).

7.25 Proposition *Let $A \subseteq M \models T_{\text{id}}$ and let $p(x)$ be a $\Delta(A)$ -type closed under logical consequences. Then the following are equivalent*

1. $p(x)$ is a prime $\Delta(A)$ -type;
2. p is a prime ideal.

Proof $1 \Rightarrow 2$ Assume 1 and suppose that the polynomial $t(x) \cdot s(x)$ belongs to p . Clearly, over $T_{\text{id}} \cup \text{Diag}(A)_M$ we have $\vdash t(x) \cdot s(x) = 0 \rightarrow t(x) = 0 \vee s(x) = 0$. As $p(x)$ is a prime $\Delta(A)$ -type, $p(x) \vdash t(x) = 0$ or $p(x) \vdash s(x) = 0$. By Proposition 7.23, the type $p(x)$ is closed under logical consequences, therefore $t(x) \in p$ or $s(x) \in p$.

$2 \Rightarrow 1$ Assume p is a prime ideal and for some $t_i(x) = 0 \in \Delta(A)$

$$p(x) \vdash \bigvee_{i=1}^n t_i(x) = 0.$$

Then

$$p(x) \vdash \prod_{i=1}^n t_i(x) = 0$$

Since $p(x)$ is closed under logical consequences, and p is a prime ideal, $t_i(x) \in p$ for some i . Hence $p(x)$ contains the equation $t_i(x) = 0$. By Corollary 4.18 this suffices to prove that $p(x)$ is a prime $\Delta(A)$ -type. \square

7.26 Exercise Let $N \models T_{\text{acf}}$ and let $p(x)$ be a $\Delta(N)$ -type where $|x| < \omega$. Prove that the following are equivalent

1. p is a proper ideal;
2. $p(x)$ has a solution in N .

\square

7.27 Exercise Let $A \subseteq M \models T_{\text{id}}$ and let $p(x)$ be a $\Delta(A)$ -type. Prove that the following are equivalent

1. $p(x)$ is a principal $\Delta(A)$ -type;
2. the ideal generated by p is finitely generated.

\square

Chapter 8

Saturation and homogeneity

The first two section introduce saturation and homogeneity and Section 3 presents the notation we shall use in the following chapters when working inside a monster model.

1 Saturated structures

Recall that a type $p(x) \subseteq L(M)$ is **finitely consistent in M** if every $\varphi(x)$, conjunction of formulas in $p(x)$, has a solution in M . When $A \subseteq M$ we write $S_x(A)$ for the set of types that are complete and finitely consistent in M . We never display M in the notation as it will always be clear from the context. When A is empty it usual to write $S_x(T)$ for $S_x(A)$ where $T = \text{Th}(M)$. We write $S(A)$ for the union of $S_x(A)$ as x ranges over all tuples of variables. Similarly for $S(T)$.

The following remark will be used in the sequel without explicit reference.

8.1 Remark Let $p(x, z) \subseteq L$, let $k : M \rightarrow N$ be an elementary map, and let $a \in (\text{dom } h)^{|z|}$. If $p(x, a)$ is finitely consistent in M , then $p(x, ka)$ is finitely consistent in N . □

8.2 Definition Let x be a single variable and let λ be an infinite cardinal. We say that a structure N is **λ -saturated** if it realizes every type $p(x)$ such that

1. $p(x) \subseteq L(A)$ for some $A \subseteq N$ of cardinality $< \lambda$;
2. $p(x)$ is finitely consistent in N .

We say that N is **saturated** tout court if it is λ -saturated and $|N| = \lambda$.

8.3 Exercise Suppose $|L| \leq \omega$ and let M be an infinite structure. Then for every non-principal ultrafilter F on ω the structure M^ω / F is a ω_1 -saturated elementary superstructure of M .

Hint: the notation is as in Chapter 3. Let \hat{c} be a tuple of length ω of elements of M^ω and let $p(x; z) = \{\varphi_i(x; z) : i < \omega\} \subseteq L$ be such that $p(x; \hat{c})$ be a type finitely consistent in M^ω / F . Without loss of generality we can assume that $\varphi_{i+1}(x; z) \rightarrow \varphi_i(x; z)$. Let $\langle X_i : i < \omega \rangle$ be a strictly decreasing chain of elements of F such that $X_{i+1} \subseteq \{j : M \models \exists x \varphi_i(x, \hat{c}_j)\}$. Let $\hat{a} \in M^\omega$ be such that $\varphi_i(\hat{a}_j, \hat{c}_j)$ holds for every $j \in X_i \setminus X_{i+1}$. Then \hat{a} realizes $p(x)$. □

8.4 Theorem Assume $|L| < \lambda$ where λ is be such that $\lambda^{<\lambda} = \lambda$. Then every structure M of cardinality λ has a saturated elementary extension of cardinality λ .

Proof We construct an elementary chain $\langle M_i : i < \lambda \rangle$ of models of cardinality λ . The chain starts with M and is the union at limit stages. Given M_i we choose as M_{i+1} any model of cardinality λ that realizes all types in $S_x(A)$ for all $A \subseteq M_i$ of

cardinality $< \lambda$. The required M_{i+1} exists because there are $2^{|L(A)|} \leq \lambda$ types in $S_x(A)$ and there are $\lambda^{<\lambda} = \lambda$ sets A .

Let N be the union of the chain. We check that N is the required extension. Let $p(x) \in S(A)$ for some $A \subseteq N$ of cardinality $< \lambda$. As λ is a regular cardinal, $A \subseteq M_i$ for some $i < \lambda$. Then M_{i+1} realizes $p(x)$, and so does N , by elementarity. \square

8.5 Theorem Assume $|L| \leq \lambda$ and let N be an infinite structure. Let \mathcal{M} be the category (see Section 6.1) that consists of models of a complete theory T and elementary maps between these. Then the following are equivalent

- 1 N is a λ -saturated structure;
- 2 N is a λ -rich model;
- 3 N realizes every type $p(x) \subseteq L(A)$, with $|x| \leq \lambda$ and $|A| < \lambda$, finitely consistent in N .

Note that it is the completeness of T which makes the category \mathcal{M} connected.

Proof $1 \Rightarrow 2$. Let $k : M \rightarrow N$ be an elementary map of cardinality $< \lambda$. It suffices to show that for every $b \in M$ there is a $c \in N$ such that $k \cup \{ \langle b, c \rangle \} : M \rightarrow N$ is an elementary map. Let a be an enumeration of $\text{dom } k$ and define $p(x; z) = \text{tp}_M(b; a)$. The required c is any element of N such that $N \models p(c; ka)$. As $p(x; a)$ is finitely consistent in M then $p(x; ka)$ is finitely consistent in N . As $|a| < \lambda$, saturation yields the required c .

$2 \Rightarrow 3$. Let $p(x)$ be as in 3. By the compactness theorem $N \preceq K \models p(b)$ for some model K and $b \in K^{|x|}$. By the downward Löwenheim-Skolem theorem there is a model $A, b \subseteq M \preceq K$ of cardinality $\leq \lambda$. (Here we use $|L|, |A|, |x| \leq \lambda$.) By 2, there is an elementary embedding $h : M \rightarrow N$ that extends id_A . Finally, as $M \models p(b)$, elementarity yields $N \models p(hb)$.

$3 \Rightarrow 1$. Trivial. \square

Two saturated structures of the same cardinality are isomorphic as soon as they are elementarily equivalent (i.e. as soon as $\emptyset : M \rightarrow N$ is an elementary map.) In fact, from Theorem 8.5 and 6.6 we obtain the following.

8.6 Corollary Every elementary map $k : M \rightarrow N$ of cardinality $< \lambda$ between saturated models of the same cardinality λ extends to an isomorphism. \square

As it turns out, we already have many examples of saturated structures.

8.7 Corollary The following models are ω -saturated

- 1 models of T_{dlo} ;
- 2 models of T_{rg} ;
- 3 models of T_{dag} with infinite rank;
- 4 models of T_{acf} with infinite degree of transcendence.

Countable models of T_{dlo} and T_{rg} and uncountable models of T_{dag} or T_{acf} are saturated.

Proof By quantifier elimination embeddings coincide with elementary embeddings. Then saturation is proved applying Theorem 8.5 and the extension lemmas proved

in Section 5 and 7. □

8.8 Exercise Let T be a consistent theory that decides the characteristic of its models. Consider the category that consists of models of T and partial isomorphism. Assume $|L| \leq \lambda$ and prove that the following are equivalent

1. Every λ -saturated model of T is λ -rich;
2. T has elimination of quantifiers. □

8.9 Exercise Let L be the language of strict order. Let T be the theory discrete linear orders without endpoints. (An order is discrete if every point has an immediate predecessor and an immediate successor.) Show that the structure $\mathbb{Q} \times \mathbb{Z}$ ordered with the lexicographic order

$$(a_1, a_2) < (b_1, b_2) \stackrel{\text{def}}{\iff} a_1 < b_1 \text{ or } a_1 = b_1 \text{ e } a_2 < b_2$$

is a saturated model of T . For every $n \in \omega$ add to the language a relation symbol $d(x, y) = n$ and add to T the axiom $d(x, y) = n \leftrightarrow \exists^{=n} z (x < z \leq y)$. Let T' be the resulting theory. Prove that T' has quantifier elimination.

Hint: answer the second question first using exercise 8.8. □

2 Homogeneous structures

Definition 6.3 introduces the notion of universal and homogeneous structures a the general context. When the morphisms of the underlying category are the elementary maps we refer to these notion as elementary homogeneity and elementary universality. Though one often omits to specify “elementary”. We repeat the definition in this particular case.

8.10 Definition A structure N is **(elementarily) λ -universal** if every $M \equiv N$ of cardinality $\leq \lambda$ there is a elementary embedding $h : M \rightarrow N$. We say **universal** tout court if it is λ -universal and of cardinality λ .

We say that N is **(elementarily) λ -homogeneous** if every elementary map $k : N \rightarrow N$ of cardinality $< \lambda$ extends to an automorphism. We say that N is **homogeneous** tout court if it is λ -homogeneous and of cardinality λ . □

As saturated structure are rich, the following theorem is an instance of Theorem 6.8.

8.11 Theorem For every structure N of cardinality $\geq |L|$ the following are equivalent

1. N is saturated;
2. N is elementarily universal and homogeneous. □

Given $A \subseteq N$ we denote by $\text{Aut}(N/A)$ the group of **A -automorphisms** of N . That is the group of automorphisms that fix A point-wise. Let a be a tuple of elements of N . The **orbit of a over A in N** is the set

$$O_N(a/A) = \{fa : f \in \text{Aut}(N/A)\}$$

When the model N is clear from the context we omit the subscript.

Orbits are particularly interesting if taken in a homogeneous structure. The following proposition is immediate but its importance cannot be overrated.

8.12 Proposition *Let N be a λ -homogeneous structure. Let $A \subseteq N$ have cardinality $< \lambda$ and $a \in N^{<\lambda}$. Then $O_N(a/A) = p(N)$, where $p(x) = \text{tp}_N(a/A)$.* \square

Finally, we want to extend the equivalence in Theorem 8.11 to λ -saturated structures. For this we only need to apply Theorem 6.23.

When the morphisms of the underlying category are the elementary maps, we say **weakly λ -saturated** for weakly λ -universal (cfr. Definition 6.21). This is not the standard definition (i.e. 2 of the proposition below) but the reader can easily verify that it is equivalent by reasoning as in the proof of theorem 8.5.

8.13 Proposition *The following are equivalent*

1. N weakly λ -saturated;
 2. N realizes every type $p(x) \subseteq L$, where $|x| < \lambda$, that is finitely consistent in N .
- \square

The following is an instance of Theorem 6.23.

8.14 Corollary *Let $|L| \leq \lambda$. The following are equivalent*

1. N is λ -saturated;
 2. N is weakly λ -saturated and weakly λ -homogeneous.
- \square

8.15 Exercise Let M be an arbitrary structure. Prove that M has an ω -homogeneous elementary extension of the same cardinality. (There is no assumption on the cardinality of the language.) \square

8.16 Exercise Let M and N be elementarily homogeneous structures of the same cardinality λ . Suppose that $M \models \exists x p(x) \Leftrightarrow N \models \exists x p(x)$ for every $p(x) \subseteq L$ such that $|x| < \lambda$. Prove that the two structures are isomorphic. \square

8.17 Exercise Let L be a language that extends that of strict linear orders with the constants $\{c_i : i \in \omega\}$. Let T be the theory that extends T_{dlo} with the axioms $c_i < c_{i+1}$ for every $i \in \omega$. Prove that T has elimination of quantifiers and is complete (it can be deduced from what is known of T_{dlo}). Exhibit a countable saturated model and a countable model that is not homogeneous. \square

3 The monster model

In this section we present some notation and terminology frequently adopted when dealing with a complete theory T . We fix a saturated structure \mathcal{U} of cardinality larger than $|L|$. We assume \mathcal{U} to be large enough that among its elementary substructures we can find any model of T we might be interested in. For this reason \mathcal{U} is nicknamed the **monster model**. We denote by κ the cardinality of \mathcal{U} when necessary we assume κ to be inaccessible.

Some terms acquire a slightly different meaning when working inside a monster model.

small/large	cardinalities smaller than κ are called small;
models	are elementary substructure of \mathcal{U} of small cardinality, they are denoted by the letters M and N , and derived symbols;
parameters	are always in \mathcal{U} ; the symbols A, B, C , etc. denote sets of parameters of small cardinality; calligraphic letters as $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc. are used for sets of arbitrary cardinality;
tuples	have length $< \kappa$ unless otherwise specified;
global types	are complete types over \mathcal{U} ; the set of global types is denoted by $S(\mathcal{U})$;
formulas	have parameters in \mathcal{U} unless otherwise specified;
definable sets	are sets of the form $\varphi(\mathcal{U})$ for some formula $\varphi(x) \in L(\mathcal{U})$; we may say A -definable if $\varphi(x) \in L(A)$;
type-definable sets	are sets of the form $p(\mathcal{U})$ for some $p(x) \subseteq L(A)$ where, as the symbol suggests, A has small cardinality;
types of tuples	are always evaluated in \mathcal{U} , we write $\text{tp}(a/A)$ for $\text{tp}_{\mathcal{U}}(a/A)$ and $a \equiv_A b$ for $\mathcal{U}, a \equiv_A \mathcal{U}, b$;
orbits	of tuples under the action of $\text{Aut}(\mathcal{U}/A)$ are denoted by $\mathcal{O}(a/A)$.

Let x be a tuple of variables. For any fixed $A \subseteq \mathcal{U}$ we introduce a topology on $\mathcal{U}^{|x|}$ that we call the **topology induced by A** or, for short, **A -topology**. (This is non standard terminology, not to be confused with the *logic* A -topology in Section 13.3.) The closed sets of the A -topology are those of the form $p(\mathcal{U})$ where $p(x) \subseteq L(A)$ is any type over A .

For $\varphi(x) \in L(A)$ the sets of the form $\varphi(\mathcal{U})$ are clopen in this topology (and vice versa by Exercise 8.23). They form both a base of closed sets and base of open sets, which makes these topologies *zero-dimensional*. By saturation, the topology induced by A is compact. Actually, saturation is equivalent to the compactness of all these topologies as A ranges over the sets of small cardinality.

These topologies are never T_0 as any pair of tuples $a \equiv_A b$ have exactly the same neighborhoods. Such pairs always exist by cardinality reasons. However it is immediate that the topology induced on the quotient $\mathcal{U}^{|x|} / \equiv_A$ is Hausdorff (this is the so-called *Kolmogorov quotient*). Indeed, this quotient corresponds to $S_x(A)$ with the topology introduced in Section 3.

The following proposition is an immediate consequence of compactness. When $A = B$ it says that the topology induced by A is *normal*: any two closed sets are separated by open sets. It could be called **mutual normality** (not a standard name) because the two closed sets belong to different topologies and the separating sets are each found in the corresponding topology.

8.18 Proposition If $p(x) \subseteq L(A)$ and $q(x) \subseteq L(B)$ are such that $p(x) \rightarrow \neg q(x)$ then there are some formulas $\varphi(x)$ and $\psi(x)$, conjunctions of formulas in $p(x)$, respectively $q(x)$, such that $\varphi(x) \rightarrow \neg\psi(x)$.

Proof The assumptions say that $p(x) \wedge q(x)$ is inconsistent. Then the formulas $\varphi(x)$ and $\psi(x)$ exist by compactness (i.e. saturation). \square

A type-definable set $p(\mathcal{U})$ is invariant over A if $f[p(\mathcal{U})] = p(\mathcal{U})$ for every $f \in \text{Aut}(\mathcal{U}/A)$. Equivalently, if $\mathcal{O}(a/A) \subseteq p(\mathcal{U})$ for every $a \models p(x)$. By homogeneity this is equivalent to requiring that $q(x) \rightarrow p(x)$ for every $a \models p(x)$ and $q(x) = \text{tp}(a/A)$. Proposition 8.21 below an important fact about invariant type-definable sets. It may clarify the proof to consider first the particular case of definable sets.

8.19 Proposition Let $\varphi(x) \in L(\mathcal{U})$ then the following are equivalent

1. $\varphi(x)$ is equivalent to some formula $\psi(x) \in L(A)$;
2. $\varphi(\mathcal{U})$ is invariant over A .

We give two proofs of this theorem as we think they are both instructive.

Proof $1 \Rightarrow 2$ Obvious.

$2 \Rightarrow 1$ From 2 and homogeneity we obtain

$$\varphi(x) \leftrightarrow \bigvee_{q(x) \rightarrow \varphi(x)} q(x)$$

where $q(x)$ above range over all types in $S_x(A)$. By compactness, we can rewrite this equivalence

$$\varphi(x) \leftrightarrow \bigvee_{\vartheta(x) \rightarrow \varphi(x)} \vartheta(x)$$

where $\vartheta(x)$ ranges over all formulas in $L(A)$. This latter equivalence says that $\neg\varphi(x)$ is equivalent to a type over A . Again by compactness we obtain

$$\varphi(x) \leftrightarrow \bigvee_{i=1}^n \vartheta_i(x)$$

for some formula $\vartheta_i(x) \in L(A)$. \square

Second proof of Proposition 8.19 $2 \Rightarrow 1$ Let $\varphi(\mathcal{U}; b)$, where $\varphi(x; z) \in L$ be invariant over A . Let $p(z) = \text{tp}(b/A)$. As $f[\varphi(\mathcal{U}; b)] = \varphi(\mathcal{U}; fb)$ for every $f \in \text{Aut}(\mathcal{U}/A)$, homogeneity and invariance yield

$$p(z) \wedge p(z') \rightarrow \forall x [\varphi(x; z) \leftrightarrow \varphi(x; z')]$$

By compactness there is a formula $\vartheta(z) \in L(A)$ such that

$$\vartheta(z) \wedge \vartheta(z') \rightarrow \forall x [\varphi(x; z) \leftrightarrow \varphi(x; z')]$$

Hence $\varphi(\mathcal{U}; b)$ is defined by the formula $\exists z [\vartheta(z) \wedge \varphi(x; z)]$, which is a formula in $L(A)$ as required. \square

8.20 Exercise Let $\varphi(x) \in L$. Prove that the following are equivalent

1. $\varphi(x)$ is equivalent to some $\psi(x) \in L_{\text{qf}}$;
2. $a \in \varphi(\mathcal{U}) \leftrightarrow fa \in \varphi(\mathcal{U})$ for every partial isomorphism $f : \mathcal{U} \rightarrow \mathcal{U}$ defined in a . \square

8.21 Proposition Let $p(x) \subseteq L(\mathcal{U})$ then the following are equivalent

1. $p(x)$ is equivalent to some type $q(x) \subseteq L(A)$;
2. $p(\mathcal{U})$ is invariant over A .

We give two proofs of this theorem as we think they are both instructive. The second proof requires Proposition 8.22 below.

Proof $1 \Rightarrow 2$ Obvious.

$2 \Rightarrow 1$ It suffices to show that for every formula $\psi(x) \in p(x)$ there is a formula $\varphi(x) \in L(A)$ such that $p(x) \rightarrow \varphi(x) \rightarrow \psi(x)$. Fix $\psi(x) \in p(x)$. By invariance, any $q(x) \in S(A)$ consistent with $p(x)$ implies $p(x)$, hence

$$p(x) \rightarrow \bigvee_{q(x) \rightarrow \psi(x)} q(x) \rightarrow \psi(x)$$

where $q(x)$ above range over all types in $S_x(A)$. By compactness we can rewrite this equivalence as follows

$$p(x) \rightarrow \bigvee_{\vartheta(x) \rightarrow \psi(x)} \vartheta(x) \rightarrow \psi(x)$$

where $\vartheta(x)$ ranges over all formulas in $L(A)$. Applying Proposition 8.18 to the first implication we obtain a finite number of formulas $\vartheta_i(x)$ such that

$$p(x) \rightarrow \bigvee_{i=1}^n \vartheta_i(x) \rightarrow \psi(x)$$

This completes the proof. □

The following easy proposition is very useful. Its proof is left to the reader. Note that it would not hold without saturation: for instance, consider \mathbb{R} as a structure in the language of strict orders and let $q(x, y) = \text{tp}(-1, 0 / \mathbb{Q}^+)$. By quantifier elimination, $-1 \equiv_{\mathbb{Q}^+} 0$. But $0 \not\models \exists y q(x, y)$.

8.22 Proposition Let $p(x, y) \subseteq L(A)$. Then $\exists y p(x, y)$ is equivalent to a type over A , namely to the type $\{\exists y \varphi(x, y) : \varphi(x, y) \text{ conjunction of formulas in } p(x, y)\}$. □

As an application we give a second proof of Proposition 8.21.

Second proof of Proposition 8.21 $2 \Rightarrow 1$ Write $p(x)$ as the type $q(x; a)$ for some $q(x; z) \subseteq L$ and some $a \in \mathcal{U}^{|z|}$. Let $s(z) = \text{tp}(a / A)$. By invariance and homogeneity the types $p(x; fa)$ for $f \in \text{Aut}(\mathcal{U} / A)$ are all equivalent. Therefore

$$\begin{aligned} p(x) &\leftrightarrow \bigvee_{f \in \text{Aut}(\mathcal{U} / A)} q(x; fa) \\ p(x) &\leftrightarrow \bigvee_{b \equiv_A a} q(x; b) \\ &\leftrightarrow \exists z [s(z) \wedge q(x; z)]. \end{aligned}$$

Hence, by Proposition 8.22, $p(x)$ is equivalent to a type over A . □

8.23 Exercise Let $p(x) \subseteq L(A)$ be such that $\neg p(x) \leftrightarrow q(x)$ for some $q(x) \subseteq L(B)$. Prove that $p(x)$ is equivalent to some conjunction of formulas in $p(x)$. Cfr. Theorem 4.10. □

8.24 Exercise Let $p(x) \subseteq L(A)$, with $|x| < \omega$. Prove that if $p(\mathcal{U})$ is infinite then it has cardinality κ . Show that this may not be true if x is an infinite tuple. \square

8.25 Exercise Let $\varphi(x, y) \in L(\mathcal{U})$. Prove that if the set $\{\varphi(a, \mathcal{U}) : a \in \mathcal{U}^{|x|}\}$ is infinite then it has cardinality κ . Does the claim remains true with a type $p(x, y) \subseteq L(A)$ for $\varphi(x, y)$? \square

8.26 Exercise Let $\varphi(x, y) \in L(\mathcal{U})$. Prove that the following are equivalent

1. there is a sequence $\langle a_i : i \in \omega \rangle$ such that $\varphi(\mathcal{U}, a_i) \subset \varphi(\mathcal{U}, a_{i+1})$ for every $i < \omega$;
2. there is a sequence $\langle a_i : i \in \omega \rangle$ such that $\varphi(\mathcal{U}, a_{i+1}) \subset \varphi(\mathcal{U}, a_i)$ for every $i < \omega$. \square

Chapter 9

Preservation theorems

A few results dating from the 1950s describe the relationship between syntactic and semantic properties of first-order formulas. These theorems states that a certain syntactic class of formulas contains, up to logical equivalence, all first-order formulas preserved under a certain algebraic relationship between structures (the existence of a particular kind of homomorphism). Criteria for quantifier-elimination follow from these theorems.

1 Lyndon-Robinson Lemma

In this section T is a consistent theory without finite models and Δ is a set of formulas closed under renaming of variables. At a first reading the reader is encouraged to assume that Δ is L_{at} .

If $C \subseteq \{\forall, \exists, \neg, \vee, \wedge\}$ is a set of connectives, we write $C\Delta$ for the closure of Δ with respect to all connectives in C . We write $\neg\Delta$ for the set containing the negation of the formulas in Δ . Warning: do not confuse $\neg\Delta$ with $\{\neg\}\Delta$.

Recall that Δ -**morphism** is a map $k : M \rightarrow N$ that preserves the truth of formulas in Δ . Below we use the following proposition without further reference.

9.1 Proposition Fix $M \models T$ and $a \in M^{|x|}$. Let $q(x) = \Delta\text{-tp}_M(a)$. Then for every $\varphi(x) \in L$ the following are equivalent

1. $N \models \varphi(ka)$ for every $k : M \rightarrow N \models T$ that is a Δ -morphism defined in a ;
2. $T \vdash q(x) \rightarrow \varphi(x)$.

Proof $2 \Rightarrow 1$ Immediate.

$1 \Rightarrow 2$ Negate 2, then there are $N \models T$ and $b \in N^{|x|}$ such that $q(b) \wedge \neg\varphi(b)$. Therefore the map $k : M \rightarrow N$, where $k = \{ \langle a, b \rangle \}$, contradicts 1. \square

The following is sometimes referred to as the Lyndon-Robinson Lemma.

9.2 Lemma For every $\varphi(x) \in L$ the following are equivalent

1. $\varphi(x)$ is equivalent over T to a formula in $\{\wedge\}\Delta$;
2. $\varphi(x)$ is preserved by Δ -morphisms between models of T .

Proof $1 \Rightarrow 2$ Immediate.

$2 \Rightarrow 1$ We claim that from 2 it follows that

$$\# \quad T \vdash \varphi(x) \leftrightarrow \bigvee \{ p(x) \subseteq \Delta : T \vdash p(x) \rightarrow \varphi(x) \}$$

Direction \leftarrow is clear. To verify \rightarrow , fix $M \models T$ and let $a \in M^{|x|}$ be such that $M \models \varphi(a)$. From 2 it follows that $\varphi(x)$ satisfy 1 of Proposition 9.1. Therefore $T \vdash q(x) \rightarrow \varphi(x)$ for $q(x) = \Delta\text{-tp}_M(a)$. Hence $q(x)$ is one of the types that occur in the disjunction in #

which is satisfied by a .

From # and compactness we obtain

$$T \vdash \varphi(x) \leftrightarrow \bigvee \{ \psi(x) \in \{\wedge\}\Delta : T \vdash \psi(x) \rightarrow \varphi(x) \}$$

Applying compactness again, we replace the infinite disjunction above with a finite one and prove 2. \square

9.3 Proposition *Let N be λ -saturated and let $k : M \rightarrow N$ be a Δ -morphism of cardinality $< \lambda$. Then the following are equivalent*

1. $k : M \rightarrow N$ is a $\{\exists\wedge\}\Delta$ -morphism;
2. for every $b \in M$ there is a $\{\exists\wedge\}\Delta$ -morphism $h : M \rightarrow N$ defined in b that extends k ;
3. for every $b \in M^{<\omega}$ there is a Δ -morphism $h : M \rightarrow N$ defined in b that extends k .

Proof $1 \Rightarrow 2$ Let a enumerate $\text{dom } k$. Define $p(x, y) = \{\exists\wedge\}\Delta\text{-tp}_M(a, b)$. By 1, $p(ka, y)$ is finitely consistent in N . By saturation there is a $b' \in N$ that realizes $p(ka, y)$. Therefore, $h : M \rightarrow N$ where $h = k \cup \{\langle b, b' \rangle\}$, witnesses 2.

$2 \Rightarrow 3$ Iterate $|b|$ -times the extension in 2.

$3 \Rightarrow 1$ From 3 it follows immediately that $k : M \rightarrow N$ is a $\{\wedge\}\Delta$ -morphism. to logical equivalence, all formulas in $\{\exists\wedge\}\Delta$ are of the form $\exists y \varphi(x, y)$ for some tuple of variables y and some $\varphi(x, y)$ in $\{\wedge\}\Delta$. Assume $M \models \exists y \varphi(a, y)$ and let b be such that $M \models \varphi(a, b)$. By 3, we can extend k to some Δ -morphism $h : M \rightarrow N$ defined in b . Then $N \models \varphi(ha, hb)$ and therefore $N \models \exists y \varphi(ka, y)$. \square

Iterating the lemma above we obtain the following.

9.4 Corollary *Let N be λ -saturated and let $|M| \leq \lambda$. Let $k : M \rightarrow N$ be a Δ -morphism of cardinality $< \lambda$. Then the following are equivalent*

1. $k : M \rightarrow N$ is a $\{\exists\wedge\}\Delta$ -morphism;
2. $k : M \rightarrow N$ extends to an $\{\exists\wedge\}\Delta$ -embedding;
3. $k : M \rightarrow N$ extends to an Δ -embedding.

\square

The following theorem is often paraphrased as follows: a formula is existential if and only if (its truth) is preserved under extensions of structures.

9.5 Theorem *For every $\varphi(x) \in L$ the following are equivalent*

1. $\varphi(x)$ is equivalent over T to a formula in $\{\exists\wedge\}\Delta$;
2. $\varphi(x)$ is preserved by Δ -embedding between models of T .

Proof $1 \Rightarrow 2$ Immediate.

$2 \Rightarrow 1$ Negate 1. By Lemma 9.2 there is a $\{\exists\wedge\}\Delta$ -morphism $k : M \rightarrow N$ between models of T that does not preserve $\varphi(x)$. We can assume that N is λ -saturated for some sufficiently large λ . By Corollary 9.4 there is a Δ -embedding $h : M \rightarrow N$ that extends k and contradicts 2. \square

A dual version of the results above is obtained replacing embeddings by epimorphisms (surjective homomorphisms) and $\{\exists\wedge\}$ by $\{\forall\vee\}$. If Δ contains the formula

$x = y$ and is closed under negation, then $k : M \rightarrow N$ is a Δ -morphism if and only if $k^{-1} : N \rightarrow M$ is a Δ -morphism. In this case the dual version follows from what proved above. Without these assumptions the results need a similar but independent proof.

9.6 Proposition *Let M be λ -saturated and let $k : M \rightarrow N$ be a Δ -morphism of cardinality $< \lambda$. Then the following are equivalent*

1. $k : M \rightarrow N$ is a $\{\forall\forall\}\Delta$ -morphism;
2. for every $c \in N$ there is a $\{\forall\forall\}\Delta$ -morphism $h : M \rightarrow N$ that extends k and such that $c \in \text{img } h$;
3. for every $c \in N^{<\omega}$ there is a Δ -morphism $h : M \rightarrow N$ that extends k and such that $c \in (\text{img } h)^{<\omega}$.

Proof Left as an exercise for the reader. To prove implication $1 \Rightarrow 2$ define $p(x, y) = \neg\{\forall\forall\}\Delta\text{-tp}_N(ka, c)$, where a is a tuple that enumerates $\text{dom } k$. From 1 we obtain that $p(a, y)$ is finitely consistent in M . Then we proceed as in the proof of Proposition 9.3. \square

9.7 Corollary *Let M be λ -saturated and let $|N| \leq \lambda$. Let $k : M \rightarrow N$ be a Δ -morphism of cardinality $< \lambda$. Then the following are equivalent*

1. $k : M \rightarrow N$ is a $\{\forall\forall\}\Delta$ -morphism;
2. $k : M \rightarrow N$ extends to an $\{\forall\forall\}\Delta$ -epimorphism;
3. $k : M \rightarrow N$ extends to an Δ -epimorphism.

\square

Finally we obtain the following.

9.8 Theorem *The following are equivalent*

1. $\varphi(x)$ is equivalent to a formula in $\{\forall\forall\}\Delta$;
2. every Δ -epimorphism between models of T preserves $\varphi(x)$.

\square

2 Quantifier elimination by back-and-forth

We say that T **admits** (or **has**) **positive Δ -elimination of quantifiers** if for every formula $\varphi(x)$ in $\{\forall, \exists, \wedge, \vee\}\Delta$ there is a formula $\psi(x)$ in $\{\wedge, \vee\}\Delta$ such that

$$T \vdash \varphi(x) \leftrightarrow \psi(x).$$

When Δ is closed under negation the attribute *positive* becomes irrelevant and will be omitted. When Δ is L_{at^\pm} or L_{qf} , we simply say that T admits elimination of quantifiers. This is by far the most common case.

Quantifier elimination is often used to prove that a theory is complete because it reduces it to something much simpler to prove. The following is an immediate consequence of the definition above with x replaced by the empty tuple.

9.9 Remark *If T has elimination of quantifiers then the following are equivalent*

1. T decides all quantifier free sentences;

2. T is complete.

Hence a theory with quantifier elimination is complete if it decides the characteristic, see Definition 4.25). \square

The following is a consequence of Lemma 9.2.

9.10 Corollary *The following are equivalent*

1. T has Δ -elimination of quantifiers;
2. every Δ -morphism between models of T is both a $\{\exists, \wedge\}\Delta$ and a $\{\forall, \vee\}\Delta$ -morphism.

Proof $1 \Rightarrow 2$ Immediate.

$2 \Rightarrow 1$ We prove by induction of syntax that Δ -morphism preserve the truth of all formulas in $\{\forall, \exists, \wedge, \vee\}\Delta$, this suffices by Lemma 9.2. Induction for the connectives \vee and \wedge is trivial. So assume as induction hypothesis that the truth of $\varphi(x, y)$ is preserved. By Lemma 9.2 $\varphi(x, y)$ is equivalent to a formula in Δ , hence by 2 the truth of $\exists y \varphi(x, y)$ and $\forall y \varphi(x, y)$ is preserved. \square

Condition 2 of the corollary above may be difficult to verify directly. The following corollary of Proposition 9.3 and 9.6 gives a back-and-forth condition which is easier to verify.

9.11 Corollary *Let $|L| \leq \lambda$. The following are equivalent*

1. T has Δ -elimination of quantifiers;
 2. for every finite Δ -morphism $k : M \rightarrow N$ between λ -saturated models of T :
 - a. for every $b \in M$ there Δ -morphism $h : M \rightarrow N$ that extends k such that $b \in \text{dom } h$.
 - b. for every $c \in N$ existe un Δ -morphism $h : M \rightarrow N$ that extends k such that $c \in \text{img } h$.
- \square

Note that when Δ contains the formula $x = y$ and is closed under negation, then $k : M \rightarrow N$ is a Δ -morphism if and only if $k^{-1} : N \rightarrow M$ is a Δ -morphism. In this case a and b are equivalent.

9.12 Exercise Let L be the language of strict orders. Let T be axiomatized by T_{ol} (see Chapter 5) and the following pair of axioms

$\text{dis}\uparrow. \exists z [x < z \wedge \neg \exists y x < y < z];$

$\text{dis}\downarrow. \exists z [z < x \wedge \neg \exists y z < y < x].$

Note that T is essentially the same theory as in Exercise 8.9 but in a different language. Let Δ be the set of formulas that contains (all alphabetic variants of) the formulas $x = y$, $x < y$, $\exists^{=n} y (x < y < z)$ for all positive integers n . Prove that T has Δ -elimination of quantifiers. \square

9.13 Exercise Let T be a complete theory without finite models in a language that consists only of unary predicates. Prove that T has elimination of quantifiers. \square

Chapter 10

Geometry and dimension

In this chapter we fix a signature L , a complete theory T without finite models, and a saturated model \mathcal{U} of inaccessible cardinality κ larger than $|L|$. Notation and implicit assumptions are as presented in Section 8.3.

1 Algebraic and definable elements

Let $a \in \mathcal{U}$ and let $\mathcal{A} \subseteq \mathcal{U}$ be some set of parameters (of arbitrary cardinality). We say that a is **algebraic over \mathcal{A}** if $\varphi(a) \wedge \exists^{=k} x \varphi(x)$ holds for some formula $\varphi(x) \in L(\mathcal{A})$ and some positive integer k . In particular when $k = 1$ we say that a is **definable over \mathcal{A}** . We write $\text{acl}(\mathcal{A})$ for the **algebraic closure of \mathcal{A}** , that is, set of all elements algebraic over \mathcal{A} . If $\mathcal{A} = \text{acl}(\mathcal{A})$, we say that \mathcal{A} is **algebraically closed**. The **definable closure of \mathcal{A}** is defined similarly and is denoted by $\text{dcl}(\mathcal{A})$.

A formula $\varphi(u) \in L(A)$ or a type $p(u) \subseteq L(A)$, where u is a finite tuple, with finitely many solutions are called an **algebraic formula**, respectively an **algebraic type**.

10.1 Proposition *For every $A \subseteq \mathcal{U}$ and every type $p(u) \subseteq L(A)$, where u is a tuple of variables finite length, the following are equivalent*

- 1 $\exists^{\leq n} u p(u)$;
- 2 $\exists^{\leq n} u \varphi(u)$ for some $\varphi(u)$ which is conjunction of formulas $p(u)$.

Proof The non trivial implication is $1 \Rightarrow 2$. Assume 1, then the following type is inconsistent

$$\bigcup_{j < i \leq n} [p(u_i) \cup \{u_i \neq u_j\}]$$

By saturation there is a formula $\varphi(u)$, conjunction of formulas in $p(u)$, such that

$$\varphi(u_0) \wedge \dots \wedge \varphi(u_n) \wedge \bigwedge_{j < i \leq n} u_i \neq u_j.$$

Which is inconsistent, therefore $\varphi(u)$ has at most n solutions. □

10.2 Exercise For every $a \in \mathcal{U}^n$ and $A \subseteq \mathcal{U}$, the following are equivalent

1. a is solution of some algebraic formula $\varphi(u) \in L(A)$, with $|u| = n$;
2. $a = a_1, \dots, a_n$ for some $a_1, \dots, a_n \in \text{acl}(A)$.

Prove also that if a is definable then so are a_1, \dots, a_n , and vice versa. □

10.3 Theorem *For every $A \subseteq \mathcal{U}$ and every $a \in \mathcal{U}$ the following are equivalent*

- 1 $a \in \text{dcl}(A)$;
- 2 $\mathcal{O}(a/A) = \{a\}$.

Proof Implication $1 \Rightarrow 2$ is obvious. As for $2 \Rightarrow 1$, recall that $\mathcal{O}(a/A)$ is the set of realizations of $\text{tp}(a/A)$, then the theorem follows from Proposition 10.1. \square

10.4 Theorem For every $A \subseteq \mathcal{U}$ and every $a \in \mathcal{U}$ the following are equivalent

- 1 $a \in \text{acl}(A)$;
- 2 $\mathcal{O}(a/A)$ is finite;
- 3 a belongs to every model containing A .

Proof $1 \Leftrightarrow 2$. This is proved as in Theorem 10.3.

$1 \Rightarrow 3$. Assume 1, then there is a formula $\varphi(x) \subseteq L(A)$ such that $\varphi(a) \wedge \exists^{=k} x \varphi(x)$ for some k . By elementarity $\exists^{=k} x \varphi(x)$ holds in every model M containing A . Again by elementarity, the k solutions of $\varphi(x)$ in M are solutions in \mathcal{U} , therefore a is one of these.

$3 \Rightarrow 2$. Assume $\mathcal{O}(a/A)$ is infinite then, by Exercise 8.24, it has cardinality κ . Hence, by cardinality reasons, $\mathcal{O}(a/A) \not\subseteq M$ for every model M containing A . Pick any $f \in \text{Aut}(\mathcal{U}/A)$ such that $fa \notin M$. Then $a \notin f^{-1}[M]$, so $f^{-1}[M]$ is a models that contradicts 3. \square

10.5 Corollary For every $A \subseteq \mathcal{U}$ and every $a \in \mathcal{U}$ the following hold

- 1 Se $a \in \text{acl } A$ then $a \in \text{acl } B$ for some finite $B \subseteq A$; *finite character*
- 2 $A \subseteq \text{acl } A$; *extensivity*
- 3 se $A \subseteq B$ allora $\text{acl } A \subseteq \text{acl } B$; *monotonicity*
- 4 $\text{acl } A = \text{acl}(\text{acl } A)$; *idempotency*
- 5 $\text{acl } A = \bigcap_{A \subseteq M} M$.

Properties 1-4 say that the algebraic closure is a closure operator with finite character.

Proof Properties 1-3 are obvious, 4 follow from 5 which in turn follows from Theorem 10.4. \square

10.6 Proposition If $f \in \text{Aut}(\mathcal{U})$ then $f[\text{acl}(A)] = \text{acl}(f[A])$ for every $A \subseteq \mathcal{U}$.

Proof We prove $f[\text{acl}(A)] \subseteq \text{acl}(f[A])$. Fix $a \in \text{acl}(A)$ and let $\varphi(x; z) \in L$ and $b \in A^{|z|}$ be such that $\varphi(x; b)$ is algebraic formula satisfied by a . By elementarity, $\varphi(x; fb)$ is algebraic and satisfied by fa . Therefore fa is algebraic over $f[A]$, which proves the inclusion.

The converse inclusion is obtained by substituting f^{-1} for f and $f[A]$ for A . \square

10.7 Exercise Let $\varphi(z) \in L(A)$ be a consistent formula. Prove that, if $a \in \text{acl}(A, b)$ for every $b \models \varphi(z)$, then $a \in \text{acl}(A)$. Prove the same claim with a type $p(z) \subseteq L(A)$ for $\varphi(z)$. \square

10.8 Exercise Let $a \in \mathcal{U} \setminus \text{acl } \emptyset$. Prove that \mathcal{U} is isomorphic to some $\mathcal{U}' \preceq \mathcal{U}$ such

that $a \notin \mathcal{U}'$. Hint: let c be an enumeration of \mathcal{U} and let $p(u) = \text{tp}(c)$ prove that $p(u) \cup \{u_i \neq a : i < |c|\}$ is realized in \mathcal{U} . \square

10.9 Exercise Let C be a finite set. Prove that if $C \cap M \neq \emptyset$ for every model M containing A , then $C \cap \text{acl}(A) \neq \emptyset$. Hint: proceed by induction of the cardinality of C . Suppose there is a $c \in C \setminus \text{acl}(A)$, then there is a saturated structure $\mathcal{U}' \preceq \mathcal{U}$ such that $A \subseteq \mathcal{U}'$ and $c \notin \mathcal{U}'$. Apply the induction hypothesis to $C' = C \cap \mathcal{U}'$ with \mathcal{U}' for \mathcal{U} . \square

10.10 Exercise Prove that for every $A \subseteq N$ there is an M such that $\text{acl } A = M \cap N$.

Hint: add the requirement $\text{acl}(A_i) \cap N \subseteq \text{acl}(A)$ to the construction used to prove the downward Löwenheim-Skolem theorem. You need to prove that every consistent $\varphi(x) \in L(A_i)$ has a solution a such that $\text{acl}(A_i, a) \cap N \subseteq \text{acl}(A)$. The required a has to realize the type

$$\{\varphi(x)\} \cup \left\{ \neg[\psi(b, x) \wedge \exists^{\leq n} y \psi(y, x)] : b \in N \setminus \text{acl}(A), \psi(y, x) \in L(A_i), n \leq \omega \right\}$$

whose consistence need to be verified. \square

10.11 Exercise Let $\varphi(x) \in L(\mathcal{U})$ and fix an arbitrary set A . Prove that the following are equivalent

1. there is some model M containing A and such that $M \cap \varphi(\mathcal{U}) = \emptyset$;
2. there is no consistent formula $\psi(z_1, \dots, z_n) \in L(A)$ such that

$$\psi(z_1, \dots, z_n) \rightarrow \bigwedge_{i=1}^n \varphi(z_i).$$

Hint: let $c = \langle c_i : i < \lambda \rangle$ be an enumeration of $N^{|x|}$, where N is any model containing A . Let $p(z) = \text{tp}(c/A)$. Prove that 2 implies the consistency of the type $p(z) \cup \{\neg \varphi(z_i) : i < \lambda\}$ and deduce the existence of the required M . \square

2 Strongly minimal theories

Finite and cofinite sets are always (trivially) definable in every structure. We say that M is a **minimal structure** all its definable subsets of arity one are finite or cofinite. Unfortunately, this notion is not elementary, i.e. it is not a property of $\text{Th}(M)$. For instance \mathbb{N} with only the order relation in the language is a minimal structure but none of its elementary extensions is. Hence the following definition: we say that M is a **strongly minimal structure** if it is minimal and all its elementary extensions are minimal.

We say that T , a consistent theory without finite models, is a **strongly minimal theory** if for every formula $\varphi(x; z) \in L$, where x has arity one, there is an $n \in \omega$ tale che

$$T \vdash \forall z \left[\exists^{\leq n} x \varphi(x; z) \vee \exists^{\leq n} x \neg \varphi(x; z) \right].$$

We show that the semantic notions match with and the syntactic one.

10.12 Proposition *The following are equivalent*

1. $\text{Th}(M)$ is a strongly minimal theory;
2. M is a strongly minimal structure;
3. M has an elementary extension which is minimal and ω -saturated.

Proof Implications $1 \Rightarrow 2 \Rightarrow 3$ are immediate, we prove $3 \Rightarrow 1$. Let $\varphi(x; z) \in L$ and fix an elementary extension N as required by 2. Let $p(z) \subseteq L$ be the following type

$$p(z) = \left\{ \exists^{>n} x \varphi(x; z) \wedge \exists^{>n} x \neg \varphi(x; z) : n \in \omega \right\}.$$

As N is minimal, $N \not\models \exists z p(z)$. By saturation $p(z)$ is not finitely consistent in M . Hence, for some n

$$M \models \forall z \left[\exists^{\leq n} x \varphi(x; z) \vee \exists^{\leq n} x \neg \varphi(x; z) \right].$$

which proves that $\text{Th}(M)$ is strongly minimal. \square

By quantifier elimination, T_{acf} and T_{dag} are strongly minimal theories.

10.13 Exercise Let T be a complete theory without finite models. Prove that the following are equivalent

1. M is minimal;
2. $a \equiv_M b$ for every $a, b \in \mathcal{U} \setminus M$.

\square

3 Independence and dimension

Throughout this section we assume that T is a complete strongly minimal theory.

When $a \notin \text{acl } B$ we say that a is **algebraically independent from** B . We say that B is an **algebraically independent set** if every $a \in B$ is independent from $B \setminus \{a\}$. Below we shall abbreviate $B \cup \{a\}$ by B, a and $B \setminus \{a\}$ by $B \setminus a$.

The following is a pivotal property of independence that holds in strongly minimal structures. It is called **symmetry** or **exchange principle**. For every B and every pair of elements $a, b \in \mathcal{U} \setminus \text{acl } B$

$$b \in \text{acl}(B, a) \Leftrightarrow a \in \text{acl}(B, b)$$

Note that when $T = T_{\text{acf}}$ this principle is the so called **Steinitz exchange lemma**.

10.14 Theorem *Independence is symmetric.*

Proof Suppose $b \notin \text{acl}(B, a)$ and $a \in \text{acl}(B, b)$. We prove that $a \in \text{acl } B$. As $a \in \text{acl}(B, b)$, there is a formula $\varphi(x, y) \in L(B)$ such that

$$\varphi(a, b) \wedge \exists^{\leq n} x \varphi(x, b).$$

As $b \notin \text{acl}(B, a)$, the formula

$$\psi(a, y) = \varphi(a, y) \wedge \exists^{\leq n} x \varphi(x, y).$$

is not algebraic. Therefore, by strong minimality, $\psi(a, y)$ has cofinitely many solutions. Hence every model containing B contains a solution of $\psi(a, y)$. As a is algebraic in any of these solutions, a belongs to every model containing B . Therefore, $a \in \text{acl } B$ by Theorem 10.4. \square

We say that $B \subseteq C$ is a **base** of C if B is an independent set and $C \subseteq \text{acl } B$. The following theorem proves that all bases have the same cardinality, which we call the **dimension** of C and denote by **dim** C . First we need the following lemma.

10.15 Lemma *If B is an independent set and $a \notin \text{acl } B$ then B, a is also an independent set.*

Proof Suppose B, a is not independent and that $a \notin \text{acl } B$. Then $b \in \text{acl}(B \setminus b, a)$ for some $b \in B$. As $a, b \notin \text{acl}(B \setminus b)$, from symmetry we obtain $a \in \text{acl}(B \setminus b, b) = \text{acl } B$. Hence B is not an independent set. \square

10.16 Corollary *For every $B \subseteq C$ the following are equivalent*

1. B is a base of C .
2. B is a maximally independent subset of C .

\square

We may write **acl**(B/A) per $\text{acl}(B \cup A)$. The notation is used to suggest to the reader that the role A in the argument is irrelevant and that it could be absorbed in the language. Independence and base **over** A are defined in the obvious way.

10.17 Theorem *Fix some arbitrary set C , then*

1. every independent set $B \subseteq C$ can be extended to a base of C ;
2. all bases of C have the same cardinality;
3. claims 1 and 2 hold over any set of parameter A .

Proof By the finite character of algebraic closure, the independent set form an inductive class. Apply Zorn lemma to obtain a maximally independent subset of C containing B . By Corollary 10.16 this set is a base of C . This prove 1 and it is immediate that the argument can be generalized as required in 3.

As for 2, assume for a contradiction that $B_1, B_2 \subseteq C$ are two bases of C and that $|B_1| < |B_2|$. First consider the case when B_2 is infinite. For each $b \in B_1$ fix a set $D_b \subseteq B_2$ such that $b \in \text{acl}(D_b)$. Let

$$D = \bigcup_{b \in B_1} D_b.$$

Then $B_1 \subseteq \text{acl } D$ and $|D| < |B_2|$. By transitivity, $C \subseteq \text{acl } D$ which contradicts the independence of B_2 . Again the argument generalizes immediately as required in 3.

Now we suppose that B_2 is finite and prove directly the generalized version. Let n be the least integer such that, for some set of parameters A , and some B_1 and B_2 bases of C over A , we have $|B_1| < |B_2| = n + 1$. Fix any $c \in B_2$. Then $B_2 \setminus c$ is a base over A, c . Note that B_1 is not independent over A, c . In fact, let $D \subseteq B_1$ be minimal such that $a \in \text{acl}(D/A)$. Note that D cannot be empty and pick an arbitrary $b \in D$. By symmetry $b \notin \text{acl}(D \setminus b/A, c)$. Hence B_1 is not an independent set over A, c . Let $D_1 \subset B_1$ be maximal independent over A, c . Then D_1 and $B_2 \setminus c$ are two bases over A, c such that $|D_1| < |B_2 \setminus c| = n$, which contradicts the minimality of n . \square

10.18 Proposition *Let k be an elementary map. Then $k \cup \{\langle b, c \rangle\}$ is also an elementary map for every $b \notin \text{acl}(\text{dom } k)$ and $c \notin \text{acl}(\text{img } k)$.*

Proof Let a be an enumeration of $\text{dom } k$. We need to show that $\varphi(b; a) \leftrightarrow \varphi(c; ka)$

holds for every $\varphi(x; z) \in L$. As k is elementary, the formulas $\varphi(x; a)$ and $\varphi(x; ka)$ are either both algebraic or both co-algebraic. As $b \notin \text{acl}(a)$ and $c \notin \text{acl}(ka)$, they are both false or both true respectively. So the proposition follows. \square

10.19 Corollary *Every bijection between independent sets is an elementary map.* \square

Finally we show that dimension classifies models of T .

10.20 Theorem *Models of T with the same dimension are isomorphic.*

Proof Let A e B be some bases of M and N respectively. By Corollary 10.19, any bijection between A and B is an elementary map. By Proposition 10.6, it extends to the required isomorphism between $\text{acl } A = M$ and $\text{acl } B = N$. \square

10.21 Corollary *Strongly minimal theories are λ -categorical for every $\lambda > |L|$.*

Proof As $|\text{acl } A| \leq |L(A)|$, all models of cardinality λ have the same dimension λ . \square

10.22 Proposition *For every model N of cardinality $\geq |L|$ the following are equivalent*

1. N is saturated;
2. $\dim N = |N|$.

Proof $2 \Rightarrow 1$. Assume 2 and let $k : M \rightarrow N$ be an elementary map of cardinality $< |N|$ and let $b \in M$. We want an extension of k defined in b . If $b \in \text{acl}(\text{dom } k)$ the required extension exists by Proposition 10.6. Otherwise, pick any $c \in N \setminus \text{acl}(\text{img } k)$, which exist as $|k| < \dim N = |N|$. By Proposition 10.18, $k \cup \{ \langle b, c \rangle \}$ is the required extension.

$1 \Rightarrow 2$. If $B \subseteq N$ is a base of N the following type is not realized in N

$$p(x) = \left\{ \neg \varphi(x) : \varphi(x) \in L(B) \text{ is algebraic} \right\}$$

Therefore, if N is saturated, $|B| = |N|$. \square

10.23 Exercise Prove that every infinite algebraically closed set is a model. \square

10.24 Exercise Prove that every model is homogeneous. \square

10.25 Exercise Prove that if $\dim N = \dim M + 1$ then there is no model K such that $M \prec K \prec N$. \square

Chapter 11

Countable models

In this chapter we fix a signature L , a complete theory T without finite models, and a saturated model \mathcal{U} of inaccessible cardinality κ larger than $|L|$. (Expect that most relevant theorems are proved under the assumption that L is countable.) Notation and implicit assumptions are as presented in Section 8.3.

1 The omitting types theorem

We say that a formula $\varphi(x)$ **isolates** $p(x)$, a type, if it is consistent and $\varphi(x) \rightarrow p(x)$. When Δ is a set of formulas, we say that Δ **isolates** $p(x)$ if some formula in Δ does. In this chapter Δ is always a set of the form $L_x(A)$ and we say A **isolates** $p(x)$ or every simply $p(x)$ **is isolated** when A is clear. We say that a model M **omits** $p(x)$ if it does not realized it.

Observe that if $p(x) \subseteq L(M)$ then M realizes $p(x)$ if and only if M isolates $p(x)$. Therefore if A isolates $p(x)$, then every model containing A realizes $p(x)$. Below we prove that the converse also holds when L and A are countable. This a famous classical theorem which is called the *omitting types theorem*.

The construction of a model M omitting $p(x)$ proceeds by stages. Along the construction we preserve the property of being *isolated* since that of being *realized* is not well behaved in sets which are not models. In the limit, when a model is obtained, these two properties coincide.

The omitting types theorem require the following lemma which is an analogue of the Kuratowski-Ulam theorem (i.e., the so-called Fubini theorem for Baire category).

11.1 Lemma *Assume $L(A)$ is countable. Let $p(x) \subseteq L(A)$ and suppose that A does not isolate $p(x)$. Then every consistent formula $\psi(z) \in L(A)$ has a solution a such that A, a does not isolate $p(x)$.*

Proof We construct sequence of formulas $\langle \psi_i(z) : i < \omega \rangle$ such that any realization a of the type $q(z) = \{\psi_i(z) : i < \omega\}$ is as required by the lemma.

To begin with, we fix an enumeration $\langle \xi_i(x; z) : i < \omega \rangle$ of all formulas in $L_{x,z}(A)$ and set $\psi_0(z) = \psi(z)$. At stage $i + 1$, if $\xi_i(x; z) \wedge \psi_i(z)$ is inconsistent, let $\psi_{i+1}(z) = \psi_i(z)$, otherwise let

$$\psi_{i+1}(z) = \psi_i(z) \wedge \exists x [\xi_i(x; z) \wedge \neg \varphi(x)]$$

for some formula $\varphi(x) \in p$ such that the resulting $\psi_{i+1}(z)$ is consistent. If possible, this guarantees that $\xi_i(x; a)$ for any $a \models q(z)$ does not isolate $p(x)$. The proof is complete if we can show that it is always possible to find a formula $\varphi(x)$ as required.

Suppose by way of contradiction that no formula makes $\psi_{i+1}(z)$ consistent, that is,

$$\xi_i(x; z) \wedge \psi_i(z) \rightarrow \varphi(x)$$

for every $\varphi(\bar{x}) \in p$. This immediately imply that

$$\exists z [\xi_i(\bar{x}; z) \wedge \psi_i(z)] \rightarrow p(\bar{x}),$$

that is, $p(\bar{x})$ is isolated by a formula in $L_{\bar{x}}(A)$, which contradicts our assumptions. \square

11.2 Theorem(Omitting types theorem) Assume $L(A)$ is countable. Then for every consistent $p(\bar{x}) \subseteq L(A)$ the following are equivalent

1. all models containing A realize $p(\bar{x})$;
2. $p(\bar{x})$ is isolated.

The theorem owes its name to the contrapositive of implication $1 \Rightarrow 2$.

Proof Implication $2 \Rightarrow 1$ is clear, we prove $1 \Rightarrow 2$. Assume A does not isolate $p(\bar{x})$. The model M is the union of a chain $\langle A_i : i < \omega \rangle$ of countable subsets of \mathcal{U} that starts with $A_0 = A$. Along the construction we require as inductive hypothesis that A_i does not isolate $p(\bar{x})$. At the end, M will not isolate $p(\bar{x})$ which for a model is equivalent to omitting.

We proceed as in the proof of the downward Löwenheim-Skolem theorem (precisely, we use the second proof 2.35). Lemma 11.1 ensure that we can satisfy any consistent formula in A_i and preserve the inductive hypothesis. \square

Gerald Sacks once famously remarked: *Any fool can realize a type but it take a model theorist to omit one.* (Though, the proof reminds me more of set theory than model theory.)

2 Prime and atomic models

We say that M is **prime over A** if $A \subseteq M$ and for every N containing A there is an elementary embedding $h : M \rightarrow N$ that fixes A . When A is empty we simply say that M is **prime**.

There is no syntactic analogue of primeness; the notion that comes most close (and works pretty well when everything is countable) is atomicity. For $a \in \mathcal{U}^{|\bar{x}|}$ we say that a is **isolated over A** if the type $p(\bar{x}) = \text{tp}(a/A)$ is isolated. Note that this equivalent to claiming that a is an isolated point in $\mathcal{U}^{|\bar{x}|}$ w.r.t. the A -topology defined in Section 8.3. We say that M is **atomic over A** if $A \subseteq M$ every $a \in M^{<\omega}$ is isolated over A .

11.3 Proposition Let \bar{b} and \bar{a} be finite tuples. Then the following are equivalent

1. A isolates \bar{b}, \bar{a} ;
2. A, \bar{a} isolates \bar{b} and A isolates \bar{a} .

Proof Let $p(\bar{x}, \bar{z}) = \text{tp}(\bar{b}, \bar{a}/A)$ and note that $p(\bar{x}, \bar{a}) = \text{tp}(\bar{x}/A, \bar{a})$ and $\exists \bar{x} p(\bar{x}, \bar{z}) = \text{tp}(\bar{a}/A)$.

$1 \Rightarrow 2$ Let $\varphi(\bar{x}, \bar{z}) \in L(A)$ be such that $\varphi(\bar{x}, \bar{z}) \rightarrow p(\bar{x}, \bar{z})$. It follows that $\varphi(\bar{x}, \bar{a}) \rightarrow p(\bar{x}, \bar{a})$ and $\exists \bar{x} \varphi(\bar{x}, \bar{z}) \rightarrow \exists \bar{x} p(\bar{x}, \bar{z})$. Therefore 2 holds.

$2 \Rightarrow 1$ Fix $\varphi(\bar{x}; \bar{z}), \psi(\bar{z}) \in L(A)$ such that $\varphi(\bar{x}; \bar{a})$ isolates $p(\bar{x}; \bar{a})$ and $\psi(\bar{z})$ isolates $\exists \bar{x} p(\bar{x}; \bar{z})$. Let $\xi(\bar{x}; \bar{z}) \in p$ be arbitrary. As $\varphi(\bar{x}; \bar{a}) \rightarrow \xi(\bar{x}; \bar{a})$, the formula

$\forall x [\varphi(x; z) \rightarrow \xi(x; z)]$ belongs to $\exists x p(x, z)$. Hence $\psi(z) \rightarrow \forall x [\varphi(x; z) \rightarrow \xi(x; z)]$. As this holds for all $\xi(x; z) \in p$, we conclude that $\psi(z) \wedge \varphi(x; z)$ isolates $p(x, z)$. \square

The straightforward direction of the proposition above yields the following useful proposition.

11.4 Proposition *If M is atomic over A then M is also atomic over A, a for every finite $a \in M^{<\omega}$.*

Proof Let $b \in M^{|x|}$ be a finite tuple. Then A isolates b, a hence A, a isolates b . \square

11.5 Proposition *Let $k : M \rightarrow N$ be an elementary map and suppose that M is atomic over $\text{dom } k$. Then for every $b \in M$ there is a $c \in N$ such that $k \cup \{ \langle b, c \rangle \} : M \rightarrow N$ is elementary.*

Proof Let $p(x; z) = \text{tp}(b; a)$ where a is an enumeration of $\text{dom } k$. And let $\varphi(x; z) \in L$ be such that $\varphi(x; a) \rightarrow p(x; a)$. Note that, by elementarity, $\varphi(x; ka) \rightarrow p(x; ka)$. Hence the required c is any solution of $\varphi(x; ka)$ in N . \square

A limiting assumption in Proposition 11.4 is that a need to be finite. Therefore the following proposition is restricted to countable models.

11.6 Proposition *Any two countable models atomic over A are isomorphic.*

Proof Easy, using Propositions 11.4 and 11.5 and back-and-forth. \square

11.7 Proposition *Assume $L(A)$ is countable. Then for every model M the following are equivalent*

1. M is countable and atomic over A ;
2. M is prime over A .

Proof $1 \Rightarrow 2$ By Propositions 11.4 and 11.5.

$2 \Rightarrow 1$ Some countable model containing A exists, as M embeds in it, M has also to be countable. Now we prove that M is atomic over A . Suppose for a contradiction that there is some $b \in M^{<\omega}$ such that $p(x) = \text{tp}(b/A)$ is not isolated. By the omitting types theorem there is a model N containing A that omits $p(x)$. Then there cannot be any A -elementary embedding of M into N . \square

11.8 Proposition *Assume $L(A)$ is countable. Then the following are equivalent*

1. there are models atomic over A ;
2. for every $|z| < \omega$, every consistent $\varphi(z) \in L(A)$ has a solution that is isolated over A .

Note that 2 says that in $\mathcal{U}^{|z|}$ isolated points are dense w.r.t. the topology defined in Section 8.3.

Proof Implication $1 \Rightarrow 2$ hold by elementarity. To prove $2 \Rightarrow 1$ we construct by induction a sequence $\langle a_i : i < \omega \rangle$. Reasoning as in (the second proof of) the downward Löwerheim-Skolem theorem ensures that $A \cup \{a_i : i < \omega\}$ is a model. To obtain an atomic model we require that $a_{|i}$ is isolated over A .

Suppose $a_{|i}$ has been defined and let $\varphi(z) \in L(A)$ be the formula that isolates $\text{tp}(a_{|i}/A)$. Let $\psi(x; z) \in L(A)$ be such that $\psi(x; a_{|i})$ is consistent (we leave to the

reader the details of the enumeration of such formulas). Then $\psi(\bar{x}; \bar{z}) \wedge \varphi(\bar{z})$ is also consistent and by assumption it has a solution $\bar{b}; \bar{c}$ that is isolated over A . As $a_{|i} \equiv_A \bar{c}$, there is an A -automorphism such that $f\bar{c} = a_{|i}$. Therefore $f\bar{b}; a_{|i}$ is a solution $\psi(\bar{x}; \bar{z})$ that is also isolated over A . Then we can set $a_i = f\bar{b}$. \square

3 Countable categoricity

Here we present some important characterizations of ω -categoricity. The second property below can be stated in different equivalent ways; for convenience, these equivalents are considered in a separate proposition. For convenience we consider the following generalization: we say that T is **ω -categorical over A** if any two countable models containing A are isomorphic.

11.9 Theorem (Engeler, Ryll-Nardzewsky, and Svenonius Theorem) *Assume $L(A)$ is countable. The following are equivalent:*

1. T is ω -categorical over A ;
2. every type $p(\bar{x}) \subseteq L(A)$ with $|\bar{x}| < \omega$ is isolated.

By 3 of Proposition 11.10 below, no theory is ω -categorical over an infinite set. The reader can easily prove that that categoricity over some finite A is equivalent to categoricity over \emptyset (see Exercise 11.11).

Proof The implication $1 \Rightarrow 2$ is an immediate consequence of the omitting types theorem. In fact, if $p(\bar{x})$ is a non-isolated A -type, then there are two countable models M and N containing A such that M realizes $p(\bar{x}) \subseteq L(A)$ while N omits it. Then M and N cannot be isomorphic over A . As for implication $2 \Rightarrow 1$, observe that 2 implies that every countable model containing A is atomic over A . But, by Proposition 11.6, countable atomic models are unique up to isomorphism. \square

11.10 Proposition *Fix a set A and a finite tuple of variables \bar{x} . The following are equivalent*

1. every A -type $p(\bar{x})$ is isolated;
2. $S_{\bar{x}}(A)$ is finite;
3. $L_{\bar{x}}(A)$ is finite up to equivalence;
4. in $\mathcal{U}^{|\bar{x}|}$ there is a finite number of orbits under $\text{Aut}(\mathcal{U}/A)$.

Proof To prove the implication $1 \Rightarrow 2$ observe that $\mathcal{U}^{|\bar{x}|}$ is the union of sets of the form $p(\mathcal{U})$ where $p \in S_{\bar{x}}(A)$. If these types are isolated then $\mathcal{U}^{|\bar{x}|}$ is the union of A -definable sets. By compactness this union has to be finite. To prove $2 \Rightarrow 1$ let $p \in S_{\bar{x}}(A)$. If $S_{\bar{x}}(A)$ is finite, $\neg p(\mathcal{U})$ is the union of finitely many type definable sets. A finite union of type definable sets is type definable. So $\neg p(\mathcal{U})$ is type definable. Hence $p(\mathcal{U})$ is isolated. We prove implication $2 \Rightarrow 3$ observe that each formula in $L_{\bar{x}}(A)$ is equivalent to the disjunction of the types in $S_{\bar{x}}(A)$ that contain this formula. If $S_{\bar{x}}(A)$ is finite, $L_{\bar{x}}(A)$ is finite up to equivalence. Implication $3 \Rightarrow 2$ is clear and equivalence $2 \Leftrightarrow 4$ follows from the characterization of orbits as type-definable sets. \square

11.11 Exercise Prove that the following are equivalent for every finite set A

1. T is ω -categorical;
2. T is ω -categorical over A . □

11.12 Exercise Prove that the following are equivalent

1. T is ω -categorical;
2. there is countable model that is both saturated and atomic. □

11.13 Exercise Assume L is countable and that T is complete. Suppose that for every finite tuple \bar{x} there is a model M that realizes only finitely many types in $S_{\bar{x}}(T)$. Prove that T is ω -categorical.

4 Small theories

Let T be, as always in this chapter, a complete theory without finite models. Fix a tuple of variable \bar{x} of length ω . We say that T is **small over A** if $S_{\bar{x}}(A)$ is countable. The most and (in a sense) the only, interesting case is when A is empty, then we say that T is **small**.

11.14 Proposition *If the theory T is small then it is small over any countable A .*

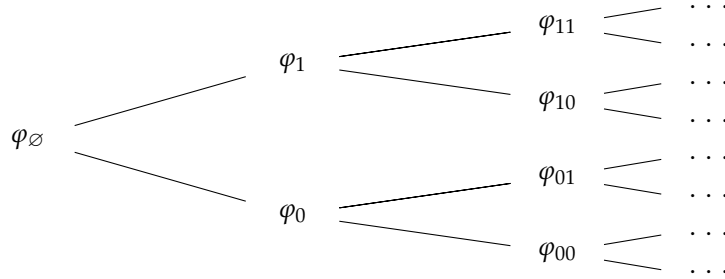
Proof Let \bar{a} be an enumeration of A . As $S_{\bar{x}}(A) = \{p(\bar{x}; \bar{a}) : p \in S_{\bar{x}; \bar{z}}(T)\}$ the proposition is immediate. □

Below we identify $S_{\bar{x}}(A)$ with $\mathcal{U}^{|\bar{x}|}/\equiv_A$

Let Δ be a set of formulas (we mainly use $\Delta = L_{\bar{x}}(A)$ in this section). A **binary tree of formulas in Δ** is a sequence $\langle \varphi_s : s \in 2^{<\omega} \rangle$ of parameter-free formulas such that

1. for each $s \in 2^\omega$ the type $p_s \stackrel{\text{def}}{=} \{\varphi_{s \upharpoonright n} : n < \omega\}$ is consistent;
2. $p_s \cup p_r$ is inconsistent for any two distinct $s, r \in 2^\omega$.

We may depict a binary tree of formulas as follows



Where branches are consistent types and distinct branches are inconsistent.

11.15 Proposition *Assume $L(A)$ is countable. The following are equivalent*

1. T is small over A ;
2. $S_{\bar{x} \upharpoonright n}(A)$ is countable for every $n < \omega$;
3. there exists a countable saturated model containing A ;

4. there is no binary tree of formulas in $L_x(A)$.

Proof $1 \Rightarrow 2$ Trivial.

$2 \Rightarrow 3$ There is a countable model M containing A that is weakly saturated (see Proposition 8.13). There is a countable homogeneous model N containing M (see Exercise 8.14). Clearly N is also weakly saturated. Then it is saturated by Corollary 8.14.

$3 \Rightarrow 4$ Clear.

$4 \Rightarrow 1$ Suppose $\mathcal{U}^{|x|}/\equiv_A$ is uncountable. We define a tree of formulas in $L_x(A)$ by induction. Let $\varphi_\emptyset(x) = \top$ and assume as induction hypothesis that $\varphi_s(\mathcal{U})/\equiv_A$ is uncountable and that $\varphi_s(x) \rightarrow \varphi_{s \upharpoonright n}(x)$ for every $n < |s|$. This will guarantee the consistency of the branches.

We claim that there is a formula $\psi(x) \in L(A)$ such that both sets $\varphi_s(\mathcal{U}) \cap \psi(\mathcal{U})/\equiv_A$ and $\varphi_s(\mathcal{U}) \cap \neg\psi(\mathcal{U})/\equiv_A$ are uncountable. If $\psi(x)$ is such a formula, we can define $\varphi_{s0}(x) = \varphi_s(x) \wedge \psi(x)$ and $\varphi_{s1}(x) = \varphi_s(x) \wedge \neg\psi(x)$ as this preserves the induction hypothesis.

First we prove that the following type is consistent

$$p(x) = \left\{ \psi(x) \in L(A) : \varphi_s(\mathcal{U}) \cap \neg\psi(\mathcal{U})/\equiv_A \text{ is countable} \right\}.$$

Suppose not, then $\neg\psi_1(x) \vee \dots \vee \neg\psi_n(x)$ holds for some $\psi_i(x) \in p$. Hence, as $\varphi_s(\mathcal{U})/\equiv_A$ is uncountable, at least one of the sets $\varphi_s(\mathcal{U}) \cap \neg\psi_i(\mathcal{U})/\equiv_A$ is uncountable, contradicting the definition of $p(x)$.

Finally note that, if the formula $\psi(x)$ required above does not exist, $p(x)$ is complete and, by the definition of $p(x)$ there are countably many types distinct from $p(x)$. This is a contradiction because we assumed $S_x(A)$ to be uncountable. \square

11.16 Proposition *A small theory has countable atomic models over every countable set A .*

Proof We prove that every formula in $L_x(A)$, where $|x| < \omega$, has a solution isolated over A . Then it suffices to apply Proposition 11.8.

Suppose for a contradiction that $\varphi(x) \in L(A)$ is consistent but has no solution isolated over A . Then there is a formula $\psi(x) \in L(A)$ such that both $\varphi(x) \wedge \psi(x)$ and $\varphi(x) \wedge \neg\psi(x)$ are consistent, otherwise $\varphi(x)$ would imply a complete type and every solution of $\varphi(x)$ would be isolated. Fix such a $\psi(x)$. Clearly neither $\varphi(x) \wedge \psi(x)$ nor $\varphi(x) \wedge \neg\psi(x)$ have a solution isolated over A . This allows to construct a tree of formulas in $L_x(A)$ and prove that T is not small over A . \square

11.17 Exercise Let $|x| = \omega$. Prove that if for all $n < \omega$ there is no binary tree of formulas in $L_{x \upharpoonright n}$ then there is no tree of formulas in L_x . \square

11.18 Exercise (Vaught) Prove that no complete theory has exactly 2 countable models (assume L is countable – though it is not really necessary).

Hint: suppose T has exactly two countable models. Then T is small and there are a countable saturated model N and an atomic model $M \subseteq N$. As T is not ω -categorical, $M \not\cong N$ and there is finite tuple a that is not isolated over \emptyset . Let K be an atomic model over a . Clearly $K \not\cong M$ and, by Exercises 11.11 and 11.12, also $K \not\cong N$. \square

5 A toy version of a theorem of Zil'ber

As an application we prove that if T is ω -categorical and strongly minimal then it is not finitely axiomatizable.

We say that T has the **finite model property** if for every sentence $\varphi \in L$ there is a finite substructure $A \subseteq \mathcal{U}$ such that

$$\text{fmp} \quad \mathcal{U} \models \varphi \Leftrightarrow A \models \varphi$$

The property is interesting because of the following proposition.

11.19 Proposition *If T has the finite model property then it is not finitely axiomatizable.*

Proof Assume fmp and suppose for a contradiction that there is a sentence $\varphi \in L$ such that $T \vdash \varphi \vdash T$. Then $A \models T$ for some finite structure A . But $T \vdash \exists^{>k} x (x = x)$ for every k . A contradiction. \square

We need the following definition. We say that $C \subseteq \mathcal{U}$ is an **homogeneous set** if for every pair of tuples $a, c \in C^{<\omega}$ such that $a \equiv c$ and for every $b \in C$ there is a $d \in C$ such that $a, b \equiv c, d$.

11.20 Lemma *Suppose L is countable. If T is ω -categorical and every finite set is contained in a finite homogeneous set, then T has the finite model property.*

Proof Assume for the moment that L is a relational language. We prove fmp also for formulas with parameters. We prove that for all n there is a finite structure $A \subseteq \mathcal{U}$ where fmp holds for all sentences $\varphi \in L(A)$ such that

$$\# \quad \text{number of parameters in } \varphi + \text{number of quantifiers in } \varphi \leq n.$$

Fix n and pick some finite set A that is homogeneous and such that all types $p(z) \subseteq L$ with $|z| \leq n$ have a realization in $A^{|z|}$. Now we prove fmp by induction on the syntax of φ .

The claim is evident for atomic formulas and the induction step is immediate for boolean connectives. As for induction step for the existential quantifier, consider the formula $\exists x \varphi(x; c)$, where $c \in A^{<n}$ and $|x| = 1$. Implication \Leftarrow of fmp follows immediately from the induction hypothesis and from the fact that, if $\exists x \varphi(x; c)$ satisfy $\#$, also $\varphi(d; c)$ satisfies it. As for \Rightarrow , assume $\mathcal{U} \models \exists x \varphi(x; c)$. Let $a, b \in A^{<\omega}$ be a solution of $\varphi(x; z)$ such that $a \equiv c$. Such a solution exists because all types with $\leq n$ variables are realized in A . By homogeneity there is a $d \in A$ such that $a, b \equiv c, d$ and therefore $A \models \varphi(d; c)$.

Finally consider the case when L contains functions symbols. Note that in fmp we may only require A to be a substructures w.r.t. symbols of L that actually occur in φ . Moreover, in $\#$ we can add the requirement that all functions symbols in φ have arity $< n$. Now note that if all types $p(z) \subseteq L$ with $|z| \leq n$ have a realization in $A^{|z|}$, then A is a substructure w.r.t. a fragment which only contains functions of arity $< n$. \square

11.21 Proposition *If T is strongly minimal, then every algebraically closed set is homogeneous.*

Proof Let A be algebraically closed and fix some $a, c \in A^{<\omega}$ such that $a \equiv c$. Let b be an element of A . Suppose first the case $b \in \text{acl } a$. Fix an $f \in \text{Aut}(\mathcal{U})$ such that $f(a) = c$. Then $d = fb$ is the required element, in fact $a, b \equiv c, d$ and $d \in \text{acl } c \subseteq A$.

Now, suppose instead that $b \notin \text{acl } a$. Then any $d \notin \text{acl } c$ satisfies $a, b \equiv c, d$. Such a d exists in A otherwise $A = \text{acl } c \neq \text{acl } a$ which contradicts $a \equiv c$. \square

From the propositions above we finally obtain the following.

11.22 Theorem *A theory which is ω -categorical and strongly minimal is not finitely axiomatizable.*

Proof If T is ω -categorical the algebraic closure of a finite set is finite. Therefore from Proposition 11.21 we infer that T satisfies the assumptions of Lemma 11.20. Hence T has the finite model property, so, by Proposition 11.19 it is not finitely axiomatizable. \square

11.23 Exercise Assume L is countable and let T be strongly minimal. Prove that the following are equivalent

1. T is ω -categorical;
2. the algebraic closure of a finite set is finite. \square

6 Notes

A famous theorem of Boris Zil'ber claims that Teorema 11.22 holds for any totally categorical theory. The same theorem has been proved independently by Cherlin, Harrington and Lachlan with a proof that uses the classification of finite simple groups. This theorem signs the birth of a subject known as *geometric stability theory* which studies in depth the geometric properties of model theories which we briefly hinted to in Chapter 10. The interested reader may consult Pillay's monograph on the subject.

[1] Anand Pillay, *Geometric stability theory*, Oxford Logic Guides, vol. 32, 1996.

Chapter 12

Definability and automorphisms

The description of first-order definability is simplified if we allow definable sets to be used as (second-order) parameters in formulas.

In this chapter we fix a signature L , a complete theory T without finite models, and a saturated model \mathcal{U} of inaccessible cardinality κ strictly larger than $|L|$. Notation and implicit assumptions are as presented in Section 8.3.

1 Many-sorted structures

A many-sorted language consists of three disjoint sets. Besides the usual L_{fun} and L_{rel} , we have a set L_{srt} of **sorts**. The language comprises also a **(many-sorted) arity function** that assigns to relations and function symbols a tuple of sorts of finite positive length

$$\text{Ar} : L_{\text{fun}} \cup L_{\text{rel}} \rightarrow L_{\text{srt}}^{<\omega} \setminus \{\emptyset\}.$$

A many sorted structure M consists of

1. For every $s \in L_{\text{srt}}$ a set M_s .
2. For every $f \in L_{\text{fun}}$ a function $f^M : M_{s_1} \times \cdots \times M_{s_n} \rightarrow M_{s_0}$, where $\text{Ar}(f) = \langle s_0, \dots, s_n \rangle$.
3. For every $r \in L_{\text{rel}}$ a relation $r^M \subseteq M_{s_0} \times \cdots \times M_{s_n}$, where $\text{Ar}(r) = \langle s_0, \dots, s_n \rangle$.

For every sort s we fix a sufficiently large a set of variables V_s . Now we define terms by induction. Along the way we assign a sort to each term.


All variables are terms of the respective sort. If t_1, \dots, t_n are terms of sorts s_1, \dots, s_n and $f \in L_{\text{fun}}$ is such that $\text{Ar}(f) = \langle s_0, \dots, s_n \rangle$ then $f t_1, \dots, t_n$ is a term of sort s_0 .

If $r \in L_{\text{rel}}$ then $\text{Ar}(r) = \langle s_0, \dots, s_n \rangle$ then $r t_0, \dots, t_n$ is a formula. Also, $t_1 = t_2$ is formula for every pair of terms. All other formulas are constructed by induction using the propositional connectives \neg and \vee and the quantifier $\exists x$ (or any other reasonable choice of logical connectives).

Truth of formulas is defined as for one-sorted languages up to one difference: we require that the witness of the quantifier $\exists x$ belongs to M_s , where s is the sort of the variable x .

Second order logic is arguably the most widely used example of many sorted structure. It may be formalized using a language with a sort n for every $n \in \omega$. The sort 0 is used for the first order elements; the sorts $n > 0$ is used for relations of arity n . For every $n > 0$ the language has a relation symbol \in_n of arity $\langle 0^n, n \rangle$, where 0^n stands for a string of 0's of length n . There are also arbitrary many functions and relation symbols of sort $\langle 0^n, 0 \rangle$ for every n .

2 The eq-expansion

 Warning: the structure \mathcal{U}^{eq} and the theory T^{eq} defined below does not coincide with the standard one. As the difference is merely cosmetic, introducing a new notation would be an overkill and we prefer to abuse the existing terminology. In Section 6 below we compare our definition with the standard one. This will also explain the symbol used.

We recall the definition of parted formulas. A **parted formula** is a triple that consists of a formula and two tuples of variables. We use a semicolon, as in $\sigma(x; z)$, to separate the two tuples of variables. Below, z is always a placeholder for parameters and x is used for the definition of a set.

Given a language L we define a many-sorted language L^{eq} which has a sort for each (parted) formula $\sigma(x; z) \in L$ and a sort **0** which we call the **home sort** or the **first-order sort**. All other sorts are called **second-order sorts**. First of all, L^{eq} contains all relations and functions of L . The many-sorted arity of a relation r is $\langle 0, \dots, 0 \rangle$, a string of 0's of length the arity of r in L . Similarly, the many-sorted arity of a function f is $\langle 0, \dots, 0 \rangle$, a string of 0's of length the arity of f in L plus 1. Moreover L^{eq} contains a relation symbol $\in_{\sigma(x; z)}$ for each sort $\sigma(x; z)$. These have arity $\langle 0, \dots, 0, \sigma(x; z) \rangle$ with $|x|$ many 0's. As there is no risk of ambiguity, we write \in without subscript.

The model \mathcal{U}^{eq} has \mathcal{U} as domain for the home sort. The domain for the sort $\sigma(x; z)$ contains the definable sets $a = \sigma(\mathcal{U}; b)$ for some $b \in \mathcal{U}^{|z|}$. The symbols in L have the same interpretation as in the one-sorted case, and \in is interpreted as set membership. We write T^{eq} for $\text{Th}(\mathcal{U}^{\text{eq}})$.

As usual L^{eq} also denotes the set of formulas constructed in this language and, if $A \subseteq \mathcal{U}^{\text{eq}}$, we write $L^{\text{eq}}(A)$ for the language and the set of formulas that use elements of A as parameters. We write $L(A)$ for the set of formulas in $L^{\text{eq}}(A)$ that contain no second-order variables (neither free nor quantified).

It is important to note right away that this expansion is a mild one: the definable subsets of the home sort of \mathcal{U}^{eq} are the same as those of \mathcal{U} . In particular, iteration of the expansion would not yield anything new.

12.1 Proposition *Let $\varphi(u; v) \in L^{\text{eq}}$ where u is a tuple of first-order variables and v a tuple of variables of sort $\sigma_i(x_i; z_i)$, with tuples of variables. Then there is a formula $\varphi'(u; z) \in L$ such that*

$$\forall v, z \left[\bigwedge_{i=1}^n v_i = \{x_i : \sigma_i(x_i; z_i)\} \rightarrow \forall u [\varphi(u; v) \leftrightarrow \varphi'(u; z)] \right]$$

where v stand for the tuple v_1, \dots, v_n and z for the concatenation of z_1, \dots, z_n . For simplicity we are assuming that the tuples z_i do not have variables in common.

When $n = 0$ the proposition asserts that L_u^{eq} and L_u have the same expressive power.

Proof By induction on syntax. When φ is atomic $\varphi' = \varphi$ unless it is of the form $t \in v$ for some tuple of terms t or it has the form $v_1 = v_2$. In the first case φ' is the formula $\sigma(t; z)$. In the second case it is the formula $\forall x [\sigma_1(x; z_1) \leftrightarrow \sigma_2(x; z_2)]$. In the latter case we assume that the variables x_1, x_2 in the sorts of v_1, v_2 have equal length otherwise we may replace $v_1 = v_2$ by \perp .

As for connectives they stay unchanged but for the quantifiers $\exists v$ where v is a second-order variable, say of sort $\sigma(x; z)$. These quantifiers are replaced with $\exists z$. \square

Proposition 12.1 implies in particular that we can always eliminate second-order quantifiers. This is an obvious fact that may be obscured by the notation: we can replace $\exists v$ by $\exists z$ if we substitute $\sigma(t; z)$ for $t \in v$ in the quantified formula.

Proposition 12.1 should convince the reader that the move from \mathcal{U} to \mathcal{U}^{eq} is almost a trivial one. Exercise 12.14 below should convince the reader that some trivial moves may be highly convenient.

12.2 Remark From Proposition 12.1 it follows that for every $A \subseteq \mathcal{U}^{\text{eq}}$, there is a $B \subseteq \mathcal{U}$ such that $L(B)$ is at least as expressive as $L(A)$. The point of \mathcal{U}^{eq} is that there might not be any $B \subseteq \mathcal{U}$ such that $L(B)$ is exactly as expressive as $L^{\text{eq}}(A)$. For instance, suppose L contains only an equivalence relation with infinitely many classes, all with infinite elements. Let a be an equivalence class. Then a is definable in $L(B)$ if and only if $B \cap a \neq \emptyset$. But no element of a is definable in $L^{\text{eq}}(\{a\})$. \square

If $\mathcal{V} \preceq \mathcal{U}$ we write \mathcal{V}^{eq} for the substructure of \mathcal{U}^{eq} that has \mathcal{V} as domain of the home sort and the set of definable sets of the form $\sigma(\mathcal{U}; b)$ for some $b \in \mathcal{V}^{|z|}$ as domain of the sort $\sigma(x; z)$. The following proposition claims that elementary substructures of \mathcal{U}^{eq} are exactly those of the form \mathcal{V}^{eq} for some $\mathcal{V} \preceq \mathcal{U}$.

12.3 Proposition *The following are equivalent for every structure \mathcal{V}^{\dagger} of signature L^{eq}*

1. $\mathcal{V}^{\dagger} \preceq \mathcal{U}^{\text{eq}}$;
2. $\mathcal{V}^{\dagger} = \mathcal{V}^{\text{eq}}$ for some $\mathcal{V} \preceq \mathcal{U}$.

Proof Implication $2 \Rightarrow 1$ is a direct consequence of Proposition 12.1. We prove $1 \Rightarrow 2$. Let \mathcal{V} be the domain of the home sort of \mathcal{V}^{\dagger} . It is clear that $\mathcal{V} \preceq \mathcal{U}$. Let $a \in \mathcal{V}^{\text{eq}}$ have sort $\sigma(x; z)$, say $a = \sigma(\mathcal{U}; b)$ for some $b \in \mathcal{V}^{|z|}$. As $\exists^1 v \forall x [x \in v \leftrightarrow \sigma(x; b)]$ holds in \mathcal{U}^{eq} , by elementarity it holds in \mathcal{V}^{\dagger} and therefore $a \in \mathcal{V}^{\dagger}$. This proves $\mathcal{V}^{\text{eq}} \subseteq \mathcal{V}^{\dagger}$. A similar argument proves the converse inclusion. \square

12.4 Proposition *Let $A \subseteq \mathcal{U}^{\text{eq}}$. Then every type $p(u; v) \subseteq L^{\text{eq}}(A)$ finitely consistent in \mathcal{U}^{eq} is realized in \mathcal{U}^{eq} . That is, \mathcal{U}^{eq} is saturated.*

Proof By Remark 12.2, there is some $B \subseteq \mathcal{U}$ and some $q(u; v) \subseteq L^{\text{eq}}(B)$ equivalent to $p(u; v)$. This already proves the proposition when v is the empty tuple. Otherwise, let $q'(u; z)$ be obtained replacing every formula $\varphi(u; v)$ in $q(u; v)$ with the formula $\varphi'(u; z)$ given by Proposition 12.1. Then $q'(u; z)$ is finitely consistent in \mathcal{U} . Assume for clarity of notation that v is a single variable of sort $\sigma(x; z)$. If $c; b \models q'(u; z)$ then $c; \sigma(\mathcal{U}; b) \models q(u; v)$. \square

Automorphisms of a many-sorted structure are defined in the obvious way: sorts are preserved as well as functions and relations. Every automorphism $f : \mathcal{U} \rightarrow \mathcal{U}$ extends to an automorphism $f : \mathcal{U}^{\text{eq}} \rightarrow \mathcal{U}^{\text{eq}}$ as follows. If $a = \sigma(\mathcal{U}; b)$ we define $fa = \sigma(\mathcal{U}; fb)$, which clearly preserves the sort. It is easy to check that $fa = \{fx : x \in a\}$, so f preserves the relation \in . Clearly, this extension is unique.

The homogeneity of \mathcal{U}^{eq} follows by back-and-forth as in the one-sorted case.

12.5 Proposition *Every elementary map $k : \mathcal{U}^{\text{eq}} \rightarrow \mathcal{U}^{\text{eq}}$ of cardinality $< \kappa$ extends to an*

automorphism of \mathcal{U}^{eq} . □

3 The definable closure in the eq-expansion

We may safely confuse automorphism of \mathcal{U} with automorphisms of \mathcal{U}^{eq} . Let $A \subseteq \mathcal{U}^{\text{eq}}$ and let a be a tuple of elements of \mathcal{U}^{eq} . We denote by $\text{Aut}(\mathcal{U}/A)$ the set of automorphisms (of \mathcal{U}^{eq}) that fix all elements of A . The symbol $\mathcal{O}(a/A)$ denotes the **orbit of a over A** . This has been defined in Section 2 and now we apply it to \mathcal{U}^{eq}

$$\mathcal{O}(a/A) \stackrel{\text{def}}{=} \{fa : f \in \text{Aut}(\mathcal{U}/A)\}.$$

By homogeneity, $\mathcal{O}(a/A) = p(\mathcal{U}^{\text{eq}})$ where $p(v) = \text{tp}(a/A)$. When $\mathcal{O}(a/A) = \{a\}$ we say that a is **invariant over A** or **A -invariant**, for short.

12.6 Definition Let $A \subseteq \mathcal{U}^{\text{eq}}$ and $a \in \mathcal{U}^{\text{eq}}$. We say that **a is definable over A** if $\varphi(a) \wedge \exists^{=1} v \varphi(v)$ holds for some formula $\varphi(v) \in L^{\text{eq}}(A)$. We write $\text{dcl}^{\text{eq}}(A)$ for the set of those $a \in \mathcal{U}^{\text{eq}}$ that are definable over A . We write $\text{dcl}(A)$ for $\text{dcl}^{\text{eq}}(A) \cap \mathcal{U}$. This is the natural generalization of the notion introduced in Section 1.

Saturation and homogeneity of \mathcal{U}^{eq} allow us to prove the following proposition with virtually the same proof of Theorem 10.3

12.7 Theorem For any $A \subseteq \mathcal{U}^{\text{eq}}$ and $a \in \mathcal{U}^{\text{eq}}$ the following are equivalent

1. a is invariant over A ;
2. $a \in \text{dcl}^{\text{eq}}(A)$. □

Note that a tuple is invariant if and only if all its components are definable. Then in Theorem 12.7 we may replace a and/or A by a tuple.

Definition 12.6 treats first- and second-order elements of \mathcal{U}^{eq} uniformly. However, for sets it can be rephrased in a more natural way.

12.8 Proposition Let $A \subseteq \mathcal{U}^{\text{eq}}$ and let $a \in \mathcal{U}^{\text{eq}}$ have sort $\sigma(x; z)$ then the following are equivalent

1. a is definable over A
2. $a = \psi(\mathcal{U})$ for some $\psi(x) \in L(A)$

Proof Implication $2 \Rightarrow 1$ is clear because extensionality is implicit in the definition of \mathcal{U}^{eq} . We prove $1 \Rightarrow 2$. Let $\varphi(v) \in L^{\text{eq}}(A)$ be a formula a is the unique solution of. Then the required formula $\psi(x)$ is $\exists v [x \in v \wedge \varphi(v)]$. □

By the characterization above, Theorem 12.7 when applied to a definable set a gives an alternative proof of Proposition 8.19.

The formula $\psi(x)$ in Proposition 12.8 need not be related to the sort $\sigma(x; z)$. For example, consider the theory of a binary equivalence relation $e(x; z)$ with two infinite classes, let a be one of these classes and let $A \neq \emptyset$ be such that $A \cap a = \emptyset$. Then a is definable over A though not by some formula of the form $e(x; b)$ for some $b \in A$. Things change radically if we replace A with a model.

12.9 Proposition Let M be a model and let a be an element of sort $\sigma(x; z)$. Then the following are equivalent

1. $a \in \text{dcl}^{\text{eq}}(M)$;
2. $a = \sigma(\mathcal{U}, b)$ for some $b \in M^{|z|}$.

In particular $M^{\text{eq}} = \text{dcl}^{\text{eq}}(M)$.

Proof Assume 1 and let $\psi(x) \in L(M)$ be such that $a = \psi(\mathcal{U})$. Such a formula exists by Proposition 12.8. Then $\exists z \forall x [\psi(x) \leftrightarrow \sigma(x; z)]$ holds in \mathcal{U}^{eq} . By elementarity it holds in M^{eq} , therefore $\exists z$ has a witness in M . This proves $1 \Rightarrow 2$, the converse implication is obvious. \square

4 The algebraic closure in the eq-expansion

Let $A \subseteq \mathcal{U}^{\text{eq}}$ and $a \in \mathcal{U}^{\text{eq}}$. We say that a is **algebraic over A** if $\varphi(a) \wedge \exists^{=k} v \varphi(v)$ holds for some formula $\varphi(v) \in L^{\text{eq}}(A)$ and some positive integer k . We write $\text{acl}^{\text{eq}}(A)$ for the set of those $a \in \mathcal{U}^{\text{eq}}$ that are algebraic over A . We write $\text{acl}(A)$ for $\text{acl}^{\text{eq}}(A) \cap \mathcal{U}$. This is the natural generalization of the notion introduced in Section 10.1.

The following proposition is proved with virtually the same proof of Theorem 10.4

12.10 Theorem For every $A \subseteq \mathcal{U}^{\text{eq}}$ and every $a \in \mathcal{U}^{\text{eq}}$ the following are equivalent

1. $\mathcal{O}(a/A)$ is finite;
2. $a \in \text{acl}^{\text{eq}}(A)$;
3. $a \in M^{\text{eq}}$ for every model such that $A \subseteq M^{\text{eq}}$.

\square

We say **finite equivalence relation** for an equivalence relation with finitely many classes. A **finite equivalence formula** or **type** is a formula, respectively a type, that defines a finite equivalence relation. Theorem 12.11 belows proves that sets algebraic over A are union of classes of a finite equivalence relations definable over A .

12.11 Theorem Let $A \subseteq \mathcal{U}^{\text{eq}}$ and let $a \in \mathcal{U}^{\text{eq}}$ be an element of sort $\sigma(x; z)$. Then the following are equivalent

1. $a \in \text{acl}^{\text{eq}}(A)$
2. for some finite equivalence formula $\varepsilon(x; y) \in L(A)$ and some $c_1, \dots, c_n \in \mathcal{U}^{|y|}$

$$x \in a \leftrightarrow \bigvee_{i=1}^n \varepsilon(x; c_i).$$

Proof Let $\varphi(v) \in L^{\text{eq}}(A)$ be an algebraic formula that has a among its solutions and define

$$\varepsilon(x; y) = \forall v \left[\varphi(v) \rightarrow [x \in v \leftrightarrow y \in v] \right]$$

If $\varphi(v)$ has n solutions, then $\varepsilon(x; y)$ has at most 2^n equivalence classes. Clearly, a is union of some these classes. \square

We write $a \stackrel{\text{Sh}}{\equiv}_A b$ for $a \equiv_{\text{acl}^{\text{eq}} A} b$. In words we say that a and b have the same

Shelah strong-type over A .

12.12 Proposition Let $A \subseteq \mathcal{U}^{\text{eq}}$ and let $a, b \in \mathcal{U}^{|x|}$. Then the following are equivalent

1. $a \stackrel{\text{Sh}}{=}_A b$;
2. $\varepsilon(a; b)$ for every finite equivalence formula $\varepsilon(x; y) \in L(A)$.

Proof Assume $\neg 2$ and let $\varepsilon(x; y)$ witness this. Let $d = \varepsilon(\mathcal{U}; b)$, then $b \in d$ and $a \notin d$. As $\varepsilon(x; y)$ is an A -invariant finite equivalence formula, $d \in \text{acl}^{\text{eq}}(A)$, and $\neg 1$ follows. Vice versa, assume $\neg 1$ and let $\varphi(x) \in L(\text{acl}^{\text{eq}}(A))$ be such that $\varphi(a) \not\leftrightarrow \varphi(b)$. Let $d = \varphi(\mathcal{U})$, then $d \in \text{acl}^{\text{eq}}(A)$. Therefore, by Proposition 12.11, the set d is union of equivalence classes of some finite equivalence formula $\varepsilon(x; y) \in L(A)$. Then $\neg \varepsilon(a; b)$ and $\neg 2$ follows. \square

12.13 Exercise Let $p(x) \subseteq L(A)$ and let $\varphi(x; y) \in L(A)$ be a formula that defines, when restricted to $p(\mathcal{U})$, an equivalence relation with finitely many classes. Prove that there is a finite equivalence relation definable over A that coincides with $\varphi(x; y)$ on $p(\mathcal{U})$. \square

12.14 Exercise Let $A \subseteq \mathcal{U}$, let a be a definable set with finite orbit over A . Without using \mathcal{U}^{eq} prove that a is union of classes of a finite equivalence relation definable over A . \square

12.15 Exercise Let T be strongly minimal and let $\varphi(x; y) \in L(A)$ with $|z| = 1$. For arbitrary $c \in \mathcal{U}$, prove that if the orbit of $\varphi(\mathcal{U}; c)$ over A is finite, then $\varphi(\mathcal{U}; c)$ is definable over $\text{acl } A$.

Hint: you can use Theorem 12.11. \square

5 Elimination of imaginaries

For the time being, we agree that *imaginary* is just another word for definable set. Though this is not formally correct (cfr. Section 6), it is morally true and helps to understand the terminology. The concept of elimination of imaginaries has been introduced by Poizat who also proved Theorem 12.22 below. A theory has elimination of imaginaries if for every $A \subseteq \mathcal{U}^{\text{eq}}$, there is a $B \subseteq \mathcal{U}$ such that $L(B)$ and $L(A)$ have the same expressive power (i.e. they are the same up to equivalence).

12.16 Definition We say that T has *elimination of imaginaries* if for every definable set a there is a formula $\varphi(x; z) \in L$ such that

$$ei \quad \exists^{=1} z \forall x \left[x \in a \leftrightarrow \varphi(x; z) \right]$$

We say that the witness of $\exists^{=1} z$ in the formula above is a *canonical parameter* of a . It is also called a *canonical name* for a . A set may have different canonical parameters for different formulas $\varphi(x; z)$.

We say that T has *weak* elimination of imaginaries if

$$wei \quad \exists^{=k} z \forall x \left[x \in a \leftrightarrow \varphi(x; z) \right]$$

for some positive integer k . □

In the formulas above we allow z to be the empty string. In this case we read ei and wei omitting the quantifiers $\exists^=1 z$, respectively $\exists^=k z$. Therefore \emptyset -definable sets have all (at least) the empty string as a canonical parameter.

To show that the notions above are well-defined properties of a theory one needs to check that they are independent of our choice of monster model. We leave this to the reader as an exercise.

12.17 Exercise Let M be an arbitrary model. Prove that ei and wei hold (in \mathcal{U}^{eq}) if and only if they hold in M^{eq} . □

We say that two tuples a and b of elements of \mathcal{U}^{eq} are **interdefinable** if $\text{dcl}^{\text{eq}}(a) = \text{dcl}^{\text{eq}}(b)$. By Theorem 12.7 this is equivalent to saying that $\text{Aut}(\mathcal{U}/a) = \text{Aut}(\mathcal{U}/b)$, that is, the automorphisms that fix a fix also b , and vice versa.

12.18 Theorem *The following are equivalent*

1. T has weak elimination of imaginaries;
2. every definable set is interdefinable with a finite set;
3. every definable set a is definable over $\text{acl}\{a\}$.

Proof $1 \Rightarrow 2$ Assume 1 and let b be the set of solutions of the formula

$$\forall x \left[x \in a \leftrightarrow \sigma(x; z) \right].$$

Hence b is finite and $b \in \text{dcl}^{\text{eq}}\{a\}$. We also have $a \in \text{dcl}^{\text{eq}}\{b\}$ because a is definable by the formula $\exists z [z \in b \wedge \sigma(x; z)]$. Therefore $\text{dcl}^{\text{eq}}\{a\} = \text{dcl}^{\text{eq}}\{b\}$.

$2 \Rightarrow 3$ Assume $\text{dcl}^{\text{eq}}\{a\} = \text{dcl}^{\text{eq}}\{b\}$ for some finite set b . The elements of b , say b_1, \dots, b_n , are the (finitely many) solutions of the formula $z \in b$. Therefore $b_1, \dots, b_n \in \text{acl}(b) = \text{acl}(a)$. Let $\varphi(v; b)$ be a formula that has a as unique solution. Then $\exists u [u = \{b_1, \dots, b_n\} \wedge \varphi(v; u)]$ is the formula that proves 3.

$3 \Rightarrow 1$ Assume 3. As wei holds trivially for all \emptyset -definable sets, we may assume $a \neq \emptyset$. Let $\sigma(x; z)$ be such that

$$\forall x \left[x \in a \leftrightarrow \sigma(x; b) \right].$$

for some tuple b of elements of $\text{acl}\{a\}$. Fix some algebraic formula $\delta(z; a)$ satisfied by b and write $\psi(z; v)$ for the formula

$$\forall x \left[x \in v \leftrightarrow \sigma(x; z) \right] \wedge \delta(z; v).$$

The following formula is clearly satisfied by b therefore, if we can prove that it has finitely many solutions, wei follows from Proposition 12.1

$$\# \quad \forall x \left[x \in a \leftrightarrow \sigma(x; z) \wedge \exists v \psi(z; v) \right].$$

We check that any c that satisfies $\#$ also satisfies $\delta(z; a)$. As a is non empty, $\psi(c; a')$ holds for some a' . By $\#$ and the definition of $\psi(z; v)$ we obtain $a' = a$ and $\delta(c; a)$. □

12.19 Theorem *The following are equivalent*

1. T has elimination of imaginaries;
2. every definable set is interdefinable with a tuple of real elements;
3. every definable set a is definable over $\text{dcl}\{a\}$.

Proof Implications $1 \Rightarrow 2 \Rightarrow 3$ are immediate. Implication $3 \Rightarrow 1$ is identical to the homologous implication in Theorem 12.18, just substitute algebraic with definable. \square

We now consider elimination of imaginaries in two concrete structures: algebraically closed fields and real closed fields. The following lemma is required in the proof of Theorem 12.21 below.

12.20 Lemma *The following is a sufficient condition for weak elimination of imaginaries*

\sharp for every $A \subseteq \mathcal{U}^{\text{eq}}$, every consistent $\varphi(z) \in L(A)$ has a solution in $\text{acl } A$.

Proof Let $a \in \mathcal{U}^{\text{eq}}$ be a definable set and let $\sigma(x; z)$ be its sort. Then $\forall x [x \in a \leftrightarrow \sigma(x; z)]$ is consistent and, by \sharp it has a solution in $\text{acl}\{a\}$. Hence weak elimination follows from Theorem 12.18. \square

12.21 Theorem *Let T be a complete, strongly minimal theory. Then, if $\text{acl } \emptyset$ is infinite, T has weak elimination of imaginaries.*

Proof If $\text{acl } \emptyset$ is infinite, $\text{acl } A$ is a model for every A (cfr. Exercise 10.23) so condition \sharp of lemma 12.20 holds by elementarity and the theorem follows. \square

12.22 Theorem *The theories T_{acf}^p have elimination of imaginaries.*

Proof By Theorem 12.21 we know that T_{acf}^p has weak elimination of imaginaries. Therefore, by Theorem 12.18 it suffices to prove that every finite set a is interdefinable with a tuple. Let $a = \{a_1, \dots, a_n\}$ where each a_i is a tuple $a_{i,1}, \dots, a_{i,m}$ of elements of \mathcal{U} . Given a we define the term

$$t_a(x; y) = \prod_{i=1}^n \left(x - \sum_{k=1}^m a_{i,k} y_k \right). \quad \text{where } y = y_1, \dots, y_m.$$

Note that (the interpretation of) the term $t_a(x; y)$ is independent on particular indexing of the set a . So, any automorphism that fixes a , fixes the $t_a(x; y)$. Now rewrite $t_a(x; y)$ as a sum of monomials and let c be the tuple of coefficients of these monomials. The tuple c uniquely determines $t_a(x; y)$ and vice versa. Therefore every automorphism that fixes a fixes c and vice versa. Hence a and c are interdefinable. \square

12.23 Exercise Let T have elimination of imaginaries and $\varphi(x; z) \in L(A)$. For arbitrary $c \in \mathcal{U}^{|z|}$, prove that if the orbit of $\varphi(\mathcal{U}; c)$ over A is finite, then $\varphi(\mathcal{U}; c)$ is definable over $\text{acl } A$. \square

6 Imaginaries, the true story

The point of the expansion to \mathcal{U}^{eq} is to add a canonical parameter for each definable set. In fact, in \mathcal{U}^{eq} every definable subset of $\mathcal{U}^{|z|}$ is the canonical parameter if itself. This allows to deal with theories without elimination of imaginaries in the most straightforward way.

The expansion to \mathcal{U}^{eq} that was originally introduced by Shelah (and still used everywhere else) is slightly different from the one introduced here. For a given set $a = \sigma(\mathcal{U}; b)$ Shelah considers the equivalence relation defined by the formula

$$\varepsilon(z; z') \stackrel{\text{def}}{=} \forall x \left[\sigma(x; z) \leftrightarrow \sigma(x; z') \right].$$

The equivalence class of b in the relation $\varepsilon(z; z')$ is what Shelah uses as canonical parameter of the set a .

Shelah's \mathcal{U}^{eq} has a sort for each \emptyset -definable equivalence relation $\varepsilon(z; z')$. The domain of the sort $\varepsilon(z; z')$ contains the classes of the equivalence relation defined by $\varepsilon(z; z')$. These equivalence classes are called **imaginaries**. Shelah's L^{eq} contains functions that map tuples in the home sort to their equivalence class.

Tortuous routes are somewhat typical of Shelah. This did not hinder him from becoming one of the most productive mathematician of his century. Lesser minds may find it useful to take a simplified route.

7 Uniform elimination of imaginaries

Let us rephrase the definition of elimination of imaginaries: for every formula $\varphi(x; u) \in L$ and every tuple $c \in \mathcal{U}^{|u|}$ there is a formula $\sigma(x; z) \in L$ such that

$$\exists^1 z \forall x \left[\varphi(x; c) \leftrightarrow \sigma(x; z) \right].$$

A priori, the formula $\sigma(x; z)$ may depend on c in a very wild manner. We say that **T has uniform elimination of imaginaries** if for every $\varphi(x; u)$ there are a formula $\sigma(x; z)$ and a formula $\rho(z)$ such that

$$\text{uei} \quad \forall u \exists^1 z \left[\rho(z) \wedge \forall x \left[\varphi(x; u) \leftrightarrow \sigma(x; z) \right] \right].$$

The role of the formula $\rho(z)$ above is may seem mysterious. It is clarified by the following propositions. In fact, uniform elimination of imaginaries is equivalent to a very natural property which in words says: every definable function is the kernel of a definable function. Recall that the kernel of the function f is the relation $fa = fb$.

Uniform elimination of imaginaries is convenient when dealing with interpretations of structure inside other structures. Let us consider a simple concrete example. Suppose \mathcal{U} is a group and let \mathcal{H} be a definable normal subgroup of \mathcal{U} . The elements of the quotient structure \mathcal{U}/\mathcal{H} are equivalent classes of a definable equivalence relation. If there is uniform elimination of imaginaries we can identify \mathcal{U}/\mathcal{H} with an actual definable subset of \mathcal{G} (the range of the function f above). Moreover, the group operation of \mathcal{G} are definable functions. As working in \mathcal{U}/\mathcal{H} may be notationally cumbersome, \mathcal{G} may offer a convenient alternative.

12.24 Proposition *The following are equivalent*

1. *T has uniform elimination of imaginaries;*
2. *for every $\varphi(x; u)$ such that $\forall u \exists x \varphi(x; u)$ there is a formula $\sigma(x; z)$ such that*

$$\forall u \exists^{=1} z \forall x \left[\varphi(x; u) \leftrightarrow \sigma(x; z) \right];$$

3. for every equivalence formula $\varepsilon(x; u) \in L$ there is given by $\sigma(x; z) \in L$ such that

$$\forall u \exists^{=1} z \forall x \left[\varepsilon(x; u) \leftrightarrow \sigma(x; z) \right].$$

Note that the definable function $fu = z$ mentioned above is defined by the formula

$$\vartheta(u; z) = \forall x \left[\varepsilon(x; u) \leftrightarrow \sigma(x; z) \right].$$

Proof $1 \Rightarrow 2$. If $\forall u \exists x \varphi(x; u)$ we can rewrite uei as

$$\forall u \exists^{=1} z \forall x \left[\varphi(x; u) \leftrightarrow \rho(z) \wedge \sigma(x; z) \right].$$

$2 \Rightarrow 3$. Clear.

$3 \Rightarrow 1$. Apply 3 to the equivalence formula

$$\varepsilon(u; v) = \forall x \left[\varphi(x; u) \leftrightarrow \varphi(x; v) \right]$$

Let $\vartheta(u; z)$ be defined as above. The reader may check that uei holds substituting for $\delta(z)$ the formula $\exists u \vartheta(u; z)$ and for $\sigma(x; z)$ the formula $\exists u [\varphi(x; u) \wedge \vartheta(u; z)]$. \square

By the following theorem, uniformity comes almost for free.

12.25 Theorem *The following are equivalent*

1. T has uniform elimination of imaginaries;
2. $\text{dcl } \emptyset$ contains at least two elements and T has elimination of imaginaries.

Proof $1 \Rightarrow 2$. Let $\varphi(x, u)$ be the formula $u_1 = u_2$. From 1 we obtain a formula $\sigma(x; z) \in L$ such that

$$\forall u_1, u_2 \exists^{=1} z \left[\rho(z) \wedge \forall x [u_1 = u_2 \leftrightarrow \sigma(x; z)] \right].$$

Therefore the formulas $\exists^{=1} z [\rho(z) \wedge \forall x \sigma(x; z)]$ and $\exists^{=1} z [\rho(z) \wedge \forall x \neg \sigma(x; z)]$ are both true. The witnesses of $\exists^{=1} z$ in these two formulas are two distinct elements of $\text{dcl } \emptyset$.

$2 \Rightarrow 1$. Assume 2 and fix a formula $\varphi(x; u)$ such that $\varphi(x, a)$ is consistent for every $a \in \mathcal{U}^{[u]}$. We prove 2 of Proposition 12.24.

Let $p(u)$ be the type that contains the formulas

$$\neg \exists^{=1} z \forall x \left[\varphi(x; u) \leftrightarrow \sigma(x; z) \right],$$

where $\sigma(x; z)$ ranges over all formulas in L . By elimination of imaginaries $p(u)$ is not consistent. Therefore, by compactness, there are some formulas $\sigma_i(x; z)$ such that

$$\# \quad \forall u \bigvee_{i=0}^n \exists^{=1} z \forall x \left[\varphi(x; u) \leftrightarrow \sigma_i(x; z) \right].$$

To prove the theorem we need to move the disjunction in front of the $\sigma_i(x; z)$.

We can assume that if $\sigma_i(\mathbf{x}; \mathbf{b}) \leftrightarrow \sigma'_i(\mathbf{x}; \mathbf{b}')$ for some $\langle \mathbf{b}, i \rangle \neq \langle \mathbf{b}', i' \rangle$ then $\sigma_i(\mathbf{x}; \mathbf{b})$ is inconsistent. Otherwise we can substitute the formula $\sigma_i(\mathbf{x}; \mathbf{z})$ with

$$\sigma_i(\mathbf{x}; \mathbf{z}) \wedge \bigwedge_{j \leq i} \neg \exists \mathbf{y} \neq \mathbf{b} \forall \mathbf{x} \left[\sigma_j(\mathbf{x}; \mathbf{y}) \leftrightarrow \sigma_i(\mathbf{x}; \mathbf{z}) \right].$$

As $\varphi(\mathbf{x}, a)$ is consistent for every a , the substitution does not break the validity of \sharp .

Fix some distinct \emptyset -definable tuples d_0, \dots, d_n of the same length (these are easy to obtain from two \emptyset -definable elements). We claim that from \sharp it follows that

$$\forall u \exists^{=1} \mathbf{z}, \mathbf{y} \forall \mathbf{x} \left[\varphi(\mathbf{x}; u) \leftrightarrow \bigvee_{i=1}^n \left[\sigma_i(\mathbf{x}; \mathbf{z}) \wedge \mathbf{y} = d_i \right] \right].$$

(The tuple \mathbf{z}, \mathbf{y} plays the role of \mathbf{z} .) We fix some a and check that the formula below has a unique solution

$$\mathfrak{b} \quad \forall \mathbf{x} \left[\varphi(\mathbf{x}; a) \leftrightarrow \bigvee_{i=1}^n \left[\sigma_i(\mathbf{x}; \mathbf{z}) \wedge \mathbf{y} = d_i \right] \right].$$

Existence follows immediately from \sharp . As for uniqueness, note that if \mathbf{b}, d_i and $\mathbf{b}', d_{i'}$ are two distinct solution of \mathfrak{b} then $\sigma_i(\mathbf{x}; \mathbf{b}) \leftrightarrow \sigma_{i'}(\mathbf{x}; \mathbf{b}')$ for some $\langle \mathbf{b}, i \rangle \neq \langle \mathbf{b}', i' \rangle$. By what assumed on $\sigma_i(\mathbf{x}; \mathbf{z})$, we obtain that $\varphi(\mathbf{x}; a)$ is inconsistent. A contradiction which proves the theorem. \square

Chapter 13

Invariant sets

In this chapter we fix a signature L , a complete theory T without finite models, and a saturated model \mathcal{U} of inaccessible cardinality κ strictly larger than $|L|$. Notation and implicit assumptions are as presented in Section 8.3.

1 Invariant sets and types

Let $\mathcal{D} \subseteq \mathcal{U}^{|z|}$, where z is a tuple of length $< \kappa$. We say that \mathcal{D} is a **set invariant over A** , or **A -invariant** if it is fixed (set-wise) by A -automorphisms. That is, $f[\mathcal{D}] = \mathcal{D}$ for every $f \in \text{Aut}(\mathcal{U}/A)$ or, yet in other words,

$$\text{is1.} \quad a \in \mathcal{D} \leftrightarrow fa \in \mathcal{D} \quad \text{for every } a \in \mathcal{U}^{|z|} \text{ and every } f \in \text{Aut}(\mathcal{U}/A),$$

which, by homogeneity is equivalent to,

$$\text{is2.} \quad a \in \mathcal{D} \leftrightarrow b \in \mathcal{D} \quad \text{for all } a, b \in \mathcal{U}^{|z|} \text{ such that } a \equiv_A b.$$

which yield the following bound on the number of invariant sets

13.1 Proposition *Let $\lambda = |L_z(A)|$. There are at most 2^{2^λ} sets $\mathcal{D} \subseteq \mathcal{U}^{|z|}$ that are invariant over A .*

Proof By is2, sets that are invariant over A are union of equivalence classes of the relation \equiv_A , that is, union of sets of the form $p(\mathcal{U})$ where $p(z) \in S(A)$. Then the number of A -invariant sets is $2^{|S_z(A)|}$. Clearly $|S_z(A)| \leq 2^\lambda$. \square

We say that \mathcal{D} is **invariant** tout court if it is invariant over some A . As we require κ to be inaccessible, there are exactly κ invariant sets.

Let $p(x) \subseteq L(\mathcal{U})$ be a consistent type. For every formula $\varphi(x; z) \in L$ we define

$$\mathcal{D}_{p,\varphi} \stackrel{\text{def}}{=} \{a \in \mathcal{U}^{|z|} : \varphi(x; a) \in p\}$$

We can read the notation in two ways. Either the tuple z has infinite length and is the same for all formulas, or it is finite and depends on φ . In most cases we opt for the first reading, then $\mathcal{D}_{p,\varphi}$ is a cylinder set (independent from all but finitely many coordinates).

Let $p(x) \subseteq L(\mathcal{U})$ be a consistent type. We say that **$p(x)$ is invariant over A** , or **A -invariant**, if for every formula $\varphi(x; z) \in L$,

$$\text{it1.} \quad \varphi(x; a) \in p \leftrightarrow \varphi(x; fa) \in p \quad \text{for every } a \in \mathcal{U}^{|z|} \text{ and every } f \in \text{Aut}(\mathcal{U}/A).$$

In other words, $p(x)$ is invariant over A if the sets $\mathcal{D}_{p,\varphi}$ are all invariant over A .

One should not confuse the notions of *invariant type* with that of *invariant type-definable set*. The first is mostly applied to global types and refers to the invariance of a family of subsets of $\mathcal{U}^{|z|}$. The latter refers to sets of the form $p(\mathcal{U}) \subseteq \mathcal{U}^{|x|}$ where $p(x)$ a type of small size.

We say that $p(x)$ **does not split over A** if

$$\text{it2.} \quad a \equiv_A b \Rightarrow \left(\varphi(x; a) \in p \Leftrightarrow \varphi(x; b) \in p \right) \quad \text{for all } a, b \in \mathcal{U}^{|z|}$$

for every formula $\varphi(x; z) \in L$. When $p(x) \in S(\mathcal{U})$ is a global type, it2 is equivalent to

$$\text{it2'}. \quad a \equiv_A b \Rightarrow \varphi(x; a) \leftrightarrow \varphi(x; b) \in p$$

By homogeneity, non splitting is equivalent to invariance. Yet another important equivalent of invariance over A is:

$$\text{it3.} \quad a \equiv_A b \Rightarrow a \equiv_{A, c} b \quad \text{for all } a, b \in \mathcal{U}^{|z|} \text{ and for all } c \models p|_{A, a, b}.$$

The easiest way to obtain invariant types is via types that are finitely satisfiable in some small set. We say that a type $p(x)$ is **finitely satisfiable in A** if $\varphi(\mathcal{U}) \cap A^{|x|} \neq \emptyset$ for every $\varphi(x)$ conjunction of formulas in p .

Note that by elementarity every type over a model is finitely satisfiable in the model, that is, if $p(x) \subseteq L(M)$ for some model M , then $p(x)$ is finitely satisfiable in M .

13.2 Proposition *Let $p(x) \in S(\mathcal{U})$ be a global type finitely satisfiable in A . Then $p(x)$ is A -invariant.*

Proof Recall that the invariance of $p(x)$ is equivalent to it2'. So, suppose $a \equiv_A b$. If for a contradiction $p(x)$ does not contain the formula $\varphi(x; a) \leftrightarrow \varphi(x; b)$, by completeness, it contains $\varphi(x; a) \nleftrightarrow \varphi(x; b)$. But $p(x)$ finitely satisfiable in A , so $\varphi(c, a) \nleftrightarrow \varphi(c, b)$ holds for some $c \in A^{|x|}$ contradicting $a \equiv_A b$. \square

13.3 Proposition *Every type $q(x) \subseteq L(\mathcal{U})$ finitely satisfiable in A has an extension to a global type finitely satisfiable in A .*

Proof Let $p(x) \subseteq L(\mathcal{U})$ be maximal among the types containing $q(x)$ and finitely satisfiable in A . We check that $p(x)$ is complete. Suppose for a contradiction that $p(x)$ contains neither $\psi(x)$ nor $\neg\psi(x)$. Then neither $p(x) \cup \{\psi(x)\}$ nor $p(x) \cup \{\neg\psi(x)\}$ is finitely satisfiable in A . This contradicts the finite satisfiability of $p(x)$. \square

When $q(x)$ finitely satisfiable in A , any type that extends $q(x)$ and is also finitely satisfiable in A is called a **coheir** of $q(x)$ (a somewhat confusing terminology).

2 Morley sequences and indiscernibles

In the following α is some infinite ordinal $\leq \kappa$ and x is a tuple of variables of length $< \kappa$. Let $p(x) \in S(\mathcal{U})$ be a global type invariant over A . We say that $c = \langle c_i : i < \alpha \rangle$ is a **Morley sequence of $p(x)$ over A** if for every $i < \alpha$

$$\text{Ms.} \quad c_i \models p|_{A, c_{\bar{i}}}(x)$$

In particular when $p(x)$ is finitely satisfiable in A we will say that c is a **coheir sequence**. We may say that c is a Morley (or coheir) sequence without specifying a global type. Then we mean that c is a Morley (respectively coheir) sequence of some invariant global type.

The following is a convenient characterization of coheir sequences

13.4 Lemma *The following are equivalent*

1. $c = \langle c_n : n < \alpha \rangle$ is a coheir sequence over A ;
2. $c_{n+1} \equiv_{A, c_{\upharpoonright n}} c_n$ and $\varphi(\mathcal{U}) \cap A^{|\mathbf{x}|} \neq \emptyset$ whenever $\varphi(\mathbf{x}) \in L(A, c_{\upharpoonright n})$ and $\varphi(c_n)$.

Proof $1 \Rightarrow 2$. Assume 1 and let $p(\mathbf{x}) \in S(\mathcal{U})$ be a global type finitely satisfiable in A and such that $c_i \models p_{\upharpoonright A, c_{\upharpoonright i}}(\mathbf{x})$. The requirement $c_{n+1} \equiv_{A, c_{\upharpoonright n}} c_n$ is clear. Now, suppose $\varphi(c_n)$ for some $\varphi(\mathbf{x}) \in L(A, c_{\upharpoonright n})$. Then $\varphi(\mathbf{x})$ belongs to $p(\mathbf{x})$ hence $\varphi(\mathcal{U}) \cap A^{|\mathbf{x}|} \neq \emptyset$ because $p(\mathbf{x})$ is finitely satisfiable in A .

$2 \Rightarrow 1$. Assume 2 and define

$$q(\mathbf{x}) = \{ \varphi(\mathbf{x}) : c_n \models \varphi(\mathbf{x}) \in L(A, c_{\upharpoonright n}) \text{ for some } n < \alpha \}.$$

Clearly c is a Morley sequence of any global type finitely satisfiable in A that extends $q(\mathbf{x})$. Hence it suffices to show that such type exists. So, we prove that $q(\mathbf{x})$ is finitely satisfiable in A and apply Proposition 13.3.

As 2 requires that every single formula in $q(\mathbf{x})$ is finitely satisfiable in A it suffices to show that $q(\mathbf{x})$ is closed under conjunctions. Let $\varphi_i(\mathbf{x}) \in L(A, c_{\upharpoonright n_i})$, for $i = 1, 2$, be two formulas in $q(\mathbf{x})$ satisfied by c_{n_i} . By the first requirement in 2, both formulas are satisfied by c_m where $m = \max\{n_1, n_2\}$. Hence their conjunction is in $q(\mathbf{x})$. \square

Let $I, <_I$ be a linear order. We call a function $a : I \rightarrow \mathcal{U}^{|\mathbf{x}|}$ an **I -sequence**, or simply a **sequence** when I is clear. Often we introduce an I -sequence by writing $a = \langle a_i : i \in I \rangle$.

If $I_0 \subseteq I$ we call $a_{\upharpoonright I_0}$ a **subsequence of a** . Most relevant are the subsets $I_0 \subseteq I$ that are well-ordered by $<_I$ and, in particular, the finite ones. When I_0 has order-type α , an ordinal, we identify $a_{\upharpoonright I_0}$ with a tuple of length α .

13.5 Definition Let $I, <_I$ be an infinite linear order and let a be an I -sequence. We say that **a is a sequence of indiscernibles over A** or, a sequence of **A -indiscernibles**, if $a_{\upharpoonright I_0} \equiv_A a_{\upharpoonright I_1}$ for every $I_0, I_1 \in [I]^n$ and $n < \omega$. As usual, $[I]^n$ is the set of subsets of I of cardinality n . \square

The condition above can be formulated in a number of clearly equivalent ways. We can require that that $\varphi(a_{i_0}, \dots, a_{i_n}) \leftrightarrow \varphi(a_{j_0}, \dots, a_{j_n})$ for every formula $\varphi(x_1, \dots, x_n) \in L(A)$ and every pairs of tuples in I^n such that $i_0 < \dots < i_n$ and $j_0 < \dots < j_n$. Or we can simply say that for every $i_0, \dots, i_n \in I$ the type $\text{tp}(a_{i_0}, \dots, a_{i_n} / A)$ only depends on the order type of i_0, \dots, i_n .

13.6 Proposition Let $p(\mathbf{x}) \in S(\mathcal{U})$ be a global A -invariant type and let $c = \langle c_i : i < \alpha \rangle$ be a Morley sequence of $p(\mathbf{x})$ over A . Then c is a sequence of indiscernibles over A .

Proof We prove by induction on $n < \omega$ that

$$\# \quad c_{\upharpoonright n} \equiv_A c_{\upharpoonright I_0} \text{ for every } I_0 \subseteq \alpha \text{ of cardinality } n.$$

For $n = 0$ the claim is trivial, then we assume $\#$ above is true and prove that

$$c_{\upharpoonright n}, c_n \equiv_A c_{\upharpoonright I_0}, c_i \text{ for every } I_0 < i < \alpha.$$

As c is Morley sequence, $c_n \equiv_{A, c_{\upharpoonright n}} c_i$ whenever $n < i$. Hence we can equivalently prove that

$$c_{\upharpoonright n}, c_i \equiv_A c_{\upharpoonright I_0}, c_i,$$

which is equivalent to

$$c_{\upharpoonright n} \equiv_{A, c_i} c_{\upharpoonright I_0}.$$

The latter holds by induction hypothesis \sharp and the invariance of $p(\mathbf{x})$ as formulated in it3 of Section 13. \square

3 From coheirs to Ramsey to indiscernibles

Let X be an infinite set and $n, k < \omega$. A total function $f : [X]^n \rightarrow k$ is also called a **coloring** of $[X]^n$ and is identified with a partition of $[X]^n$ into finitely k classes. A set $H \subseteq X$ such that f is constant on $[H]^n$ is called **monochromatic set**. In the literature it may be called a **homogeneous set**, but we already use this word with a different meaning. The following is a most famous theorem which here we prove it in unusual way.

13.7 Proposition (Ramsey Theorem) *For every infinite set X , every $n, k < \omega$ and every coloring $f : [X]^n \rightarrow k$ there is an infinite monochromatic set.*

Proof Let L be a language that contains k relation symbols r_0, \dots, r_{k-1} of arity n . Given $f : [X]^n \rightarrow k$ as above we define now a structure with domain X . The interpretation the relation symbols is

$$r_i^X = \{a_0, \dots, a_{n-1} : \{a_0, \dots, a_{n-1}\} \in [X]^n \text{ and } f(\{a_0, \dots, a_{n-1}\}) = i\}.$$

Then

$$\# \quad X \models \forall x_0, \dots, x_{n-1} \left[\bigwedge_{0 \leq i < j < n} x_i \neq x_j \rightarrow \bigvee_{i < k} r_i(x_0, \dots, x_{n-1}) \right].$$

We may assume that X is an elementary substructure of some large saturated model \mathcal{U} . Pick any $p(\mathbf{x}) \in S(\mathcal{U})$ finitely satisfied in X , not realized in X and let $c = \langle c_i : i < \omega \rangle$ be a coheir sequence of $p(\mathbf{x})$. By $\#$ and indiscernibility, all tuples of n distinct elements of c satisfy the same relation, which we assume to be r_0 . The theorem is proved if we can find in X a sequence $a = \langle a_i : i < \omega \rangle$ with the same property. (Note parenthetically that no element of c is in X .) We construct a_i by induction on i as follows.

Assume as induction hypothesis that all subsequences of $a_{\upharpoonright i}, c_{\upharpoonright n}$ of length n satisfy r_0 . Our goal is to find a_i such as the same property holds for $a_{\upharpoonright i}, a_i, c_{\upharpoonright n}$. By the indiscernibility of c , the property holds for $a_{\upharpoonright i}, c_{\upharpoonright n}, c_n$. And this can be written by a formula $\varphi(a_{\upharpoonright i}, c_{\upharpoonright n}, c_n)$. As c is a coheir sequence, by Lemma 13.4 we can find $a_i \in X$ such that $\varphi(a_{\upharpoonright i}, c_{\upharpoonright n}, a_i)$. So, as the order is irrelevant, $a_{\upharpoonright i}, a_i, c_{\upharpoonright n}$ satisfies the induction hypothesis. \square

Let $I, <_I$ be an infinite linear order and let a be an I -sequence. Fix a sequence of distinct variables $x = \langle x_i : i < \omega \rangle$. We write $p(\mathbf{x}) = \text{EM-tp}(a/A)$ and say that $p(\mathbf{x})$ is the **Ehrenfeucht-Mostowski type of a over A** if $p(\mathbf{x})$ is the set of formulas $\varphi(x_{\upharpoonright n}) \in L(A)$ such that $\varphi(a_{\upharpoonright I_0})$ holds for every $I_0 \in [I]^n$. Note that if a is A -indiscernible then $\text{EM-tp}(a/A)$ is a complete type, and vice versa. Moreover, if a and b are two A -indiscernible I -sequences with the same Ehrenfeucht-Mostowski type of a over A , then $a \equiv_A b$.

13.8 Theorem(Ehrenfeucht-Mostowski) *Let $I, <_I$ and $J, <_J$ be two infinite linear orders such that $|J| \leq \kappa$. Then for every sequence $a = \langle a_i : i \in I \rangle$ there is an J -sequence of A -indiscernibles c such that $\text{EM-tp}(a/A) \subseteq \text{EM-tp}(c/A)$.*

Proof To begin with, we prove the theorem assuming that both I and J have order type ω . Then any realization of the following type yields a J -sequence of A -indiscernibles.

$$q(x) \cup \left\{ \varphi(x_{\upharpoonright n}) \leftrightarrow \varphi(x_{\upharpoonright J_0}) : \varphi(x_{\upharpoonright n}) \in L(A), J_0 \in [\omega]^n, n < \omega \right\}.$$

We will prove that any finite subset of the type above is realized by a subsequence of a . First note that any infinite subsequence of a realizes $q(x)$, so we only need to pay attention to the set on the right. We prove that for every $\varphi_0, \dots, \varphi_{m-1} \in L_x(A)$ and every $n < \omega$ there is an infinite $H \subseteq I$ such that $a_{\upharpoonright H}$

$$\# \quad \left\{ \varphi_i(x_{\upharpoonright n}) \leftrightarrow \varphi_i(x_{\upharpoonright I_0}) : I_0 \in [\omega]^n \text{ and } i < m \right\}.$$

Let $f : [I]^n \rightarrow \mathcal{P}(m)$ map $I_0 \in [I]^n$ to the set $\{i < m : \varphi_i(a_{\upharpoonright I_0})\}$. By Ramsey theorem with 2^m colors there is some infinite monochromatic set $H \subseteq I$. As I has order type ω , so does H , and it is immediate that $a_{\upharpoonright H}$ realizes $\#$, as required to prove the theorem with I and J of order type ω .

Now we note that the argument above easily generalizes to arbitrary linear order I . Only, in this case we cannot guarantee that H has order type ω . However, as H is infinite, for every $k < \omega$ there is a any subset of H of size k that realizes the restriction of $\#$ to $n < k$ and $I_0 < k$. This suffices to prove the theorem with I arbitrary and J of order type ω .

Now we extend the result to arbitrary J of cardinality $\leq \kappa$. From the argument above, there is a sequence of indiscernibles $\langle b_i : i < \omega \rangle$ such that $\text{EM-tp}(a/A) \subseteq \text{EM-tp}(b/A)$. Fix a J -sequence of variables $z = \langle z_i : i \in J \rangle$ and define

$$p(z) = \left\{ \varphi(z_{\upharpoonright J_0}) \in L(A) : \varphi(b_{\upharpoonright n}) \text{ for some } J_0 \subseteq [J]^n, n < \omega \right\}.$$

This type is clearly finitely consistent. Every c that realizes $q(z)$ is a J -sequence as required by the theorem. \square

13.9 Proposition Let $c = \langle c_i : i \in I \rangle$ be a sequence of A -indiscernibles. Then c is indiscernible over some model M containing A .

Proof Fix an arbitrary model M containing A . By Theorem 13.8 there is an I -sequence of M -indiscernibles b such that $\text{EM-tp}(c/M) \subseteq \text{EM-tp}(b/M)$. As c is an A -indiscernible sequence $a \equiv_A b$, so some $h \in \text{Aut}(\mathcal{U}/A)$ takes b to c coordinatewise. Therefore $hb = c$ is indiscernible over $h[M]$. \square

13.10 Exercise Let $a = \langle a_i : i \in I \rangle$ be an A -indiscernible sequence and let $I \subseteq J$ with $|J| \leq \kappa$. Then there is an A -indiscernible sequence $c = \langle c_i : i \in J \rangle$ such that $c_{\upharpoonright I} = a$. \square

13.11 Exercise Let $a \equiv_A b$ and define $q(x) = \text{tp}(a/A) = \text{tp}(b/A)$. Prove that if $q(x)$ extends to a global A -invariant type $p(x) \in S(\mathcal{U})$ then there is a sequence $c = \langle c_i : i < \omega \rangle$ such that a, c and b, c are both sequences of A -indiscernibles. \square

4 Expansions

Given $\mathcal{D} \subseteq \mathcal{U}^{|z|}$, we will occasionally find it convenient to expand the language L with a predicate for $z \in \mathcal{D}$. We denote by $L(\mathcal{D})$ such expansion. We denote

by $\langle \mathcal{U}, \mathcal{D} \rangle$ the corresponding expansion of \mathcal{U} . If $\mathcal{C}, \mathcal{D} \subseteq \mathcal{U}^{|\mathcal{z}|}$ write $\mathcal{C} \equiv_A \mathcal{D}$ for $\langle \mathcal{U}, \mathcal{C} \rangle \equiv_A \langle \mathcal{U}, \mathcal{D} \rangle$. As usual, $L(\mathcal{D})$ denotes also the set of formulas in that language, and we write $L(A; \mathcal{D})$ for the set of formulas with parameters from A .

When \mathcal{z} finite tuple, it is straightforward how to define such an expansions. But when \mathcal{z} is an infinite tuple, though the intuition stays essentially the same, a more involved definition is required because first-order logic does not allow infinitary predicates. We think of $L(\mathcal{D})$ for a two sorted language. We call the two sorts the home sort and the \mathcal{z} -sort. The canonical expansion of \mathcal{U} to an $L(\mathcal{D})$ -structure is denoted by $\langle \mathcal{U}, \mathcal{D} \rangle$. The domain of the home sort is \mathcal{U} , the domain of the \mathcal{z} -sort is $\mathcal{U}^{|\mathcal{z}|}$. Besides the symbols of L , there is a function symbol π_i for every $i < |\mathcal{z}|$. These functions have many-sorted arity $\langle \mathcal{z}, 0 \rangle$ (see Section 12.1 for the notation) which are interpreted as the projection to the i -coordinate. There is also a predicate of many-sorted arity $\langle \mathcal{z} \rangle$ interpreted as \mathcal{D} .

The definitions above are generalized in the obvious way if for \mathcal{D} we take a tuple $\langle \mathcal{D}_i : i < \lambda \rangle$ of subsets of $\mathcal{U}^{|\mathcal{z}_i|}$.

13.12 Proposition *If $\mathcal{D} \subseteq \mathcal{U}^{|\mathcal{z}|}$ is invariant over A then every A -indiscernible sequence is indiscernible in the language $L(A; \mathcal{D})$.*

Proof Let $c = \langle c_i : i \in I \rangle$ be an A -indiscernible sequence. For every $I_0, I_1 \in I^{[n]}$ there is an $f \in \text{Aut}(\mathcal{U}/A)$ such that $fc_{\upharpoonright I_0} = c_{\upharpoonright I_1}$. Hence

$$\varphi(c_{\upharpoonright I_0}; \mathcal{D}) \leftrightarrow \varphi(fc_{\upharpoonright I_0}; f[\mathcal{D}]) \leftrightarrow \varphi(c_{\upharpoonright I_1}; \mathcal{D}). \quad \square$$

13.13 Exercise Prove that if $\mathcal{C} \subseteq \mathcal{U}^{|\mathcal{z}|}$ is type-definable over B then $\equiv_{A,B}$ implies $\equiv_{A;\mathcal{C}}$. \square

Chapter 14

Lascar invariant sets

In this chapter we fix a signature L , a complete theory T without finite models, and a saturated model \mathcal{U} of inaccessible cardinality κ strictly larger than $|L|$. Notation and implicit assumptions are as presented in Section 8.3.

1 Lascar strong types

Let $\mathcal{D} \subseteq \mathcal{U}^{|\mathbf{z}|}$, where \mathbf{z} is a tuple of length $< \kappa$. The **orbit of \mathcal{D} over A** is the set

$$o(\mathcal{D}/A) \stackrel{\text{def}}{=} \left\{ f[\mathcal{D}] : f \in \text{Aut}(\mathcal{U}/A) \right\}$$

So, \mathcal{D} is invariant over A when $o(\mathcal{D}/A) = \{\mathcal{D}\}$. We say that **\mathcal{D} is Lascar invariant over A** if it is invariant over every model $M \supseteq A$. Recall that this means that if $\mathbf{a} \equiv_M \mathbf{c}$ for some model M containing A then $\mathbf{a} \in \mathcal{D} \leftrightarrow \mathbf{c} \in \mathcal{D}$.

14.1 Proposition *Let $\lambda = |L_{\mathbf{z}}(A)|$. There are at most 2^{2^λ} sets $\mathcal{D} \subseteq \mathcal{U}^{|\mathbf{z}|}$ that are Lascar invariant over A .*

Proof Let N be a model containing A of cardinality $\leq \lambda$. Every set that is Lascar invariant over A is invariant over N . As $|L_{\mathbf{z}}(N)| = \lambda$ the bound follows from Proposition 13.1. \square

14.2 Theorem *For every \mathcal{D} and every $A \subseteq M$ the following are equivalent:*

1. \mathcal{D} is Lascar invariant over A ;
2. every set in $o(\mathcal{D}/A)$ is M -invariant;
3. $o(\mathcal{D}/A)$ has cardinality $\leq 2^{2^{|L(A)|}}$;
4. $o(\mathcal{D}/A)$ has cardinality $< \kappa$;
5. $\mathbf{c}_0 \in \mathcal{D} \leftrightarrow \mathbf{c}_1 \in \mathcal{D}$ for every A -indiscernible sequence $\langle \mathbf{c}_i : i < \omega \rangle$.

Proof The implication $1 \Rightarrow 2$ is clear because all sets in $o(\mathcal{D}/A)$ are Lascar invariant over A . When $|M| \leq |L(A)|$ implication $2 \Rightarrow 3$ follows from the bounds discussed in Section 13. So, we may temporarily add this assumption on M as, once the proof of the proposition is finished, it is easily seen to be redundant. Implication $3 \Rightarrow 4$ holds because κ is strong limit.

We now prove $4 \Rightarrow 5$. Assume $\neg 5$. Then we can find an A -indiscernible sequence $\langle \mathbf{c}_i : i < \kappa \rangle$ such that $\mathbf{c}_0 \in \mathcal{D} \not\leftrightarrow \mathbf{c}_1 \in \mathcal{D}$. Define

$$E(\mathbf{u}; \mathbf{v}) \Leftrightarrow \mathbf{u} \in \mathcal{C} \leftrightarrow \mathbf{v} \in \mathcal{C} \text{ for every } \mathcal{C} \in o(\mathcal{D}/A).$$

Then $E(\mathbf{u}; \mathbf{v})$ is an A -invariant equivalence relation. As $\neg E(\mathbf{c}_0; \mathbf{c}_1)$, indiscernibility over A implies that $\neg E(\mathbf{c}_i; \mathbf{c}_j)$ for every $i < j < \kappa$. Then $E(\mathbf{u}; \mathbf{v})$ has κ equivalence classes. As κ is inaccessible, this implies $\neg 4$.


We prove $5 \Rightarrow 1$. Assume 5 and fix any $a \equiv_M b$ for some model M containing A . It suffices to prove that $a \in \mathcal{D} \leftrightarrow b \in \mathcal{D}$. Let $p(z) \in S(\mathcal{U})$ be a global coheir of $\text{tp}(a/M) = \text{tp}(b/M)$. Let $c = \langle c_i : i < \omega \rangle$ be a Morley sequence of $p(z)$ over M, a, b . Then both a, c and b, c are A -indiscernible sequences. So from 5 we obtain $a \in \mathcal{D} \leftrightarrow c_0 \in \mathcal{D} \leftrightarrow b \in \mathcal{D}$. \square

The following corollary easily follows from Theorem 14.2 and Proposition 13.12. As an exercise the reader may wish to prove it using Proposition 13.9.

14.3 Corollary *The following are equivalent*

1. \mathcal{D} is Lascar invariant over A ;
2. every sequence of A -indiscernibles is $L(A; \mathcal{D})$ -indiscernible. \square

Given a tuple $a \in \mathcal{U}^{|z|}$, we write $\mathcal{L}(a/A)$ for the intersection of all sets containing a that are Lascar invariant over A . Clearly $\mathcal{L}(a/A)$ is Lascar invariant over A . We say that a and b have the same **Lascar strong type over A** if $a \in \mathcal{D} \leftrightarrow b \in \mathcal{D}$ for every set \mathcal{D} that is Lascar invariant over A or, in other words, if $\mathcal{L}(a/A) = \mathcal{L}(b/A)$. The symbol used is $a \equiv_A^L b$.

 **Warning:** the symbol $\mathcal{L}(a/A)$ is not standard. The relation $a \equiv_A^L b$ is sometime denoted by $a E_{L/A} b$ or by $\text{L-stp}(a/A) = \text{L-stp}(b/A)$.

14.4 Proposition *The relation \equiv_A^L is the finest equivalence relation with $< \kappa$ classes that is invariant over A . \square*

Proof Clearly \equiv_A^L is an equivalence relation invariant over A . Each equivalence class is Lascar invariant over A , hence the number of equivalence classes is bounded by the number of Lascar invariant sets over A . To see that \equiv_A^L is the finest of such equivalences. Suppose \mathcal{D} is an equivalence class of an A -invariant equivalence relation with $< \kappa$ classes. Then $\mathcal{o}(\mathcal{D}/A)$ has also cardinality $< \kappa$. Then \mathcal{D} is Lascar invariant and as such is union of classes of the relation \equiv_A^L . \square

Let $p(x) \in S(\mathcal{U})$ be global type; we say that **p is Lascar invariant over A** if the sets $\mathcal{D}_{p,q}$ defined in Section 13 are all Lascar invariant over A .

14.5 Proposition *Let $p(x) \in S(\mathcal{U})$ be a global type. Then the following are equivalent*

1. $p(x)$ is Lascar invariant over A ;
2. every A -indiscernible sequence $c = \langle c_i : i < \omega \rangle$ is A, a -indiscernible for every $a \models p|_c$.

Here we either assume $|z| \geq \omega$ or in 2 we quantify over all $|z| < \omega$.

Proof $2 \Rightarrow 1$ If $p(x)$ is not Lascar invariant over A then $c_0 \in \mathcal{D}_{p,q} \not\leftrightarrow c_1 \in \mathcal{D}_{p,q}$ for some A -indiscernible sequence $c = \langle c_i : i < \omega \rangle$ and some $\varphi(x; z) \in L$. Then $p(x)$ contains the formula $\varphi(x; c_0) \not\leftrightarrow \varphi(x; c_1)$. Hence, c is not indiscernible over any realization of $p(x)|_{c_0, c_1}$.

$1 \Rightarrow 2$ Assume 1 and fix an A -indiscernible sequence $c = \langle c_i : i < \omega \rangle$ and some $a \models p|_c$. We need to prove that for every formula $\varphi(x; z') \in L$, where $z' = z_1, \dots, z_n$,

$$\varphi(a; c_0, \dots, c_{n-1}) \leftrightarrow \varphi(a; c_{i_1}, \dots, c_{i_{n-1}}).$$

holds for every $i_0 < \dots < i_{n-1} < \omega$. Suppose not and let m be any integer larger than i_{n-1} . Then the following equivalences cannot both be true

$$\begin{aligned} \varphi(a; c_m, \dots, c_{m+n-1}) &\leftrightarrow \varphi(a; c_0, \dots, c_{n-1}); \\ \varphi(a; c_m, \dots, c_{m+n-1}) &\leftrightarrow \varphi(a; c_{i_0}, \dots, c_{i_{n-1}}). \end{aligned}$$

If the first is false, define $c'_k = c_{km}, \dots, c_{km+n-1}$ for all $k < \omega$. Otherwise, do this only for positive k and set $c'_0 = c_{i_0}, \dots, c_{i_{n-1}}$. In either cases $\langle c'_k : k < \omega \rangle$ is a sequence of A -indiscernibles and $c'_0 \in \mathcal{D}_{p,\varphi} \not\leftrightarrow c'_1 \in \mathcal{D}_{p,\varphi}$. This contradicts 1. \square

2 The Lascar graph and Newelski's theorem

Here we study Lascar strong types from a different viewpoint. The **Lascar graph over A** has an arc between all pairs $a, b \in \mathcal{U}^{|\mathcal{Z}|}$ such that $a \equiv_M b$ for some model M containing A . We write $d_A(a, b)$ for the distance between a and b in the Lascar graph over A . Let us spell this out: $d_A(a, b) \leq n$ if there is a sequence a_0, \dots, a_n such that $a = a_0$, $b = a_n$, and $a_i \equiv_{M_i} a_{i+1}$ for some models M_i containing A . We write $d_A(a, b) < \infty$ if a and b are in the same connected component of the Lascar graph over A .

14.6 Proposition For every $a \in \mathcal{U}^{|\mathcal{Z}|}$

$$\mathcal{L}(a/A) = \{c : d_A(a, c) < \infty\}.$$

Proof To prove inclusion \supseteq it suffices to show that every Lascar A -invariant set containing a contains the set on the r.h.s. Let \mathcal{D} be Lascar A -invariant, and let $b \in \mathcal{D}$. Then \mathcal{D} contains also every c such that $b \equiv_M c$ for some model M containing A . That is, \mathcal{D} contains every c such that $d_A(b, c) \leq 1$. It follows that \mathcal{D} contains every c such that $d_A(a, c) < \infty$.

To prove inclusion \subseteq we prove the set on the r.h.s. is Lascar A -invariant. Suppose the sequence a_0, \dots, a_n , where $a_0 = a$ and $a_n = c$, witnesses $d_A(a, c) \leq n$ and suppose that $c \equiv_M b$ for some M containing A , then the sequence a_0, \dots, a_n, b witnesses $d_A(a, b) \leq n + 1$. \square

We write $\text{Autf}(\mathcal{U}/A)$ for the subgroup of $\text{Aut}(\mathcal{U}/A)$ that is generated by the automorphisms that fix point-wise some model M containing A . (The “f” in the symbol stands for *fort*, the French word for *strong*.) It is easy to verify that $\text{Autf}(\mathcal{U}/A)$ is a normal subgroup of $\text{Aut}(\mathcal{U}/A)$. The following is a corollary of Proposition 14.6.

14.7 Corollary The following are equivalent

1. $a \stackrel{\text{L}}{\equiv}_A b$
2. $a = fb$ for some $f \in \text{Autf}(\mathcal{U}/A)$.

\square

14.8 Exercise Prove that the equivalence relation $a \stackrel{\text{L}}{\equiv}_A b$ is the transitive closure of the relation: there is a sequence $\langle c_i : i < \omega \rangle$ indiscernible over A such that $c_0 = a$ and $c_1 = b$. \square

It may not be immediately obvious that the relation $d_A(z, y) \leq n$ is type-definable.

14.9 Proposition For every $n < \omega$ there is a type $p_n(z, y) \subseteq L(A)$ equivalent to $d_A(z, y) \leq n$.

Proof It suffices to prove the proposition with $n = 1$. Let $\lambda = |L(A)|$ and fix a tuple of distinct variables $w = \langle w_i : i < \lambda \rangle$, then $p_1(\mathbf{z}, \mathbf{y}) = \exists w p(w, \mathbf{z}, \mathbf{y})$ where

$$p(w, \mathbf{z}, \mathbf{y}) = q(w) \cup \left\{ \varphi(\mathbf{z}, w) \leftrightarrow \varphi(\mathbf{y}, w) : \varphi(\mathbf{z}, w) \in L(A) \right\}$$

and $q(w) \subseteq L(A)$ is a consistent type with the property that all its realizations enumerate a model containing A .

It remains to verify that such a type exist. Let $\langle \psi_i(x, w_{|i}) : i < \lambda \rangle$ be an enumeration of the formulas in $L_{x,w}(A)$, where x is a single variable. Let

$$q(w) = \left\{ \exists x \psi_i(x, w_{|i}) \rightarrow \psi_i(w_i, w_{|i}) : i < \lambda \right\}.$$

Any realization of $q(w)$ satisfy the Tarski-Vaught test therefore it enumerates a model containing A . Vice versa it is clear that we can realize $q(w)$ in any model containing A . \square

We conclude this section with a theorem of Ludomir Newelski.

The following notions apply generally to any group G acting on some set \mathcal{X} and to any set $\mathcal{D} \subseteq \mathcal{X}$. Below we always have $G = \text{Autf}(\mathcal{U}/A)$ and $\mathcal{X} = \mathcal{U}^{|\mathcal{Z}|}$. We say that **\mathcal{D} is drifting** if for every finitely many $f_1, \dots, f_n \in G$ there is a $g \in G$ such that $g[\mathcal{D}]$ is disjoint from all the $f_i[\mathcal{D}]$. We say that **\mathcal{D} is quasi-invariant** if for every finitely many $f_1, \dots, f_n \in G$ the sets $f_i[\mathcal{D}]$ have non-empty intersection. We say that a formula or a type is drifting or quasi-invariant if the set it defines is.

The union of drifting sets need not be drifting. However, by the following lemma it cannot be quasi-invariant.

14.10 Lemma *The union of finitely many drifting sets is not quasi-invariant.*

Proof It is convenient to prove an apparently more general claim. If $\mathcal{D}_1, \dots, \mathcal{D}_n$ are all drifting and \mathcal{L} is such that for some finite $F \subseteq G$

$$\# \quad \mathcal{L} \subseteq \bigcup_{f \in F} f[\mathcal{D}_1 \cup \dots \cup \mathcal{D}_n],$$

then \mathcal{L} is not quasi-invariant. (The statement is slightly awkward since a superset of a quasi-invariant set must be quasi-invariant.)

The claim is vacuously true for $n = 0$. Let n be positive, let $\mathcal{C} = \mathcal{D}_1 \cup \dots \cup \mathcal{D}_{n-1}$, and assume the claim holds for $n - 1$. Since \mathcal{D}_n is drifting there is a $g \in G$ such that $g[\mathcal{D}_n]$ is disjoint from $f[\mathcal{D}_n]$ for every $f \in F$, which implies that

$$\mathcal{L} \cap g[\mathcal{D}_n] \subseteq \bigcup_{f \in F} f[\mathcal{C}].$$

Hence for every $h \in G$ there holds

$$hg^{-1}[\mathcal{L}] \cap h[\mathcal{D}_n] \subseteq \bigcup_{f \in F} hg^{-1}f[\mathcal{C}].$$

Rewriting $\#$ as

$$\mathcal{L} \subseteq \bigcup_{f \in F} f[\mathcal{C}] \cup \bigcup_{h \in F} h[\mathcal{D}_n],$$

we observe that

$$\mathcal{L} \cap \bigcap_{h \in F} hg^{-1}[\mathcal{L}] \subseteq \bigcup_{f \in F} f[\mathcal{C}] \cup \bigcup_{h \in F} \bigcup_{f \in F} hg^{-1}f[\mathcal{C}].$$

By the induction hypothesis, the r.h.s. cannot be quasi invariant. Hence neither is \mathcal{L} , proving the claim and with it the lemma. \square

The following is a consequence of Baire's category theorem. We sketch a proof for the convenience of the reader.

14.11 Lemma *Let $p(x) \subseteq L(B)$ and $p_n(x) \subseteq L(A)$, for $n < \omega$, be consistent types such that*

$$1. \quad p(x) \rightarrow \bigvee_{n < \omega} p_n(x)$$

Then there is an $n < \omega$ and a formula $\varphi(x) \in L(A)$ consistent with $p(x)$ such that

$$2. \quad p(x) \wedge \varphi(x) \rightarrow p_n(x)$$

Proof Negate 2 and choose inductively for every $n < \omega$ a formula $\psi_n(x) \in p_n(x)$ such that $p(x) \wedge \neg\psi_0(x) \wedge \dots \wedge \neg\psi_n(x)$ is consistent. By compactness, we contradict 1. \square

Finally we can prove Newelski's theorem.

14.12 Theorem(Newelski) *For every $a \in \mathcal{U}^{|z|}$ the following are equivalent*

1. $\mathcal{L}(a/A)$ is type-definable;
2. $\mathcal{L}(a/A) = \{c : d_A(a, c) < n\}$ for some $n < \omega$.

Proof Implications $2 \Rightarrow 1$ holds by Proposition 14.9. We prove $1 \Rightarrow 2$. Suppose $\mathcal{L}(a/A)$ is type-definable, say by the type $l(z)$. Let $p(z, y)$ be some consistent type (to be defined below) such that $p(z, y) \rightarrow l(z) \wedge l(y)$. Then, in particular

$$p(z, y) \rightarrow \bigvee_{n < \omega} d_A(z, y) < n.$$

By Proposition 14.9 and Lemma 14.11, there is some $n < \omega$ and some $\varphi(z, y) \in L(A)$ consistent with $p(z, y)$ such that

$$\#_1 \quad p(z, y) \wedge \varphi(z, y) \rightarrow d_A(z, y) < n.$$

Below we define $p(z, y)$ so that for every $\psi(z, y) \in L(A)$

$$\#_2 \quad p(z, a) \wedge \psi(z, a) \quad \text{is non-drifting whenever it is consistent.}$$

Drifting and quasi-invariance are relative to the action of $\text{Autf}(\mathcal{U}/A)$ on $\mathcal{U}^{|z|}$. Then, in particular, $p(z, a) \wedge \varphi(z, a)$ is non-drifting and the theorem follows. In fact, by non-drifting, there are some $a_0, \dots, a_k \in \mathcal{L}(a/A)$ such that every set $p(\mathcal{U}, c) \cap \varphi(\mathcal{U}, c)$ for $c \in \mathcal{L}(a/A)$ intersects some $p(\mathcal{U}, a_i) \cap \varphi(\mathcal{U}, a_i)$. Let m be such that $d_A(a_i, a_j) \leq m$ for every $i, j \leq k$. From $\#_1$ we obtain that $d_A(a, c) \leq m + 2n$. As $c \in \mathcal{L}(a/A)$ is arbitrary, the theorem follows.

The required type $p(z, y)$ is union of a chain of types $p_\alpha(z, y)$ defined as follows

$$p_0(z, y) = l(z) \cup l(y);$$

$$\#_3 \quad p_{\alpha+1}(z, y) = p_\alpha(z, y) \cup \left\{ \neg\psi(z, y) \in L(A) : p_\alpha(z, a) \wedge \psi(z, a) \text{ is drifting} \right\};$$

$$p_\alpha(z, y) = \bigcup_{n < \alpha} p_n(z, y) \quad \text{for limit } \alpha.$$

Clearly, the chain stabilizes at some stage $\leq |L(A)|$ yielding a type which satisfies \sharp_2 . So we only need to prove consistency. We prove that $p_\alpha(z, a)$ is quasi-invariant (so, in particular, consistent). Suppose that $p_n(z, a)$ is quasi-invariant for every $n < \alpha$ but, for a contradiction, $p_\alpha(z, a)$ is not. Then for some $f_1, \dots, f_k \in \text{Autf}(\mathcal{U}/A)$

$$p_\alpha(z, a) \cup \bigcup_{i=1}^k p_\alpha(z, f_i a)$$

is inconsistent. By compactness there is some $n < \alpha$ and some $\psi_i(z, y)$ as in \sharp_3 such that

$$p_n(z, a) \rightarrow \neg \bigwedge_{j=1}^m \bigwedge_{i=1}^k \neg \psi_j(z, f_i a)$$

As $p_n(z, a)$ is quasi-invariant, from Lemma 14.10 we obtain that $p_n(z, f_i a) \wedge \psi_j(z, f_i a)$ is non-drifting for some i, j . Clearly we can replace $f_i a$ with a , then this contradicts the construction of $p_\alpha(z, y)$ and proves the theorem. \square


14.13 Exercise Let \mathcal{L} be quasi-invariant and let \mathcal{D} be drifting, prove that $\mathcal{L} \setminus \mathcal{D}$ is quasi-invariant. \square

3 Kim-Pillay types

Given a tuple $a \in \mathcal{U}^{|z|}$, we write $\mathcal{K}(a/A)$ for the intersection of all type-definable sets containing a that are Lascar invariant over A . Or, more concisely, the intersection of all sets that are type-definable over a model containing A . We call $\mathcal{K}(a/A)$ the **Kim-Pillay strong type over A** . Clearly, $\mathcal{L}(a/A) \subseteq \mathcal{K}(a/A) \subseteq \mathcal{O}(a/A)$.

Clearly $\mathcal{K}(a/A)$ is Lascar invariant over A . It also easy to see that $\mathcal{K}(a/A)$ is type-definable. In fact, by invariance, we can assume that all types in the intersection above are over M , for any fixed model containing A . Hence $\mathcal{K}(a/A)$ is the minimal type-definable set containing a and closed under the relation $\stackrel{L}{\equiv}_A$. It follows that if $b \in \mathcal{K}(a/A)$ then $\mathcal{K}(b/A) \subseteq \mathcal{K}(a/A)$.

If $\mathcal{K}(a/A) = \mathcal{K}(b/A)$, we say that a and b have the same **Kim-Pillay strong type over A** . We abbreviate this by $a \stackrel{\text{KP}}{\equiv}_A b$. In other words, $a \stackrel{\text{KP}}{\equiv}_A b$ holds when $a \in \mathcal{D} \leftrightarrow b \in \mathcal{D}$ for every type-definable set \mathcal{D} that is Lascar invariant over A .

 **Warning:** the symbol $\mathcal{K}(a/A)$ is not standard. The symbol $a \stackrel{\text{KP}}{\equiv}_A b$ is not unusual, but others write $\text{KP-stp}(a/A) = \text{KP-stp}(b/A)$ or $a E_{\text{KP}/A} b$.

14.14 Proposition Fix some $a \in \mathcal{U}^{|z|}$ and some $A \subseteq \mathcal{U}$. Then there is a type $e(z; w) \subseteq L(A)$ such that $\mathcal{K}(b/A) = e(\mathcal{U}; b)$ for all $b \in \mathcal{O}(a/A)$ and $e(z; w)$ defines an equivalence relation on $\mathcal{O}(a/A)$.

Proof Notice that $\mathcal{K}(a/A)$ is type-definable over A, a . In fact, if $f \in \text{Aut}(\mathcal{U}/A, a)$ and \mathcal{D} is a set containing a that is type-definable and Lascar invariant over A , then so is $f[\mathcal{D}]$. Therefore $\mathcal{K}(a/A)$ is invariant over A, a . As $\mathcal{K}(a/A)$ is type-definable, invariance implies that it is type-definable over A, a . Let $e(z; w) \subseteq L(A)$ be such that $\mathcal{K}(a/A) = e(\mathcal{U}; a)$.

We prove that $\mathcal{K}(b/A) = e(\mathcal{U}; b)$ for all $b \in \mathcal{O}(a/A)$. Let $f \in \text{Aut}(\mathcal{U}/A)$ be such that $f a = b$. If \mathcal{D} is a type-definable Lascar invariant over A , then so is $f[\mathcal{D}]$.

Therefore, f is a bijection between type-definable sets that are Lascar invariant over A and contain a and analogous sets containing b . Then $f[\mathcal{K}(a/A)] = \mathcal{K}(b/A)$ and $\mathcal{K}(b/A) = e(\mathcal{U}; b)$ follows.

We prove that $e(b; \mathcal{U}) \cap \mathcal{O}(a/A)$ is Lascar invariant over A . Let $f \in \text{Autf}(\mathcal{U}/A)$ and $c \in \mathcal{O}(a/A)$. Then $e(b; fc)$ is equivalent to $e(f^{-1}b; c)$ which in turn is equivalent to $e(b; c)$, by the invariance of $e(\mathcal{U}; c)$.

Finally we are ready to prove that $e(z; w)$ defines a symmetric relation on $\mathcal{O}(a/A)$. From what proved above, $e(b; \mathcal{U}) \cap \mathcal{O}(a/A)$ is a type-definable Lascar invariant set containing b and therefore it contains $\mathcal{K}(b/A)$. We conclude that for all $b, c \in \mathcal{O}(a/A)$

$$e(c; b) \leftrightarrow c \in e(\mathcal{U}; b) \leftrightarrow c \in \mathcal{K}(b/A) \rightarrow e(b; c)$$

Reflexivity is clear; we prove transitivity. As remarked above, $\mathcal{K}(b/A) \subseteq \mathcal{K}(c/A)$, for all $b \in \mathcal{K}(c/A)$ or equivalently $e(b; c)$. Hence if $e(b; c)$ then

$$e(d; b) \rightarrow d \in \mathcal{K}(b/A) \leftrightarrow d \in \mathcal{K}(c/A) \rightarrow e(d; c).$$

Which completes the proof. \square

14.15 Corollary For every $a, b \in \mathcal{U}^{|\mathcal{Z}|}$ and $A \subseteq \mathcal{U}$ the following are equivalent

1. $a \in \mathcal{K}(b/A)$;
2. $b \in \mathcal{K}(a/A)$;
3. $\mathcal{K}(a/A) = \mathcal{K}(b/A)$.

\square

The following useful lemma is the key ingredient in the proof of Theorem 14.17.

14.16 Lemma Let $p(z) \subseteq L(A)$ and let $e(z; w) \subseteq L(A)$ define a bounded equivalence relation on $p(\mathcal{U})$. Then there is a type $e'(z; w) \subseteq L(A)$ which defines a bounded equivalence relation (on \mathcal{U}) and refines $e(z; w)$ on $p(\mathcal{U})$.

Proof Let $\mathcal{C} = \langle \mathcal{C}_i : i < \lambda \rangle$ enumerate the partition of $p(\mathcal{U})$ induced by $e(z; w)$. Note that each \mathcal{C}_i is type-definable over A, a for any $a \in \mathcal{C}_i$. If $x = \langle x_i : i < \lambda \rangle$, we write $x \in \mathcal{C}$ for the type that is the conjunction of $x_i \in \mathcal{C}_i$ for $i < \lambda$.

We claim that the required type is

$$e'(a; b) \stackrel{\text{def}}{=} \exists x' \in \mathcal{C} \exists x'' \in \mathcal{C} \quad a, x' \equiv_A b, x''$$

As the type above is invariant over A , we can assume that $e'(a; b)$ is type-definable over A . Moreover, it is clearly a reflexive and symmetric relation, so we only check it is transitive. Suppose $a, x' \equiv_A b, x''$ and $b, y' \equiv_A c, y''$ for some $x', x'', y', y'' \in \mathcal{C}$. Let z be such that $a, x', z \equiv_A b, x'', y'$ then $z \in \mathcal{C}$ and $a, z \equiv_A c, y''$.

The relation $e'(a; b)$ clearly refines the equivalence defined by $e(z; w)$ when restricted to $p(\mathcal{U})$. To prove that it is bounded fix some $c \in \mathcal{C}$ and note that $e'(a; b)$ is refined by $a \equiv_{A, c} b$, which is bounded. \square

Finally, we have the following.

14.17 Theorem Denote by $e_A(z; w)$ be the finest bounded equivalence relation on $\mathcal{U}^{|\mathcal{Z}|}$ that is type-definable over A . Then for every $a, b \in \mathcal{U}^{|\mathcal{Z}|}$ the following are equivalent

1. $a \stackrel{\text{KP}}{\equiv} b$

2. $e_A(a; b)$. □

Proof As $a \stackrel{\text{KP}}{\equiv} b$ is equivalent to $a \in \mathcal{K}(b/A)$, it suffices to prove that $e_A(\mathcal{U}; b) = \mathcal{K}(b/A)$ for every $b \in \mathcal{U}^{|z|}$.

\subseteq The orbit of $e_A(\mathcal{U}; b)$ under $\text{Aut}(\mathcal{U}/A)$ has cardinality $< \kappa$. Hence, by Theorem 14.2, it is Lascar invariant over A . As $\mathcal{K}(b/A)$ is the least of such sets, $\mathcal{K}(b/A) \subseteq e_A(\mathcal{U}; b)$.

\supseteq Let $e(z; w)$ be the type-definable equivalence relation given by Proposition 14.14. The orbit of $\mathcal{K}(b/A)$ under $\text{Aut}(\mathcal{U}/A)$ has cardinality $< \kappa$. Hence the equivalence relation that $e(z; w)$ defines on $\mathcal{O}(b/A)$ is bounded. By Lemma 14.16 there is a type-definable bounded equivalence relation $e'(z; w) \subseteq L(A)$ that refines $e(z; w)$ on $\mathcal{O}(b/A)$. As $e'(z; w)$ is refined by $e_A(z; w)$, we obtain $e_A(\mathcal{U}; b) \subseteq e(\mathcal{U}; b) = \mathcal{K}(b/A)$. □

The **logic A -topology**, or simply the **logic topology** when A is empty, is the topology on $\mathcal{U}^{|z|}$ whose closed sets are the type-definable Lascar A -invariant sets. It is clearly a compact topology. Its Kolmogorov quotient, that is $\mathcal{U}^{|z|} / \stackrel{\text{KP}}{\equiv}$, is Hausdorff. In fact, by Corollary 14.15, the open sets $\neg\mathcal{K}(a/A)$ and $\neg\mathcal{K}(b/A)$ separate $b \not\stackrel{\text{KP}}{\equiv} a$.

14.18 Exercise Let $\mathcal{S}(a/A)$ be the intersection of all definable sets that contain a and are Lascar invariant over A . Prove that $\mathcal{S}(a/A) = \{b : b \stackrel{\text{Sh}}{\equiv}_A a\}$. □

14.19 Exercise Prove that the clopen sets in the logic A -topology are definable over $\text{acl}^{\text{eq}} A$. □

4 Notes

The original proof of Newelski theorem is rather long and complex. A simplified proof appears in Rodrigo Peláez's thesis (due, reportedly, essentially to Newelski himself). The proof here is a streamlined version the latter.


Chapter 15

Stable formulas (tentative, a scribbling block)

1 Externally definable sets

Let $\mathcal{C}, \mathcal{D} \subseteq \mathcal{U}^{|z|}$. The set $\mathcal{D} \cap A^{|z|}$ is called the **trace** of \mathcal{D} over A . We write $\mathcal{C} =_A \mathcal{D}$ if \mathcal{C} and \mathcal{D} have the same trace on A .

We say that \mathcal{D} is **externally definable** if there are a global type $p \in S_x(\mathcal{U})$ and a formula $\varphi(x; z)$ such that $\mathcal{D} = \{a : \varphi(x; a) \in p\}$. Equivalently, a set \mathcal{D} is externally definable if it is the trace over \mathcal{U} of a set which is definable in some elementary extension of \mathcal{U} . Precisely, \mathcal{D} is the trace on \mathcal{U} of a set of the form $\varphi(b'; \mathcal{U}')$ where $\langle \mathcal{U}', \mathcal{D}' \rangle$ is elementary extension of $\langle \mathcal{U}, \mathcal{D} \rangle$ and $b' \in \mathcal{U}'^{|x|}$. This explains the terminology.

 We prefer to deal with external definability in a different, though equivalent, way. This is not the most common approach.

15.1 Definition We say that \mathcal{D} is **approximable** by the formula $\varphi(x; z)$ if for every finite B there is a $b \in \mathcal{U}^{|x|}$ such that $\varphi(b; \mathcal{U}) =_B \mathcal{D}$. If in addition we have that $\varphi(b; \mathcal{U}) \subseteq \mathcal{D}$, we say that \mathcal{D} is **approximable from below**. If $\mathcal{D} \subseteq \varphi(b; \mathcal{U})$ we say that \mathcal{D} is **approximable from above**. We may call $\varphi(x; z)$ the **sort** of \mathcal{D} . □

Approximability from below is an adaptation to our context of the notion of *having an honest definition* in [1]. It will be discussed in Chapter ??.

The following proposition is clear by compactness.

15.2 Proposition For every \mathcal{D} the following are equivalent:

1. \mathcal{D} is approximable;
2. \mathcal{D} is externally definable. □

15.3 Example Every definable set is trivially approximable. However if we fix the sort, this is not true any more. Let $T = T_{\text{dlo}}$. Then $\mathcal{D} = \{z \in \mathcal{U} : a \leq z \leq b\}$ is approximable both from below and from above by the formula $x_1 < z < x_2$ though it is not definable by this formula.

Now, let $T = T_{\text{rg}}$. Then every $\mathcal{D} \subseteq \mathcal{U}$ is approximable and, when \mathcal{D} has small infinite cardinality, it is approximable from above but not from below, see Exercise 5.22. □

In Definition 15.1, the sort $\varphi(x; z)$ is fixed (otherwise any set would be approximable) but this requirement of uniformity may be dropped if the sets B are allowed to be infinite.

15.4 Proposition For every \mathcal{D} the following are equivalent:

1. \mathcal{D} is approximable;

2. for every $B \subseteq \mathcal{U}$ of cardinality $\leq |T|$ there is $\psi(z) \in L(\mathcal{U})$ such that $\psi(\mathcal{U}) =_B \mathcal{D}$.

Similarly, the following are equivalent:

3. \mathcal{D} is approximable from below;

4. for every $B \subseteq \mathcal{D}$ of cardinality $\leq |T|$ there is $\psi(z) \in L(\mathcal{U})$ such that $B \subseteq \psi(\mathcal{U}) \subseteq \mathcal{D}$.

Proof To prove $2 \Rightarrow 1$ assume 2 and negate 1 for a contradiction. For each formula $\psi(x; z) \in L$ choose a finite set B such that $\psi(b; \mathcal{U}) \neq_B \mathcal{D}$ for every $b \in \mathcal{U}^{|x|}$. Let C be the union of all these finite sets. Clearly $|C| \leq |T|$. By 2 there are a formula $\varphi(x; z)$ and a tuple c such that $\varphi(c; \mathcal{U}) =_C \mathcal{D}$, contradicting the definition of C .

The implication $1 \Rightarrow 2$ is obtained by compactness and the equivalence $3 \Leftrightarrow 4$ is proved similarly. \square

15.5 Remark Let $\mathcal{C} \subseteq \mathcal{U}^{|z|}$. If \mathcal{D} is approximated by $\varphi(x; z)$ then so is any \mathcal{C} such that $\mathcal{C} \equiv \mathcal{D}$, see Section 13.4 for the definition. The same holds for approximability from below and from above. In fact, if the set \mathcal{D} is approximable by $\varphi(x; z)$ then for every n

$$\forall z_1, \dots, z_n \exists x \bigwedge_{i=1}^n [\varphi(x; z_i) \leftrightarrow z_i \in \mathcal{D}].$$

So the same holds for any $\mathcal{C} \equiv \mathcal{D}$. As for approximability from below, add the conjunct $\forall z [\varphi(x; z) \rightarrow z \in \mathcal{D}]$ to the formula above, and similarly for approximability from above. \square

15.6 Proposition Suppose \mathcal{C}, \mathcal{D} are both approximated by $\psi(x; z) \in L(A)$ and invariant over A . If $\mathcal{C} \equiv_A \mathcal{D}$ then there is an A -automorphism of \mathcal{U} such that $f[\mathcal{C}] = \mathcal{D}$.

Proof Let $\langle \mathcal{U}', \mathcal{C}', \mathcal{D}' \rangle$ be a saturated elementary extension of $\langle \mathcal{U}, \mathcal{C}, \mathcal{D} \rangle$ of larger cardinality. Let $c' \in \mathcal{U}^{|x|}$ be such that $\varphi(c'; \mathcal{U}') \cap \mathcal{U}^{|z|} = \mathcal{C}$. By saturation and elementarity, there is a $d' \equiv_A c'$ such that $\varphi(d'; \mathcal{U}') \cap \mathcal{U}^{|z|} = \mathcal{D}$. Then there is an

or some \square

15.7 Lemma Suppose $\mathcal{B} \subseteq \mathcal{U}^{|z|}$ and $\mathcal{A} \subseteq \mathcal{U}^{|x|}$ are approximated by $\varphi(x; z)$ and $\varphi(z; x)^*$, respectively. Then the following are equivalent

1. $\varphi(x; z) \wedge x \in \mathcal{A}$ approximates \mathcal{B}
2. $\varphi(z; x)^* \wedge z \in \mathcal{B}$ approximates \mathcal{A}

Proof It suffices to prove $1 \Rightarrow 2$ because the converse implication is equivalent. Let $\langle \mathcal{U}', \mathcal{A}', \mathcal{B}' \rangle$ be a saturated elementary extension of $\langle \mathcal{U}, \mathcal{A}, \mathcal{B} \rangle$ of larger cardinality. Assume 1 then, by the remark above and saturation, there is an $a' \in \mathcal{A}'$ such that $\varphi(a'; \mathcal{U}') =_{\mathcal{U}} \mathcal{B}'$. Again by saturation and the remark above, there is a $b' \in \mathcal{U}'^{|z|}$ such that $\varphi(\mathcal{U}'; b') =_{\mathcal{U}, a'} \mathcal{A}'$ which implies $\varphi(a'; b')$. Hence $b' \in \mathcal{B}'$ and 2 follows. \square

We say that a set \mathcal{D} approximable by $\varphi(x; z)$ is stationary over A if it is invariant over A and for any other such set \mathcal{C} if $\mathcal{C} \equiv_A \mathcal{D}$ then $\mathcal{C} = \mathcal{D}$. Note that the notion of stationary depends only on $\varphi(x; z)$ and $A^{|z|} \cap \mathcal{D}$.

2 The ladder property

The **ladder order** of a parted formula $\varphi(x; z)$ is the largest ordinal n such that for some $a_i \in \mathcal{U}^{[x]}$ and $b_i \in \mathcal{U}^{[z]}$

Though we never made it explicit, by *formula* most often we mean a pair $\varphi(x)$ that consists of a formula and a tuple of variables. In the following we need to deal with two sorts of variables: those that are placeholder for parameters, and those used to define a set. To make things more complicated here and there we shall switch the role of the variables.

15.8 Definition A **2-parted formula** is a triple $\varphi(x; z)$ consistig of a formula and two tuples of variables. The variables actually occurring in φ are all among x, z . We write $\varphi(z; x)^*$ for the parted formula where the order of the two tuples is switched. The asterisk will avoid confusion with substitution.

3

Let $D \subseteq A$ be the trace on A of a definable set, then it is the trace of some set \mathcal{A} definable over $\text{acl}^{\text{eq}} A$, and such a set \mathcal{A} is unique.

The following is a version of Harrinton's mysterious Lemma (cfr. [2, Lemma 8.3.4]).

15.9 Lemma Let $\varphi(x; z) \in L$ be a stable formula and suppose $\mathcal{B} \subseteq \mathcal{U}^{[z]}$ and $\mathcal{A} \subseteq \mathcal{U}^{[x]}$ are approximated by $\varphi(x; z)$ and $\varphi(z; x)^*$, respectively. Then the following are equivalent

1. $\varphi(x; z) \wedge x \in \mathcal{A}$ approximates \mathcal{B}
2. $\varphi(z; x)^* \wedge z \in \mathcal{B}$ approximates \mathcal{A}

Proof It suffices to prove $1 \Rightarrow 2$ because the converse implication is equivalent. Assume 1 and negate 2 for a contradiction. As \mathcal{A} is approximable, if not 2, then $\varphi(x; z) \wedge z \notin \mathcal{B}$ approximates \mathcal{A} . Pick a_0 arbitrarily, then recursively find $b_i \notin \mathcal{B}$ and $a_i \in \mathcal{A}$ such that

$$\begin{aligned} \varphi(\mathcal{U}; b_i) &=_{a_0, \dots, a_{i-1}} \mathcal{A} \\ \varphi(a_i; \mathcal{U}) &=_{b_0, \dots, b_i} \mathcal{B} \end{aligned}$$

Then

$$\begin{aligned} i < j &\Rightarrow \varphi(a_i; b_j) \\ j \leq i &\Rightarrow \neg \varphi(a_i; b_j) \end{aligned}$$

Which contradicts the stability of $\varphi(x; z)$. □

We say that $\mathcal{D} \subseteq \mathcal{U}^{[z]}$ is **approximable** by $\varphi(x; z)$ if for every finite $B \subseteq \mathcal{U}^{[z]}$ there is a $a \in \mathcal{U}^{[x]}$ such that $\varphi(a; B) = \mathcal{D} \cap B$. For $A \subseteq \mathcal{U}$, we say **approximable over A** if we can further require that $a \in A^{[x]}$. We say **almost over A** if $a \in M^{[x]}$ for every model M containing A .

15.10 Proposition If \mathcal{D} is approximable from A , then it is invariant over A . If it is approximable almost over A , then it is Lascar invariant over A .

Proof Suppose for a contradiction that $f[\mathcal{D}] \neq \mathcal{D}$ for some $f \in \text{Aut}(\mathcal{U}/A)$. Pick some b such that $b \in \mathcal{D} \nleftrightarrow fb \in \mathcal{D}$. Clearly $\{b, fb\} \cap \mathcal{D}$ cannot be the trace on

$\{b, fb\}$ of a set definable over A . The same argument apply to prove Lascar invariance. \square

Tentativo

15.11 Corollary *Under the same assumptions of Lemma 15.9 the following are mutually exclusive*

1. $\varphi(x; z) \wedge z \notin \mathcal{B}$ approximates \mathcal{A} ;
2. $\varphi(x; z) \wedge z \in \mathcal{B}$ approximates \mathcal{A} .

\square

15.12 Proposition *Let $A \subseteq \mathcal{U}$ and $\mathcal{D} \subseteq \mathcal{U}^{|\mathcal{Z}|}$. Suppose that some stable formula $\varphi(x; z) \in L$ approximates $\mathcal{D} \subseteq \mathcal{U}^{|\mathcal{Z}|}$. Then the following are equivalent:*

1. \mathcal{D} is invariant over $\text{acl}^{\text{eq}} A$;
2. \mathcal{D} is definable over $\text{acl}^{\text{eq}} A$.

Proof Only $1 \Rightarrow 2$ requires a proof. Assume 1. By stability, \mathcal{D} is definable. By invariance, $\mathcal{D} \in M^{\text{eq}}$ for every model M containing A and 2 follows. \square

15.13 Proposition *For every M and every externally definable set \mathcal{D} is some $\mathcal{B} \equiv_M \mathcal{D}$ that is approximable in M .*

Proof

\square

4 Notes e references

- [1] Artem Chernikov and Pierre Simon, *Externally definable sets and dependent pairs*, Israel J. Math. **194** (2013), no. 1, 409–425.
- [2] Katrin Tent and Martin Ziegler, *A course in model theory*, Lecture Notes in Logic, vol. 40, Association for Symbolic Logic, Cambridge University Press, 2012.

Chapter 16

Some combinatorics

In this chapter we expose some results in combinatorics that are relevant to model theory. Here we only deal with set-systems i.e., pairs that consist of a set Ω and a collection of subsets $\Delta \subseteq \mathcal{P}(\Omega)$. There is no other structure, and we not even mention first-order formulas. So the chapter may be read independently from the rest of the notes. Eventually the results presented here will be applied to describe externally definable sets in stable and NIP theories. Precisely, in Section 2 we deal implicitly with stable formulas and in the other section with NIP formulas.

1 A general setting

Let Ω be any set and $\Delta \subseteq \mathcal{P}(\Omega)$ be two fixed non-empty sets. The pair $\langle \Omega, \Delta \rangle$ is called a **set-system** and sometimes may be denoted simply by Δ . In some literature set-systems go under the names of *hypergraphs*, *range spaces*, *concept classes*, and possibly more. Here we occasionally call **edges** the elements of Δ , typically we use symbol φ for edges.

For $B \subseteq \Omega$, we call the set $\varphi \cap B$ the **trace of φ on B** . The set traces on B of the elements of Δ is called the **trace on B of the set-system Δ** and is denoted by

$$\Delta \restriction B \stackrel{\text{def}}{=} \{ \varphi \cap B : \varphi \in \Delta \}.$$

A **sub-set-system of $\langle \Omega, \Delta \rangle$** is a pair $\langle B, D \rangle$ where $B \subseteq \Omega$ and $D \subseteq \Delta \restriction B$. Sub-set-systems of the form $\langle B, \Delta \restriction B \rangle$ are also known as a **full sub-set-systems**.

Let \leq_I be an order relation on a set I . We say that Δ **realizes \leq_I** if there is a full sub-set-system $\langle B, D \rangle$ such that \leq_I is isomorphic to \subseteq_D , where the latter is the usual subset relation on the elements of D . We repeat this definition with a more explicit notation as it is necessary below. We say that Δ realizes \leq_I if there is a sequence $\langle \varphi_i : i \in I \rangle$ of elements of Δ such that for some $B \subseteq \Omega$

$$\varphi_i \cap B \subseteq \varphi_j \cap B \Leftrightarrow i \leq_I j$$

If Δ does not realize \leq_I we say that Δ **omits \leq_I** . Mathematicians appreciate simplicity, so we are mainly interested in set-systems that omits complicated orders.

Typically Δ is not closed under Boolean operations and it may be interesting to know how the complexity of the set-systems increase when closing Δ under union, intersections, etc. Here is the notation that we use. We write $\bigwedge \Delta$ for the set whose elements are of the form $\varphi_1 \cap \varphi_2$ for some $\varphi_i \in \Delta$. We write $\{\bigwedge\}\Delta$ for the closure of Δ under intersection. The sets $\bigvee \Delta$ and $\{\bigvee\}\Delta$ are defined in a similar way. Note that $\{\bigvee\}\{\bigwedge\}\Delta$ coincides with $\{\bigvee \bigwedge\}\Delta$, the closure of Δ under union and intersection. Also, we write $\neg \Delta$ for set whose elements are $\neg \varphi$, that is $\Omega \setminus \varphi$, for $\varphi \in \Delta$. You should not confuse $\neg \Delta$ with $\{\neg\}\Delta$ which is the closure of Δ under complementation, $\Delta \cup \neg \Delta$.

We are also interested in the subsets of Ω that are close to Δ in a sense that we know

make precise. We say that $D \subseteq \Omega$ is **approximable by Δ** if for every finite $B \subseteq \Omega$ there is some $\varphi \in \Delta$ such that $B \cap \varphi = B \cap D$. We say that $D \subseteq \Omega$ is **approximable from below** in Δ if we can further require that $\varphi \subseteq D$. If instead we require $D \subseteq \varphi$ then we say that D is **approximable from above**. Clearly D is approximable from above if and only if $\neg D = \Omega \setminus D$ is approximable from below in the set-system $\neg\Delta$.

It may be useful to rephrase the definition above in topological terms. Endow $\mathcal{P}(\Omega)$ with the topology that has as basic open sets those of the form $\{D \subseteq \Omega : F \subseteq D \subseteq C\}$ where F is finite and C co-finite. This is also a base of closed sets. Up to identifying $\mathcal{P}(\Omega)$ with 2^Ω , this is the product topology on 2^Ω which is a compact, zero-dimensional topology. Let $\bar{\Delta}$ denote the closure Δ in this topology. Then D is approximable in Δ if and only if $D \in \bar{\Delta}$.

For every $a \in \Omega$ define $\Delta_a = \{\varphi \in \Delta : a \in \varphi\}$. Let $\Omega^* = \Delta$ and $\Delta^* = \{\Delta_a : a \in \Omega\}$. We say that $\langle \Omega^*, \Delta^* \rangle$ is the **dual system** of $\langle \Omega, \Delta \rangle$. The operation of duality is more natural is represented in the following way. We now write Δ for a pure set. To every relation $R \subseteq \Omega \times \Delta$ we associate a set-system. The edges are $\{x : xRa\}$ for $a \in \Delta$. We say that R is extensional if $\{x : xRa\} = \{x : xRb\}$ only if $a = b$. If R is not extensional we can factor by the obvious equivalence relation and obtain an extensional relation with the same associated set-system. So, suppose that R is extensional. It is easy to check that the set-system associated to $R^{-1} \subseteq \Delta \times \Omega$ is dual to the one associated to R .

2 The combinatorial bare bones of stability

We write \leq_ω for the usual order on ω and \leq_ω^{-1} for the reverse order. For $k < \omega$ we write \leq_k for the natural order on the set $k = \{0, \dots, k-1\}$. In this section we prove that if Δ omits \leq_k for some $k < \omega$ then $\bar{\Delta} \subseteq \{\vee \wedge\} \Delta$.

16.1 Lemma *Suppose Δ omits \leq_ω then*

1. $\wedge \Delta$ omits \leq_ω .
2. $\vee \Delta$ omits \leq_ω .

And the same holds with \leq_ω^{-1} for \leq_ω .

Proof It suffices to prove claim 1 both for \leq_ω and \leq_ω^{-1} . In fact, Δ omits \leq_ω if and only if $\neg\Delta$ omits \leq_ω^{-1} .

Suppose $\wedge \Delta$ realizes \leq_ω . That is, for some B and some $\langle \varphi_i, \psi_i \in \Delta : i < \omega \rangle$

$$a. (\varphi_i \cap \psi_i) \cap B \subseteq (\varphi_j \cap \psi_j) \cap B \Leftrightarrow i \leq j$$

We can assume that $\varphi_0 \cap \psi_0$ is non empty, otherwise we shift the sequence by 1. Choose $\{a_i : i < \omega\} \subseteq B$ recursively: $a_0 \in \varphi_0 \cap \psi_0$ and $a_{i+1} \in (\varphi_{i+1} \cap \psi_{i+1}) \setminus (\varphi_i \cap \psi_i)$. We obtain

$$b. a_i \in \varphi_j \cap \psi_j \Leftrightarrow i \leq j$$

Hence for every pair $j < i < \omega$ at least one the following obtains

$$c. a_i \notin \varphi_j \quad \text{or} \quad a_i \notin \psi_j$$

By Ramsey's Theorem there is an infinite $I \subseteq \omega$ such that $i, j \in I$ satisfies the same (say, the first) inequality in c whenever $j < i$. Then

d. $a_i \in \varphi_j \Leftrightarrow i \leq j$ for every $i, j \in I$

So, if we define $C = \{a_i : i \in I\}$ we may conclude

$$\varphi_i \cap C \subseteq \varphi_j \cap C \Leftrightarrow i \leq j \quad \text{for every } i, j \in I.$$

As I is infinite, this suffices to show that Δ realizes \leq_ω .

As for \leq_ω^{-1} the proof is identical but for the choice of $\{a_i : i < \omega\} \subseteq B$. In this case we pick some elements $a_i \in (\varphi_i \cap \psi_i) \setminus (\varphi_{i+1} \cap \psi_{i+1})$. \square

A finitary version of Lemma 16.1 can be proved along the same idea.

16.2 Proposition *For every $n < \omega$ there is a $k < \omega$ such that for every Δ that omits \leq_k*

1. $\wedge \Delta$ omits \leq_n ;
2. $\vee \Delta$ omits \leq_n .

Proof In this finite version of the lemma 1 and 2 are equivalent. If we repeat the proof of Lemma 16.1 with k for ω , we need that every function $f : [k]^2 \rightarrow 2$ is constant in $[I]^2$ for some $I \subseteq k$ of cardinality n . Then the finite Ramsey theorem yields the required k . \square

16.3 Theorem *If Δ omits \leq_k then*

1. every set that is approximable in Δ is approximable from below in $\wedge^k \Delta$;
2. every set that is approximable in Δ is approximable from above in $\vee^k \Delta$;
3. every set approximable in Δ belongs to $\vee^h \wedge^k \Delta$ for some $h < \omega$.

Proof

1. Let $D \subseteq \Omega$ be approximable in Δ . We need to show that for every finite $C \subseteq D$ there is a $\varphi \in \wedge^k \Delta$ such that $C \subseteq \varphi \subseteq D$. Let $B_0 = C$ and define recursively $C \subseteq \varphi_i \in \Delta$ and a chain of finite sets $B_i \subseteq \Omega$. As D is approximable, there is a $\varphi_i \in \Delta$ such that

$$\varphi_i \cap B_i = D \cap B_i.$$

If $\varphi_0 \cap \dots \cap \varphi_i \subseteq D$ terminate the construction: we have as required

$$C \subseteq \varphi_0 \cap \dots \cap \varphi_i \subseteq D$$

Otherwise pick some arbitrary $a_i \in \varphi_0 \cap \dots \cap \varphi_i \setminus D$ and let $B_{i+1} = B_i \cup \{a_i\}$.

Let

$$B = \bigcup_{i \in \omega} B_i$$

and notice that

$$\varphi_i \cap B \subseteq \varphi_j \cap B \Leftrightarrow j \leq i.$$

As Δ omits \leq_k , the construction terminates at some stage $< k$.

2. Easy, similar to 2.
3. Let $D \subseteq \Omega$ be approximable in Δ . Then by 1 it is approximable from below in $\wedge^k \Delta$. So, by Lemma 16.2, it omits \leq_h for some $h < \omega$. Therefore, by claim 2, it belongs to $\vee^h \wedge^k \Delta$. \square

3 Vapnik-Chervonenkis dimension

In the following sections we deal with set-systems of finite Vapnik-Chervonenkis dimension.

16.4 Definition The *Vapnik-Chervonenkis dimension of Δ* , abbreviated by *VC-dimension*, is the largest cardinality of a finite set $B \subseteq \Omega$ such that $\Delta \restriction B = \mathcal{P}B$. If such a maximum does not exist we say that Δ has infinite VC-dimension. In lengthier words, the VC-dimension is the largest k such that there are $\langle a_i \in \Omega : i < k \rangle$ and $\langle \varphi_J : J \subseteq k \rangle$ such that

$$a_i \in \varphi_J \iff i \in J \quad \text{for every } i < k \text{ and every } J \subseteq k.$$

The dimension is 0 when Δ contains only one element. Set-systems with dimension 1 may contain a collection of disjoint sets or chain of sets. The set of intervals of reals has dimension 2, the set of squares in \mathbb{R}^2 has dimension 3.

The following lemma has been proved independently by Shelah, Sauer, and Vapnik-Chervonenkis around 1970. Shelah was working in model theory while Sauer and Vapnik-Chervonenkis were in statistics.

16.5 Proposition(Sauer's Lemma) If Δ has finite VC-dimension k then $|\Delta \restriction B| \leq \sum_{i=0}^k \binom{n}{i}$ for every finite set $B \subseteq \Omega$ of cardinality $n > k$.

Proof We proceed by induction on k . If Δ has null VC-dimension, both sides of the inequality are 1. Now assume the lemma is true for $k - 1$, we claim that for every $n > k$

$$|\Delta \restriction B| \leq \sum_{i=0}^k \binom{n}{i}.$$

We proceed by induction on $n = |B|$. If $n = k$ the claim says $|\Delta \restriction B| \leq 2^n$ which is clearly true. Fix some $a \in B$ and define

$$\Gamma = \{C \subseteq B \setminus a : C, C \cup \{a\} \in \Delta \restriction B\}.$$

where we use $B \setminus a$ as a shorthand for $B \setminus \{a\}$. We may assume that Γ is non empty otherwise the claim is obvious. Observe that

$$|\Delta \restriction B| = |\Delta \restriction (B \setminus a)| + |\Gamma|.$$

Given that Δ has VC-dimension k , it is clear that Γ has VC-dimension at most $k - 1$. We apply the induction hypothesis on n to $B \setminus a$ and the induction hypothesis on k to the set-system Γ . Note that $\Gamma = \Gamma \restriction (B \setminus a)$. We obtain

$$\begin{aligned} |\Delta \restriction B| &\leq \sum_{i=0}^k \binom{n-1}{i} + \sum_{i=0}^{k-1} \binom{n-1}{i} \\ &= \binom{n-1}{0} + \sum_{i=1}^k \binom{n-1}{i} + \binom{n-1}{i-1} \\ &= \binom{n-1}{0} + \sum_{i=1}^k \binom{n}{i} \\ &\leq \sum_{i=0}^k \binom{n}{i} \end{aligned}$$

which proves the claim, and with it, the lemma. \square

The following is called the **shatter function**

$$S(n, \Delta) \stackrel{\text{def}}{=} \max \left\{ |\Delta \restriction B| : B \subseteq \Omega, |B| = n \right\}$$

Next corollary states an important dichotomy. It says that the shatter function grows exponentially unless the VC-dimension is finite. In this case the growth is only polynomial.

16.6 Corollary *For every set Δ one of the following obtains*

1. *the VC-dimension is infinite and $S(n, \Delta) = 2^n$ for every positive integer n ;*
2. *the VC-dimension is k and $S(n, \Delta) \leq n^k$ for every integer $n > k$.*

Proof If VC-dimension is infinite claim 1 is obvious. So suppose Δ has VC-dimension is k and observe that when $k > 1$, Sauer's lemma

$$S(n, \Delta) \leq \sum_{i=0}^k \binom{n}{i} \leq \sum_{i=0}^k \frac{n^i}{i!} \leq n^k.$$

When $k = 0$ the claim is obvious. When $k = 1$ we obtain the bound $n + 1$. We can get rid of the extra 1 with a minor (otherwise uninteresting) improvement of Sauer's lemma. \square

The dimension of the dual set-system $\langle \Omega^*, \Delta^* \rangle$ is called **VC*-dimension of Δ** . Unraveling the definition we obtain that the VC*-dimension is the largest k for which there is no tuple $\varphi_0, \dots, \varphi_{k-1} \in \Delta$ and $\langle \varphi_I : I \subseteq k \rangle$ such that

$$a_I \in \varphi_j \Leftrightarrow j \in I \quad \text{for every } I \subseteq k \text{ and every } j \in k.$$

16.7 Proposition *If Δ has VC-dimension $< k$ then its VC*-dimension is $< 2^k$.*

Proof Assume that the VC*-dimension of Δ is $\geq 2^k$. Then, replacing k with $\mathcal{P}(k)$ in the definition above we obtain two sequences $\langle \varphi_I : I \subseteq k \rangle$ and $\langle a_{\mathcal{J}} : \mathcal{J} \subseteq \mathcal{P}(k) \rangle$ such that

$$\mathfrak{h} \quad a_{\mathcal{J}} \in \varphi_I \Leftrightarrow I \in \mathcal{J} \quad \text{for every } I \subseteq k \text{ and every } \mathcal{J} \subseteq \mathcal{P}(k).$$

For $j \in k$ define $\mathcal{J}_j = \{I \subseteq k : j \in I\}$. As a particular case of \mathfrak{h} we obtain

$$a_{\mathcal{J}_j} \in \varphi_I \Leftrightarrow I \in \mathcal{J}_j \quad \text{for every } j \in k.$$

which is equivalent to

$$a_{\mathcal{J}_j} \in \varphi_I \Leftrightarrow j \in I \quad \text{for every } I \subseteq k \text{ and every } j \in k.$$

Now, defining $a_j = a_{\mathcal{J}_j}$ we obtain \mathfrak{h} of Definition 16.4, hence Δ has VC-dimension $\geq k$. \square

16.8 Exercise Prove that if Δ has finite VC-dimension then also $\wedge \Delta$ has finite VC-dimension. Hint: apply Corollary 16.6.

4 Epsilon-approximations

Let μ be some probability measure on Ω that makes all sets in Δ measurable. We say that a finite set B is an **ϵ -approximation of μ** or **ϵ -sample** if for every $\varphi \in \Delta$

$$\star \quad \left| \mu(\varphi) - \frac{|B \cap \varphi|}{|B|} \right| \leq \varepsilon$$

The uniform probability measures on a finite set Ω admits ε -approximations for arbitrarily small ε . In fact, Ω itself is a 0-approximation. This is not true if the measure is non-uniform (see Exercise 16.15). This is unfortunate for the intended applications. So we introduce a weaker notion of approximation.

A **finite multi-set** is a map $B : \Omega \rightarrow \omega$ that has finitely many non-zero values. The **cardinality of B** is defined as

$$|B| = \sum_{a \in \Omega} B(a)$$

When B is a multi-set and $\varphi \in \Delta$ we read $B \cap \varphi$ as the product of B with the characteristic function of φ . We say that $B : \Omega \rightarrow \omega$ is a **multi-set ε -approximation** of μ if inequality \star holds with the reading adapted to multi-sets. It is clear that every measure on a finite set admits multi-set ε -approximations for arbitrarily small ε or, when μ is rational-valued, even a 0-approximation.

For the moment we work with (plain) approximations. Given ε we are interested in the least n such that some ε -approximations of cardinality n exist. The idea is to start with a large approximation and reduce size at the cost of slightly enlarging ε . We now introduce a powerful technique for achieving this.

Let $B \subseteq \Omega$ have finite cardinality n . We call a total map $c : B \rightarrow \{-1, +1\}$ a **coloring of B** . For any $\varphi \in \Delta \cup \{\Omega\}$ we define

$$c_\varphi \stackrel{\text{def}}{=} \sum_{a \in B \cap \varphi} c(a).$$

The **(relative) discrepancy of $\Delta|B$** is

$$\delta_B \stackrel{\text{def}}{=} \min_{c: B \rightarrow \{\pm 1\}} \max_{\varphi \in \Delta \cup \{\Omega\}} \frac{|c_\varphi|}{n}$$

The next lemma demonstrate that discrepancy can be applied to reduce the size of an approximation.

16.9 Lemma *Let $B \subseteq \Omega$ be an ε -approximation of μ of cardinality n with discrepancy δ_B . Then there is an $(\varepsilon + 2\delta_B)$ -approximation $B^+ \subseteq B$ of cardinality $\leq n/2$.*

Proof Fix $c : B \rightarrow \{\pm 1\}$ such that $|c_\varphi| \leq n\delta_B$ per ogni $\varphi \in \Delta \cup \{\Omega\}$ and define $B^\pm = c^{-1}[\pm 1]$ and $n^\pm = |B^\pm|$. Then

$$c_\varphi = |B^+ \cap \varphi| - |B^- \cap \varphi|.$$

In particular $c_B = n^+ - n^-$. We may assume that $n^+ \leq n/2$, otherwise swap $+1$ and -1 , then $c_B < 0$ and we have

$$\begin{aligned} 1. \quad \frac{|B \cap \varphi|}{n} &= \frac{2|B^+ \cap \varphi|}{n} - \frac{c_\varphi}{n} \\ &\leq \frac{|B^+ \cap \varphi|}{n^+} + \delta_B \end{aligned}$$

We also have

$$\begin{aligned} 2. \quad \frac{|B \cap \varphi|}{n} &= \frac{2|B^+ \cap \varphi| - c_\varphi}{2n^+} \cdot \left(1 - \frac{c_B}{2n^+}\right)^{-1} \\ &\geq \left(\frac{|B^+ \cap \varphi|}{n^+} - \delta_B\right) (1 + \delta_B)^{-1} \end{aligned}$$

$$\begin{aligned}
&\geq \left(\frac{|B^+ \cap \varphi|}{n^+} - \delta_B \right) (1 - \delta_B) \\
&\geq \frac{|B^+ \cap \varphi|}{n^+} - 2\delta_B
\end{aligned}$$

Combining 1 and 2 we obtain

$$\left| \frac{|B \cap \varphi|}{n} - \frac{|B^+ \cap \varphi|}{n^+} \right| \leq 2\delta_B$$

hence

$$\left| \mu(\varphi) - \frac{|B^+ \cap \varphi|}{n^+} \right| \leq \varepsilon + 2\delta_B$$

as claimed by the lemma. \square

The lemma above tells that approximations with small discrepancy are useful, but as yet we have no clue as to finding one. It is immediate that if $\Delta \restriction B = \mathcal{P}(B)$ then δ_B is large (i.e. close to $1/2$). We are going to prove that when $\Delta \restriction B$ is small in comparison with B , then the discrepancy of B is not too large. We use a probabilistic argument to prove this bound, hence a brief digression into probability theory. The following inequality is a classical tool in this context.

16.10 Lemma(Chernoff bound) *For $i = 1, \dots, n$ let X_i be independent identically distributed random variables such that $\Pr(X_i = \pm 1) = 1/2$. Then for every $\delta \in \mathbb{R}^+$*

$$\Pr\left(\sum_{i=1}^n X_i > n\delta\right) \leq \exp(-\frac{n}{2}\delta^2)$$

Proof Let $Y = \sum_{i=1}^n X_i$ and let $t \in \mathbb{R}^+$ be arbitrary. Then

$$\begin{aligned}
\# \quad \Pr(Y > n\delta) &= \Pr(e^{tY} > e^{tn\delta}) \\
&\leq \frac{\mathbb{E}(e^{tY})}{e^{tn\delta}}
\end{aligned}$$

In fact, the equality follows because the exponential is an increasing function and the inequality is Markov's inequality, which says that $a \Pr(X > a) \leq \mathbb{E}(X)$ for every a and is immediate to verify. Now observe that

$$\begin{aligned}
\mathbb{E}(e^{tX_i}) &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} \\
&= \sum_{i=0}^{\infty} \frac{t^i}{i!} + \sum_{i=0}^{\infty} \frac{(-t)^i}{i!} \\
&= \sum_{i=0}^{\infty} \frac{t^{2i}}{(2i)!} \\
&\leq \sum_{i=0}^{\infty} \frac{(t^2/2)^i}{i!} \\
&= e^{t^2/2}
\end{aligned}$$

From this, by independence we have

$$\mathbb{E}(e^{tY}) = \prod_{i=1}^n e^{t^2/2} = e^{nt^2/2}$$

Applying this to \sharp inequality we obtain $\Pr(Y > n\delta) \leq e^{nt^2/2 - tn\delta}$ and finally the Chernoff inequality is obtained substituting δ for t . \square

16.11 Lemma *Let $B \subseteq \Omega$ and suppose that $n \leq |B|$ and $|\Delta \restriction B| \leq m$ is finite. Then*

$$\delta_B \leq \frac{1}{n} \sqrt{2n \ln(3m)}$$

Proof Fix a tuple $\langle a_i : i < n \rangle$ of distinct elements of B . Let C be the set of colorings of B . Imagine C as a probability space where each $c : B \rightarrow \{\pm 1\}$ is obtained by flipping independently n times a fair coin. For each $i < n$ define the random variables $X_i(c) = c(a_i)$. Then apply Lemma 16.10 to obtain

$$\Pr(c_\varphi \geq n_\varphi \delta) \leq e^{-n_\varphi \delta^2 / 2} \quad \text{where } n_\varphi = |\varphi|.$$

By symmetry

$$\Pr(|c_\varphi| \geq n_\varphi \delta) \leq 2e^{-n_\varphi \delta^2 / 2}$$

Substitute $\sqrt{(2/n_\varphi) \ln(3m)}$ for δ

$$\Pr(|c_\varphi| \geq \sqrt{2n_\varphi \ln(3m)}) \leq \frac{2}{3m}$$

Hence

$$\Pr(|c_\varphi| \geq \sqrt{2n \ln(3m)}) \leq \frac{2}{3m}$$

As $|\Delta \restriction B| \leq m$,

$$\Pr(\exists \varphi \in \Delta \quad |c_\varphi| \geq \sqrt{2n \ln(3m)}) \leq \frac{2}{3} < 1$$

Therefore

$$\Pr(\forall \varphi \in \Delta \quad |c_\varphi| < \sqrt{2n \ln(3m)}) > 0$$

In particular, for such a φ there exist some c such that $|c_\varphi| < \sqrt{2n \ln(3m)}$ for all $\varphi \in \Delta$. So, finally

$$\delta_B \leq \frac{1}{n} \sqrt{2n \ln(3m)}$$

which proves the require bound. \square

16.12 Theorem *Fix some positive $\varepsilon < 1$. Let Δ have VC-dimension $1 < k < \omega$. Let μ be some probability measure that admits an $\varepsilon/2$ -approximation (of arbitrary size). Then there is an ε -approximation of cardinality*

$$\sharp \leq 2^8 \frac{k}{\varepsilon^2} \ln \frac{1}{\varepsilon^2}.$$

Note that the bound claimed by the theorem only depends on the dimension of Δ and it is independent of μ .

Proof For $i = 0, \dots, h$ we construct a decreasing chain B_i of ε_i -approximations. By assumption we can require B_0 is an $\varepsilon_0 = \varepsilon/2$ -approximation. We denote by n_i and δ_i the cardinality, respectively the discrepancy, of B_i . Without loss of generality we can assume that n_0 is a power of two, so by lemma 16.9, we can require that $\varepsilon_{i+1} \leq \varepsilon_i + 2\delta_i$ and $n_i = 2^{-i}n_0$. Then

$$\varepsilon_h = \frac{\varepsilon}{2} + 2 \sum_{i=1}^h \delta_i$$

The construction stops at the least h such that $\varepsilon < \varepsilon_h + 2\delta_h$. So we have “only” have to prove that (independently of n_0) this h satisfies

$$\mathfrak{h}\mathfrak{h} \quad 2^{-h}n_0 \leq 2^8 \frac{k}{\varepsilon^2} \ln \frac{1}{\varepsilon^2}.$$

We may rewrite the condition $\varepsilon < \varepsilon_h + 2\delta_h$ as

$$\varepsilon < 4 \sum_{i=1}^{h+1} \delta_i$$

To get rid of some annoying square roots that will appear soon, we substitute the latter inequality with

$$\varepsilon^2 < 2^5 \sum_{i=1}^{h+1} \delta_i^2$$

We do not know δ_i but Lemma 16.11, together with Sauer’s Lemma 16.5, gives an upper bound

$$\begin{aligned} \delta_i^2 &\leq \frac{2}{n_i} \ln n_i^k \\ &= \frac{2^{i+1}}{n_0} (k \ln 2^{-i} + k \ln n_0) \\ &\leq \frac{2^{i+1}}{n_0} k \ln n_0 \end{aligned}$$

Then we obtain

$$\sum_{i=1}^{h+1} \delta_i^2 \leq 2^{h+2} k \frac{\ln n_0}{n_0}$$

The theorem is complete as any h that satisfies the inequality

$$\varepsilon^2 < 2^{h+7} k \frac{\ln n_0}{n_0}$$

verifies $\mathfrak{h}\mathfrak{h}$ (and in particular the h at which the constructions stops). The verification of is immediate, it suffices to substitute for ε^2 the r.h.s. of the inequality above and recall that we can assume that n_0 is sufficiently large. \square

16.13 Corollary *Let Δ be a finite set-system of VC-dimension $1 < k < \omega$. Then for every positive $\varepsilon < 1$, every probability measure μ admits a multi-set ε -approximation of cardinality bounded by \mathfrak{h} of Theorem 16.12.*

Proof Without loss of generality we can assume that μ is rational valued. Then there is a uniform probability measure μ' on some finite set Ω' and a surjection $f : \Omega' \rightarrow \Omega$ such that $\mu(f^{-1}\varphi) = \mu(\varphi)$. By Theorem 16.12 μ' admits an ε -approximation B' with cardinality by \mathfrak{h} , for every positive $\varepsilon < 1$. We now define a multi-set B such that $B(a) = |B' \cap f^{-1}\varphi|$ for every $a \in \Omega$. As $|B'| = |B|$ and $|B' \cap f^{-1}\varphi| = |B \cap \varphi|$, this is the required multi-set ε -approximation. \square

The following corollary is used in Theorem 16.16 below. Though it is sufficient for our application, stronger bound are known (essentially, we can replace ε for ε^2).

16.14 Corollary Let Δ be a finite set-system of VC-dimension $1 < k < \omega$. Then for every positive $\varepsilon < 1$, every probability measure μ admits a multi-set ε -net of cardinality bounded by \mathfrak{h} of Theorem 16.12.

16.15 Exercise Suppose that $\{a\}, \{b\} \in \Delta$ for some $a, b \in \Omega$. Show that, if μ and ε are such that $0 < \varepsilon < \mu(a)$ and $2\varepsilon < \mu(b) - \mu(a)$, then μ admits no ε -approximation. \square

5 A piercing problem

After Pierre Simon, after Jiri Matousek, after Noga Alon and Daniel Kleitman.

Let $q \leq p < \omega$ we say that Δ has the (p, q) -property if out of every p sets in Δ some q have non empty intersection. It has the dual (p, q) -property if for every p point in Ω some q are belong to the same set in Δ .

The following is a particular case of a theorem of Matousek (based on a proof of Noga Alon and Daniel Kleitman). The proof by Pierre Simon is simpler. [I will add details, references, and much more.]

16.16 Theorem Let Ω be finite and let Δ have VC-dimension $k < \omega$. There are some integers q and h that depends only on k such that, if every $B \subseteq \Omega$ of cardinality q is a subset of some $\varphi \in \Delta$, then Ω is covered by some $X \subseteq \Delta$ of cardinality h .

This theorem is often stated in the dual form which explains why it is sometimes called a piercing (or Helly-type) theorem: there are some integers q and h that depends only on k such that, if every q sets in Δ have non-empty intersection, then there is a $B \subseteq \Omega$ of cardinality h that intersects every $\varphi \in \Delta$. See Exercise 16.18.

Proof Let $\langle a_i : i < m \rangle$ and $\langle \varphi_j : j < n \rangle$ enumerate Ω and Δ without repetitions. Recall that the dual system is $\Omega^* = \Delta$ and $\Delta^* = \{\Delta_i : i < m\}$, where $\Delta_i = \{\varphi \in \Delta : a_i \in \varphi\}$. By Proposition 16.7 the dual system has VC-dimension 2^k . Choose some $\varepsilon = 1/2$. From Corollary 16.14 we obtain an h such that every probability measure μ^* on Δ admits an ε -net of cardinality h . As ε is fixed, this h only depends on k . Restating this explicitly there is a set $X \subseteq \Omega^*$ of cardinality h such that

$$\# \quad |\mu^*(\Delta_i)| > \varepsilon \Rightarrow X \cap \Delta_i \neq \emptyset$$

So, if the probability measure on Δ is such that $\mu^*(\Delta_i) > \varepsilon$ for all $i < m$, then $X \subseteq \Delta$ cover Ω as required by the theorem. So we only need to find such a probability measure.

For $i < m$ and $j < n$ let $P_{i,j} = 1$ if $a_i \in \varphi_j$ and 0 otherwise. Suppose we can find some $x_j \in \mathbb{R}$ such that for every $i < m$

$$\flat \quad \sum_{j < n} (P_{i,j} - \varepsilon) x_j > 0$$

Then we can assume that all x_j are positive. In fact,

$$\sum_{j < n} P_{i,j} \geq 1$$

(because every a_i belongs to some φ_j) so we can translate solutions of any positive quantity. We can also rescale solutions so, setting $\mu^*(\varphi_j) \stackrel{\text{def}}{=} x_j$, we obtain a well-defined probability measure

$$\mu^*(\Delta_i) = \sum_{j < n} P_{i,j} x_j > \varepsilon$$

We only have to prove that equation \flat has a solution. It suffices to show that clause 2 of Farkas' Lemma cannot obtain, that is, for every $\lambda_i \in \mathbb{R}_+$ there are some $x_j \in \mathbb{R}_+$ such that

$$\flat \quad \sum_{i < m} \lambda_i \sum_{j < n} (P_{i,j} - \varepsilon) x_j > 0.$$

As we can assume the λ_i add to 1, setting $\mu(a_i) \stackrel{\text{def}}{=} \lambda_i$, we obtain a probability measure on Ω . Note that for every j we have

$$\mu(\varphi_j) = \sum_{i < m} \lambda_i P_{i,j}.$$

Apply Corollary 16.14 once again to the set-system $\langle \Omega, \neg\Delta \rangle$. There is a q such that for every probability measure μ on Ω there is a set $B \subseteq \Omega$ of cardinality q such that

$$\sharp\sharp \quad \mu(\neg\varphi_j) > \varepsilon \Rightarrow B \not\subseteq \varphi_j$$

As ε is fixed, q only depends on k . By assumption, there is a \check{j} such that $B \subseteq \varphi_{\check{j}}$ and therefore $\mu(\varphi_{\check{j}}) \geq 1 - \varepsilon$. Let \check{x}_j be 1 for $j = \check{j}$ and 0 otherwise. We claim that the tuple $\langle x_j : j < n \rangle$ is a solution of \flat . In fact

$$\sum_{i < m} \lambda_i \sum_{j < n} (P_{i,j} - \varepsilon) \check{x}_j = \sum_{i < m} \lambda_i (P_{i,\check{j}} - \varepsilon) = \mu(\varphi_{\check{j}}) - \varepsilon \geq 1 - \varepsilon > 0.$$

This concludes the proof. \square

The following is a classical result. The proof may be found in any introductory text of convex analysis or linear programming. There are many (sometimes non equivalent) ways to state it, we recommend Terence Tao's exposition which may be found in his mathematical blog (we have reversed strict and weak inequalities, but the proof is identical).

16.17 Proposition(Farkas' Lemma) For $i < m$ let $P_i : \mathbb{R}^n \rightarrow \mathbb{R}$ be affine linear functions. Then the following are equivalent

1. $\bigwedge_{i < m} P_i(x) > 0$ for some $x \in \mathbb{R}^n$;
2. there are no $\lambda_i \in \mathbb{R}_+$ such that $\sum_{i < m} \lambda_i P_i(x) < 0$ for every $x \in \mathbb{R}^n$. \square

16.18 Exercise Let k, q and h be some integers as in Theorem 16.16 and let Δ be a set-system with dual VC-dimension k . Suppose that any q sets in Δ have non-empty intersection and prove that there is some set $B \subseteq \Omega$ of cardinality h that intersects every $\varphi \in \Delta$. \square

6 Appendix: Farkas' lemma

Below, an affine version of Farkas' Lemma tailored to our purposes.

16.19 Proposition Fix some $v_1, \dots, v_n, u \in \mathbb{Q}^k$ and let $r_1, \dots, r_n, s \in \mathbb{Q}$. Let

$$X(r_1, \dots, r_n) = \{x \in \mathbb{Q}^k : r_i \leq v_i \cdot x, \text{ for every } i = 1, \dots, n\}.$$

Then the following are equivalent

1. $s \leq u \cdot x$ for all $x \in X(r_1, \dots, r_n)$;

2. there exist $q_1, \dots, q_n \in \mathbb{Q}^+$ such that $\sum_{i=1}^n q_i v_i = u$.

Proof Implication $2 \Rightarrow 1$ is immediate. To prove $1 \Rightarrow 2$ assume 1. We claim that

3. $0 \leq u \cdot x$ for all $x \in X(0, \dots, 0)$.

Suppose not and let $x \in X(0, \dots, 0)$ be such that $u \cdot x < 0$. Let $y \in X(r_1, \dots, r_n)$. Then $y + ax \in X(r_1, \dots, r_n)$ for every $a \in \mathbb{Q}^+$. From 1 we obtain $s \leq u \cdot (y + ax)$ which, for a is sufficiently large, is a contradiction.

From 3 it follows that if $x \cdot v_i = 0$ for $i = 1, \dots, n$ then $0 = u \cdot x$. We can assume without loss of generality that v_1, \dots, v_n are linearly independent. Then by linear algebra $\sum q_i v_i = u$ for some $q_i \in \mathbb{Q}$. Now fix i and verify that q_i is non-negative. Let x_i such that $x_i \cdot v_j = 1$, if $i = j$, and 0 otherwise. Then $0 \leq x_i \cdot u = q_i$. \square