London Mathematical Society
Lecture Note Series 244

Model Theory of Groups and Automorphism Groups

Edited by
David M. Evans

CAMBRIDGE
UNIVERSITY PRESS

LONDON MATHEMATICAL SOCIETY LECTURE NOTE SERIES

Managing Editor: Professor J.W.S. Cassels, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England

The titles below are available from booksellers, or, in case of difficulty, from Cambridge University Press.

46 p-adic analysis: a short course on recent work, N. KOBLITZ
50 Commutator calculus and groups of homotopy classes, H.J. BAUES
59 Applicable differential geometry, M. CRAMPIN & F.A.E. PIRANI
66 Several complex variables and complex manifolds II, M.J. FIELD
69 Representation theory, I.M. GELFAND *et al*
86 Topological topics, I.M. JAMES (ed)
87 Surveys in set theory, A.R.D. MATHIAS (ed)
88 FPF ring theory, C. FAITH & S. PAGE
89 An F-space sampler, N.J. KALTON, N.T. PECK & J.W. ROBERTS
90 Polytopes and symmetry, S.A. ROBERTSON
92 Representation of rings over skew fields, A.H. SCHOFIELD
93 Aspects of topology, I.M. JAMES & E.H. KRONHEIMER (eds)
94 Representations of general linear groups, G.D. JAMES
96 Diophantine equations over function fields, R.C. MASON
97 Varieties of constructive mathematics, D.S. BRIDGES & F. RICHMAN
98 Localization in Noetherian rings, A.V. JATEGAONKAR
99 Methods of differential geometry in algebraic topology, M. KAROUBI & C. LERUSTE
100 Stopping time techniques for analysts and probabilists, L. EGGHE
104 Elliptic structures on 3-manifolds, C.B. THOMAS
105 A local spectral theory for closed operators, I. ERDELYI & WANG SHENGWANG
107 Compactification of Siegel moduli schemes, C.-L. CHAI
109 Diophantine analysis, J. LOXTON & A. VAN DER POORTEN (eds)
110 An introduction to surreal numbers, H. GONSHOR
113 Lectures on the asymptotic theory of ideals, D. REES
114 Lectures on Bochner-Riesz means, K.M. DAVIS & Y.-C. CHANG
116 Representations of algebras, P.J. WEBB (ed)
118 Skew linear groups, M. SHIRVANI & B. WEHRFRITZ
119 Triangulated categories in the representation theory of finite-dimensional algebras, D. HAPPEL
121 Proceedings of *Groups - St Andrews 1985*, E. ROBERTSON & C. CAMPBELL (eds)
122 Non-classical continuum mechanics, R.J. KNOPS & A.A. LACEY (eds)
128 Descriptive set theory and the structure of sets of uniqueness, A.S. KECHRIS & A. LOUVEAU
129 The subgroup structure of the finite classical groups, P.B. KLEIDMAN & M.W. LIEBECK
130 Model theory and modules, M. PREST
131 Algebraic, extremal & metric combinatorics, M.-M. DEZA, P. FRANKL & I.G. ROSENBERG (eds)
132 Whitehead groups of finite groups, ROBERT OLIVER
133 Linear algebraic monoids, MOHAN S. PUTCHA
134 Number theory and dynamical systems, M. DODSON & J. VICKERS (eds)
135 Operator algebras and applications, 1, D. EVANS & M. TAKESAKI (eds)
137 Analysis at Urbana, I, E. BERKSON, T. PECK, & J. UHL (eds)
138 Analysis at Urbana, II, E. BERKSON, T. PECK, & J. UHL (eds)
139 Advances in homotopy theory, S. SALAMON, B. STEER & W. SUTHERLAND (eds)
140 Geometric aspects of Banach spaces, E.M. PEINADOR & A. RODES (eds)
141 Surveys in combinatorics 1989, J. SIEMONS (ed)
144 Introduction to uniform spaces, I.M. JAMES
145 Homological questions in local algebra, JAN R. STROOKER
146 Cohen-Macaulay modules over Cohen-Macaulay rings, Y. YOSHINO
148 Helices and vector bundles, A.N. RUDAKOV *et al*
149 Solitons, nonlinear evolution equations and inverse scattering, M. ABLOWITZ & P. CLARKSON
150 Geometry of low-dimensional manifolds 1, S. DONALDSON & C.B. THOMAS (eds)
151 Geometry of low-dimensional manifolds 2, S. DONALDSON & C.B. THOMAS (eds)
152 Oligomorphic permutation groups, P. CAMERON
153 L-functions and arithmetic, J. COATES & M.J. TAYLOR (eds)
155 Classification theories of polarized varieties, TAKAO FUJITA
156 Twistors in mathematics and physics, T.N. BAILEY & R.J. BASTON (eds)
158 Geometry of Banach spaces, P.F.X. MÜLLER & W. SCHACHERMAYER (eds)
159 Groups St Andrews 1989 volume 1, C.M. CAMPBELL & E.F. ROBERTSON (eds)
160 Groups St Andrews 1989 volume 2, C.M. CAMPBELL & E.F. ROBERTSON (eds)
161 Lectures on block theory, BURKHARD KÜLSHAMMER
162 Harmonic analysis and representation theory, A. FIGA-TALAMANCA & C. NEBBIA
163 Topics in varieties of group representations, S.M. VOVSI
164 Quasi-symmetric designs, M.S. SHRIKANDE & S.S. SANE
166 Surveys in combinatorics, 1991, A.D. KEEDWELL (ed)
168 Representations of algebras, H. TACHIKAWA & S. BRENNER (eds)
169 Boolean function complexity, M.S. PATERSON (ed)
170 Manifolds with singularities and the Adams-Novikov spectral sequence, B. BOTVINNIK
171 Squares, A.R. RAJWADE

172 Algebraic varieties, GEORGE R. KEMPF
173 Discrete groups and geometry, W.J. HARVEY & C. MACLACHLAN (eds)
174 Lectures on mechanics, J.E. MARSDEN
175 Adams memorial symposium on algebraic topology 1, N. RAY & G. WALKER (eds)
176 Adams memorial symposium on algebraic topology 2, N. RAY & G. WALKER (eds)
177 Applications of categories in computer science, M. FOURMAN, P. JOHNSTONE & A. PITTS (eds)
178 Lower K- and L-theory, A. RANICKI
179 Complex projective geometry, G. ELLINGSRUD *et al*
180 Lectures on ergodic theory and Pesin theory on compact manifolds, M. POLLICOTT
181 Geometric group theory I, G.A. NIBLO & M.A. ROLLER (eds)
182 Geometric group theory II, G.A. NIBLO & M.A. ROLLER (eds)
183 Shintani zeta functions, A. YUKIE
184 Arithmetical functions, W. SCHWARZ & J. SPILKER
185 Representations of solvable groups, O. MANZ & T.R. WOLF
186 Complexity: knots, colourings and counting, D.J.A. WELSH
187 Surveys in combinatorics, 1993, K. WALKER (ed)
188 Local analysis for the odd order theorem, H. BENDER & G. GLAUBERMAN
189 Locally presentable and accessible categories, J. ADAMEK & J. ROSICKY
190 Polynomial invariants of finite groups, D.J. BENSON
191 Finite geometry and combinatorics, F. DE CLERCK *et al*
192 Symplectic geometry, D. SALAMON (ed)
193 Computer algebra and differential equations, E. TOURNIER (ed)
194 Independent random variables and rearrangement invariant spaces, M. BRAVERMAN
195 Arithmetic of blowup algebras, WOLMER VASCONCELOS
196 Microlocal analysis for differential operators, A. GRIGIS & J. SJÖSTRAND
197 Two-dimensional homotopy and combinatorial group theory, C. HOG-ANGELONI, W. METZLER & A.J. SIERADSKI (eds)
198 The algebraic characterization of geometric 4-manifolds, J.A. HILLMAN
199 Invariant potential theory in the unit ball of $\mathbf{C}^n$, MANFRED STOLL
200 The Grothendieck theory of dessins d'enfant, L. SCHNEPS (ed)
201 Singularities, JEAN-PAUL BRASSELET (ed)
202 The technique of pseudodifferential operators, H.O. CORDES
203 Hochschild cohomology of von Neumann algebras, A. SINCLAIR & R. SMITH
204 Combinatorial and geometric group theory, A.J. DUNCAN, N.D. GILBERT & J. HOWIE (eds)
205 Ergodic theory and its connections with harmonic analysis, K. PETERSEN & I. SALAMA (eds)
206 An introduction to noncommutative differential geometry and its physical applications, J. MADORE
207 Groups of Lie type and their geometries, W.M. KANTOR & L. DI MARTINO (eds)
208 Vector bundles in algebraic geometry, N.J. HITCHIN, P. NEWSTEAD & W.M. OXBURY (eds)
209 Arithmetic of diagonal hypersurfaces over finite fields, F.Q. GOUVÊA & N. YUI
210 Hilbert C*-modules, E.C. LANCE
211 Groups 93 Galway / St Andrews I, C.M. CAMPBELL *et al*
212 Groups 93 Galway / St Andrews II, C.M. CAMPBELL *et al*
214 Generalised Euler-Jacobi inversion formula and asymptotics beyond all orders, V. KOWALENKO, N.E. FRANKEL, M.L. GLASSER & T. TAUCHER
215 Number theory 1992–93, S. DAVID (ed)
216 Stochastic partial differential equations, A. ETHERIDGE (ed)
217 Quadratic forms with applications to algebraic geometry and topology, A. PFISTER
218 Surveys in combinatorics, 1995, PETER ROWLINSON (ed)
220 Algebraic set theory, A. JOYAL & I. MOERDIJK
221 Harmonic approximation, S.J. GARDINER
222 Advances in linear logic, J.-Y. GIRARD, Y. LAFONT & L. REGNIER (eds)
223 Analytic semigroups and semilinear initial boundary value problems, KAZUAKI TAIRA
224 Computability, enumerability, unsolvability, S.B. COOPER, T.A. SLAMAN & S.S. WAINER (eds)
225 A mathematical introdcution to string theory, S. ALBEVERIO, J. JOST, S. PAYCHA, S. SCARLATTI
226 Novikov conjectures, index theorems and rigidity I, S. FERRY, A. RANICKI & J. ROSENBERG (eds)
227 Novikov conjectures, index theorems and rigidity II, S. FERRY, A. RANICKI & J. ROSENBERG (eds)
228 Ergodic theory of $\mathbf{Z}^d$ actions, M. POLLICOTT & K. SCHMIDT (eds)
229 Ergodicity for infinite dimensional systems, G. DA PRATO & J. ZABCZYK
230 Prolegomena to a middlebrow arithmetic of curves of genus 2, J.W.S. CASSELS & E.V. FLYNN
231 Semigroup theory and its applications, K.H. HOFMANN & M.W. MISLOVE (eds)
232 The descriptive set theory of Polish group actions, H. BECKER & A.S. KECHRIS
233 Finite fields and applications, S. COHEN & H. NIEDERREITER (eds)
234 Introduction to subfactors, V. JONES & V.S. SUNDER
235 Number theory 1993–94, S. DAVID (ed)
236 The James forest, H. FETTER & B. GAMBOA DE BUEN
237 Sieve methods, exponential sums, and their applications in number theory, G.R.H. GREAVES, G. HARMAN & M.N. HUXLEY (eds)
238 Representation theory and algebraic geometry, A. MARTSINKOVSKY & G. TODOROV (eds)
239 Clifford algebras and spinors, P. LOUNESTO
240 Stable groups, FRANK O. WAGNER
241 Surveys in combinatorics, 1997, R. BAILEY (ed)
242 Geometric Galois actions I, L. SCHNEPS & P. LOCHAK (eds)
243 Geometric Galois actions II, L. SCHNEPS & P. LOCHAK (eds)
244 Model theory of groups and automorphism groups, D. M. EVANS (ed)

London Mathematical Society Lecture Note Series. 244

Model Theory of Groups and Automorphism Groups

Blaubeuren, August 1995

Edited by

David M. Evans
University of East Anglia

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521589550

First published 1997

A catalogue record for this publication is available from the British Library

ISBN-13 978-0-521-58955-0 paperback
ISBN-10 0-521-58955-X paperback

Transferred to digital printing 2005

Contents

Preface

The articles in this volume represent the invited lectures at the RESMOD Summer School on Model Theory of Groups and Automorphism Groups held in Blaubeuren, Germany, from 31 July to 5 August 1995. This was an EC-funded meeting directed at graduate students and researchers in Model Theory and Algebra and consisted mainly of invited lectures surveying various recent interactions between model theory and and other branches of mathematics, notably group theory.

RESMOD is the acronym for the European Human Capital and Mobility Network on Model Theory and Applications coordinated by the Équipe de Logique Mathématique at Université Paris 7. The programme committee for the meeting consisted of Wilfrid Hodges, Daniel Lascar and Dugald Macpherson. The meeting took place at the Heinrich Fabri Institut of the University of Tübingen, and the local organisers were Ulrich Felgner and Frieder Haug.

David M. Evans,
School of Mathematics,
University of East Anglia,
Norwich NR4 7TJ,
England.
February 1997.

Introduction

The articles in this volume demonstrate the wide variety of interactions between algebra (particularly group theory) and current research in model theory. On the one hand, the analysis of direct questions about the first-order theories of classes of algebraic structures requires an interplay between model-theoretic and algebraic methods, and often such questions also evolve into ones which are interesting from a purely algebraic viewpoint. More indirectly, the model-theoretic analysis of classes of structures using some of the latest developments of model theory (particularly stability theory) has recently resulted in a wave of new applications of model theory to other parts of mathematics.

Alongside these developments there has been considerable interaction between model theory and the study of infinite permutation groups. Automorphism groups of model-theoretically interesting structures have provided a rich supply of examples and problems for the permutation group theorists, and the study of automorphism groups has been a crucial tool in certain model-theoretic questions.

Readers can judge for themselves the extent to which the articles in this volume fit into this pattern, but I shall give a brief sketch of them, emphasising the interactions between model theory and other parts of mathematics.

The article by Evans, Ivanov and Macpherson is a survey largely concerned with a question that originated in studying the fine detail of totally categorical structures, but which is now seen (and studied) as a problem about infinite permutation groups. The techniques in the papers by Lascar and Evans are model-theoretic in flavour, but the applications are to the study of the automorphism group of the field of complex numbers, and the papers are written without using model-theoretic terminology. The papers by Chatzidakis and Hodges form a survey of recent work on the model theory of pseudo-finite fields and in particular give surprising applications of these results (due to Hrushovski and Pillay) to the subgroup structure of Chevalley groups over prime fields.

Cameron's paper draws together strands from model theory, permutation groups and combinatorics. It studies a graded algebra which can be associated to any countable $\aleph_0$-categorical structure, and which is also significant in enumerative combinatorics.

The papers by Boffa, Oger and Chiswell are surveys of various aspects of the model theory of particular classes of groups. The questions considered start out as model-theoretic ones (equivalence of formulas, elementary equivalences at various levels of quantifier complexity *et cetera*), but also develop into questions which are interesting from a purely group-theoretic viewpoint. The techniques are a mixture of group theory and model theory (notably ultraproducts). The paper by Pfander follows on from Chiswell's article and gives new results on the finite presentability of groups with the same existential and universal theory as the non-abelian free groups.

The paper by Burke and Prest is a contribution to the theory of modules: an area where model-theoretic methods have had a significant impact on the algebraic

theory. Finally, the paper by Kim is a survey of recent work (of Kim and Pillay) on Shelah's notion of *simplicity* of a first-order theory. In such theories a good notion of independence (*forking*) can be described and various unstable algebraic structures have recently been shown to have simple theories (for example, pseudo-finite fields).

In the remainder of this introduction I will give some background material and pointers to the literature which may be helpful to the non-specialist reader. This will be very brief, not least because there are already several excellent swift introductions to the area in print: for example, the opening sections of [1] and (more comprehensively) the article [6].

Section 2 below is based on notes of Dugald Macpherson originally prepared as an appendix to the 'Finite covers' paper in this volume.

1 Model theory

The books by Chang and Keisler [2] and Hodges [5] give a thorough treatment of model theory excluding stability theory. A good introduction to the latter can be found in the the book by Pillay [7], and [8] has many of the more recent developments.

1.1 First-order languages and structures

In a first-order language one has an alphabet of symbols and certain finite sequences of these symbol (the formulas of the language) are the objects of interest. The symbols are connectives $\wedge$ (*and*), $\vee$ (*or*), $\neg$ (*not*); quantifiers $\forall$ and $\exists$; punctuation (parentheses and commas); variables; and constant, relation and function symbols, with each of the last two coming equipped with a finite 'arity' specifying how many arguments it has. The number of these constant, relation and function symbols (together with their arities) is referred to as the *signature* of the language.

The *terms* of the language are built inductively. Any variable or constant symbol is a term and if f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is also a term (all terms are built in this way).

Now we can build the *formulas* of the language. Again, this is done inductively. If R is an n-ary relation symbol in the language and $t_1, \ldots, t_n$ are terms then $R(t_1, \ldots, t_n)$ is a formula (an *atomic* formula). If ϕ, ψ are formulas and x a variable, then $(\phi) \wedge (\psi)$, $(\phi) \vee (\psi)$, $\neg(\phi)$, $\forall x(\phi)$, $\exists x(\phi)$ are formulas (of higher 'complexity'). A formula not involving any quantifiers is called *quantifier free* or *open*. There is a natural notion of a *free variable* in a formula, and when we write a formula as $\phi(x_1, \ldots, x_m)$ we mean that its free variables are amongst the variables $x_1, \ldots, x_m$. A formula with no free variables is called a *sentence*. For more details the reader could consult ([2], Section 1.3) or ([5], Section 2.1).

If L is a first-order language then an *L-structure* consists of a set M equipped with a constant (that is, a distinguished element of M), n-ary relation (that is, a subset of M^n), and n-ary function $M^n \to M$ for each constant symbol and

n-ary relation and function symbol in L. If $\phi(x_1, \ldots, x_m)$ is an L-formula and $a_1, \ldots, a_m \in M$ then one can 'read' $\phi(a_1, \ldots, a_m)$ as a statement about the behaviour of $a_1, \ldots, a_m$ and these constants, relations and functions (interpreting each constant, relation or function symbol as the corresponding constant, relation or function of M), which is either true or false. If it is true, then we write

$$M \models \phi(a_1, \ldots, a_m).$$

All of this can of course be made completely precise (defined inductively on the complexity of ϕ): see ([2], Section 1.3) and ([5], Section 2.1) again. We shall always have $=$ as a binary relation symbol in L and interpret it as true equality in any L-structure.

If Φ is a set of L-sentences and M an L-structure we say that M is a *model* of Φ (and write $M \models \Phi$) if every sentence in Φ is true in M. If there is a model of Φ we say that Φ is *consistent*. The set of L-sentences true in M is called the *theory* of M. Two L-structures M_1 and M_2 are *elementarily equivalent* if they have the same theory. This is written as $M_1 \equiv M_2$. Thus in this case the structures M_1 and M_2 cannot be distinguished using the language L. The following basic result of model theory shows that one should not expect first-order languages to be able to completely describe infinite structures.

Theorem 1.1 (Löwenheim-Skolem) *Let L be a first-order language with signature of cardinality λ. Let μ, ν be cardinals with $\mu, \nu \geq \max(\lambda, \aleph_0)$, and suppose M_1 is an L-structure with cardinality μ. Then there exists an L-structure M_2 elementarily equivalent to M_1 and of cardinality ν.*

The 'upward' part of this result (where $\nu \geq \mu$) follows easily from the fundamental theorem of model theory:

Theorem 1.2 (The Compactness Theorem) *Let L be a first-order language and Φ a set of L-sentences. If every finite subset of Φ is consistent, then Φ is consistent.*

The original version of this is due to Gödel (1931). Proofs (using a method due to Henkin (1949)) can be found in ([2], Theorem 3.2.2) and ([5], Theorem 6.1.1). Algebraists may prefer the proof using ultraproducts ([2], Corollary 4.1.11) and the theorem of Los ([2], Theorem 4.1.9, or [5], Theorem 9.5.1).

If M, N are L-structures with $M \subseteq N$ and the distinguished relations, functions (and constants) of N extend those of M, then we say that M is a *substructure* of N. If also for every L-formula $\phi(x_1, \ldots, x_m)$ and $a_1, \ldots, a_m \in M$ we have

$$M \models \phi(a_1, \ldots, a_m) \Leftrightarrow N \models \phi(a_1, \ldots, a_m)$$

then we say that M is an *elementary* substructures of N (and that N is an elementary extension of M) and write $M \preceq N$. A stronger version of the Löwenheim-Skolem Theorem (1.1) is true: the smaller of M_1, M_2 may be taken to be an elementary substructure of the larger. Proofs can be found in ([2], Theorems 3.1.5 and 3.1.6) and ([5], Corollaries 3.1.5 and 6.1.4).

1.2 Definable sets; types

Suppose L is a first-order language and M an L-structure. Let $n \in \mathbb{N}$. A subset A of M^n is called (parameter) *definable* if there exist $b_1, \ldots, b_m \in M$ and an L-formula $\phi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ with

$$A = \{\bar{a} \in M^n : M \models \phi(\bar{a}, \bar{b})\}.$$

If the parameters $\bar{b}$ can be taken from the subset $X \subseteq M$ then A is said to be X-definable. The union of the finite X-definable subsets of M is called the *algebraic closure* of X, denoted by $\mathrm{acl}(X)$, and the union of the X-definable singleton subsets of M is the *definable closure* of X, denoted by $\mathrm{dcl}(X)$. It is not hard to check that both of these are indeed closure operations on M.

So the definable subsets of M^n are the ones which can be described using L-formulas (and parameters). Conversely one could take a particular n-tuple $\bar{a} \in M^n$ and a set of parameters $A \subseteq M$ and ask what the language L can say about $\bar{a}$ (in terms of A and M). This gives the notion of the *type* of $\bar{a}$ over A, which by definition is

$$\mathrm{tp}_M(\bar{a}/A) = \{\phi(x_1, \ldots, x_n, b_1, \ldots, b_m) : b_1, \ldots, b_m \in A,\ M \models \phi(\bar{a}, \bar{b})\}$$

(the subscript M is dropped if this is clear from the context). It is sometimes useful to consider the type of $\bar{a}$ (over A) using only certain L-formulas. For example, for the *quantifier free type* of $\bar{a}$ over A one takes only quantifier free ϕ in the above definition. It is also possible to define the type of an infinite sequence of elements of M. The reader can consult ([5], Section 6.3) or ([2], Section 2.3) for further details here.

More generally, a (complete) n-type over A is a set of L-formulas with parameters from A equal to $\mathrm{tp}_N(\bar{a}/A)$ for some elementary extension N of M and some $\bar{a} \in M^n$. There is no reason to suppose, for arbitrary M and A, that this type should be *realised* in M, that is, there exists $\bar{a}' \in M^n$ with $\mathrm{tp}_M(\bar{a}'/A) = \mathrm{tp}_N(\bar{a}/A)$. For example, this would clearly be impossible if $A = M$ and $\bar{a} \notin M^n$. However, it can happen that for some infinite cardinal κ if $|A| < \kappa$ then every complete n-type over A is realised in M: in this case M is called *κ-saturated,* and if $\kappa = |M|$ then M is *saturated.* The reader should consult ([2], Section 2.3) and then ([2], Chapter 5) and ([5], Chapter 10) for more on this subject as the need arises.

1.3 Interpreted structures; imaginary elements

Some structures can be built out of others in a definable way. The classical example is the construction of the field of rational numbers from the ring of integers. Another example is algebraic groups over a particular field.

Formalising this leads to the notion of an *interpretation* of one structure in another. Suppose K and L are first-order languages, M a K-structure and N an L-structure. We say that N is *interpretable* in M if for some $n \in \mathbb{N}$ there exist:

1. a $\emptyset$-definable subset D of M^n;

2. a $\emptyset$-definable equivalence relation E on D;

3. a bijection $\gamma : N \to D/E$

such that for every $\emptyset$-definable subset R of N^m the subset of M^{mn} given by

$$\hat{R} = \{(\bar{a}_1, \ldots, \bar{a}_m) \in (M^n)^m : (\gamma^{-1}(\bar{a}_1/E), \ldots, \gamma^{-1}(\bar{a}_m/E)) \in R\}$$

is $\emptyset$-definable in M.

Thus the set N can be identified with a $\emptyset$-definable subset of M^n factored by a $\emptyset$-definable equivalence relation, and with this identification all of the L-structure on N can be derived from the K-definable structure on M. There is a considerable amount of redundancy in the definition: it is only necessary to have $\emptyset$-definability of $\hat{R}$ when R is a distinguished constant or relation, or the graph of a distinguished function.

If E is simply equality on D then we say that N is *definable* in M. If also $D = M$ then we say that N is a *reduct* of M (so N just consists of M with some of its definable structure forgotten). It is also possible to formulate a notion of interpretation using parameters. The reader should consult ([5], Section 5.3) for further information on interpretations.

Equivalence classes in D/E as above are referred to as *imaginary* elements of M. Taking the set of all imaginary elements (as D and E range over all $\emptyset$-definable sets and equivalence relations) gives us the set M^{eq}. We wish to regard this as a first-order structure, so we extend the language K of M in a canonical way (to a first-order language K^{eq}), and part of the K^{eq}-theory of M^{eq} describes how the imaginary elements correspond in a $\emptyset$-definable way to the original K-structure M. The reader can consult ([5], Section 4.3) for the precise details of how to do all of this. Once we have this concept, it makes sense to extend notions such as 'parameter definable', 'types', 'algebraic closure' etc. to subsets of M^{eq}. Again we refer the reader to ([5]) for further details if the need arises.

2 Permutation groups

Most of what is said here can be found in more detail in ([5], Section 4.1), ([1], Chapters 1 and 2) and ([6]). A general reference on permutation groups which usefully contains material on automorphism groups of infinite structures is [3].

2.1 Actions and orbits

Let X be any set. The group of all permutations of X is called the *symmetric group* on X and is denoted by $\mathrm{Sym}(X)$. A *permutation group* on X is a subgroup of this. The image of the element $x \in X$ under the permutation $g \in \mathrm{Sym}(X)$ is denoted by gx. More generally, an *action* of a group G on X is a function $\alpha : G \times X \to X$ such that for all $x \in X$ and $g, h \in G$ we have $\alpha(1, x) = x$ and $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$. It is easy to see that this is equivalent to the map $g \mapsto \alpha(g, -)$ being a homomorphism from G into $\mathrm{Sym}(X)$. Thus each element of

G induces a permutation of X and a product of elements in the group induces the corresponding product of permutations. Henceforth, we shall also denote $\alpha(g,x)$ by gx if the action is clear from the context. (It should be noted that some people write their actions on the right, and so would write xg instead of gx, with corresponding changes needed for other pieces of notation. This rarely causes confusion.)

Given an action of a group G on a set X the *orbits* are the equivalence classes under the equivalence relation $\sim$ on X, where $x_1 \sim x_2$ if there is $g \in G$ with $gx_1 = x_2$. We say that the action is *transitive* on X if there is a unique orbit. One way of manufacturing transitive actions of a group G is via *coset spaces*. Let H be a subgroup of G and let $Y = \{gH : g \in G\}$ be the set of left cosets of H in G. Define an action of G on Y by setting $\alpha(g_1, g_2H) = g_1g_2H$. Clearly this is a transitive action. However, in a strong sense this gives us all transitive actions of G. Suppose G acts transitively on a set X. Let $x \in X$ and let $H = \{g \in G : gx = x\}$ (the *stabiliser* of x in G, usually denoted by G_x). Then the map $\theta : Y \to X$ given by $\theta(gH) = gx$ is a well-defined bijection and for all $y \in Y$ and $g \in G$ we have $\theta(gy) = g\theta(y)$. Thus the actions of G on X and Y are *equivalent*. This is the *Orbit-Stabiliser Theorem*.

Out of any given action of a group G on a set X we can produce various other actions of G. For example, if $Y \subseteq X$ is a union of G-orbits, then one can simply restrict the action to Y. Also, suppose there is a G-invariant equivalence relation on X. Then one can consider the action of G on the set of equivalence classes. Next, suppose k is a positive integer. Then there is an induced action of G on X^k (given by $g(x_1, \ldots, x_k) = (gx_1, \ldots, gx_k)$), and an equally natural action on $X^{\{k\}}$, the set of k-sets from X. We say that G is *k-transitive* if, in the first action, all k-tuples of distinct elements lie in the same orbit. We say G is *k-homogeneous* if it is transitive on $X^{\{k\}}$. The original action is *highly* transitive (or homogeneous) if it is k-transitive (or k-homogeneous) for all $k \in \mathbb{N}$.

2.2 Automorphism groups and topological groups

Suppose L is a first-order language and M an L-structure. By an *automorphism* of M we mean a permutation of M which preserves each of the distinguished constants, relations and functions of M. The set of these forms a subgroup of $\mathrm{Sym}(M)$, called the *automorphism group* of M, and is denoted by $\mathrm{Aut}(M)$. It is clear that if $A \subseteq M$, then $\mathrm{Aut}(M/A) = \{g \in \mathrm{Aut}(M) : ga = a \; \forall a \in A\}$ is a subgroup of $\mathrm{Aut}(M)$ which stabilises any A-definable subset of M^k. Furthermore, if $\bar{b} \in M^k$ and $g \in \mathrm{Aut}(M/A)$ then $\mathrm{tp}_M(\bar{b}/A) = \mathrm{tp}_M(g\bar{b}/A)$. More subtly, if M is saturated then the converse is also true: if $\bar{b}$ and $\bar{b}'$ have the same type over A and $|A| < |M|$ then $\bar{b}$ and $\bar{b}'$ are in the same $\mathrm{Aut}(M/A)$-orbit (for example, see ([5], Corollary 10.4.12), or the proof of ([2], Theorem 2.3.9) if M is countable). Note that any element of $\mathrm{Aut}(M)$ induces an automorphism of M^{eq}.

Conversely, if G is a permutation group on a set X then there is a natural first-order structure with domain X, on which G acts as a group of automorphisms (with, for each $n \in \mathbb{N}$, the same orbits on n-tuples as the full automorphism group). For each orbit Ω of G on X^n (as n ranges through $\mathbb{N}$) introduce an n-ary relation

symbol R_Ω, interpreted on X by the orbit Ω. The corresponding language is known as the *canonical language*, and the structure on X as the *canonical structure*.

Suppose that X is any set. Then there is a natural topology on $\mathrm{Sym}(X)$ which makes it into a topological group (so multiplication and inversion are continuous maps). The open sets are unions of cosets of pointwise stabilisers of finite subsets of X. We then make any permutation group G on X into a topological group by giving it the relative topology. To put this another way, if $g \in G$ then the cosets $gG_{(F)}$ as F ranges over the finite subsets of X form a basis of open neighbourhoods of g in G, where $G_{(F)} = \{h \in G : hx = x \ \forall x \in F\}$. Clearly this topology is Hausdorff. In fact, as any open coset is closed, the topology is totally disconnected. It it separable if and only if X is countable and discrete if and only if $G_{(F)} = \{1\}$ for some finite $F \subseteq X$. It is not hard to show that a closed subgroup G of $\mathrm{Sym}(X)$ is compact if and only if all of its orbits on X are finite.

For us, the most important fact about this topology will be that a subgroup of $\mathrm{Sym}(X)$ is closed if and only if it is the full group of automorphisms of a first-order structure with domain X. In fact, if $G \leq \mathrm{Sym}(X)$ then the automorphism group of the canonical structure of G on X is the closure of G in $\mathrm{Sym}(X)$.

If X is countable the topology is metrisable: enumerate X as $(x_n : n \in \mathbb{N})$, and define a metric d on $\mathrm{Sym}(X)$ by putting, for distinct $g, h \in \mathrm{Sym}(X)$, the distance $d(g,h)$ to be $1/m$ where m is as large as possible such that g agrees with h, and g^{-1} with h^{-1}, on x_l for all $l < m$. Thus, $\mathrm{Sym}(X)$ becomes a complete metric space with a countable basis of open sets (a *Polish space*).

2.3 $\aleph_0$-categoricity

For saturated structures M there is a strong connection between what is definable in a first-order way and the automorphism group: over small subsets of M orbits equate to types. For countable $\aleph_0$-categorical structures the connection is even stronger, and automorphism groups of $\aleph_0$-categorical structures are probably the most widely studied class of infinite permutation groups.

If M an L-structure and κ an infinite cardinal we say that M is *κ-categorical* if its theory has a model of size κ and all such are isomorphic. The case $\kappa = \aleph_0$ (that is, countably infinite κ) has a group-theoretic formulation. Say that a permutation group G on an infinite set X is *oligomorphic* if it has finitely many orbits on X^k for all positive integers k. Then the theorem of Engeler, Ryll-Nardzewski, and Svenonius asserts that, for a countably infinite structure M, the following are equivalent:

1. M is $\aleph_0$-categorical;
2. $\mathrm{Aut}(M)$ acts oligomorphically on M;
3. for every $n \in \mathbb{N}$ there are only finitely many n-types over $\emptyset$ (realised in elementary extensions of M).

Proofs can be found in ([2], Theorem 2.3.13) or ([5], Theorem 7.3.1). There is a translation between the group-theoretic and model-theoretic terminology in this

case. The countable model M is saturated so (realisations in M of) n-types over a finite subset A of M are exactly $\mathrm{Aut}(M/A)$-orbits on M^n. But as there are only finitely many of these, each of them is actually A-definable. So a subset of M^n is A-definable if and only if it is invariant under $\mathrm{Aut}(M/A)$. It then follows from the Orbit-Stabiliser Theorem that $a \in M$ is in the algebraic closure of A if and only if $\mathrm{Aut}(M/A \cup \{a\})$ is of finite index in $\mathrm{Aut}(M/A)$. The same is true in M^{eq}. Moreover stabilisers of elements of M^{eq} are exactly the open subgroups of $\mathrm{Aut}(M)$ (use ([1], 1.2, Exercise 4)).

Obvious examples of countable $\aleph_0$-categorical structures include a pure set, the set of unordered pairs from a pure set (with a natural induced graph structure, two 2-sets adjacent if they intersect in a singleton), the rationals as an ordered set and the countable atomless boolean algebra. The paper [4] is a survey of various ways of constructing $\aleph_0$-categorical structures and classification results relating to them. The book [1] contains a large amount of information about automorphism groups of these structures.

References

[1] Peter J. Cameron, *Oligomorphic Permutation Groups,* London Mathematical Society Lecture Notes Series, 152, Cambridge University Press, Cambridge, 1990.

[2] C. C. Chang, H. Jerome Keisler, *Model Theory* (3rd edition), North-Holland, Amsterdam, 1990.

[3] John D. Dixon, Brian Mortimer, *Permutation Groups,* Springer Graduate Texts in Mathematics 163, Springer, New York, 1996.

[4] David M. Evans, 'Examples of $\aleph_0$-categorical structures', in *Automorphisms of First-Order Structures,* eds. R. Kaye and D. Macpherson, pp. 33–72, Oxford University Press, Oxford, 1994.

[5] Wilfrid Hodges, *Model Theory,* Cambridge University Press, Cambridge, 1993.

[6] Richard Kaye and Dugald Macpherson, 'Models and groups', in *Automorphisms of First-Order Structures,* eds. R. Kaye and D. Macpherson, pp. 3–31, Oxford University Press, Oxford, 1994.

[7] Anand Pillay, *An Introduction to Stability Theory,* Oxford Logic Guides 8, Oxford University Press, Oxford, 1983.

[8] Anand Pillay, *Geometric Stability Theory,* Oxford Logic Guides 32, Oxford University Press, Oxford, 1996.

Finite Covers

David M. Evans, Dugald Macpherson, Alexandre A. Ivanov

Contents

0 Outline of the Notes

These notes examine a technique for building new structures from simpler ones. The original motivation for this construction is Zil'ber's 'ladder theorem' (Theorem 1.6.4 here), which describes how totally categorical structures are built from strictly minimal sets by a sequence of covers. Similar results exist for several other classes of structures, such as $\aleph_1$-categorical structures, $\aleph_0$-categorical ω-stable structures, and smoothly approximated structures.

We will concentrate on finite covers of countable $\aleph_0$-categorical structures, and we often describe structures entirely by their automorphism groups, without reference to any particular language (in the $\aleph_0$-categorical case this is justified by the Ryll-Nardzewski theorem). The following terminology is convenient.

If Ω is a set then we regard $\mathrm{Sym}(\Omega)$, the symmetric group on Ω, as a topological group with a base of open sets being given by cosets of pointwise stabilisers of finite subsets of Ω. Then a *permutation structure* is a pair $\langle W; G\rangle$ where W is a non-empty set (the *domain*), and G is a closed subgroup of $\mathrm{Sym}(W)$ (the group of *automorphisms*). We shall usually write $G = \mathrm{Aut}(W)$ and refer simply to 'the permutation structure W.' If A is a subset of W and B a subset of W (or more generally of some set on which $\mathrm{Aut}(W)$ is acting in an obvious way), then $\mathrm{Aut}(A/B)$ denotes the permutations of A which extend to elements of $\mathrm{Aut}(W)$ fixing every element of B. We shall write permutations on the left of the elements of W.

Permutation structures are all obtained by taking automorphism groups of first-order structures on W, and we often regard a first-order structure as a permutation structure without explicitly saying so. When we do this, the group of automorphisms for the permutation structure is, of course, just the automorphism group of the first-order structure. We can now define a finite cover (a model-theoretic definition is given in 1.1.2).

Definition 0.0.1 If C, W are permutation structures, then a finite-to-one surjection $\pi : C \to W$ is a *finite cover* if its fibres form an $\mathrm{Aut}(C)$-invariant partition of C, and the induced map $\mu : \mathrm{Aut}(C) \to \mathrm{Sym}(W)$ given by $\mu(g)w = \pi(g\pi^{-1}(w))$ for $g \in \mathrm{Aut}(C)$ and $w \in W$ has image $\mathrm{Aut}(W)$. We refer to μ as the *restriction* map. The *kernel* of the finite cover is $\ker\mu = \mathrm{Aut}(C/W)$.

The main problem which concerns us is:

> **The Cover Problem:** For a given $\aleph_0$-categorical structure W, describe its finite covers.

An overview of how the material in this paper relates to this problem can be found in Section 1.7, after we have given the basic definitions, examples and results. For the rest of this section, we simply describe the structure of these notes and highlight some of the principal results in each section.

Section 1 first gives the basic definitions and some 'naturally occuring' examples of covers. We discuss notions closely related to finite covers, notably symmetric extensions, and give some of the basic theory, sometimes in this wider context. Finally we review some of the model-theoretic background to the cover problem.

Three general constructions of finite covers are described in Section 2: free covers, digraph coverings and coverings of two-graphs. We show that free covers are uniquely determined by choice of fibre and binding groups, and so we have a satisfactory classification of these. The material on digraph and two-graph coverings is suggested by ideas from topology (covering spaces) and finite combinatorics. Both constructions provide examples of finite covers with finite kernels.

In Section 3 we give some preliminary results on finite covers. We then give various ways of dividing up the general cover problem and make various reductions which show that we should focus on some special types of covers (minimal, superlinked and abelian kernel). Any finite cover is an expansion of a free finite cover with the same fibre and binding groups, and we aim for classification up to conjugacy within the automorphism group of this free cover. On the other hand, any finite cover is a reduct of a minimal cover. We show that the kernel of a minimal cover of an $\aleph_0$-categorical structure is nilpotent, and thereby reduce certain problems to consideration of finite covers with abelian kernels.

Finite covers whose kernels are finite are analysed in Section 4. We show that in some cases these can all be described in terms of digraph coverings. In some other cases not covered by these results, a careful analysis of the example of a vector space covering its projective space provides a different answer. The techniques and notions in Section 4 parallel very clearly some ideas from stability theory (strong types, stationarity, and distinguished extensions of types). In Section 5 we amplify further on this, and consider the results of Section 4 from this viewpoint.

In Section 6 we consider finite covers with abelian kernels. Following the approach of Ahlbrandt and Ziegler we divide the problem into two parts: describe the possibilities for the kernels, then work out what the possible covers can be with each particular kernel. For the first part we outline how Pontriagin duality can sometimes be useful. For the second part, we describe the construction of the cohomology group H^1_c which parametrises the covers with a given kernel. We use these ideas to show how results on the cohomology and representation theory of finite groups can be used in our context.

Section 7 contains further results which can be used to calculate cohomology groups. These are all standard results from cohomology of discrete groups, adapted to our purposes. We show how these can be used to prove finiteness of H^1_c, given additional constraints on W.

Section 8 contains some open problems and questions which occurred to us during the writing of this paper.

1 Introduction to covers

1.1 Definitions

We give the basic definitions associated with permutation structures and finite covers. We suggest that the reader skims over them quickly and refers back when necessary.

1.1.1 Permutation structures

If W_1 and W_2 are sets of the same cardinality then any bijection $\phi : W_1 \to W_2$ induces an isomorphism $f_\phi : \mathrm{Sym}(W_1) \to \mathrm{Sym}(W_2)$. We say that permutation structures $\langle W_1; G_1 \rangle$ and $\langle W_2; G_2 \rangle$ are *isomorphic* if for some bijection ϕ we have $f_\phi(G_1) = G_2$. (As pointed out to us by Martin Ziegler, this produces a slight conflict in terminology: the group of isomorphisms from a permutation structure $\langle W; G \rangle$ to itself is actually the normaliser in $\mathrm{Sym}(W)$ of G, so it might be more correct to refer to *this* as the 'automorphism group of the permutation structure,' rather than G.)

Two permutation structures are *bi-interpretable* if their automorphism groups are isomorphic as topological groups. If the permutation structures arise from countable $\aleph_0$-categorical structures there is a model-theoretic interpretation of this notion due to G. Ahlbrandt and M. Ziegler ([2]: see also Section 7 of [42]). The following useful observation is due to E. Hrushovski ([36]).

Lemma 1.1.1 *A permutation structure $\langle W; G \rangle$ such that G has finitely many orbits on W is bi-interpretable with a transitive permutation structure $\langle W_1; G_1 \rangle$.*

Proof. Let $\bar{x}$ be a finite tuple of elements from W containing (at least) one element from each G-orbit. Let W_1 be the orbit under G of $\bar{x}$. We get a natural continuous, injective homomorphism $G \to \mathrm{Sym}(W_1)$, and it is easy to see that the image G_1 of this is closed in $\mathrm{Sym}(W_1)$. The inverse map $G_1 \to G$ is also continuous, and so we have the result. □

Related to this construction is the notion of a *Grassmannian* of a transitive permutation structure W. First recall that if W has the property that $\mathrm{Aut}(W/X)$ has finitely many finite orbits for all finite subsets X of W then we define the *algebraic closure* $\mathrm{acl}(X)$ of X to be the union of the finite $\mathrm{Aut}(W/X)$-orbits. This is a closure operation on the finite subsets of W. If A is a finite algebraically closed subset of W then the Grassmannian $\mathrm{Gr}(W; A)$ is the permutation structure having domain $W_A = \{gA : g \in \mathrm{Aut}(W)\}$ and automorphism group those permutations induced on this set by $\mathrm{Aut}(W)$. To see that this is a closed subgroup of $\mathrm{Sym}(W_A)$ observe that, as in Lemma 1.1.1, the group of permutations induced by $\mathrm{Aut}(W)$ on the orbit of an enumeration of A is closed, and there is an invariant finite-to-one map from this orbit to W_A, so what we want follows from Lemma 1.4.2. If $\mathrm{Aut}(W)$ acts faithfully on W_A then $\mathrm{Gr}(W; A)$ is bi-interpretable with W.

We shall frequently employ the following terminology. Suppose C_0 and C are permutation structures with the same domain, and $\mathrm{Aut}(C) \leq \mathrm{Aut}(C_0)$. Then we say that C_0 is a *reduct* of C, or C is an *expansion* of C_0. We use the adjective *proper* to indicate that $\mathrm{Aut}(C) < \mathrm{Aut}(C_0)$.

1.1.2 Finite covers

We first give the model-theoretic definition of *finite cover*. In practice, however, we will use the group-theoretic translation of this given in the opening remarks

(0.0.1).

Definition 1.1.2 Let C and W be first-order structures. A finite-to-one surjection $\pi : C \to W$ is a *finite cover* of W if there is a 0-definable equivalence relation E on C whose classes are the fibres of π, and any relation on W^n (respectively, C^n) which is 0-definable in the 2-sorted structure (C, W, π) is already 0-definable in W (respectively, C).

Observe that a finite cover $\pi : C \to W$ induces a homomorphism

$$\mu : \mathrm{Aut}(C) \to \mathrm{Aut}(W),$$

given by putting $\mu(g)(w) = \pi(g\pi^{-1}(w))$ for all $g \in \mathrm{Aut}(C)$ and $w \in W$. In fact, if W is countable $\aleph_0$-categorical, then the above definition of a finite cover is equivalent to saying that the fibres of π are the classes of an $\mathrm{Aut}(C)$-invariant equivalence relation on C, and the map $\mathrm{Aut}(C) \to \mathrm{Aut}(W)$ induced by π has image $\mathrm{Aut}(W)$ (Lemma 1.4.2 below ensures that Definition 1.1.2 implies the surjectivity), and so this agrees with what was given as Definition 0.0.1. We refer to μ as the *restriction* homomorphism.

Suppose that $\pi : C \to W$ is a finite cover. Then $\mathrm{Aut}(C)$ has a normal subgroup K, the *kernel* of the cover, defined by

$$K := \{g \in \mathrm{Aut}(C) : \pi(x) = \pi(gx) \text{ for all } x \in C\},$$

(so also the kernel of the restriction homomorphism $\mathrm{Aut}(C) \to \mathrm{Aut}(W)$). We have a short exact sequence

$$1 \to K \to \mathrm{Aut}(C) \xrightarrow{\mu} \mathrm{Aut}(W) \to 1.$$

The cover *splits* if K has a closed complement in $\mathrm{Aut}(C)$, that is, there is a closed subgroup H of $\mathrm{Aut}(C)$ such that $KH = \mathrm{Aut}(C)$ and $K \cap H = 1$. Equivalently, C is a reduct of a cover of W with trivial kernel (namely, a structure with automorphism group H).

For each $a \in W$ let $C(a)$ denote the fibre above a, that is $\{x \in C : \pi(x) = a\}$. We also define, for any $a \in W$, the *fibre group* of the cover at a as the permutation group induced by $\mathrm{Aut}(C)$ on $C(a)$. The *binding group* at a is a normal subgroup of the fibre group, and is the permutation group induced on a fibre $C(a)$ by the kernel K. Clearly, if $\mathrm{Aut}(W)$ acts transitively on W then all of the fibre groups are isomorphic as permutation groups, as are the binding groups. We refer to these as the fibre and binding groups of the cover. If these are unequal, we say that the cover is *twisted.*

We mention some special kinds of covers. We say that $\pi : C \to W$ is *free* if $\mathrm{Aut}(C/W) = \Pi_{w \in W} \mathrm{Aut}(C(w)/W)$, that is, the kernel is the full direct product of the binding groups (so as big as possible). At the other extreme, the cover is *trivial* if its kernel $\mathrm{Aut}(C/W)$ is the trivial group (this differs from the terminology in [3] and [4] where 'trivial' means 'split'). A *principal* cover $\pi : C \to W$ is a free finite

cover where the fibre and binding groups at each point are equal. So the kernel of a principal cover is the direct product of all the fibre groups.

If C, C' are permutation structures with the same domain and $\pi : C \to W$ and $\pi' : C' \to W$ are finite covers with $\pi(c) = \pi'(c)$ for all $c \in C = C'$ then we say that π' is a *covering expansion* of π if $\mathrm{Aut}(C') \leq \mathrm{Aut}(C)$.

We say that finite covers $\pi_1 : C_1 \to W$ and $\pi_2 : C_2 \to W$ are *isomorphic* if there exists a bijection $\phi : C_1 \to C_2$ which sends the set of fibres of π_1 to the set of fibres of π_2 and such that the induced map $f_\phi : \mathrm{Sym}(C_1) \to \mathrm{Sym}(C_2)$ (as in Section 1.1.1) sends $\mathrm{Aut}(C_1)$ to $\mathrm{Aut}(C_2)$. If additionally $\phi(\pi_1^{-1}(w)) = \pi_2^{-1}(w)$ for all $w \in W$ then we say that π_1 and π_2 are isomorphic *over* W.

1.2 Examples

We give some examples of cover-like constructions, which will be important later.

Example 1. This is the crudest kind of cover. Let W be any structure, and C have domain $W \times \{0,1\}$, and regard C as a cover of W via the map $(w,i) \overset{\pi}{\mapsto} w$. Assume there is no other structure on W. If W is an L-structure, we may regard C as an $L \cup \{E\}$-structure, where E is a binary relation interpreted as the equivalence relation given by the fibres of π, and each relation of L holds of a tuple in C if and only if it holds of its image under π. Then C is a cover (a 'double cover') of W, and $\mathrm{Aut}(C) = Z_2 \,\mathrm{Wr}\, \mathrm{Aut}(W)$ (the unrestricted wreath product in its natural imprimitive action, with Z_2 denoting the cyclic group of order 2). The kernel is the Cartesian product Z_2^W, and the fibre and binding groups are both Z_2, acting regularly.

Example 2. Let D be a pure set and W be the set of 2-subsets of D. Regard W as a structure with automorphism group $\mathrm{Sym}(D)$, acting naturally (for example, we can view W as a graph, two 2-sets being adjacent if they intersect in a singleton). Let C be the set of *ordered* 2-subsets of D. Let $\pi : C \to W$ by the map $(x,y) \mapsto \{x,y\}$. Then C is a double cover of W. The kernel is trivial, and the automorphism group of C is just $\mathrm{Sym}(D)$, so in particular the cover splits. The fibre group is Z_2, but the binding group is trivial. As these are unequal, this is an example of a *twisted* cover.

Example 3. Let V be an infinite-dimensional vector space over a finite field F_q, let $V^* := V \setminus \{0\}$, and let PV be the corresponding projective space (which has as domain the set of 1-dimensional subspaces of V). We regard V^* as a structure with automorphism group $\mathrm{GL}(V)$ (the group of invertible linear transformations $V \to V$) and PV as a structure with automorphism group $\mathrm{PGL}(V)$ (the quotient of $\mathrm{GL}(V)$ by the central subgroup of linear transformations acting as scalars). The map $V^* \to PV$ given by $v \mapsto \langle v \rangle$ is a cover of PV. The kernel is just the centre of $\mathrm{GL}(V)$ (equal to the group of scalar transformations, isomorphic to the multiplicative group F_q^* of F_q), and the cover is non-split. The fibre group and binding group are both F_q^*.

Example 4. Let W be the structure in Example 2. We build a non-split cover $\pi : C_1 \to W$ with fibre groups Z_4 and binding groups Z_2. The cover C_1 has two sorts: the set C of Example 2, and also a set C_2 such that, for each $\{a,b\} \in W$, $C_2(\{a,b\})$ is a 4-vertex digraph which is a cycle. There is also a symmetric binary relation holding between $C(\{a,b\})$ and $C_2(\{a,b\})$, such that one opposite pair from the 4-cycle are joined to (a,b), the other opposite pair to (b,a). This is a free finite cover with binding groups Z_2 and fibre group Z_4. It is non-split. Essentially, the reason this construction exists is that the stabiliser in $\mathrm{Aut}(W)$ of an element of W is isomorphic to $Z_2 \times \mathrm{Sym}(\omega)$, so in particular has a proper closed normal subgroup of finite index. The quotient group is cyclic of order 2, and the cyclic group of order 4 is a non-split extension of this. More details, and the general construction, will be described in the section on free covers (2.1).

All the above examples are totally categorical. We now give an example of an unstable, non-split $\aleph_0$-categorical cover.

Example 5. We describe the imprimitive homogeneous directed graph $\hat{\mathbb{Q}}$ from [13]. The vertices of the graph are the points $\{e^{2\pi i\theta} : \theta \in \mathbb{Q}\}$ of the unit circle in the complex plane. There is a directed edge from vertex x to vertex y if and only if the angle at the origin from x to y, measured in a clockwise direction, is strictly between 0 and π. It is shown in [13] that this directed graph is homogeneous (in the sense that any isomorphism between finite subgraphs extends to an automorphism of the graph). It is clearly imprimitive: non-adjacent vertices must be diametrically opposite, so non-adjacency is an invariant equivalence relation (with classes of size 2). Let W denote the set of pairs of opposite vertices and $\sigma : \hat{\mathbb{Q}} \to W$ the natural map. By 1.4.2 the group Σ of permutations induced by $\mathrm{Aut}(\hat{\mathbb{Q}})$ on W is closed, so σ can be thought of as a finite cover.

It is easy to see that the kernel of the cover has order 2 (the non-trivial element of the kernel interchanges every point with its opposite). We claim that it is non-split. Suppose H is a closed subgroup of $\mathrm{Aut}(\hat{\mathbb{Q}})$ of finite index. Then H intersects every point-stabiliser in a closed subgroup of finite index. But a point stabiliser is isomorphic to $\mathrm{Aut}(\mathbb{Q})$ and this has no proper closed subgroups of finite index (this follows, for example, from the determination of all normal subgroups of $\mathrm{Aut}(\mathbb{Q})$ by G. Higman in [30]). So H contains the stabiliser of every point of $\hat{\mathbb{Q}}$ and this easily implies that $H = \mathrm{Aut}(\hat{\mathbb{Q}})$.

It is worth considering what W is here. It is not hard to see that $\mathrm{Aut}(W)$ is highly homogeneous on W, that is, transitive on the set of k-sets from W, for all $k \in \mathbb{N}$ (for any k-subset of W take representatives in one half of $\hat{\mathbb{Q}}$). Moreover, $\mathrm{Aut}(W)$ is 2- but not 3-transitive on W, and has no proper closed subgroup of finite index. So a theorem of Peter Cameron ([9]) implies that W is the countable dense circular ordering, that is, the ternary relation Cr on $\mathbb{Q}$ defined by:

$$Cr(x,y,z) \leftrightarrow (x < y < z) \vee (y < z < x) \vee (z < x < y).$$

We end with another totally categorical example (Example 3 of [4]).

Example 6. Let G be the abelian group $Z_4^{(\omega)}$ (the direct sum of $\aleph_0$ copies of Z_4). This has as socle $2G$, the subgroup of elements of order at most 2. Let $W := 2G \setminus \{0\}$, the infinite dimensional projective space of F_2. For each $a \in W$, let $F_a := \{x \in G : 2x = a\}/\{0, a\}$. The F_a are the fibres of a cover $C \to W$, with corresponding 'binding groups' $2G/\{0, a\}$. Then C is an affine cover of W (a notion defined in Section 1.3.1), and G is in the definable closure of C. This cover is non-split.

1.3 Related Notions

There are several notions closely related to finite covers.

1.3.1 Affine covers

We first define a more general notion of cover, which includes finite covers, and then define affine covers. This is taken from [4], but is implicit in [35].

Definition 1.3.1 A structure M is a *cover* of W if

(a) W is a 0-definable subset of M,

(b) every relation on W^n 0-definable in M is 0-definable in W,

(c) every relation on W^n which is definable in M with parameters is definable with parameters in W,

(d) there is a 0-definable surjection $\pi : M \setminus W \to W$,

(e) there is a 0-definable family G_a ($a \in W$) of groups (the *structure groups*) living in W^{eq}, which act regularly on the fibres $M(a) = \pi^{-1}(a)$, their action being a-definable in M.

The cover is *affine* if the fibres are infinite.

Because of the reference to the regular action of the structure groups, this definition appears to conflict with the earlier definition of a finite cover. However, it will be shown in Lemma 3.2.1(a) that any finite cover is bi-interpretable with a regular finite cover, that is, one with regular fibre groups. A regular finite cover can be regarded as a cover in the above sense, with structure groups adjoined formally. In these notes, we work with finite rather than affine covers, so can disregard the structure groups. An example of an affine cover is Example 6 of Section 1.2.

1.3.2 Symmetric extensions

Symmetric extensions form a convenient generalisation of finite covers, developed by Hodges and Pillay [34].

Suppose that L is a sublanguage of a language L^+, and that there is a unary relation symbol $P \in L^+ \setminus L$. Let M be an L^+-structure, and let N denote the substructure with domain $P^M = \{a \in M : M \models P(a)\}$ of the reduct of M to L. In this situation, we say M is a *relativised expansion* of N. Clearly every

automorphism of M induces an automorphism of N, and indeed restriction is a homomorphism $\mu : \mathrm{Aut}(M) \to \mathrm{Aut}(N)$. We say M is a *symmetric extension* of N if it is a relativised expansion of N and μ is surjective. The *kernel* of the symmetric extension is $\ker \mu = \mathrm{Aut}(M/N)$. If additionally M is algebraic over N (in the language L^+) then we say that it is a symmetric *algebraic* extension of N. Clearly, if $\pi : C \to W$ is a finite cover (in the sense of Definition 1.1.2) then the structure (C, W, π) is a symmetric algebraic extension of W. If M is a symmetric extension of N then an expansion M_1 of M is called a *symmetric expansion* of M if M_1 is also a symmetric extension of N.

Certain model-theoretic properties pass from a structure to its covers or symmetric extensions. An easy application of the Ryll-Nardzewski Theorem yields that any finite cover of a countable $\aleph_0$-categorical structure is $\aleph_0$-categorical. The following theorem (Theorem 9 of [34]) is a generalisation of this for symmetric extensions. Its proof uses the Ryll-Nardzewski Theorem and Lemma 1.4.3 below, together with a syntactical characterisation of the 'one-cardinal' condition.

Theorem 1.3.2 *Let M be a countable symmetric extension of N which is one-cardinal over N (that is, $|P^{M'}| = |M'|$ for every M' elementarily equivalent to M). If N is $\aleph_0$-categorical, then so is M.* □

Corollary 1.3.3 (Theorem 10 of [34]) *Let N be totally categorical, and M be a countable symmetric extension of N which is one-cardinal over N. Then M is totally categorical.*

Proof. There is a strongly minimal formula ϕ in N, and its relativisation $\phi(N)$ to P is strongly minimal in M. Furthermore, M is one-cardinal over N, and N is one-cardinal over $\phi(N)$, so M is one-cardinal over $\phi(N)$. It follows from a theorem of Erimbetov (Theorem 4 of [19]) that M is uncountably categorical. Also, M is $\aleph_0$-categorical by Theorem 1.3.2. □

Further results of this kind can be found in Kikyo and Tsuboi ([43]).

1.4 Topological arguments

We first record a triviality.

Lemma 1.4.1 *Let M be an infinite structure, and P a subset of M invariant under* $\mathrm{Aut}(M)$. *Let $\phi : \mathrm{Aut}(M) \to \mathrm{Sym}(P)$ be the homomorphism induced by restriction, and give* $\mathrm{Im}(\phi)$ *the topology induced from* $\mathrm{Sym}(P)$. *Then ϕ is continuous.*

Proof. Let $H := \mathrm{Im}(\phi)$. A typical basic open set of H is $H_{(F)} = \{h \in H : hf = f \; \forall f \in F\}$ for some finite $F \subset P$. The preimage of this under ϕ is $\mathrm{Aut}(M/F)$, which is also open. □

The next result is crucial to our use of topological arguments. In its full generality it appears as Lemma 1.1 of [24]. It was proved in [27], assuming C countable.

Lemma 1.4.2 *Let C be a permutation structure, W a set, and $\pi : C \to W$ be a finite-to-one surjection whose fibres form an $\mathrm{Aut}(C)$-invariant partition of C. Then*

(a) *The restriction map $\mu : \mathrm{Aut}(C) \to \mathrm{Sym}(W)$ maps closed subgroups of $\mathrm{Aut}(C)$ to closed subgroups of $\mathrm{Sym}(W)$.*

(b) *If $\pi : C \to W$ is a finite cover, then the restriction map μ sends open subgroups of $\mathrm{Aut}(C)$ to open subgroups of $\mathrm{Aut}(W)$, so is an open map.*

Proof. We prove the results assuming C is countable.

(a) Let $(\hat{g}_i : i \in \omega)$ be a sequence of automorphisms of C such that $(g_i : i \in \omega)$ converges to $h \in \mathrm{Sym}(W)$, where $g_i = \mu(\hat{g}_i)$. By continuity of μ (see Lemma 1.4.1), it suffices to show that $(\hat{g}_i : i \in \omega)$ has a convergent subsequence. Enumerate W as $W := (w_i : i \in \omega)$. By thinning out the $\hat{g}_i$, we may assume that $g_i(w_j) = g_{i'}(w_j)$ and $g_i^{-1}(w_j) = g_{i'}^{-1}(w_j)$ whenever $i, i' \geq j$, so in particular $\hat{g}_i, \hat{g}_{i'} : C(w_j) \to C(hw_j)$ and $\hat{g}_i^{-1}, \hat{g}_{i'}^{-1} : C(w_j) \to C(h^{-1}w_j)$. It follows by the pigeon-hole principle that for each j there is infinite $I \subseteq \omega$ such that $\hat{g}_i \mid C(w_j) = \hat{g}_{i'} \mid C(w_j)$ and $\hat{g}_i^{-1} \mid C(w_j) = \hat{g}_{i'}^{-1} \mid C(w_j)$ whenever $i, i' \in I$. Apply this repeatedly to obtain the subsequence.

(b) By a result of Evans [22], the open subgroups of $\mathrm{Aut}(C)$ and $\mathrm{Aut}(W)$ are precisely the closed subgroups of countable index, so the result follows from (a). □

In the countable case, this result has the following generalisation (Lemma 5 of [34]).

Lemma 1.4.3 *If M is a countable symmetric extension of N then the restriction map $\mu : \mathrm{Aut}(M) \to \mathrm{Aut}(N)$ is open.* □

Corollary 1.4.4 ([34], Lemma 6) *Let M be a countable symmetric extension of N with restriction map $\mu : \mathrm{Aut}(M) \to \mathrm{Aut}(N)$ and kernel K. Then the natural group isomorphism $\mu_K : \mathrm{Aut}(M)/K \to \mathrm{Aut}(N)$ is a homeomorphism (where $\mathrm{Aut}(M)/K$ is given the quotient topology). This is true without the countability assumption if M arises from a finite cover of N.*

Proof. This follows from Lemmas 1.4.1 and 1.4.3 (or 1.4.2). □

1.5 Kernels

1.5.1 Kernels of symmetric extensions

We shall discuss kernels in the more general context of symmetric extensions. Suppose that N is a fixed first-order structure and that M_0 is a fixed symmetric extension of N (in the context of finite covers, it might be a principal cover or a free finite cover). Let $\mu : \mathrm{Aut}(M_0) \to \mathrm{Aut}(N)$ be the restriction map, and let K_0 denote its kernel. So we have the short exact sequence

$$1 \to K_0 \to \mathrm{Aut}(M_0) \xrightarrow{\mu} \mathrm{Aut}(N) \to 1$$

of continuous maps, and if M_0 is countable, or if the symmetric extension arises from a finite cover, these are also open maps. We consider expansions M of M_0 such that $\mu(\mathrm{Aut}(M)) = \mathrm{Aut}(N)$ (*symmetric* expansions). The *kernel* of such an M is just $K := \mathrm{Aut}(M) \cap K_0$. To classify such expansions of M_0, we must do the following:

1. Classify those subgroups K of K_0 which can occur as kernels.
2. For a given kernel K, classify the expansions M with kernel K such that $\mu(\mathrm{Aut}\, M) = \mathrm{Aut}\, N$.

This is essentially the setting of Ahlbrandt and Ziegler ([3] and [4]) except that they assume that M_0 is the principal cover of N with respect to a family of structure groups $(G_a : a \in N)$, and in particular $K_0 \cong \prod(G_a : a \in N)$.

We begin with an easy observation, important for Corollary 3.3.2. If M is a symmetric extension of N we say that M is *algebraic* over N if all orbits of the kernel $\mathrm{Aut}(M/N)$ are finite (this is the group-theoretic analogue of the model-theoretic condition that every element of M is in the algebraic closure of N). Clearly this property holds if M arises from a finite cover of N.

Lemma 1.5.1 *Let M be a symmetric extension of N, with kernel K. Then*

(a) K is closed,

(b) if M is a symmetric algebraic extension of N, then K is compact.

Proof. (a) This follows since K is the automorphism group of the expansion of M obtained by naming all elements of N. Alternatively, K is the kernel of a continuous homomorphism.

(b) Since the orbits of K are finite, K is a subgroup of a direct product of finite groups. By Tychonoff's Theorem the latter is compact, so by (a), K is also compact. □

The connection between M and K is given by the following result. For finite covers with abelian kernel Lemma 3.1.5 shows that the situation is much more straightforward if the restriction map μ splits.

Proposition 1.5.2 *Suppose M_0 is countable. Let H be a subgroup of* $\mathrm{Aut}(M_0)$ *such that $\mu(H) = \mathrm{Aut}(N)$, and let $K := H \cap K_0$. Then H is closed if and only if*

(a) K is closed,

(b) the isomorphism $\mu_K : H/K \to \mathrm{Aut}(N)$ given by $\mu_K(hK) = \mu(h)$ for $h \in H$ is a homeomorphism.

Proof. If H is closed, then trivially $K = H \cap K_0$ is closed (as K_0 is closed by Lemma 1.5.1), and μ_K is a homeomorphism by Corollary 1.4.4. The converse is proved in Theorem 11 of [34]. □

A countable $\aleph_0$-categorical (or permutation) structure M is said to have the *small index property* if every subgroup of $\mathrm{Aut}(M)$ of index less than $2^{\aleph_0}$ is open. Since the small index property gives a characterisation of the open subgroups of $\mathrm{Aut}(M)$, this says that the topology on $\mathrm{Aut}(M)$ is determined by the abstract group structure. It is now known that many $\aleph_0$-categorical structures (for example, $(\mathbb{Q}, <)$, the random graph, and all $\aleph_0$-categorical ω-stable structures, and hence all totally categorical structures) have the small index property (see [31] for the last two cases and further references). The following extension of Proposition 1.5.2 was proved in [34]. We are assuming that M_0 is countable.

Proposition 1.5.3 *Suppose that the structure N has the small index property. Let H be a subgroup of $\mathrm{Aut}(M_0)$ such that $\mu(H) = \mathrm{Aut}(N)$, and suppose that $K := K_0 \cap H$ is closed. Then μ_K is a homeomorphism. In particular, clause (b) of Proposition 1.5.2 can be omitted.*

Proof. It suffices to show that μ_K is open. The open subgroups of H/K form a base of neighbourhoods of the identity, and are images of open subgroups of H, so have countable index in H/K. Since μ_K is an isomorphism, their images have countable index in $\mathrm{Aut}(N)$, so are open by the small index property. □

In Section 6, we discuss further the problem of determining the possible abelian kernels of finite covers of a structure W.

1.6 The model-theoretic context

We discuss some model-theoretic motivation for looking at covers. We concentrate on the totally categorical case, though some of the ideas apply more generally, and are used, for example, in the analysis of smoothly approximated structures in [16]. The remarks below are an amalgam of [33], [35] and [36], and in particular [48]. We shall assume familiarity with standard model-theoretic terminology (see [12], for example).

The class of countable totally categorical structures is closed under taking finite (and affine) covers. Thus an understanding of the process of taking covers is a necessary part of an understanding of the fine-detail of this class. However, the connections run much deeper. First, recall the following well-known characterisation (due to Baldwin and Lachlan [6]).

Theorem 1.6.1 *A countably infinite structure M is totally categorical if and only if it satisfies each of the following three conditions.*

(i) $\mathrm{Aut}(M)$ *has finitely many orbits on M^n for all $n \in \mathbb{N}$ (this is equivalent to $\aleph_0$-categoricity).*

(ii) *M is non-two-cardinal; that is, whenever M' is a proper elementary extension of M and $\phi(x)$ is a formula with parameters in M with infinitely many solutions in M,*

$$\{x \in M : M \models \phi(x)\} \text{ is a proper subset of } \{x \in M' : M' \models \phi(x)\}.$$

(iii) *M has finite Morley rank (or Cantor-Bendixson rank).* □

If M is an $\aleph_0$-categorical structure, then a set $X \subset M^{\mathrm{eq}}$ is *strictly minimal* if it is 0-definable, strongly minimal, and there is no proper non-trivial 0-definable equivalence relation on X. It is easily seen that any strictly minimal set has doubly transitive automorphism group. There is a classification of strictly minimal sets, due independently to Cherlin [15], Mills, and Zil'ber [50]. (There is now a concise model-theoretic proof of this classification, due to Hrushovski [37], and a self-contained geometric proof due to Evans [21] but, modulo the classification of finite simple groups, perhaps the simplest is still that due to Cherlin and Mills, which goes via the classification of finite 2-transitive groups.) The result is this. If X is a strictly minimal set with automorphism group G, then one of the following holds.

1. X is a pure set, and $G = \mathrm{Sym}(X)$ (the *disintegrated* case).
2. $X = \mathrm{PG}(V)$ (projective space over an infinite dimensional vector space V over a finite field F_q) and $\mathrm{PGL}(\aleph_0, q) \leq G \leq \mathrm{P\Gamma L}(\aleph_0, q)$ (the *projective* case)
3. $X = \mathrm{AG}(V)$ (affine space over V as above) and $\mathrm{AGL}(\aleph_0, q) \leq G \leq \mathrm{A\Gamma L}(\aleph_0, q)$ (the *affine* case).

The strictly minimal sets of types (1) and (2) are said to be *modular*, since the familiar dimension formula

$$\dim A + \dim B = \dim A \cap B + \dim \mathrm{acl}(A \cup B)$$

holds for any two algebraically closed sets. Affine strictly minimal sets are *locally modular*, since when we localise (name a point v, and factor out the relation of being inter-algebraic modulo v) we obtain a modular strictly minimal set (projective space). It is important that one can also pass from an affine space to a modular strictly minimal set *without* naming a parameter: form a projective space over the same field, whose points are parallel classes of lines of the affine space. We state now a crucial theorem from [15] (a consequence of their Coordinatisation Theorem, together with Shelah's Finite Equivalence Relation Theorem).

Theorem 1.6.2 *If M is a countable totally categorical structure, then there is a 0-definable strictly minimal set in M^{eq}.* □

By the above remarks, the strictly minimal set can be assumed to be modular. It is also shown in [15] that given any two modular strictly minimal sets in M there is a unique 0-definable bijection between them (this can be proved from the non-two-cardinal property by a group-theoretic argument).

If P, Q are 0-definable sets in M^{eq} we define Q to be a *precover* of P if there are

(a) a partition of $Q \setminus P$ into a 0-definable family $\{H_{\bar{a}} : \bar{a} \in P\}$,

(b) a 0-definable family $\{G_{\bar{a}} : \bar{a} \in P\}$ of groups (the *structure groups*) living in P^{eq},

(c) a regular $\bar{a}$-definable action of each $G_{\bar{a}}$ on $H_{\bar{a}}$.

The precover is *finite* if the structure groups are finite. Thus essentially a precover is like a cover except that the fibres correspond to tuples in the base structure, not singletons.

Finally, we recall a key result from [15], the following consequence of the Coordinatisation Theorem for totally categorical structures.

Theorem 1.6.3 *For every $a \in M^{\mathrm{eq}}$ there is a sequence $a_0, \ldots, a_n = a$ such that each a_i lies in $\mathrm{acl}(a)$ and each type $\mathrm{tp}(a_0/\emptyset)$, $\mathrm{tp}(a_{i+1}/a_0 \ldots a_i)$ is algebraic or strictly minimal.* □

We now state Zil'ber's 'ladder theorem.'

Theorem 1.6.4 (Zil'ber) *Let M be totally categorical. Then there is a 0-definable modular strictly minimal set D and a sequence*

$$D = M_0 \subset M_1 \subset \ldots \subset M_n$$

such that each M_{i+1} is a precover of M_i and M is in the definable closure of M_n. Furthermore if D is disintegrated then all the precovers are finite, and if D is a projective space over F_q, then all the structure groups are finite or F_q-vector spaces, and the structure groups live in D^{eq}. □

It follows that any totally categorical structure can be built from a 0-definable modular strictly minimal set by a sequence of covers of Grassmannians. A proof of Theorem 1.6.4, and the analogous result in the uncountably categorical case, can be found in [51]. By Theorems 1.6.3 and 1.6.2, and the classification of strictly minimal sets, the key observation in the proof of Theorem 1.6.4 is the following:

> If $\bar{c} \in M_{i-1}$, $a \in M$, and $\mathrm{tp}(a/\bar{c})$ is algebraic or strictly minimal, then there are precovers M_i of M_{i-1} and M_{i+1} of M_i, each of the required kind, such that $a \in \mathrm{dcl}(M_{i+1})$.

Details can be found at the end of Section 3 of [48].

Via an analysis of covers, Hrushovski [36] gave a detailed structure theory for the class $\mathcal{D}$ of totally categorical structures whose co-ordinatising strictly minimal set is disintegrated. Any member of $\mathcal{D}$ is in the algebraic closure of such a strictly minimal set (this follows essentially by Theorem 1.6.4 and the fact that no infinite group can be interpreted in such a structure). He explicitly described the members of a certain subclass $\mathcal{C}$ of $\mathcal{D}$, and then showed that any member of $\mathcal{D}$ can be expanded to a member of $\mathcal{C}$ by naming finitely many constants.

An analogous program was developed in [27] for non-disintegrated totally categorical structures. The results there apply in many unstable situations, but one of

the more quotable theorems is the following (Theorem 3.2 of [27]). The connection with our context is that the intermediate structures (the M_i and $M_{i,0}$) arise from a sequence of cover-like constructions (M_i is, in some sense, a free finite cover over $M_{i,0}$). The proof involves an analysis of finite covers with finite kernels of modular strictly minimal sets.

Theorem 1.6.5 *Let M be a countable totally categorical structure of Morley rank n which lies in the algebraic closure of a 0-definable modular strictly minimal set. Then there are 0-definable subsets*

$$M_0, D \subseteq M_{1,0} \subseteq M_1 \subseteq M_{2,0} \subseteq \ldots \subseteq M_{n,0} \subseteq M_n$$

of M^{eq} such that

(a) M_0 is finite, $\mathrm{acl}^{eq}(\emptyset) = \mathrm{dcl}^{eq}(M_0)$, and M_i has Morley rank i,

(b) $\mathrm{Aut}(M_{1,0}/M_0 \cup D)$ is nilpotent by finite-abelian,

(c) for $2 \leq i \leq n$, $\mathrm{Aut}(M_{i,0}/M_{i-1})$ is nilpotent, and for $1 \leq i \leq n$, $\mathrm{Aut}(M_i/M_{i,0})$ is a direct product of finite groups,

(d) $M \subseteq M_n$.□

1.7 An overview

So far, apart from the results of Ahlbrandt and Ziegler ([3, 4]), most of the successes have been with finite rather than affine covers, and for the remainder of these notes we will be concerned mainly with finite covers of countable $\aleph_0$-categorical structures (although we often present results in more generality). The material is, for the most part, taken from the papers of Ahlbrandt and Ziegler ([3, 4]), Evans ([23, 24, 25]), Evans and Hrushovski ([27]), Hodges and Pillay ([34]), Ivanov ([38, 39]), and Ivanov and Macpherson ([40]). Some of the material does not appear elsewhere. The material on free covers (Section 2.1), and the presentation of Pontriagin duality (Section 6.3) is due to Evans, but is undoubtedly well-known to others.

Recall that the main problem is:

> **The Cover Problem:** For a given $\aleph_0$-categorical structure W, describe its finite covers.

We are usually satisfied with only a partial solution to this for any particular W. Indeed, the only W for which a complete solution is known are the highly homogeneous structures (see Section 3.1.3, for results due to Ivanov [39] and Ziegler [48]). The successes of the the theory have been in proving splitting results (showing that for certain W all finite covers split), and describing finite covers with finite or abelian kernel.

The results which we present can be divided into three types.

1. Constructions: describing abstract covering constructions which generalise familiar examples.

2. Reduction theorems: reducing the analysis of finite covers to that of simpler kinds.

3. Classification results: classifying finite covers of certain familiar $\aleph_0$-categorical structures.

The constructions in Section 2 show how non-split finite covers can arise from non-split finite extensions of finite homomorphic images of point stabilisers (Lemma 2.1.5) and, more subtly, from combinatorial properties of 0-definable binary and ternary relations on W (Sections 2.2 and 2.3). In these latter cases the kernels of the covers are finite.

The main reduction result is that any finite cover has a minimal covering expansion, the kernel of which is nilpotent (Corollary 3.3.2 and Lemma 3.3.4). It then follows that the splitting problem reduces to consideration of finite covers with abelian kernel (Corollary 3.3.5). For structures with trivial algebraic closure the situation is more straightforward, and one can reduce to consideration of finite covers with finite kernels (Theorem 3.5.1). Classification results for finite covers with finite kernels are given in Section 4.

To analyse finite covers with abelian kernel we use cohomological methods first introduced into this subject by Ahlbrandt and Ziegler ([4], and Section 6 here). These methods have been used (together with results from representation theory and cohomology of finite groups) to give very precise information about finite covers of Grassmannians of strictly minimal sets (see Sections 6.3.2 and 6.5), as well as information which is rather more qualitative (cf. Theorem 7.2.11). They can also be used to provide information about finite covers with finite kernels (Section 7.2).

2 General constructions

2.1 Free covers

2.1.1 Existence and Uniqueness

Suppose $\pi : C \longrightarrow W$ is a finite cover. Then we have the following data:

- the base structure W
- for every $w \in W$, the fibre group $F(w) = \mathrm{Aut}(C(w)/w)$
- for every $w \in W$, the binding group $B(w) = \mathrm{Aut}(C(w)/W)$.

Here, $B(w)$ is a normal subgroup of $F(w)$, and these should both be regarded as permutation groups on the fibre $C(w) = \pi^{-1}(w)$.

First, we indicate what is needed to ensure that there is *some* finite cover with the given data.

Lemma 2.1.1 *Suppose $\pi : C \longrightarrow W$ is a finite cover. Then, for every $w \in W$, there is a continuous epimorphism $\chi_w : \mathrm{Aut}(W/w) \to F(w)/B(w)$.*

Proof. Let $g \in \mathrm{Aut}(W/w)$. Then there exists $h \in \mathrm{Aut}(C/w)$ which extends g. Suppose h' also extends g. Then $h^{-1}h' \in \mathrm{Aut}(C/W)$ and so $(h|C(w))B(w) = (h'|C(w))B(w)$. So if we define $\chi_w(g) = (h|C(w))B(w)$, we get a well-defined homomorphism $\chi_w : \mathrm{Aut}(W/w) \to F(w)/B(w)$, which is clearly onto. To see that χ_w is continuous, note that its kernel consists of those $g \in \mathrm{Aut}(W/w)$ which extend to an element of $Aut(C)$ inducing an element of $B(w)$ on $C(w)$. Thus $\ker \chi_w$ is the image (under the restriction map $\mathrm{Aut}(C) \to \mathrm{Aut}(W)$) of an open subgroup of $\mathrm{Aut}(C)$. By 1.4.2, this implies that $\ker \chi_w$ is an open subgroup of $\mathrm{Aut}(W/w)$. As $F(w)/B(w)$ is finite, this means that χ_w is continuous. □

We refer to the epimorphisms χ_w as the *canonical homomorphisms* of the cover.

Recall that a finite cover $\pi : C \longrightarrow W$ is *free* if (with the above notation)

$$\mathrm{Aut}(C/W) = \prod_{w \in W} B(w).$$

The next lemma (which we include for completeness, but which we shall not really use), shows that free covers are completely determined by the fibre and binding groups, together with the canonical homomorphisms. The existence part comes from [23] (Lemma 4.4), but it is really just an elaboration of the construction of free covers in [36]. The reader might find it useful to refer back to Example 4 in Section 1.2 of the free cover with fibre group Z_4 and binding group Z_2 of the Grassmannian of 2-sets from a disintegrated set.

Lemma 2.1.2 *Let W be a transitive permutation structure and F a permutation group on a finite set X. Let $w_0 \in W$, suppose B is a normal subgroup of F and suppose $\chi : \mathrm{Aut}(W/w_0) \to F/B$ is a continuous epimorphism. Then there exists a free finite cover $\sigma : M \to W$ with fibre and binding groups at w_0 equal to F and B, and such that the canonical epimorphism χ_{w_0} from 2.1.1 is equal to χ. With these properties, σ is determined uniquely (up to isomorphism over W).*

Proof. The proof is in a series of steps. The notation is cumulative.

Step 1. Let C be the set of (left) cosets of $\ker \chi$ in $\mathrm{Aut}(W)$. It is easy to show that the group of permutations which $\mathrm{Aut}(W)$ induces on this is closed, and so we may consider C as a permutation structure with automorphism group isomorphic to $\mathrm{Aut}(W)$. Furthermore, the map $\eta : C \to W$ given by $\eta(g \ker \chi) = gw_0$ is a finite cover with trivial kernel.

Step 2. Let $Z = C(w_0)$ and consider $Y = X \cup Z$ as a finite (permutation) structure with F as its automorphism group: the action of $f \in F$ on $c \in Z$ is given via χ as $f(c) = (\chi^{-1}(fB))(c)$. For each $w \in W$ let $g_w \in \mathrm{Aut}(W)$ be such that $g_w w = w_0$ (and suppose g_{w_0} is the identity). Then g_w maps $C(w)$ to Z and so induces an embedding $\sigma_w : C(w) \to Y$.

Step 3. We now build a finite cover $\sigma' : M' \to W$, which we describe as a first-order structure.

The domain of M' is the disjoint union of W, C and $W \times Y$. The structure on M' consists of $\sigma' : M' \longrightarrow W$, which is made up of the identity map on W, η on C, and the projection map to the first coordinate on $W \times Y$. We also have an injection $\tau : C \longrightarrow W \times Y$ given by $\tau(c) = (\eta(c), \sigma_{\eta(c)}(c))$ (so τ has image $W \times Z$). The remaining structure on M' is made up of the original structures on W and C, and for each n-ary relation R on Y we have an n-ary relation R^+ on $W \times Y$ given by

$R^+((w_1, y_1), \ldots, (w_n, y_n))$ iff $w_i = w_j$ for all i, j and $R(y_1, \ldots, y_n)$.

To show that $\sigma' : M' \longrightarrow W$ is a finite cover, it will suffice to show that any automorphism of C extends to an automorphism of M'. So let $g \in \mathrm{Aut}(C)$ and $w \in W$. Then g induces (via τ) a bijection from $\{w\} \times Z$ to $\{gw\} \times Z$, and so an automorphism of Z. This extends to an automorphism $\alpha(w, g)$ of Y. If we also denote by $\alpha(w, g)$ the induced map from $\{w\} \times Y$ to $\{gw\} \times Y$, then $g \cup \bigcup_{w \in W} \alpha(w, g)$ is an automorphism of M' extending g.

Step 4. Note that $\mathrm{Aut}(M'/C) = \prod_{w \in W} \mathrm{Aut}(\{w\} \times Y / \{w\} \times Z)$. Let σ be the restriction of σ' to $M = W \times X \subseteq M'$ considered as permutation structure with $\mathrm{Aut}\, M'$ acting. Then $\sigma : M \to W$ is a free finite cover with the required data at

Clearly there is equality in the first coordinate here. For the second:

$$\begin{aligned}\hat{g}_{h\eta(c)}h\hat{g}^{-1}_{\eta(c)}(\sigma_{\eta(c)}(c)) &= \hat{g}_{h\eta(c)}h\hat{g}^{-1}_{\eta(c)}g_{\eta(c)}(c)\\ &= \hat{g}_{h\eta(c)}hc\\ &= \sigma_{\eta(hc)}(hc),\end{aligned}$$

as required.

So the cover ν embeds in the cover σ' with image M, and the map β induces an embedding of $\mathrm{Aut}(N)$ into $\mathrm{Aut}(M)$. Thus M can be regarded as a reduct of N. But N is free with the same binding groups as M and so it follows that M can have no more automorphisms than N. □

In practice, what we shall use is the following, which shows that any finite cover is an expansion of a free finite cover with the same data.

Lemma 2.1.3 *Every finite cover $\pi : C \longrightarrow W$ is an expansion of a free finite cover with the same fibre groups, binding groups and canonical homomorphisms as in π.*

Proof. This can be deduced from the proof of 2.1.2, but we give a quicker, independent proof.

Let $\Gamma = \mathrm{Aut}(C)$ and $K = \mathrm{Aut}(C/W)$ and $B(w) = \mathrm{Aut}(C(w)/W)$. Let $H = \prod_{w\in W} B(w) \leq \mathrm{Sym}(C)$. We claim

(i) H is normalised by Γ;

(ii) $H\Gamma$ is a closed subgroup of $Sym(C)$;

(iii) $H\Gamma$ is the automorphism group of a free finite cover $\pi_0 : C_0 \longrightarrow W$ which is a reduct of C.

Here, (i) is routine; (ii) follows from (i) and compactness of H (cf. the proof of 3.1.4), and (iii) follows immediately from (ii).

As $\mathrm{Aut}(C_0) = H\Gamma$, it follows that the binding groups of π and π_0 are the same. Also, for any $w \in W$ we have

$$Aut(C_0(w)/w) = (H|C(w))(\Gamma|C(w)) = \mathrm{Aut}(C(w)/w)$$

so the same is true of the fibre groups. The definition of the canonical homomorphisms shows that they coincide in π and π_0. □

2.1.2 Splitting

We prove the following. For free covers with abelian kernel, Shapiro's lemma (7.1.4) provides a more precise analysis.

Lemma 2.1.4 *Suppose W is a transitive permutation structure and $\pi : M \to W$ a free finite cover with fibre group F and binding group B. If F splits over B then the cover π also splits.*

Proof. Let $w \in W$. By 3.2.1 we may assume that the fibre group $F(w)$ acts regularly on $M(w)$. Let $H(w)$ be a complement to $B(w)$ in $F(w)$ and let $\mathcal{T}(w)$ be an $H(w)$-orbit. This is a transversal of the $B(w)$-orbits on $M(w)$. For each $w' \in W$ choose in $M(w')$ an $\mathrm{Aut}(M)$-translate $\mathcal{T}(w')$ of $\mathcal{T}(w)$ and let $\mathcal{T}$ be the union of these. Let T be the setwise stabiliser in $\mathrm{Aut}\, M$ of $\mathcal{T}$.

We claim that T is a complement to the kernel of π in $\mathrm{Aut}(M)$. By regularity, it is enough to show that any element g of $\mathrm{Aut}(W)$ extends to an element of T. Let $\hat{g} \in \mu^{-1}(g)$ and for each $w' \in W$ choose an $x_{w'} \in \mathcal{T}(w')$. So $\hat{g}x_{w'} \in M(gw')$. For each $w' \in W$ there is a $k_{w'} \in \mathrm{Aut}(M(gw')/W)$ such that $k_{w'}\hat{g}x_{w'} \in \mathcal{T}(gw')$. As π is a free cover it follows that there exists $k \in \mathrm{Aut}(M/W)$ such that $k\hat{g}x_{w'} \in \mathcal{T}(gw')$ for all $w' \in W$. It will suffice to show that $\hat{g}' = k\hat{g} \in T$. Note that $\mu(\hat{g}') = g$ and $\hat{g}'$ has the property that it sends each $x_{w'}$ to an element of $\mathcal{T}$. Let $x' \in \mathcal{T}(w')$ and suppose $\hat{g}'x' \notin \mathcal{T}(gw')$. Let $f \in \mathrm{Aut}(M)$ send $\mathcal{T}(w')$ to $\mathcal{T}(gw')$. So, if $h = f^{-1}\hat{g}'$, then $h \in \mathrm{Aut}(M/w')$, $hx_{w'} \in \mathcal{T}$ and $hx' \notin \mathcal{T}$. But by regularity, there is a unique element of $F(w')$ taking $x_{w'}$ to x', and this stabilises $\mathcal{T}(w')$. This is a contradiction. □

The following is a variation on Theorem 4.5 of [23]. It may be possible to improve on this: indeed, it may be that the converse of 2.1.4 is true.

Lemma 2.1.5 *Let W be a transitive permutation structure. Let $w \in W$ and let $\mathrm{Aut}^\circ(W/w)$ be the intersection of the closed subgroups of finite index in $\mathrm{Aut}(W/w)$. Suppose that $G = \mathrm{Aut}(W/w)/\mathrm{Aut}^\circ(W/w)$ is finite and non-trivial. Then there exists a non-split free finite cover of W.*

Proof. We need the following fact from finite group theory (see Remark 2.1.6 below): there exists a non-split finite extension of G

$$1 \to B \to F \to G \to 1$$

(in fact, B can be taken to be elementary abelian). Now construct a free finite cover $\sigma : M \to W$ using 2.1.2, where the fibre group is F (in its regular representation) and binding group B and the kernels of the canonical homomorphisms are the groups $\mathrm{Aut}^\circ(W/w)$. Suppose, for a contradiction, that H is a closed complement to $K = \mathrm{Aut}(M/W)$. Thus $\mathrm{Aut}(M) = KH$, and $\mathrm{Aut}(M/w) = KH_w$, where H_w denotes the stabiliser in H of w. Restricting this equation to $M(w)$ we get that $F = BT$ where T is the restriction of H_w to $M(w)$. Thus $|T| \geq |F/B| = |G|$. But T is a finite, continuous homomorphic image of H_w, and the restriction map gives an isomorphism from H_w to $\mathrm{Aut}(W/w)$, so $|T| \leq |G|$. Thus $|T| = |G|$ and $T \cap B$ is the identity subgroup. So F splits over B, a contradiction. □

Remark 2.1.6 The 'fact' about finite groups we used in the above is well-known (cf. [17], p. 211), but we can give a self-contained proof of it using the construction of free covers (which, of course, works also for finite permutation structures). In the notation of the proof of 2.1.5, let p be a prime dividing $|G|$ and $X \leq G$ a subgroup of order p. Let W_1 be the set of left cosets of X in G with G acting (it

is not important here that this action might not be faithful). Let $w_0 = X$. Form a free finite cover $\pi_1 : C_1 \to W_1$ with fibre group Z_{p^2} and binding group Z_p and with X as the domain of the canonical homomorphism at w_0. We claim this is non-split (and so $\mathrm{Aut}(C_1)$ is a non-split extension of G by an elementary abelian p-group). As in the proof of 2.1.5, if π_1 is split then the fibre group at w_0 is the product of the binding group at w_0 with a homomorphic image of X. But this is clearly impossible.

2.2 Digraph coverings

In this section we summarise the theory of coverings of digraphs, as developed in [23]. This will provide us with examples of finite covers with finite kernels. The construction is very closely related to the topological notion of a covering space. In a later section (Section 4) we shall see that for some structures, digraph coverings give all of the non-split finite covers with finite kernels.

Definition 2.2.1 **(i)** A digraph on a set L is an irreflexive binary relation (usually denoted R) on L which is either a symmetric relation, or an antisymmetric relation. If (L, R) is a digraph and $a \in L$, put

$$a^+ := \{a' \in L : R(a, a')\}$$

and

$$a^- := \{a' \in L : R(a', a)\}.$$

A *path* in (L, R) is a sequence $x_0, \ldots, x_n$ such that for each $0 \leq i \leq n-1$, $R(x_i, x_{i+1})$ or $R(x_{i+1}, x_i)$ holds. The digraph is *connected* if any two vertices are linked by a path.

(ii) Let $\mathcal{S}$ denote the set of paths in L. We say that $p_1, p_2 \in \mathcal{S}$ are *elementarily homotopic* if one can be obtained from the other by one of the following operations (*elementary homotopies*):

(a) replace a consecutive triple abc with $R(a,b), R(b,c), R(a,c)$ by ac,
(b) replace a consecutive triple abc with $R(b,a), R(c,a), R(c,b)$ by ac,
(c) replace a consecutive triple aba by a.

Two paths are *homotopic* if one can be obtained from the other by a sequence of elementary homotopies. This is an equivalence relation on $\mathcal{S}$, and the homotopy class of p is denoted $[p]$. A connected digraph is *simply connected* if any two paths with the same endpoints are homotopic.

(iii) Let (A, R), (B, R') be two digraphs, both symmetric or both antisymmetric. A map $\sigma : A \to B$ is a *homomorphism* if, for every $b, b' \in \mathrm{Im}(\sigma)$, we have $R'(b, b')$ if and only if there are $a, a' \in A$ such that $R(a, a')$ and $\sigma(a) = b$, $\sigma(a') = b'$. A surjective homomorphism $\sigma : A \to B$ is a *covering* of digraphs if, for each $a \in A$, σ induces digraph isomorphisms $a^+ \to \sigma(a)^+$ and $a^- \to \sigma(a)^-$.

(iv) Let (L, R) be a connected digraph, fix $x_0 \in L$, let $\mathcal{P}$ denote the members of $\mathcal{S}$ which start at x_0, and U be the corresponding set of homotopy classes. Define $\sigma : U \to L$ by

$$\sigma([x_0 \dots x_r]) = x_r.$$

We can make U into a digraph with edge relation S (and σ into a covering) by putting $S([p_1], [p_2])$ if and only if $R(\sigma([p_1]), \sigma([p_2]))$ and $[p_2] = [p_1\sigma([p_2])]$. We call $\sigma : U \to L$ the *universal covering* of L.

It is shown in Lemma 5.3 of [24] that a universal covering $\sigma : U \to L$ is indeed a covering, that it has the expected universal property with respect to all coverings of L, and that it is simply connected and indeed is up to isomorphism the *unique* simply connected covering of L.

With $\sigma : U \to L$ as above, let $\Gamma(L, R)$ denote the subgroup of $\mathrm{Aut}(U)$ which maps fibres of σ to fibres of σ. The induced map $\Gamma(L, R) \to \mathrm{Aut}(L)$ is surjective and its kernel $\Delta(L, R)$ has the following characterisation. Fix $x_0 \in L$ as above, and let $p = [x_0 \dots x_r x_0] \in \sigma^{-1}(x_0)$. Define the *deck transformation* $d_p : U \to U$ by

$$d_p([x_0 x_1' \dots x_r']) = [x_0 \dots x_r x_0 x_1' \dots x_r'].$$

Then $\Delta(L, R)$ is the set of all deck transformations. Taking 'quotients' of σ by normal subgroups of finite index in Δ gives finite covers of L with finite kernels (see the 'converse' part of the statement of 1.14 in [24], or Corollary 4.2.3 here, for a precise formulation). We illustrate this with the following example, taken from [23].

Example 2.2.2 Among the homogeneous directed graphs there is a tournament (that is, a digraph such that between any two vertices there is a directed edge) which can be described as follows. Let the vertex set of D be any countable dense subset of the unit circle, with no two antipodal points. Put $x \to y$ if it is faster to go clockwise from x to y than to go anticlockwise. For example, take D as having vertex set

$$\{e^{i\theta} : \theta \in \mathbb{Q} \cap [0, 2\pi)_{\mathbb{R}}\},$$

in the complex plane, with an edge $e^{i\theta_1} \to e^{i\theta_2}$ if and only if the angle at 0 subtended by the circular arc from $e^{i\theta_1}$ to $e^{i\theta_2}$ is less than π. The finite covers of D with finite kernels can be described in terms of quotients of the universal covering $\tau : V \to D$, in a natural sense. The domain of V may be considered as the subset

$$\{2\pi n + q : n \in \mathbb{Z}, q \in \mathbb{Q} \cap [0, 2\pi)_{\mathbb{R}}\}$$

of $\mathbb{R}$, and there is an arc $x \to y$ in V if and only if $0 < y - x < \pi$. It is easily seen that V is simply connected. Let $\tau : V \to D$ be the map $x \mapsto e^{ix}$. Observe that τ is surjective, and that $u \to w$ in D if and only if there are $x, y \in V$ such that $\tau(x) = u$, $\tau(y) = w$, and $x \to y$. Then τ is a covering, and so is the universal covering of D.

For any $n \in \mathbb{N}$, let V_n be the set of cosets $V/2\pi n\mathbb{Z}$ in the additive group $\mathbb{R}/2\pi n\mathbb{Z}$, with the natural map $\nu_n : V \to V_n$. The map τ factors through ν_n to give a finite-to-one map $\tau_n : V_n \to D$. Now, ν_n induces a digraph structure on V_n (we have $u \to w$ in V_n if and only if there are $x, y \in V$ with $x \to y$ and $\nu_n(x) = u$, $\nu_n(y) = w$) together with an equivalence relation (the fibres of τ_n). Then τ_n is a non-split finite cover of D, and its fibre groups and binding groups are isomorphic to Z_n.

Note that the group $\Delta(V, \to)$ is isomorphic to $\mathbb{Z}$. The kernel of τ_n is a natural homomorphic image of this group.

2.3 Coverings of two-graphs

A *two-graph* T on a set X is a set of 3-subsets of X with the property that, for any 4-set $Y \subseteq X$, an even number of the 3-subsets of Y belong to T. Given any graph on X with edge set R, the set of triples carrying an odd number of edges of R is a two-graph on X. Any two-graph (X, T) arises in this way: let $x \in X$ and take as R the set of $\{y, z\}$ such that $\{x, y, z\} \in T$. The operation of *switching* a graph on X with respect to a partition of X into two parts replaces all edges between the parts with nonedges and all nonedges with edges, leaving edges and nonedges within each part unaltered. Any pair of graphs on X give the same two-graph if and only if they lie in the same switching class (that is, each can be obtained from the other by switching). See [10] for more details.

Any graph (X, R) determines a 'double covering' (X^*, R^*) of the complete graph on X as follows (this is not a cover or covering in any of the senses we have previously used, hence the quote marks). Let $X^* = \{x^+, x^- : x \in X\}$, where $(x^+, y^+), (x^-, y^-) \in R^*$ if and only if $(x, y) \in R$, and $(x^+, y^-), (x^-, y^+) \in R^*$ if and only if (x, y) is not in R (in [10] the 'double covering' is considered under the complement of our R^*). Under R^* the transversals X^+ and X^- are copies of $W_0 = (X, R)$. Switching corresponds to interchanging the labels x^+ and x^- for some points x ([10]). So we can identify the set of switching operations with F_2^X/F_2 (that is, characteristic functions modulo 2 of subsets of X, factored out by the constant functions to identify a set and its complement).

The triples from X inducing two triangles in the double covering form the two-graph T corresponding to R. It is easily seen that $M = (X^*, R^*)$ is a cover of $W = (X, T)$ under the natural map $\pi : X^* \to X$. Indeed, if $\alpha \in \mathrm{Aut}(W)$ then the graphs (X, R) and $(X, \alpha R)$ are in the same switching class so there exists $k \in F_2^X$ such that $k\alpha \in \mathrm{Aut}(M)$ (where, of course, we define $\alpha(x^+) = (\alpha(x))^+$ etc.). Moreover, the kernel of π is of order 2 (the non-identity element interchanges x^+ and x^- for every $x \in X$).

We can now produce several non-split covers. The most natural one is built in the following way.

Example 2.3.1 Let (X, R) be the countable, universal, homogeneous graph (the 'random graph'). Let $W = (X, T)$ be the corresponding two-graph, and let $\pi : M \to W$ be the double cover constructed as above, with $M = (X^*, R^*)$. We show that this does not split.

Let $G = \mathrm{Aut}(W)$. Let $x \in X$ and $Y = \{y \in X : (x,y) \in R\}$. Switching with respect to Y we get a graph on X with x as an isolated vertex and a copy of the random graph on $X \setminus \{x\}$. It follows that G is transitive on W and G_x, the stabiliser in G of x, is isomorphic to the automorphism group of the random graph (on $X \setminus \{x\}$). In particular, G acts 2-transitively on X and point stabilisers have no proper closed subgroups of finite index.

Suppose, for a contradiction, that π splits. Let H be a closed complement to $\mathrm{Aut}(M/W)$ in $\mathrm{Aut}(M)$. So H is the automorphism group of a trivial covering expansion of π. By Lemma 2.1.1 and the previous paragraph, the fibre group of this covering expansion is trivial, and so H has two orbits on X^*, and acts 2-transitively on each of these. This implies that either (X^*, R^*) or its complement is bipartite. But neither of these is the case: it is easy to see that both the complete graph and the null graph on three vertices can be embedded in (X^*, R^*). This is the contradiction, and so π is non-split. (We will give a different explanation for the non-splitting in Section 6.6.)

The notion of a *switching presentation* in [39] is a generalisation of this construction. It allows us to build non-split covers in several other cases. A typical example of this kind is as follows.

Example 2.3.2 Let (W, Cr) be the countable dense circular ordering and $<$ be an order on W inducing Cr (see Example 5 in Section 1.2). For a set $A = \{a_1, ..., a_k\}$ we construct a cover $\pi : W \times A \to W$ by defining a circular order on $W \times A$. First, extend the natural orders $<$ on all $W \times \{a_i\}$ as follows:

$$W \times \{a_1\} < W \times \{a_2\} < ... < W \times \{a_k\}.$$

Now take the corresponding circular order Cr_C on $W \times A$. This relation and the projection π define a non-split transitive cover of (W, Cr) with the kernel isomorphic to Z_k. It is shown in [39] that all finite covers of (W, Cr) may be considered as reducts of tuples of such covers.

A circular order is an example of an *oriented two-graph* (see p.117 in [10]). Moreover, Example 5 of Section 1.2 just describes the corresponding double covering of the dense circular ordering (notice that this is the above example for A consisting of two elements). This slightly explains why the examples of this subsection are amalgamated in a general construction in [39].

3 Reductions and special classes of covers

3.1 Split covers

Recall that a finite cover $\pi : C \to W$ splits if there is a closed complement to $\mathrm{Aut}(C/W)$ in $\mathrm{Aut}(C)$. Equivalently, there is a covering expansion of π with trivial kernel. Thus to determine the split covers of W we need first to describe the trivial finite covers.

3.1.1 Trivial covers

Suppose W is transitive. Let $w \in W$ and let H be a closed subgroup of finite index in $\mathrm{Aut}(W/w)$. Let C be the set of left cosets of H in $\mathrm{Aut}(W)$. The group of permutations induced by $\mathrm{Aut}(W)$ on this (by left multiplication) is closed and $\pi : C \to W$ given by $\pi(gH) = gw$ (for $g \in \mathrm{Aut}(W)$) is a transitive finite cover of W with trivial kernel and fibre group isomorphic to the group of permutations induced on the cosets of H in $\mathrm{Aut}(W/w)$ by $\mathrm{Aut}(W/w)$. It is easy to see that any transitive, trivial finite cover of W is isomorphic (over W) to such a coset space (take H as the stabiliser of a point in the fibre above w). In particular, if $\mathrm{Aut}(W/w)$ has no proper closed subgroup of finite index then any trivial finite cover $\pi : C \to W$ has the property that each $\mathrm{Aut}(C)$-orbit intersects each fibre of π in exactly one element. The following terminology is introduced in [39].

Definition 3.1.1 Let W be a permutation structure. The finite cover $\pi : C \to W$ is *strongly split* if there is a covering expansion in which all fibre groups are trivial.

Using this terminology, the above remarks give:

Lemma 3.1.2 *If W is a transitive permutation structure then all split finite covers of W are strongly split if and only if $\mathrm{Aut}(W/w)$ has no proper closed subgroup of finite index (for $w \in W$).* □

Example 3.1.3 Let $k \geq 2$ and let W be the Grassmannian of k-sets from a disintegrated set (see Section 1.1.1). Let $w \in W$ and note that $\mathrm{Aut}(W/w)$ has a closed subgroup of index 2 (those elements inducing an even permutation on the set w). Thus we get a trivial finite cover $\pi : C \to W$ with fibres of size 2 and fibre group cyclic of order 2. This has a reduct $\pi_0 : C_0 \to W$ which is free and has kernel Z_2^W (this is Example 1 of Section 1.2). Clearly this is strongly split, and so there are at least two isomorphism classes of trivial covering expansions of π_0.

3.1.2 Kernels of split covers

Lemma 3.1.4 *Suppose C, W are permutation structures and $\pi : C \to W$ is a split finite cover with kernel K. Let T be a closed complement to K in $\mathrm{Aut}(C)$. Then any closed subgroup H of K which is normalised by T is a kernel of some split covering expansion of π.*

Proof. Let $\Gamma = \mathrm{Aut}(C)$. Then TH is a subgroup of Γ whose intersection with K is H. The lemma follows once we have shown that TH is actually a closed subgroup of Γ (because the covering expansion we want can be taken to have automorphism group TH). This is a general fact about topological groups. The map $\Gamma \times K \to \Gamma \times K$ given by $(g, k) \mapsto (gk, k)$ is a homeomorphism, and compactness of K implies that projection onto the first coordinate in $\Gamma \times K$ is a closed map. So the image of $T \times H$ under the composition of these maps is closed in Γ. □

Note that if K in the above is abelian then any subgroup of K which is normalised by T is actually normal in $\mathrm{Aut}(C)$. Combining this observation with Lemma 2.1.4 we obtain:

Lemma 3.1.5 *Let $\pi : C \to W$ be a free finite cover with abelian kernel in which the fibre groups split over the binding groups. Then a subgroup K of $\mathrm{Aut}(C/W)$ is the kernel of some covering expansion of π if and only if K is a closed normal subgroup of $\mathrm{Aut}(C)$.* □

As a special case of this we get the following result. Lemma 2.1 of [4] gives a proof of this for all principal covers (not just finite covers) with abelian kernel.

Proposition 3.1.6 *Suppose $\pi : M_0 \to N$ is a principal finite cover of N with abelian kernel K_0. Then*

(i) π is split;

(ii) conjugation in $\mathrm{Aut}(M_0)$ gives an action of $\mathrm{Aut}(N)$ on K_0;

(iii) with this action, a subgroup K of K_0 is the kernel of a covering expansion of M_0 if and only if the following hold.

(a) K is invariant under $\mathrm{Aut}(N)$,

(b) K is closed in K_0.

Proof. (i) As the cover is principal, all fibre and binding groups are equal, so this follows from 2.1.4 (the transitivity assumption on the base structure is not essential).

(ii) This follows from the assumption that K_0 is abelian (cf. 6.2.1).

(iii) This follows from the above and Lemma 3.1.5. □

3.1.3 Specific examples

We consider the case where W is the rationals considered as an ordered set. Suppose $\pi : C \to W$ is a finite cover. Note that as $\mathrm{Aut}(W)$ is transitive on W all fibres have the same size (n, say) and all the fibre groups are isomorphic to a subgroup F of $\mathrm{Sym}(n)$. It follows from 2.1.1 and the fact that $\mathrm{Aut}(W/w)$ has no proper closed subgroups of finite index (for $w \in W$) that π is untwisted and so is an expansion of a free (indeed, principal) cover $\pi_0 : C_0 \to W$ with fibre group F. In other words, $\mathrm{Aut}(C)$ can be identified with a closed subgroup of the wreath product $\mathrm{Aut}(C_0) = F \,\mathrm{Wr}_W \,\mathrm{Aut}(W)$. It is shown in [39] that π must be split (see also Theorem 4.3.5 and Theorem 5.1.4 here) and so by the above remarks is strongly split. Thus we may assume $\mathrm{Aut}(C) = KG$ where $K = \mathrm{Aut}(C/W)$ and G is the natural copy of $\mathrm{Aut}(W)$ inside the above wreath product. So the cover problem is reduced to describing the possibilities for K. But these are precisely the closed G-invariant subgroups of $\mathrm{Aut}(C_0/W) = F^W$ (by Lemma 3.1.4). The following result from ([39], Theorem 2.10) determines these precisely. A similar result for finite covers of a disintegrated set was obtained previously by Martin Ziegler ([49]).

Lemma 3.1.7 *With the above notation, let $w \in W$ and let*

$$H = \{f(w) : f \in K, f(w') = 1 \ \forall w' \neq w\}.$$

Then H is independent of the choice of w, and is a normal subgroup of F. The kernel K is equal to

$$K_H = \{f \in F^W : \forall x, y \ f(x)H = f(y)H\}. \square$$

Of course, every normal subgroup H of F gives rise to such a kernel. It is also shown in [39] that these are the only possibilities for kernels if W is any *highly homogeneous* structure (that is, $\mathrm{Aut}(W)$ is transitive on the set of k-sets from W, for all finite k). Essentially, the reason is that any such W is a reduct of the rationals, by a theorem of Peter Cameron [9]. Furthermore, it is a consequence of the main results of [38] that no new kernels can be obtained if we consider binary expansions of a principal cover of a strictly minimal set. However, a projective space admits finite covers with more complicated kernels (see particularly Section 6.3.2) and the general problem of the description of the kernels of finite covers of strictly minimal sets is still open, although much can be said for abelian kernels (cf. Section 6.3.2).

3.2 Regular covers and simple covers

We shall say that a finite cover $\pi : C \to W$ is *regular* if the fibre group at each $w \in W$ acts regularly on $C(w)$ (that is, transitively and with trivial point stabiliser). The following lemma (Lemma 1.8 of [27]) enables us for some purposes to reduce to considering only regular covers. The second part enables us sometimes to reduce to the case when the fibre groups are simple. We say a cover $\pi' : C' \to W$ *factors through* π if there is a map $\pi^* : C' \to C$ such that $\pi' = \pi\pi^*$. The hypothesis below that W is transitive will not be unduly restrictive, because of Lemma 1.1.1.

Lemma 3.2.1 *Let W be a transitive permutation structure and $\pi : C \to W$ a finite cover.*

- *(a) There is a regular finite cover $\pi' : C' \to W$ which factors through π, with $\mathrm{Aut}(C) = \mathrm{Aut}(C')$.*
- *(b) If π is a regular finite cover with fibre group $G(w) := \mathrm{Aut}(C(w)/w)$ and $H < G(w)$ then there is a permutation structure C/H and finite covers $\pi_1 : C \to C/H$ and $\pi_2 : C/H \to W$ such that $\pi = \pi_2\pi_1$, the fibre group of π_1 is isomorphic to H, and that of π_2 isomorphic to $G(w)/\bigcap(gHg^{-1} : g \in G(w))$.*

If W and C are countable $\aleph_0$-categorical structures then C/H is a sort in C^{eq} and π_1 and π_2 are 0-definable in C^{eq}.

Proof. (a) Let $w \in W$, let $\bar{b}$ be an enumeration of $C(w)$, and let C' be the orbit of $\bar{b}$ under $\mathrm{Aut}(C)$. The group of permutations induced by $\mathrm{Aut}(C)$ on C' is closed.

Let $\pi_1 : C' \to C$ be the map taking each element of C' to its first coordinate. Then $\pi' = \pi\pi_1$ has the required properties.

(b) Fix $w \in W$, pick $c \in C(w)$ and let Δ be the H-orbit of c and C/H the set of $\mathrm{Aut}(C)$-translates of Δ. Then C/H is an $\mathrm{Aut}(C)$-invariant partition of C refining that given by π. Again, the group of permutations induced on C/H by $\mathrm{Aut}(C)$ is closed. Define, for $c' \in C$, $\pi_1(c')$ to be the translate Δ' of Δ containing c', and $\pi_2(\Delta')$ to be the element w' such that $\Delta' \subseteq C(w')$. □

With the above notation, it follows that if $1 = N_1 < \cdots < N_m = G(w)$ is a composition series of $G(w)$ then $C \to C/N_2 \to \cdots \to W$ is a factorisation of π into regular finite covers with simple fibre groups $N_2/N_1, \ldots, N_m/N_{m-1}$ respectively. We refer to a finite cover where all the fibre groups (and hence all binding groups) are simple as a *simple* finite cover. We now give some results due to E. Hrushovski on the possibilities for the kernels of such covers. These are taken from [27].

Definition 3.2.2 Let $\pi : C \to W$ be a simple finite cover in which all the binding groups are isomorphic to a simple group G. Suppose $w, w_1, \ldots, w_n \in W$. We say that w *depends on* $w_1, \ldots, w_n$ (in the *cover pregeometry* determined by π) if $\mathrm{Aut}(C(w)/W \cup \bigcup_{i=1}^n C(w_i))$ is trivial. Write wE_Cw_1 to indicate that w depends on w_1 in the pregeometry.

Note that $\mathrm{Aut}(C(w)/W \cup \bigcup_{i=1}^n C(w_i))$ is a normal subgroup of $\mathrm{Aut}(C(w)/W)$ so is either trivial or $\mathrm{Aut}(C(w)/W)$. The following is from ([27], Lemma 5.7 and Theorem 5.8).

Theorem 3.2.3 *With the notation as in the above definition:*

(i) The dependence relation is $\mathrm{Aut}(W)$*-invariant and satisfies the exchange condition.*

(ii) The relation E_C *is an equivalence relation.*

(iii) If G *is abelian and* w *depends on* $w_1, \ldots, w_n$ *then the equivalence class* w/E_C *is stabilised by* $\mathrm{Aut}(W/w_1, \ldots, w_n)$.

(iv) If G *is non-abelian and* w *depends on* $w_1, \ldots, w_n$ *then* w *depends on* w_i *for some* $i \in \{1, \ldots, n\}$.

Proof. (i) Invariance under $\mathrm{Aut}(W)$ is clear. Suppose w depends on $w_1, \ldots, w_n$ and w', but not on $w_1, \ldots, w_n$. For the exchange condition, we must show that w' depends on $w_1, \ldots, w_n, w$.

Let $G_1 = \mathrm{Aut}(C(w)/W \cup \bigcup_{i=1}^n C(w_i))$, $G_2 = \mathrm{Aut}(C(w')/W \cup \bigcup_{i=1}^n C(w_i))$ and $H = \mathrm{Aut}(C(w) \cup C(w')/W \cup \bigcup_{i=1}^n C(w_i))$. Then G_1 and G_2 are isomorphic to G. Moreover $H \leq G_1 \times G_2$ projects onto both coordinates, and the kernel of projection onto the second coordinate is trivial. So H is isomorphic to G, and it follows that the kernel of projection onto the first coordinate is trivial, which is what is required.

(ii) This follows immediately from (i).

(iii) Let $g \in \mathrm{Aut}(W/w_1, \ldots, w_n)$ and $w' = gw$. We must show that $w'E_Cw$. Let $\hat{g} \in \mathrm{Aut}(C)$ extend g. Then conjugation by $\hat{g}$ shows that

$$\mathrm{Aut}(\bigcup_{i=1}^n C(w_i)/W \cup C(w)) = \mathrm{Aut}(\bigcup_{i=1}^n C(w_i)/W \cup C(w')).$$

Now suppose there exists $h \in \mathrm{Aut}(C/W \cup C(w))$ with $h|C(w') \neq 1$. By the above, there exists $h' \in \mathrm{Aut}(C/W \cup C(w'))$ with the same restriction to $\bigcup_{i=1}^{n} C(w_i)$ as h. Then $h'h^{-1}$ is trivial on $W \cup \bigcup_{i=1}^{n} C(w_i)$ and non-trivial on $C(w')$. This contradicts dependence of w' on $w_1, \ldots, w_n$.

(iv) See ([27], Lemma 5.7(ii)). □

Corollary 3.2.4 *Suppose W is a primitive permutation structure with trivial algebraic closure and $\pi : C \to W$ is a simple finite cover. Then either the kernel of π is finite, or π is free.*

Proof. First note that either the binding groups are trivial (in which case π is a trivial cover), or they are equal to the fibre groups and hence simple. Primitivity implies that the equivalence relation E_C is the universal relation, or equality. In the former case, $\mathrm{Aut}(C/W \cup C(w))$ is trivial for all $w \in W$. In particular, the kernel of π is finite. So suppose we are in the latter case. If the binding groups are non-abelian then Theorem 3.2.3(iv) shows that π is free. If the binding groups are abelian then triviality of algebraic closure combined with Theorem 3.2.3(iii) again shows that π is free. □

3.3 Minimal covers

3.3.1 Existence of minimal covers

Suppose that $\pi : C \to W$ is a finite cover. Then we say that π is *minimal* if any proper expansion of C induces new structure on W (this is elsewhere referred to as a *maximal* cover: [35], for example). Equivalently, if $\mu : \mathrm{Aut}(C) \to \mathrm{Aut}(W)$ is the restriction map, π is minimal if, for every proper closed subgroup G_1 of $\mathrm{Aut}(C)$, $\mu(G_1) \neq \mathrm{Aut}(W)$. Clearly, if the collection of proper subcovers (that is, proper expansions of C inducing no new structure on W) has the descending chain condition, then there is a minimal cover. In Theorem 5.4 of [27], it is shown that if W is countable, $\aleph_0$-categorical and has a 'nice enumeration' (a technical combinatorial condition due to Ahlbrandt and Ziegler [2]) then any finite cover of W does indeed have the descending chain condition on subcovers (see Section 7.2.2 for more on this). If we are just interested in minimal covers, however, then the following group-theoretic result of Cossey, Kegel and Kovács [17] gives something stronger.

Proposition 3.3.1 *Let Γ and Σ be Hausdorff topological groups, and let $\mu : \Gamma \to \Sigma$ be a continuous epimorphism with compact kernel K. Then there is a closed subgroup $G \leq \Gamma$ such that $\mu(G) = \Sigma$, and for every proper closed subgroup H of G we have $\mu(H) < \Sigma$.*

Proof. Let $\mathcal{F}$ be the set of closed subgroups H of Γ such that $\mu(H) = \Sigma$. By Zorn's Lemma, it suffices to show that if $(H_i : i \in I)$ is a chain (under inclusion) of members of $\mathcal{F}$ then $H := \bigcap(H_i : i \in I)$ is in $\mathcal{F}$. Clearly H is closed. Let $g \in \Sigma$. Then there is $h_0 \in \Gamma$ such that $\mu(h_0) = g$. Now $\{h \in \Gamma : \mu(h) = g\}$ is just the compact set h_0K. Each set $H_i \cap h_0K$ is non-empty, so by compactness $H \cap h_0K$ is non-empty, and it follows that $H \in \mathcal{F}$. □

Corollary 3.3.2 *Let W be a permutation structure.*

(a) Let C be a symmetric algebraic extension of W. Then C has a maximal expansion which is a symmetric extension of W.

(b) Let $\pi : C \to W$ be a finite cover. Then there is an expansion of C which is a minimal cover of W.

Proof. (a) Apply Lemma 1.5.1(b) and the above Propsition.

(b) This follows from (a), since a finite cover is a special case of a symmetric algebraic extension. □

Corollary 3.3.3 *Let W be a permutation structure. Then the following are equivalent:*

(a) every finite cover of W splits;

(b) every minimal finite cover of W is trivial.

Proof. Suppose that (a) holds, and let $\pi : C \to W$ be a minimal finite cover with restriction map μ. By (a), its kernel K has a closed complement G. By minimality, $G = \mathrm{Aut}(C)$, so $K = 1$.

Conversely, assume (b), and let $\pi : C \to W$ be a finite cover. Then by Corollary 3.3.2 and (b), there is a minimal cover expanding C with automorphism group G and trivial kernel. Then G is a closed complement to the kernel of π. □

Remark. Corollary 3.3.2 tells us more than just Corollary 3.3.3. It says that once we know the minimal covers of W and the possible kernels, we know all the finite covers of W. For if $\pi : C \to W$ is a finite cover, then $\mathrm{Aut}(C) = K.\mathrm{Aut}(M)$, where M is a minimal cover of W which is an expansion of C, and K is the kernel of π.

3.3.2 Frattini covers

We shall state our next result in the broader language of topological groups. So we shall say that a continuous epimorphism of Hausdorff topological groups $\phi : G \to H$ is a *Frattini cover* if for every proper closed subgroup G_1 of G we have $\phi(G_1) \neq H$. (The reason for the terminology is that if G and H are profinite then ϕ is Frattini if and only if $\ker(\phi)$ is contained in the Frattini subgroup of G, that is, the intersection of the maximal open subgroups of G – see Section 20.6 of [28]). A version of the 'Frattini argument' for finite groups yields the following. It uses results from Section 20.10 of [28], where the Sylow theory for profinite groups is presented. (If p is a prime, a closed subgroup of a profinite group is a Sylow p-subgroup if the index of every open subgroup containing it is coprime to p, and no closed subgroup properly contained in it has this property.)

Lemma 3.3.4 ([25]) *(a) Suppose that $\phi : G \to H$ is a Frattini cover of Hausdorff topological groups such that $K := \ker \phi$ is profinite. Then K is pronilpotent (that is, an inverse limit of nilpotent groups).*

(b) If $\pi : C \to W$ is a minimal finite cover and the fibres of π are of bounded size, then the kernel of π is nilpotent.

(c) Suppose W is a permutation structure such that $\mathrm{Aut}(W)$ has finitely many orbits on W. Let $\pi : C \to W$ be a finite cover of W. Then there is a covering expansion of π with nilpotent kernel.

Proof. (a) It suffices to show that for each prime p the group K has a unique Sylow p-subgroup, for then each Sylow subgroup is normal in K, so the Sylow subgroups commute. Let P be a Sylow p-subgroup of K. Then, as P is closed, so is $N_G(P)$. Let $g \in G$. Then P^g is a Sylow p-subgroup of K, so by Proposition 20.43 of [28] there is $k \in K$ such that $P^g = P^k$. Hence $gk^{-1} \in N_G(P)$, so we obtain $G = KN_G(P)$. As ϕ is a Frattini cover and $N_G(P)$ is closed, it follows that $N_G(P) = G$, that is, P is normal in G. Now K is topologically generated by its Sylow subgroups (by 20.43(d) of [28]) and is a commuting product of them, and as each of the Sylow subgroups is pronilpotent, so is K.

(b) From (a) and Lemma 1.5.1 we get that the kernel K of π is pronilpotent. But K is a subdirect product of finite groups of bounded size, so is therefore nilpotent.

(c) follows from Corollary 3.3.2 and (b). □

We omit the proof of the following (it makes use of 3.3.4). It enables us for some purposes just to work with covers with abelian kernel, where the machinery of Section 6 can be used.

Corollary 3.3.5 ([25]) *Let W be a permutation structure. Then the following are equivalent:*

(a) every finite cover of W splits;

(b) every finite cover of W with elementary abelian kernel splits. □

3.3.3 Amalgamation

The following lemma shows how minimal finite covers can be amalgamated. It is really nothing more than the fibre-product construction and Lemma 20.30 of [28], adapted to our purposes.

Lemma 3.3.6 *Suppose $\pi_1 : C_1 \to W$ and $\pi_2 : C_2 \to W$ are minimal finite covers of permutation structures. Then there is a minimal finite cover $\sigma : M \to W$ which factors through both π_1 and π_2. If C_1 and C_2 are transitive, then M can be taken to be transitive. If π_1 and π_2 have finite kernel, then σ can be taken to have finite kernel.*

Proof. Let

$$M_0 = \{(c_1, c_2) \in C_1 \times C_2 : \pi_1(c_1) = \pi_2(c_2)\}.$$

Then

$$\Gamma_0 = \{(g_1, g_2) : g_i \in \mathrm{Aut}(C_i),\ g_1|W = g_2|W\}$$

is a closed subgroup of $\mathrm{Sym}(M_0)$. There is a natural continuous epimorphism $\rho_0 : \Gamma_0 \to \mathrm{Aut}(W)$ and this has compact kernel. So by 3.3.1 there is a closed subgroup Γ of Γ_0 such that the restriction of ρ_0 to Γ is a Frattini cover of $\mathrm{Aut}(W)$. Let M be the permutation structure with domain M_0 and automorphism group Γ. If $\sigma : M \to W$ is the map $\sigma(c_1, c_2) = \pi_1(c_1) = \pi_2(c_2)$, then σ is a minimal finite cover. Clearly σ factors through π_1 and π_2. The fact that any automorphism of C_1 or C_2 lifts to an element of $\mathrm{Aut}\, M$ follows from minimality of π_1 and π_2. For a transitive cover we can use 3.2.1. If π_1 and π_2 have finite kernels, then it is clear that the kernel of σ is also finite. □

3.4 Irreducibility conditions

We say that the automorphism group of a permutation structure M (or the structure itself) is *irreducible* if $\mathrm{Aut}(M)$ has no proper closed subgroups of finite index. Equivalently, for countable $\aleph_0$-categorical M, if $\mathrm{acl}^{\mathrm{eq}}(\emptyset) = \mathrm{dcl}^{\mathrm{eq}}(\emptyset)$. Most of the general results on describing the finite covers of a permutation structure W will assume some sort of irreducibility conditions: usually on W itself, but also sometimes on various point stabilisers. This may seem unduly restrictive, so some further comments on these hypotheses are in order.

Firstly, many of the familiar $\aleph_0$-categorical structures (such as all the primitive homogeneous graphs and digraphs) have the property that $\mathrm{Aut}(W/X)$ is irreducible for all finite algebraically closed X (see Lemma 3.3 and Theorem 4.1(a) of [23] for some examples of how to verify this). Even if W is not irreducible then it is often possible to consider an expansion of W which is, and use knowledge of the finite covers of this to determine information about the finite covers of W. An example of this is the treatment of the highly homogeneous betweenness and separation relations in [39].

A second reason for the irreducibility conditions is that without them it is easy to construct slightly exotic finite covers. For example, results in Section 2.1 show how to construct non-split free covers when point stabilisers are not irreducible. The following simple lemma shows that if W itself is not irreducible then constructing non-split covers is even easier.

Lemma 3.4.1 *Let W be a permutation structure which is not irreducible. Suppose $\mathrm{Aut}(W)$ has a smallest closed subgroup of finite index. Then there exists a non-split finite cover of W with finite kernel.*

Proof. Let $\phi : \mathrm{Aut}(W) \to F$ be a continuous group epimorphism with F finite and as large as possible. Let

$$1 \to A \to G \xrightarrow{\psi} F \to 1$$

be a non-split finite extension of F (see Remark 2.1.6). Consider G acting faithfully on some finite set X. Let $C = W \times X$ and consider this as the domain of a permutation structure with automorphism group

$$\Gamma = \{(\gamma, g) \in \mathrm{Aut}(W) \times G : \phi(\gamma) = \psi(g)\}.$$

Projection to the first coordinate gives a finite cover $\pi : C \to W$, which is easily seen to be non-split. Indeed, the kernel of the cover is $K = \{(1, \alpha) : \alpha \in \ker \psi\}$, which is obviously isomorphic to A. Suppose for a contradiction that $\Gamma = KH$, where the restriction map is an isomorphism between H and $\mathrm{Aut}(W)$. Applying π_2, projection to the second co-ordinate, gives that $\pi_2(H)$ is a supplement to $\ker \psi$ in G. But the maximum size for a continuous finite homomorphic image of H is $|F|$. So $\pi_2(H)$ is a complement to $\ker \psi$ in G, a contradiction. □

In Section 7 of the notes, we indicate how results which are qualitative rather than quantitative can be obtained if we weaken the irreducibility conditions to G-finiteness: a notion due to Lascar ([46]), and which in many ways seems to be the most attractive level of generality one could aim for.

Definition 3.4.2 We say that a permutation structure W (or its automorphism group $\mathrm{Aut}(W)$) is *G-finite* if for all finite $X \subseteq W$ there is a smallest closed subgroup of finite index in $\mathrm{Aut}(W/X)$.

3.5 Superlinked covers

The cover $\pi : C \to W$ is said to be *superlinked* if it has finite kernel. This terminology was introduced in [27].

Recall that a permutation structure W is said to have *trivial algebraic closure* if the pointwise stabiliser in $\mathrm{Aut}(W)$ of any finite $A \subseteq W$ has no finite orbits on $W \setminus A$. Many homogeneous relational structures have this property, but non-trivial Grassmannians do not. For structures W with trivial algebraic closure, the following theorem gives good reasons for restricting to superlinked covers. The key point is the description of the kernels of simple finite covers of such structures given in Corollary 3.2.4.

Theorem 3.5.1 ([23], Lemma 2.5) *Suppose that W is a primitive permutation structure with trivial algebraic closure, and that each point stabiliser $\mathrm{Aut}(W/w)$ is irreducible. If every superlinked finite cover of W with simple fibre groups splits, then every finite cover of W splits.*

Proof. By Corollary 3.3.2, it suffices to show that if $\pi : C \to W$ is a minimal finite cover then it is trivial, so suppose for a contradiction that it is non-trivial. By Lemma 3.2.1(a) we may suppose that π is regular, and by Lemma 3.2.1(b) π can be factored as $\pi_2\pi_1$, where $\pi_1 : C \to C_1$ and $\pi_2 : C_1 \to W$ and the latter has (non-trivial) simple fibre groups. As $\mathrm{Aut}(W/w)$ is irreducible, $\mathrm{Aut}(C_1(w)/w) = \mathrm{Aut}(C_1(w)/W)$ (by Lemma 2.1.1), so these groups are non-trivial and π_2 is non-trivial. It follows from Corollary 3.2.4 that π_2 is free or superlinked. So in either case, π_2 splits (by Lemma 2.1.4 in the fomer case, and by hypothesis in the latter). Thus there is a closed subgroup H of $\mathrm{Aut}(C)$ with $H \cap \mathrm{Aut}(C/W) = \mathrm{Aut}(C/C_1)$ and the image of H under the restriction map to W being $\mathrm{Aut}(W)$. As π_2 is non-trivial, this contradicts minimality of π. □

4 Finite covers with finite kernels

At the end of the previous section we showed how the splitting question for the finite covers of certain permutation structures reduced to the special case of *superlinked* covers, where the kernel is finite. In this section, which is mainly a commentary on [24], we analyse these types of covers in detail. Initial reductions show that we should consider two sorts of superlinked covers: locally trivial and locally transitive ones. In the nice cases, the first sort can be described in terms of digraph coverings, and the second type can be thought of as being like a vector space covering its projective space.

4.1 Elementary reductions

We consider a finite cover $\pi : C \to W$ of permutation structures. Since any $\aleph_0$-categorical structure is biinterpretable with a transitive one (Lemma 1.1.1) we usually assume that W is transitive. Moreover since in this case we can replace π by a regular finite cover (Lemma 3.2.1) we also often assume that C is transitive (and say that π is a transitive finite cover).

As explained in Section 3.4, irreducibility is a useful assumption on C, and we shall work mostly with irreducible C and W. In fact, sometimes irreducibility of $\mathrm{Aut}(C)$ follows from that of $\mathrm{Aut}(W)$.

Lemma 4.1.1 *Let W be irreducible, and let $\pi : C \to W$ be a minimal cover. Then C is irreducible.*

Proof. Let H be a closed subgroup of $\mathrm{Aut}(C)$ of finite index, and let $\mu(H)$ be its image in $\mathrm{Aut}(W)$ under restriction. By Lemma 1.4.2, this is a closed subgroup of $\mathrm{Aut}(W)$ of finite index, so equals $\mathrm{Aut}(W)$. Hence by minimality of π, we have $H = \mathrm{Aut}(C)$. □

A non-trivial bonus from our irreducibility assumption is the following. The example to keep in mind is Example 3 of 1.2 (a vector space over a finite field covering a projective space).

Lemma 4.1.2 *Suppose that $\pi : C \to W$ is an irreducible, superlinked finite cover. Then*

(a) $\mathrm{Aut}(C/W)$ *is central in* $\mathrm{Aut}(C)$*;*

(b) the cover is split if and only if it is trivial.

Proof. (a) As $\mathrm{Aut}(C/W)$ is a finite normal subgroup of $\mathrm{Aut}(C)$ its centraliser in $\mathrm{Aut}(C)$ is a closed subgroup of finite index in $\mathrm{Aut}(C)$, so, by irreducibility, equal to $\mathrm{Aut}(C)$.

(b) Since the kernel is finite, any complement is a closed subgroup of finite index in $\mathrm{Aut}(C)$. □

In Section 2.2, we showed how quotients of the universal covering of a digraph could give examples of minimal superlinked finite covers. These covers $\pi : C \to W$ satisfy the first two of the following conditions:

Definition 4.1.3 Suppose $\pi : C \to W$ is a (transitive) finite cover. Then:

- π is *untwisted* if the fibre groups and binding groups are equal;
- π is *locally trivial* with respect to a certain $\mathrm{Aut}(W)$-orbit R on W^2 if whenever $(x, w) \in R$, then $\mathrm{Aut}(C(x)/C(w)) = 1$;
- π is *locally transitive* with respect to a certain $\mathrm{Aut}(W)$-orbit R on W^2 if whenever $(x, w) \in R$, then $\mathrm{Aut}(C/C(w))$ is transitive on $C(x)$.

The example to bear in mind for a locally transitive finite cover is that of a vector space covering its projective space. See Example 2 of Section 1.2 for a twisted (trivial) cover. These examples are in some sense typical, as we shall see.

There is a lemma (Lemma 1.6 of [24] - omitted here) which enables us under standard hypotheses (transitive, irreducible, superlinked) to factor a cover as an trivial cover followed by an untwisted cover. The conditions above are on particular covers, not on W alone, but by the following lemma, irreducibility assumptions on one and two point stabilisers in $\mathrm{Aut}(W)$ often ensure that the first two conditions in the above hold. Example 4 of 1.2, and Lemma 2.1.5, indicate how reducibility of the point stabiliser can give twisted non-split finite covers of W in a canonical way. The following are taken from [24].

Lemma 4.1.4 *Let $\pi : C \to W$ be a transitive finite cover whose kernel is central in $\mathrm{Aut}(W)$. Then*

(a) for any $c \in C$, $\mathrm{Aut}(C/c, W) = 1$,

(b) if $\mathrm{Aut}(W/w)$ is irreducible, then the cover is untwisted, and

(c) if $x, w \in W$ and $\mathrm{Aut}(W/x, w)$ is irreducible, then

$$\mathrm{Aut}(C(x)/C(w)) = 1.$$

Proof. (a) Straightforward, and omitted (see ([24], Lemma 1.3)).

(b) This follows from Lemma 2.1.1.

(c) Let $\mu : \mathrm{Aut}(C) \to \mathrm{Aut}(W)$ be restriction. By Lemma 1.4.2 and our irreducibility assumption,

$$\mu(\mathrm{Aut}(C/C(x), C(w))) = \mathrm{Aut}(W/x, w).$$

Let $g \in \mathrm{Aut}(C/x, C(w))$. There is $g' \in \mathrm{Aut}(C/C(x), C(w))$ such that $\mu(g) = \mu(g')$. It follows by (a) that $g'^{-1}g = 1$, so $g \mid C(x) = 1$, as required. □

Of course, this result is applicable to irreducible superlinked finite covers, because of Lemma 4.1.2.

We now record a factorisation lemma, which enables us to separate out the locally trivial part of a cover (to which we can apply Theorem 4.2.2) and (when the locally trivial part is trivial) obtain a locally transitive cover.

Lemma 4.1.5 ([24], Lemma 3.1) *Let W be a transitive irreducible permutation structure. Let $\pi : C \to W$ be an untwisted, transitive, irreducible, superlinked finite cover and $x, w \in W$ be distinct. Then there are finite covers $\pi_1 : C \to C'$ and $\pi_2 : C' \to W$ such that the following hold.*

(a) $\pi = \pi_2\pi_1$.

(b) $\mathrm{Aut}(C'(x)/C'(w)) = 1$.

(c) if $s \in C'(x)$ and $t \in C'(w)$ then $\mathrm{Aut}(C/C(t), s)$ is transitive on $C(s)$.

Sketch Proof. We just give the construction of π_1 and π_2. Let $K := \mathrm{Aut}(C/W)$. As π is untwisted, the natural map $j : K \to \mathrm{Aut}(C(x))$ is an isomorphism. Put $K_1 := j^{-1}(\mathrm{Aut}(C(x)/C(w)))$. Let C' be the set of K_1-orbits on C, and define π_1, π_2 by

$$\pi_1(c) = K_1c \text{ and } \pi_2(K_1c) = \pi(c).$$

Also let $\mathrm{Aut}(C')$ be the group of permutations of C' induced by $\mathrm{Aut}(C)$. It is easy to check that these covers satisfy our conditions. □

4.2 Graphic triples and digraphs

We give a combinatorial condition which holds of many homogeneous structures and enables us to control the superlinked finite covers via digraph coverings.

If W is a permutation structure, A a finite subset of W, and $n \in \mathbb{N}$, then by an *n-type* of W over A we mean an $\mathrm{Aut}(W/A)$-orbit P on W^n (and if $A = \emptyset$ we refer to this simply as an n-type). For $\bar{x} \in W^n$ we write $P(\bar{x})$ to indicate that $\bar{x} \in P$.

Definition 4.2.1 Suppose that L is an irreducible transitive permutation structure with a 3-type P and 2-types Q, R. We say that (P, Q, R) is a *graphic triple of types* if the following hold.

1. $P(w, x, y)$ implies w, x, y are distinct and $Q(w, x), Q(w, y), R(x, y)$;
2. R (as a digraph relation) is connected;
3. if $(w, x, y), (w', x, y) \in P$ and H is a closed subgroup of $\mathrm{Aut}(L/x, y)$ of finite index then w, w' lie in the same H-orbit;
4. either of the following holds:
   (a) if $R(x, y), R(x, z), R(y, z)$ then there is $w \in L$ such that $P(w, x, y)$, $P(w, y, z)$ and $P(w, x, z)$;
   (b) $P := \{(w, x, y) \in L^3 : R(w, x), R(w, y), R(x, y)\}$.

Usually we construct graphic triples via *strong types* (see Definition 4.3.1 below), and then we have $Q = R$.

The main theorem on graphic triples is the following. It describes the Q-locally trivial finite covers of a structure L with a graphic triple (P, Q, R) in terms of digraph coverings of $(L; R)$. Note that connectedness of R implies connectedness

of the relation Q, so such a finite cover is necessarily superlinked. The 'untwisted' assumption seems unavoidable, in view of constructions like Example 4 of Section 1.2.

Theorem 4.2.2 ([24], Theorem 1.13) *Let L be a transitive, irreducible permutation structure with a graphic triple of types (P,Q,R). Suppose that $\tau : C \to L$ is an untwisted, transitive, irreducible finite cover such that whenever $x, w \in L$ with $Q(w,x)$ then $\mathrm{Aut}(C(x)/C(w)) = 1$. Then there is an $\mathrm{Aut}(C)$-invariant digraph relation R' on C such that $\tau : (C,R') \to (L,R)$ is a covering of digraphs, and an automorphism g of the pure digraph (C,R') is an automorphism of C if and only if g maps τ-fibres to τ-fibres and induces an automorphism of L.*□

We omit the proof here, but observe that R' is defined as follows:

> For $a,b \in C$, we have $R'(a,b)$ if and only if there is $w \in L$ such that $P(w,\tau(a),\tau(b))$ and a,b lie in the same $\mathrm{Aut}(C/C(w))$-orbit.

(Hypothesis 3 in the above definition ensures that this is independent of choice of w.)

The theorem has the following corollary.

Corollary 4.2.3 ([24], Corollary 1.14) *Let L be an irreducible transitive permutation structure with a graphic triple (P,Q,R). Let $\sigma : (U,R'') \to (L,R)$ be the universal covering of digraphs, let Δ be the group of deck transformations, normal in the subgroup Γ of all elements of $\mathrm{Aut}(U,R'')$ which map fibres to fibres and induce automorphisms of L.*

- *(a) Let $\tau : C \to L$ be an untwisted, irreducible transitive finite cover such that whenever $Q(w,x)$ holds in L we have $\mathrm{Aut}(C(x)/C(w)) = 1$. Then there is a surjective group homomorphism $\Gamma \to \mathrm{Aut}(C)$ whose kernel H is contained in Δ, such that $\Delta/H \cong \mathrm{Aut}(C/L)$.*
- *(b) Conversely, if H is a closed normal subgroup of Γ of finite index in Δ such that Γ/H is irreducible, then Δ/H is isomorphic to the kernel of an irreducible finite cover $\tau_H : C_H \to L$.*□

Again, we omit the proof. In (b), the set C_H is just the set of H-orbits on U. There is a digraph relation R' on C_H, with $R'(a,b)$ holding if and only if there are $u \in a$ and $v \in b$ such that $R''(u,v)$. The structure on C_H is that preserved by those permutations of C_H which preserve R', induce automorphisms of L, and preserve the set of fibres of the natural map to L.

Example. Recall Example 2.2.2. There, for each $n > 1$ a construction was given of a non-split finite cover V_n of the countable homogeneous local order D. It follows from Theorem 4.2.2 and Lemma 4.1.4 that the finite covers V_n described there are the only transitive irreducible superlinked finite covers of D.

The irreducible superlinked covers of the countable homogeneous circular order can be described similarly ([24], Example 2.9). A different approach, due to Ivanov, can be found in [39].

4.3 Strong types

We introduce a very weak (and in general non-symmetric) notion of independence, which is convenient for handling covers. From it we can construct graphic triples, but (Theorem 4.3.3) the digraph coverings which arise are all trivial.

Definition 4.3.1 Suppose that W is a permutation structure. A *strong type* over W is a function p which assigns to each finite $A \subseteq W$ a 1-type over A, denoted by $p \mid A$, subject to the following coherence conditions.

(i) for all $A \subset W$, $A \cap (p \mid A) = \emptyset$;
(ii) if $A \subseteq A'$ then $p \mid A' \subseteq p \mid A$;
(iii) if $g \in \mathrm{Aut}(W)$ then $g(p \mid A) = p \mid gA$.

Model-theoretically, a strong type over W may be regarded as a non-algebraic complete 1-type q with W as a set of parameters, such that for any a realising q (in an elementary extension of W), $g \in \mathrm{Aut}(W)$ and $\bar{b} \in W$ we have $\mathrm{tp}(a, \bar{b}) = \mathrm{tp}(a, g\bar{b})$. We give some examples.

(a) If W is the random graph, then there are two strong types, corresponding to adjacency to everything in A, or to non-adjacency. Similarly, if $W = (\mathbb{Q}, <)$, then there are two strong types, corresponding to being greater than everything in A, or to being less than everything in A.

(b) If W is a vector space (or its projective space) then we have a strong type by taking as $p \mid A$ the vectors (respectively, points) linearly independent from A.

(c) More generally, if W is stable, saturated and irreducible, then any element independent from W (over $\emptyset$) realises a strong type over W.

(d) If $(\mathbb{Q}, T)$ is the countable dense linear betweenness relation without endpoints (on the rationals), then there is no strong type over W. Suppose p were such a thing, and let $a, b \in \mathbb{Q}$. Then (without loss of generality) $p|\{a, b\}$ consists of all $x \in \mathbb{Q}$ greater than a and b. Let g be an automorphism interchanging a and b. Applying this gives a contradiction to (iii). A similar argument shows that the countable dense circular order has no strong type (for once any parameter is named, all pairs acquire an orientation, so again (iii) is violated).

(e) The countable, universal, homogeneous local order D has no strong type. For any $x \in L$ the only posssibility for $p|\{x\}$ would be the in-vertices, or the out-vertices of x (by homogeneity). But no point dominates, or is dominated by a cycle, and so (ii) is impossible to satisfy with either of these.

Remark. There is a related notion due to Ivanov. A structure W is said to have *orthogonal copies* if the conditions of Definition 4.3.1 hold with respect to *all* types, not just 1-types. Theorem 4.3.5 below can also be deduced from Ivanov's

results on structures having orthogonal copies: see [39] for more details, as well as Theorem 5.1.4 here.

The following lemma (whose proof we omit) enables us to construct graphic triples from strong types.

Lemma 4.3.2 ([24], Lemma 2.2) *Let p be a strong type over W, let A be a finite subset of W, and H be a closed normal subgroup of finite index in* $\mathrm{Aut}(W/A)$. *Then H is transitive on $p \mid A$.* □

Theorem 4.3.3 ([24]) *Let W be a transitive irreducible permutation structure, and p be a strong type over W. Define*

$$R = \{(x, y) \in W^2 : x \in p \mid \{y\}\},$$

$$P = \{(w, x, y) \in W^3 : R(x, y) \text{ and } w \in p \mid \{x, y\}\}.$$

Then

(a) (P, R, R) is a graphic triple of types for W;

(b) any two paths in (W, R) with the same endpoints are homotopic.

Proof. (a) Since W is transitive, P and R are types. Conditions (1) and (4)(a) of the definition of graphic triple are immediate, as is (2) (the digraph has diameter two). For (3), suppose that H is a closed subgroup of $\mathrm{Aut}(W)$ of finite index in $\mathrm{Aut}(W/x, y)$. Then by the last lemma, $\{w : P(w, x, y)\}$ is an H-orbit.

(b) Let $x_0 \ldots x_k$ be a path in (W, R). Pick $w \in p \mid \{x_0, \ldots, x_k\}$. In particular, $R(w, x_i)$ for $i = 0, \ldots, k$. It suffices to show that $x_0 \ldots x_k$ is homotopic to $x_0 w x_k$. This reduces inductively to showing that $x_0 x_1$ is homotopic to $x_0 w x_1$. If $R(x_0, x_1)$ we have elementary homotopies

$$x_0 x_1 \sim x_0 w x_0 x_1 \sim x_0 w x_1,$$

and if $R(x_1, x_0)$ we have

$$x_0 x_1 \sim x_0 x_1 w x_1 \sim x_0 w x_1. \square$$

From this and Theorem 4.2.2, we obtain

Corollary 4.3.4 ([24], Corollary 2.4) *Let W be a transitive irreducible permutation structure with a strong type p and corresponding graphic triple (P, R, R). Let $\pi : C \to W$ be an untwisted irreducible finite cover such that whenever $R(x, y)$ holds in W, we have* $\mathrm{Aut}(C(x)/C(y)) = 1$. *Then π is trivial.* □

We give an application of Corollary 4.3.4 to finite covers of some familiar homogeneous structures. There is a slightly more general statement (phrased more group-theoretically) in Theorem 2.7 of [24]. Parts (i)-(iv) are also proved (in a different way) in ([23], Theorem 4.1).

Theorem 4.3.5 *Suppose that W is one of the following countable $\aleph_0$-categorical structures.*

i. a pure set;
ii. $(\mathbb{Q}, <)$;
iii. any homogeneous graph with primitive automorphism group;
iv. a Henson digraph (that is, a homogeneous digraph whose set of finite substructures is the collection of all finite digraphs which do not embed any of a given set of finite tournaments);
v. a vector space over a finite field.

Then any superlinked finite cover of W splits. □

In cases (i)-(iv) Theorem 4.3.5 combined with Lemma 3.5.1 shows that any finite cover splits. This generalises earlier results of Ziegler [49], and Ivanov ([38] and [39]) on finite covers of a pure set and some other structures. Results of Ivanov in Section 5 also provide a self-contained proof of 4.3.5.

4.4 A vector space covering its projective space

The most obvious algebraic example of a finite cover is Example 3 from Section 1.2: the map $\pi : V \setminus \{0\} \to \mathrm{PG}(V)$ given by $v \mapsto \langle v \rangle$, where V is a countable-dimensional vector space over a finite field. This example is not covered by the above theory, since it is not locally trivial. In fact, it is *locally transitive*, that is, the pointwise stabiliser of any one fibre is transitive on any other fibre. We now describe some of the results of Section 3 of [24], which subsume this example. The key feature of the projective space which gives rise to the non-split superlinked covering (namely the vector space) is that the stabiliser of two points has the multiplicative group of the field as a continuous homomorphic image (the stabiliser acts regularly on the remaining points of the line through the two points).

In the rest of this section we assume W to be an irreducible, countable, transitive, $\aleph_0$-categorical structure, whose 1-point stabilisers are irreducible, and we also suppose that W has a strong type p and a corresponding graphic triple (P, R, R). We shall sketch a classification of all the transitive irreducible superlinked finite covers of W. By Lemma 2.1.1, all such covers are untwisted.

Fix $x, w \in W$ such that $R(w, x)$. First, let $\pi : M \to W$ be a transitive, irreducible superlinked finite cover. Let $\pi = \pi_2\pi_1$ be the factorisation in Lemma 4.1.5, relative to the pair (x, w). Then by Corollary 4.3.4, π_2 is trivial, so $\pi = \pi_1$. In particular, $\mathrm{Aut}(M/M(w))$ is transitive on $M(x)$. Let $K := \mathrm{Aut}(M/W)$. As π is untwisted, it follows by Lemma 4.1.4(a) that K is isomorphic to $\mathrm{Aut}\, M(x)$ and acts regularly on $M(x)$. From this and the last paragraph, it follows that by comparing actions on $M(x)$ we get a surjective homomorphism from $\mathrm{Aut}(M/M(w), x)$ to K. It is easy to check that restriction to W gives an isomorphism $\mathrm{Aut}(M/M(w), x) \to \mathrm{Aut}(W/w, x)$. By composing these maps we get an epimorphism

$$\phi_{w,x} : \mathrm{Aut}(W/w, x) \to K.$$

If $g \in \mathrm{Aut}(W)$ and $h \in \mathrm{Aut}(W/w, x)$, then

$$\phi_{gw,gx}(ghg^{-1}) = \phi_{w,x}(h). \tag{1}$$

Lemma 4.4.1 (Lemma 3.3 of [24]) *In the above situation, the following hold.*

(a) Suppose that $R(w,x)$, $R(w,y)$, and $R(x,y)$ hold. Then for any element g of $\mathrm{Aut}(W/w,x,y)$ we have

$$\phi_{w,y}(g) = \phi_{w,x}(g)\phi_{x,y}(g).$$

(b) If R is symmetric, then for $(w,x) \in R$ we have $\phi_{w,x} = \phi_{x,w}^{-1}$. □

Definition 4.4.2 Let W be as above, and K a finite abelian group. Then a family $(\phi_{w,x} : (w,x) \in R)$ of continuous epimorphisms $\phi_{w,x} : \mathrm{Aut}(W/w,x) \to K$ is called a *conjugate system of homomorphisms* for (W, R, K) if conditions (1) and condition (a) of Lemma 4.4.1 hold. If the family arises from a cover π as above, it is said to be *associated with* (π, R).

The following theorem is an amalgam of Theorems 3.5, 3.7 and 3.8 of [24] (where it is not assumed that one-point stabilisers in $\mathrm{Aut}(W)$ are irreducible). In the proof, the above observations are reversed, and a finite cover is constructed from a conjugate system of homomorphisms.

Theorem 4.4.3 *Suppose that W is a countable, transitive $\aleph_0$-categorical structure with a strong type p and an associated graphic triple (P, R, R), and let K be a finite abelian group. There is a one-to-one correspondence between conjugate systems of homomorphisms for (W, R, K), and untwisted, transitive, irreducible, locally transitive, superlinked finite covers $\pi : M \to W$ with kernel isomorphic to K.* □

There are generalisations of this result due to Jeffrey Koshan ([44], [45]) not assuming that W has a strong type. These are complicated to state in full generality, but one result assumes only that for all $x, y, z \in W$ there exists $w \in W$ such that x, y, z lie in the same $\mathrm{Aut}(W/w)$-orbit. Given this, and irreducibility of W and point stabilisers in W, Koshan describes explicitly all irreducible superlinked finite covers of W.

5 The cover problem and independence

5.1 Strongly determined types

The techniques and notions in the previous section parallel very clearly some ideas from stability theory (strong types, stationarity, and distinguished extensions of types). Indeed, the stability-theoretic notions motivated many of the definitions there, although to avoid presuming specialist knowledge, we did not present them in this way. In this section we amplify further on this, and consider the results of Section 4 from this viewpoint. We use standard model-theoretic terminology and

results throughout. The following theorem serves as the source of many of the ideas presented in this section, and shows that assuming stability simplifies considerably the problems considered in the previous section.

Theorem 5.1.1 ([24], Lemma 2.5) *Let M, W be transitive, irreducible, stable saturated structures. If $\tau : M \to W$ is an irreducible transitive finite cover with local triviality for independent pairs, then it is one-to-one.*

Proof. Let $x, w \in W$ be independent (over $\emptyset$) and $a, b \in M(x)$. By transitivity of M, $tp(a/\emptyset) = tp(b/\emptyset)$. Since $M(w) \subseteq acl(w)$, both a and b are independent from $M(w)$. If a and b are distinct then saturation and local triviality of M imply that they have distinct types over $M(w)$. Thus there are two distinct non-forking extensions of $tp(a/\emptyset)$ over $M(w)$. But then the Finite Equivalence Relation Theorem implies that M must be reducible, a contradiction. □

The main point of the above proof is the existence of some independent w. Also, we needed the Finite Equivalence Relation Theorem, because we considered a finite cover. So finite equivalence relations are unavoidable in considerations of this kind. To some extent, the irreducibility assumptions of the previous section finesse this: irreducibility of W means precisely that there are no non-trivial invariant finite equivalence relations on any $\mathrm{Aut}(W)$-orbit of W^n for all finite n.

One can ask whether it is possible to generalise this argument introducing weaker variants of independence. The notion of a strong type in the previous section is one way of doing this. We now give a more refined version of this, not assuming irreducibility of W. First, we introduce some notation and terminology. Denote by $\mathrm{Aut}^\circ(W)$ the intersection of all closed subgroups of finite index in $\mathrm{Aut}(W)$ and refer to the elements of this as *strong maps*. In the countable, $\aleph_0$-categorical case, this is precisely the group of automorphisms preserving all 0-definable finite equivalence relations (on W^n, for all n). From now on we assume that W is countable and $\aleph_0$-categorical (but we should mention here that the material of this section largely extends to the general case [40]). The following is a generalisation of the notion of strong type introduced earlier (4.3.1).

Definition 5.1.2 Suppose that W is a permutation structure. A *strongly determined n-type* over W is a function ρ which assigns to each finite $A \subseteq W$ a complete n-type over A (whose set of realisations is denoted by $\rho \mid A$) subject to the following coherence conditions.

(i) for all $A \subset W$, $A^n \cap (\rho \mid A) = \emptyset$;

(ii) if $A' \subseteq A$ then $\rho \mid A \subseteq \rho \mid A'$;

(iii) if $g \in \mathrm{Aut}^\circ(W)$ then $g(\rho \mid A) = \rho \mid gA$.

A strongly determined n-type is a *strong* n-type if the condition (iii) of the definition holds for all automorphisms of W.

Here we view $\rho \mid A$ as a set of elements independent from A. A strongly determined n-type over W may be regarded as a non-algebraic complete n-type q

with W as a set of parameters, such that for any $\bar{a}$ realising q, any tuple $\bar{b}$ from W, and any strong map $g \in \mathrm{Aut}^\circ(W)$, the tuples $\bar{a}\bar{b}$ and $\bar{a}g(\bar{b})$ meet the same classes of 0-definable finite equivalence relations and realise the same type over $\emptyset$. We get this type over W (denoted by $\rho \mid W$) by compactness. It is worth noting here that $\rho \mid W$ is a type definable almost over $\emptyset$ ([40]). Also, any type over W definable almost over $\emptyset$ induces a strongly determined type.

Examples of strong 1-types were given in Section 4.3. Recall that if W is the countable linear dense betweenness relation without endpoints, then there is no strong 1-type over W, since any element reversing the order will violate condition (iii). Similarly, if W is the countable dense circular order, there is no strong 1-type, for once any parameter is named, all pairs acquire an orientation, so again (iii) is violated. There is no strongly determined 1-type for the circular order, but there are (exactly two) strongly determined 1-types for the dense betweenness relation. In the latter case one can say that we have a strongly determined type of multiplicity 2. More generally, define the multiplicity of a strongly determined type ρ as the number of all conjugates of $\rho \mid W$ under $\mathrm{Aut}(W)$ (the conjugate of ρ by $g \in \mathrm{Aut}(W)$ is defined as ${}^g\rho \mid A = g(\rho \mid g^{-1}A)$). A strong type is exactly a strongly determined type of multiplicity 1.

Theorem 5.1.3 ([40]) *Let W be an $\aleph_0$-categorical transitive structure such that $\mathrm{Aut}^\circ(W)$ is of finite index in $\mathrm{Aut}(W)$. Let ρ be a strongly determined n-type over W and $\pi : C \to W$ be a finite cover of W. For $\bar{a} \in \rho \mid \emptyset$ let $\bar{b}$ be an enumeration of $C(\bar{a})$. Then $tp(\bar{b}/\emptyset)$ extends to a strongly determined type over C.*

Proof. We work in an ω_1-saturated elementary extension of C. Choose $\bar{b}'$ of the type of $\bar{b}$ such that $\pi(\bar{b}') \in \rho \mid W$. Let $\bar{a}' = \pi(\bar{b}')$ and G be the group of automorphisms of C which extend to elementary maps fixing $\bar{b}'$. Note that this is a closed subgroup of $\mathrm{Aut}(C)$ as elementary maps are finitely-determined.

By definition of ρ, any element of $\mathrm{Aut}^\circ(W)$ extends to an elementary map $W \cup \bar{a}' \to W \cup \bar{a}'$. Moreover we claim that any automorphism in the kernel of π extends to an elementary map $C \cup \bar{a}' \to C \cup \bar{a}'$. To see this it suffices to show that for any finite tuple $\bar{w}$ in W there exists a tuple $\bar{a}_1$ of elements of W with $tp(\bar{a}'/C(\bar{w})) = tp(\bar{a}_1/C(\bar{w}))$. But $tp(\bar{a}'/C(\bar{w}))$ is determined by $tp(\bar{a}'/\bar{w}')$ for some finite tuple $\bar{w}'$ (by openness of the restriction mapping), so taking $\bar{a}_1 \in \rho \mid \bar{w}'$ works.

A similar argument and König's lemma shows that any elementary map $W \cup \bar{a}' \to W \cup \bar{a}'$ which induces an element of $\mathrm{Aut}^\circ(W)$ on W extends to an elementary map $C \cup \bar{a}' \to C \cup \bar{a}'$. It follows that the group H of automorphisms of C extending to elementary maps fixing $\bar{a}'$ contains all those inducing strong maps of W (so is of finite index in $\mathrm{Aut}(C)$). Clearly G is of finite index in H, and so in particular contains all strong maps of C. It is now easy to deduce that $tp(\bar{b}'/C)$ is a strongly determined type over C. □

The next result is similar to Theorem 5.1.1. To eliminate local transitivity (cf. part (c) of 4.1.4) we assume here that $acl = dcl$ and W has weak elimination of imaginaries: for every $c \in W^{eq}$ there is a finite $A \subseteq W \cap acl^{eq}(c)$ such that

$c \in dcl^{eq}(A)$. Note that these imply that $\mathrm{Aut}(W/X)$ does not have proper closed subgroups of finite index, for all finite $X \subseteq W$. Thus any strongly determined type is a strong type.

Theorem 5.1.4 ([40]) *Assume that W is a countable, transitive $\aleph_0$-categorical structure satisfying the above assumptions and any type over $\emptyset$ extends to a strongly determined type over W. Let $\pi : C \to W$ be a superlinked finite cover of W and E be the finest 0-definable finite equivalence relation on C. Then each E-class has a single intersection with any fibre of π. In particular, π is untwisted and split.*

Proof. Let $\bar{a}$ be a tuple from W such that every $\alpha \in \mathrm{Aut}(C/W)$ fixing $\bar{b} = C(\bar{a})$ pointwise is trivial. One can check that weak elimination of imaginaries guarantees that for any $c \in C$ the type of $(c, \bar{b})$ over $acl(c, \bar{b}) \cap W$ implies $tp(c\bar{b}/W)$. Since the latter has a unique realisation extending $\bar{b}$, we get that $c \in dcl(\bar{b} \cup (acl(c, \bar{b}) \cap W))$. Since $dcl = acl$ in W, $c \in dcl(\bar{b} \cup \{\pi(c)\})$.

By Theorem 5.1.3 there exists a strongly determined type ρ extending $tp(\bar{b}/\emptyset)$. For distinct c and c' satisfying $\pi(c) = \pi(c')$ consider the types $\rho \mid \{c\}$ and $\rho \mid \{c'\}$. We may assume that $\bar{b} \in \rho \mid \{c, c'\}$. Since $c, c' \in dcl(\bar{b} \cup \{\pi(c)\})$, the types of $c\bar{b}$ and $c'\bar{b}$ are different. Thus if $\alpha \in \mathrm{Aut}(C)$ sends c to c', then $\alpha(\rho \mid \{c\}) \neq \rho \mid \{\alpha(c)\}$. Hence $\alpha \notin \mathrm{Aut}^\circ(C)$ and the elements c and c' meet distinct E-classes.

On the other hand, as W is irreducible, each E-class meets each fibre. Now it is easy to see that adding unary predicates for the E-classes we get a cover of W whose fibre group is trivial. □

Remark 5.1.5 It follows from Lemma 3.1.2 that the cover π in the above is actually strongly split.

5.2 Universal covers

We now return to digraph coverings. As we noted in Section 2.2.1, if W is a connected digraph then one can define the (topological) *universal digraph covering* $U \to W$ and the corresponding *group of deck transformations* Δ such that the kernels of natural superlinked finite covers of W can be realised as homomorphic images of Δ.

This inspires the following definition. A map $U \to W$ inducing a symmetric extension of W is a *universal covering* for a class Ω of finite covers of W if for every $M \to W$ from Ω there exists a finite cover N of U with trivial kernel which also covers M and the corresponding diagram for N, M, U, W is commutative. In this definition the natural isomorphism $\mathrm{Aut}(U) \to \mathrm{Aut}(N)$ induces a surjective homomorphism $\nu : \mathrm{Aut}(U) \to \mathrm{Aut}(M)$. Moreover, if for some $\alpha \in \mathrm{Aut}(U)$ we have $\nu(\alpha) \in \mathrm{Aut}(M/W)$ then $\alpha \in \mathrm{Aut}(U/W)$. Thus the kernel of any cover from Ω is a homomorphic image of the kernel of ν.

The main obstacle here is that $U \to W$ is not necessarily a finite cover. The following shows that nevertheless we can sometimes get finiteness.

Theorem 5.2.1 ([39]) *Let W be a transitive irreducible permutation structure which has no proper irreducible superlinked finite cover. If W covers a structure W_0 (by a superlinked finite cover $\pi : W \to W_0$) then W covers it universally for the class of all superlinked minimal finite covers of W_0.*

Proof. Let $\pi_1 : C_1 \to W_0$ be a minimal superlinked cover. Since W is irreducible, C_1 is irreducible too (Lemma 4.1.1). By Lemma 3.3.6 there is a minimal superlinked finite cover $\sigma : M \to W_0$ which factors through π_1 and π. By assumption M is a trivial cover of W, so we have the result. □

As we noted in Section 5.1, if W satisfies the assumptions of Theorem 5.1.4 then the assumptions of Theorem 5.2.1 hold. The best example of such W and W_0 is a vector space over its projective space. This case was analysed in [27] (see also Theorem 4.4.3 here). It is worth noting that the ideas that we use for strong types are from the analysis in [27] concerning the situation of a vector space over its projective space. In [27] they are applied to a classification of all superlinked irreducible covers of the projective space over a finite field F. Their kernels are exactly the groups of the form F^*/H where H is a subgroup of the multiplicative group F^*.

5.3 Highly homogeneous structures

Another good example for Theorem 5.1.4 is the ordering of rationals. It is shown in [39] that any finite cover of the rationals has a split covering expansion with trivial fibre group (see also 4.3.5 here). This is the main tool in the description of finite covers of highly homogeneous structures given in [39] (cf. Section 3.1.3).

The method of [39] can be described as follows. Let N be a reduct of an $\aleph_0$-categorical structure N_0 (as we noted above, $(\mathbb{Q}, <)$ is an expansion of any highly homogeneous structure) and suppose $G_0 \leq \mathrm{Aut}(N)$ is a finitely generated subgroup such that $\mathrm{Aut}(N) = \langle \mathrm{Aut}(N_0), G_0 \rangle$. Let $M \to N$ be a finite cover of N such that the corresponding cover of N_0 (obtained from M by adding the N_0-relations on N) has a splitting expansion with trivial fibre group. This defines an isomorphism from $\mathrm{Aut}(N_0)$ onto some $H_M \leq \mathrm{Aut}(M)$. The main point of the method is to find an explicitly described class K_N of covers of N such that for every $M \to N$ there exists a finite $G_1 \leq \mathrm{Aut}(M)$ such that the restriction map induces a surjection $G_1 \to G_0$ and the group $\langle H_M, G_1 \rangle$ defines a cover from K_N. This reduces the problem to a characterisation of the kernels.

The highly homogeneous structures, apart from $(\mathbb{Q}, <)$ and the highly transitive one, are the dense linear betweenness relation, the dense circular order and the dense separation relation (see [9]). It is a very helpful fact that in the first two cases we can choose $G_0 \leq \mathrm{Aut}(N)$ such that the intersection $G_0 \cap \mathrm{Aut}(N_0)$ is trivial and every automorphism of N can be presented in the form $\alpha \cdot g \cdot \beta$, where $g \in G_0$ and $\alpha, \beta \in \mathrm{Aut}(N_0)$.

There are several observations which make us think that this method can be applied in more general cases. The first one is that for any open subgroup H of

the automorphism group of a $\aleph_0$-categorical structure N there exist $g_1, \ldots, g_n$ such that $\mathrm{Aut}(N) = (H \cup \{g_1, \ldots, g_n\})^3$. This is a consequence of the fact that if W is a transitive permutation structure and H is a point stabiliser of $\mathrm{Aut}(W)$ then the number of double cosets HgH ($g \in \mathrm{Aut}(W)$) is equal to the number of H-orbits on W (see page 21 of [42]).

On the other hand, in the G-finite case (see 3.4.2) if any type of N over some $\bar{c}$ extends to a strongly determined type then adding some tuple of constants extending $\bar{c}$ we get an expansion where every type over $\emptyset$ extends to a strong type. We can now regard this expansion as a natural candidate for the structure N_0 described above, and if we have some reduction to the superlinked case (such as Lemma 3.5.1), then Theorem 5.1.4 can sometimes guarantee a complete characterisation of the finite covers of this expansion. Then the group generated by the above $g_1, \ldots, g_n$ becomes a natural candidate for G_0. Some examples illustrating this can be found in [39].

6 Symmetric extensions with abelian kernels

6.1 A strategy

Suppose $\pi : C \longrightarrow W$ is a finite cover. Then for each $w \in W$ we have the following data (the *canonical data* of the cover):

- the fibre group $F(w) = \mathrm{Aut}(C(w)/w)$
- the binding group $B(w) = \mathrm{Aut}(C(w)/W)$
- the canonical homomorphisms $\chi_w : \mathrm{Aut}(W/w) \to F(w)/B(w)$ (see Lemma 2.1.1).

Here, $B(w)$ is a normal subgroup of $F(w)$, and these should both be regarded as permutation groups on the fibre $C(w) = \pi^{-1}(w)$. Recall our

Basic Problem: Describe all the finite covers with these data.

We have encountered various results which emphasise the importance of analysing finite covers with abelian kernels: for example, the nilpotence of the kernel of a minimal finite cover (Lemma 3.3.4), the consequent reduction of the splitting problem to the abelian case (Corollary 3.3.5), and the fact that the kernel of an irreducible superlinked finite cover is abelian (Lemma 4.1.2). This section is a commentary on a strategy for carrying out this analysis which was developed in the papers [3] and [4] by Gisela Ahlbrandt and Martin Ziegler, and refined in the paper [34] by Wilfrid Hodges and Anand Pillay.

Assume from now on that the binding groups $B(w)$ are abelian. The strategy is this:

1. Compute the possible kernels of covers $\pi : C \longrightarrow W$ with the given data (–these will be closed subgroups of $\prod_{w \in W} B(w)$);

2. For each possible kernel from (1), parametrise the covers with that as kernel.

We first put the problem in the slightly wider context of symmetric expansions with abelian kernel.

6.2 Symmetric expansions of symmetric extensions

Suppose M_0 is a symmetric extension of W with abelian kernel K_0. So K_0 is a closed normal subgroup of $\Gamma_0 = \mathrm{Aut}(M_0)$ and we have the short exact sequence

$$1 \to K_0 \to \Gamma_0 \xrightarrow{\mu} G \to 1$$

where μ is restriction to W, and $G = \mathrm{Aut}(W)$. We wish to classify (up to isomorphism over W) expansions M of M_0 which are also symmetric extensions of W. Thus we are interested in closed subgroups Γ of Γ_0 with $\mu(\Gamma) = G$. Call these *full* subgroups of Γ_0. Now consider Γ_0 acting on K_0 by conjugation. As K_0 is abelian, K_0 is in the kernel of this action, and so we get an action of $G = \Gamma_0/K_0$ on K_0. From now on we shall usually write K_0 additively, with the G-action on the left. Thus $gk = hkh^{-1}$, for $g \in G$, $k \in K_0$ and any $h \in \mu^{-1}(g)$. This makes K_0 into a G-module (or perhaps more accurately, a $\mathbf{Z}G$-module). With this notation, we have the following basic fact.

Lemma 6.2.1 *Suppose either that M_0 is countable, or that M_0 is a finite cover of W. Then K_0 is a topological G-module.*

Proof. The extra hypotheses on M_0 ensure that the restriction map μ is an open mapping (see Lemmas 1.4.2 and 1.4.3). Give $G \times K_0$ the product topology. What is being claimed is that the G-action $\alpha : G \times K_0 \to K_0$ is continuous. Let $\Gamma_0 = \mathrm{Aut}(M_0)$ and consider the map $\beta : \Gamma_0 \times K_0 \to K_0$ given by conjugation. This is continuous, and if $O \subseteq K_0$ is open, then $\alpha^{-1}(O) = (\mu \times 1)\beta^{-1}(O)$, where $\mu \times 1 : \Gamma_0 \times K_0 \to G \times K_0$ is the obvious map. This is open, so α is continuous, as required. □

Now suppose Γ is a closed, full subgroup of Γ_0. Then $K = K_0 \cap \Gamma$, which we refer to as the *kernel* of Γ, is a closed G-submodule of K_0. We can now divide Part 1 of our strategy as:

(1a). Determine the closed G-submodules of K_0.
(1b). Determine which submodules can appear as kernels of closed, full subgroups of Γ_0.

We shall concentrate on (1a) here, essentially because we do not know anything about (1b), beyond what has already been stated in Sections 1.5 and 3.1.2. If the original symmetric extension is *split* and arises from a finite cover then Lemma 3.1.4 shows that any closed G-submodule of K_0 appears as a kernel. Before giving some examples, we give a duality which is useful in analyzing submodules of kernels of finite covers.

6.3 Applications of Pontriagin duality

6.3.1 The duality

All the material in this section is to be found in Chapter 6 of Pontriagin's book [47]. Throughout, A will be a Hausdorff topological abelian group which is either discrete or compact.

The *dual* or *character group* of A, written A^*, is the set of continuous homomorphisms (*characters*) from A into S, the multiplicative group of the complex numbers of modulus 1. This is a group under the operation of pointwise multiplication: if $f, g \in A^*$ and $a \in A$ then $(fg)(a) = f(a)g(a)$. Moreover, we can regard A^* as a topological group in the following way. For natural numbers k, let $S_k = \{e^{2\pi i\theta} : |\theta| < 1/k\}$. Then the collection of subsets of A^* of the form $\{f \in A^* : f(X) \subseteq S_k\}$, where X is a compact subset of A and $k \in \mathbb{N}$, forms a complete system of neighbourhoods of the identity in A^* (see Definition 34 of [47]). The following result gives a more convenient way of thinking of this topology, and comes from the proof of Theorem 36 in [47].

Lemma 6.3.1 *If A is compact, then A^* is discrete. If A is discrete, then A^* is compact. Moreover in the latter case, the topology on A^* coincides with its topology as a subspace of the product space S^A.* □

Examples 1. The finite abelian group Z_n (with the discrete topology) is self-dual.

2. If p is a prime, X any set, and $A = (Z_p)^X$ with the product topology, then a non-trivial continuous homomorphism $\chi : A \to S$ has as its image the set of p-th roots of unity, and its kernel is open. So there exists a finite subset $x_1, \ldots, x_n$ of X such that for all $f \in A$, $\chi(f)$ is determined by $f(x_1), \ldots, f(x_n)$. It is now easy to see that A^* is naturally isomorphic to Z_pX, the discrete group of formal (finite) linear combinations of elements of X with coefficients in Z_p (identify Z_p with the group of p-th roots of unity in S and let χ correspond to $\Sigma_{x\in X}\chi(f_x)x$, where f_x is the characteristic function at x).

Conversely, if A is Z_pX, then A^* is Z_p^X with the product topology.

Evaluation at $a \in A$ gives a character of A^*, so there is a natural map $\omega : A \to A^{**}$. The fundamental result of Pontriagin ([47], Theorem 39) is that ω is an isomorphism of topological groups. The most important consequence of this that we shall use is that there is an inclusion-reversing correspondence between the closed subgroups of A and A^*. If C is a closed subgroup of A the corresponding subgroup of A^* is

$$\Phi = \{f \in A^* : f(C) = 1\},$$

the annihilator of C in A^*. The dual of C is naturally isomorphic to A^*/Φ (Theorem 41 of [47]), and the dual of A/C is naturally isomorphic to Φ (Theorem 37 of [47]).

We can also phrase these results in terms of exact sequences. Suppose that A, B are topological groups which are either both discrete or both compact. Note that if $\theta : A \longrightarrow B$ is a continuous homomorphism of topological abelian groups,

then there is naturally a dual continuous homomorphism $\theta^* : B^* \longrightarrow A^*$ (given by $(\theta^* f)(a) = f(\theta(a))$, for $f \in B^*$ and $a \in A$). The following is then an easy consequence of the results just cited.

Theorem 6.3.2 *If*

$$0 \to C \xrightarrow{\gamma} A \xrightarrow{\theta} B \to 0$$

is an exact sequence of topological abelian groups, either all discrete or all compact, then the dual sequence

$$0 \to B^* \xrightarrow{\theta^*} A^* \xrightarrow{\gamma^*} C^* \to 0$$

is also exact. □

We observed an instance of the following result in Example 2 above. It is a paraphrase of Theorem 45 of [47].

Theorem 6.3.3 *The dualising operation sends a direct product of compact abelian groups to the direct sum of the duals of the factors, and vice versa.* □

Suppose now that G is a topological group and Y is a compact topological G-module. Then Y^* is a G-module with action given by $(gf)(y) = f(g^{-1}y)$. It is easy to see that a closed subgroup X of Y is G-invariant if and only if its annihilator in Y^* is G-invariant. Thus we have:

Theorem 6.3.4 *The lattices of closed G-invariant subgroups of Y and Y^* are dual.* □

We now return to the problem of classifying the possible kernels of a finite cover $\pi : C \longrightarrow W$, having been given the fibre and binding groups $F(w)$ and $B(w)$ (and the canonical homomorphisms), where the $B(w)$ are abelian. According to our strategy, we consider the G-module $K_0 = \prod_{w \in W} B(w)$ (where $G = \mathrm{Aut}(W)$), the kernel of the free cover with the given data (see Section 2.1), and we want to know the closed G-submodules of K_0. (If W is transitive, then (for any $w \in W$) the module K_0 is the G-module coinduced from the $\mathrm{Aut}(W/w)$-module $B(w)$, with the action being given via the canonical homomorphism (cf. Section III.5 of [8]).) Now K_0 is compact and so by Pontriagin duality, this problem is equivalent to determining all the G-submodules of the direct sum

$$K_0^* = \bigoplus_{w \in W} B(w)^*.$$

Care is needed here in writing down the G-action. If W is transitive, then (for any $w \in W$) this module is the G-module induced from the $\mathrm{Aut}(W/w)$-module $B(w)^*$, with the action being given via the canonical homomorphism (cf. Section III.5 of [8]). The simplest situation is where the fibre groups and binding groups are all cyclic of order p. Then $K_0 = F_p^W$, and K_0^* is the permutation module F_pW (F_p is the field of p elements, considered as a trivial G-module). The correspondence

between submodules of the two modules given by the duality is exactly that given by annihilation in the natural pairing

$$F_p^W \times F_pW \longrightarrow F_p$$

$$(f, \Sigma_{w\in W} a_w w) \longmapsto \Sigma_{w\in W} a_w f(w).$$

6.3.2 Examples

Example 6.3.5 Suppose W is a disintegrated set (so $G = \mathrm{Aut}(W) = \mathrm{Sym}(W)$). Let $F(w) = B(w) = Z_p$, for all $w \in W$ and some prime p. Thus $K_0 = F_p^W$ and K_0^* is the permutation module F_pW. We claim that the only proper, non-trivial submodule of this is $A = \langle w - w' : w, w' \in W\rangle$. Indeed, let $x = a_1w_1 + \ldots + a_rw_r$ be a non-zero element of F_pW, where the a_i are non-zero, and the w_i are distinct. Consider the submodule $\langle x\rangle_{F_pG}$ generated by x. If $r = 1$ then (by transitivity of G on W) this is the whole of F_pG. If $r \geq 2$, take $g \in G$ fixing $w_1, \ldots, w_{r-1}$ and moving w_r. Then $x - gx = a_r(w_r - gw_r)$, and this is a generator for A (for example, because G is 2-transitive on W). So $\langle x\rangle_{F_pG}$ contains A. As A is of codimension 1 in F_pW, this establishes the claim.

So now, by duality, it follows that the only proper, non-trivial closed submodule of $K_0 = F_p^W$ is the finite submodule consisting of the constant functions.

Remarks. 1. The above argument works assuming only that $\mathrm{Aut}(W)$ is primitive on W, and that definable closure in W is trivial: see Theorem 2.1 of [11].

2. The result can of course be proved directly, without mentioning duality, by making use of the cover pregeometry (cf. Corollary 3.2.4). But conversely the cover pregeometry can be described very explicitly using the duality. Suppose $\pi : C \to W$ is a simple finite cover with all fibre groups isomorphic to F_p. Then its kernel K can be identified with a subgroup of F_p^W. Let $\Phi(K)$ be the annihilator of K in F_pW (as above). It is easy to show for $w, w_1, \ldots, w_n \in W$ that w is dependent on $w_1, \ldots, w_n$ in the cover pregeometry determined by π if and only if w is linearly dependent on $w_1, \ldots, w_n$ modulo $\Phi(K)$ (that is, working in the quotient space $F_pW/\Phi(K)$).

Example 6.3.6 The following results are due to Ahlbrandt and Ziegler, and are to be found in the paper [3]. However, our presentation of the results is slightly different.

Let $V = V(\aleph_0, 2)$ be a countably infinite dimensional vector space over the field F_2 with 2 elements, and $G = GL(\aleph_0, 2)$ its automorphism group. Then G is transitive on W, the non-zero vectors in V (which we can also identify with the set of one-dimensional subspaces of V). We consider finite covers of W where the fibre and binding groups are cyclic of order 2. (Ahlbrandt and Ziegler also consider the more difficult case of affine covers of W where the structure groups are isomorphic to $V/\langle w\rangle$, for $w \in W$). So, we wish to determine the closed, G-invariant subgroups of F_2^W. By duality, this is equivalent to determining the G-invariant subgroups of the permutation module F_2W. We now describe these.

For any vector space X of dimension $n \leq \omega$ over a finite field F_q, and finite k with $0 \leq k \leq n$, let $[[X]]^k$ be the set of k-dimensional subspaces of X, considered as a permutation structure with $GL(X)$ acting. If $l \leq k$ there is a natural homomorphism of $GL(X)$-modules

$$\beta_{k,l} : F_q[[X]]^k \longrightarrow F_q[[X]]^l$$

given by

$$\beta_{k,l}(w) = \Sigma_{w' \in [[w]]^l} w'$$

for $w \in [[X]]^k$. (Of course, this can also be defined for permutation modules over a different field.) It is easy to show that if $l \leq k' \leq k$ then there is a non-zero $t \in F_q$ such that $\beta_{k,l} = t\beta_{k',l}\beta_{k,k'}$ and so $im(\beta_{k,l}) \leq im(\beta_{k',l})$. So as submodules of $F_q[[X]]^1$ we have $im(\beta_{k,1})$ for $1 \leq k \leq n$, and the intersections of these with $ker(\beta_{1,0})$. The main result of Part 1 of [3] is that for finite n and $q = 2$, these are the only $GL(X)$-submodules of $F_2[[X]]^1$. It is then a straightforward matter to deduce that the same holds for the infinite dimensional case: any element of the permutation module has its support contained in a finite dimensional subspace, and so we can read off from the finite case that it must generate some $im(\beta_{k,1})$ or its intersection with $ker(\beta_{1,0})$. Now, we can identify W with the set of 1-dimensional subspaces of V, and so we know all the G-invariant subgroups of F_2W.

From the duality, we now get the following (the notation is that of [4]).

Theorem 6.3.7 (Ahlbrandt-Ziegler, [3]) *The closed, G-invariant subgroups of F_2^W consist of the submodules*

$$Pol_k = \{f \in F_2^W : \Sigma_{x \in w \setminus \{0\}} f(x) = 0 \, \forall w \in [[V]]^{k+1}\}$$

and the sums of these with the one dimensional submodule of constant functions. □

With the benefit of hindsight, the above submodules of $F_2[[X]]^1$ are recognisable as 'well-known' objects from the theory of error-correcting codes: they are the Reed-Muller codes. On closer examination of the coding-theoretic literature (notably [1] and [18]), D. Gray discovered that the above result is a special case of a result of Delsarte (see Theorem 8 of [18]): for arbitrary prime p, the $GL(n,p)$-invariant subgroups of $F_p[[V(n,p)]]^1$ are given by the images of the $\beta_{k,1}$ maps, and their intersections with $ker(\beta_{1,0})$. There is, of course a corresponding result for the infinite dimensional case, and a dual version. So Theorem 6.3.7 holds for all primes, not just for the prime 2. More details can be found in [26].

Example 6.3.8 We report some work of D. Gray ([29]), which is (in spirit) similar to that of the previous example, and deals with the case where $W = [D]^k$ is the permutation structure of k-sets from a disintegrated set D (so the automorphism group here is $G = \mathrm{Sym}(D)$). Let F be any field, and consider the FG-permutation modules $F[D]^k$, for $k < \omega$. Again, for $k \geq l$, there are natural FG-module homomorphisms

$$\beta_{k,l} : F[D]^k \longrightarrow F[D]^l$$

given by

$$\beta_{k,l}(w) = \Sigma_{w' \in [w]^l} w'$$

for $w \in [D]^k$.

Theorem 6.3.9 (D. Gray, [29]) *If D is infinite, the FG-submodules of $F[D]^k$ are given by intersections of kernels $ker(\beta_{k,l})$, where $0 \leq l \leq k$.* □

The proof uses the representation theory of the finite symmetric groups, as developed in the book of G. D. James [41]. In fact, there is an effective algorithm for determining the complete submodule lattices in the above (they depend only on k and the characteristic p of F, and the algorithm involves only the checking of whether certain binomial coefficients are divisible by p). For our purposes, the main consequence of Gray's results is (by duality) a determination of the possible kernels of a finite cover of $W = [D]^k$, with fibre and binding groups of order p.

6.4 Derivations and H^1_c

In this section, we follow [4] and [34] (particularly Section 5) rather closely. We develop some algebraic machinery for attacking the second part of our strategy: for a fixed symmetric extension C_0 of W with abelian kernel, parametrise the expansions with a particular kernel K which are still symmetric extensions of W. In our applications, C_0 will be a free finite cover with given canonical data, and we shall parametrise covering expansions with kernel K up to conjugacy of their automorphism groups in $\mathrm{Aut}(C_0)$. This is *a priori* a finer classification than classification up to isomorphism over W. However this distinction does not create any practical difficulties.

Definition 6.4.1 If G is a group and A is a G-module, then a *derivation* from G to A is a map $d : G \longrightarrow A$ which satisfies $d(gh) = d(g) + gd(h)$ for all $g, h \in G$. Denote the set of all these by $Der(G, A)$. Note that the sum of two derivations is again a derivation, so $Der(G, A)$ is in fact an abelian group. An *inner* derivation is a derivation of the form d_a (for $a \in A$) where $d_a(g) = ga - a$ for all $g \in G$. The inner derivations form a subgroup of $Der(G, A)$. The quotient group is denoted by $H^1(G, A)$, and is referred to as the *first cohomology group* of G on A. If A is a topological G-module then the continuous derivations form a subgroup $Der_c(G, A)$ of the group of all derivations, and this clearly contains all the inner derivations. We denote the quotient group of $Der_c(G, A)$ by the inner derivations by $H^1_c(G, A)$.

Suppose now that C_0 is a symmetric extension of W with abelian kernel K_0. Thus we have a surjection $\mu : \mathrm{Aut}(C_0) \longrightarrow \mathrm{Aut}(W)$ whose kernel is K_0. As we have already remarked, conjugation in $\mathrm{Aut}(C_0)$ (denoted generally by $g^h = hgh^{-1}$ for $g, h \in \mathrm{Aut}(C_0)$) gives an action of $G = \mathrm{Aut}(W)$ on K_0 which makes K_0 into a G-module, which we write additively. In practice, we will always work in situations where K_0 is actually a topological G-module (see Lemma 6.2.1).

Suppose K is a G-invariant subgroup of K_0 such that there exists $H_0 \leq \mathrm{Aut}(C_0)$ with $H_0 \cap K_0 = K$ and $\mu(H_0) = G$. Note that $G \cong H_0/H_0 \cap K_0 \leq \mathrm{Aut}(C_0)/K$. So

there is an embedding $\sigma_0 : G \longrightarrow \mathrm{Aut}(C_0)/K$ given by $\sigma_0(g) = (\mu^{-1}(g) \cap H_0)K$. Denote by $\overline{\mu}$ the homomorphism $\mathrm{Aut}(C_0)/K \rightarrow \mathrm{Aut}(W)$ induced by μ. Clearly $\overline{\mu}\sigma_0 = 1_G$, the identity on G. In fact, this establishes a bijection between the set of subgroups H of $\mathrm{Aut}(C_0)$ which satisfy $\mu(H) = G$ and $H \cap K_0 = K$, and the set of embeddings $\sigma : G \rightarrow Aut(C_0)/K$ satisfying $\overline{\mu}\sigma = 1_G$.

Now, suppose $\sigma : G \longrightarrow Aut(C_0)/K$ is any embedding such that $\overline{\mu}\sigma = 1_G$. Let $d_\sigma : G \rightarrow K_0/K$ be defined by $d_\sigma(g) = \sigma(g)\sigma_0(g)^{-1}$ (note that this is indeed an element of K_0/K, not just an element of $\mathrm{Aut}(C_0)/K$). On the other hand, given a derivation $d : G \rightarrow K_0/K$, define $\sigma_d : G \rightarrow \mathrm{Aut}(C_0)/K$ by $\sigma_d(g) = d(g)\sigma_0(g)$. We have the following result from [34] (Proposition 16):

Lemma 6.4.2 *(i) The map $\sigma \mapsto d_\sigma$ is a bijection between the set of embeddings $\sigma : G \longrightarrow \mathrm{Aut}(C_0)/K$ such that $\overline{\mu}\sigma = 1_G$, and $Der(G, K_0/K)$. Its inverse is the map $d \mapsto \sigma_d$.*

(ii) If C_0 is countable and H_0 is a closed subgroup of $\mathrm{Aut}(C_0)$, then the map in (i) induces a bijection between the subset of continuous maps σ and $Der_c(G, K_0/K)$.

Proof. (i) We check that if σ is an embedding $G \rightarrow Aut(C_0)/K$, then d_σ is a derivation. Remember that the action of G on K_0 (and so on K_0/K) is given by conjugation in $\mathrm{Aut}(C_0)$. So, for $g \in G$ and $a \in K_0/K$ we have

$$ga = \sigma_0(g)a\sigma_0(g)^{-1} = \sigma(g)a\sigma(g)^{-1}$$

computed in $\mathrm{Aut}(C_0)/K$. Thus,for $g, h \in G$,

$$d_\sigma(gh) = \sigma(gh)\sigma_0(gh)^{-1} = \sigma(g)\sigma(h)\sigma_0(h)^{-1}\sigma_0(g)^{-1}$$

$$= (\sigma(h)\sigma_0(h)^{-1})^{\sigma(g)}\sigma(g)\sigma_0(g)^{-1}$$

$$= \sigma(g)\sigma_0(g)^{-1} + \sigma_0(g)d_\sigma(h)\sigma_0(g)^{-1} = d_\sigma(g) + gd_\sigma(h).$$

(ii) The extra hypotheses imply that K is a closed subgroup of K_0 and that σ_0 is continuous (by Corollary 1.4.4). The statement now follows immediately from the definitions and (i). □

So now:

1. we have a bijection between $Der(G, K_0/K)$ and the set of subgroups H of $\mathrm{Aut}(C_0)$ which satisfy $\mu(H) = G$ and $H \cap K_0 = K$;

2. if C_0 is countable and H_0 is closed, the above bijection sets up a bijection between $Der_c(G, K_0/K)$ and the set of *closed* subgroups H of $\mathrm{Aut}(C_0)$ which satisfy $\mu(H) = G$ and $H \cap K_0 = K$.

According to our strategy for classifying symmetric expansions of a given symmetric extension up to isomorphism over the base structure, what we are interested in is classifying groups H as in (2) above *up to conjugacy in* $\mathrm{Aut}(C_0)$. The next lemma is Proposition 17 of [34]. We omit its proof (it is a simple exercise). Denote the derivation corresponding to the subgroup H in (1) above by d_H. Then:

Lemma 6.4.3 *Let H and H' be subgroups of* $\mathrm{Aut}(C_0)$ *such that* $\mu(H) = \mu(H') = G$ *and* $H \cap K_0 = H' \cap K_0 = K$. *Then H and H' are conjugate in* $\mathrm{Aut}(C_0)$ *if and only if d_H and $d_{H'}$ differ by an inner derivation.* □

Corollary 6.4.4 ([34], Corollary 18) *Suppose C_0 is countable and K is a closed, G-invariant subgroup of K_0 such that there is a closed subgroup H_0 of* $\mathrm{Aut}(C_0)$ *with* $\mu(H_0) = G$ *and* $H_0 \cap K_0 = K$. *Then there is a bijection between the cohomology group $H^1_c(G, K_0/K)$ and the set of* $\mathrm{Aut}(C_0)$*-conjugacy classes of closed subgroups H of* $\mathrm{Aut}(C_0)$ *which satisfy* $\mu(H) = G$ *and* $H \cap K_0 = K$. □

The following is one of the most important consequences of this. We state it for finite covers, but one could formulate a similar result in terms of symmetric extensions.

Corollary 6.4.5 *Let W be a countable permutation structure. Suppose $\pi_0 : C_0 \to W$ is a split finite cover with abelian kernel K_0 and K is a closed* $\mathrm{Aut}(W)$*-invariant subgroup of K_0. Then there is a covering expansion of π_0 with kernel K. If $H^1_c(\mathrm{Aut}(W), K_0/K) = \{0\}$, then it is unique (up to isomorphism over W). In particular, any covering expansion of π_0 with kernel K is split.*

Proof. Existence of a split covering expansion $\pi : C \to W$ of π_0 with kernel K follows from Lemma 3.1.4. By Corollary 6.4.4 and our assumption, the automorphism group of any covering expansion of π_0 with K as kernel is conjugate in $\mathrm{Aut}(C_0)$ to $\mathrm{Aut}(C)$. In particular, the finite cover is isomorphic over W to π and is split. □

Remarks 6.4.6 1. A re-statement of Corollary 6.4.4 is that the group $H^1_c(G, K_0/K)$ parametrises isomorphism classes of symmetric expansions of the symmetric extension C_0 which have kernel K.

2. The bijection obtained in Corollary 6.4.4 depends on the choice of H_0, or, equivalently, the choice of σ_0. If there is a continuous splitting $\tau : G \longrightarrow \mathrm{Aut}(C_0)$ of μ, then, of course, we can take $H_0 = KT$, where $T = im(\tau)$. Equivalently, we let $\sigma_0(g) = \tau(g)K$ (this is the approach taken in [4], where C_0 is a split (principal) cover of W).

3. The fact that the bijection from the Corollary depends on the choice of σ_0 means that we do not have a natural parametrisation. Note however that what we have constructed is an *action* of $H^1_c(G, K_0/K)$ on the set of $\mathrm{Aut}(C_0)$-conjugacy classes of closed subgroups H of $\mathrm{Aut}(C_0)$ which satisfy $\mu(H) = G$ and $H \cap K_0 = K$. This is *regular* (that is, transitive and with all non-identity elements acting without fixed points), and is independent of the choice of σ_0. We thank Martin Ziegler for pointing this out to us.

4. At this point, we should mention that there is a subtle difference between the classification we are proposing here and what is proposed in [4] and [34]. The approach there considers a principal cover $\pi : C_0 \longrightarrow W$ with abelian kernel (– so the necessary assumption is that the fibre groups are abelian), and classifies covering expansions up to conjugacy in $\mathrm{Aut}(C_0)$. The strategy we are following

instead involves considering a free finite cover $\pi : C_0 \to W$ with abelian kernel (– so we are assuming that the binding groups are abelian, but the fibre groups could be non-abelian). We then classify covering expansions up to conjugacy in this $\mathrm{Aut}(C_0)$. So our approach is a little more general, but, of course, it may produce a slightly different answer. We doubt, however, that this should present any real problems.

5. Recall that a (countable) structure W has the *small index property* if any subgroup of index less than $2^{\aleph_0}$ in $\mathrm{Aut}(W)$ is open. Thus, the topology on $\mathrm{Aut}(W)$ can be recovered solely from the abstract group structure of $\mathrm{Aut}(W)$. Proposition 1.5.3, due to Hodges and Pillay ([34]) show that (with the hypotheses of 6.4.4) if W has the small index property, then any derivation in $Der(G, K_0/K)$ is continuous, and so $H^1_c(G, K_0/K) = H^1(G, K_0/K)$.

Experience suggests that the groups H^1 and H^1_c should be small (and quite likely to be zero) for many respectable $\aleph_0$-categorical structures. Anything other than this situation should be regarded as exotic (we shall say more on this in Section 7). But of course, the question is how can one compute the groups H^1 and H^1_c? In some cases, the hard work has already been done by group theorists and sometimes their results can be used or adapted. An example of this will be given in Section 6.5. In Section 7 we shall adapt some of the well-known machinery from general group cohomology to our purposes and show how it can sometimes be used to calculate (or at least estimate) H^1 and H^1_c. To start off with here is a very easy, but useful lemma.

Lemma 6.4.7 *Let Γ be a topological group and M a continuous Γ-module. Let N be a closed submodule of M and suppose that $H^1_c(\Gamma, M/N)$ and $H^1_c(\Gamma, N)$ are trivial. Then $H^1_c(\Gamma, M)$ is trivial.*

Proof. Suppose $d : \Gamma \longrightarrow M$ is a continuous derivation. By composing with the natural map $M \longrightarrow M/N$ we get a continuous derivation $\Gamma \longrightarrow M/N$ and so, by assumption, there exists $a \in M$ such that $(d - d_a)(g) \in N$ for all $g \in \Gamma$. But then, as $d - d_a$ is also continuous, there exists $b \in N$ such that $d - d_a = d_b$. So $d = d_{a+b}$ is inner. □

6.5 Finite covers of $V(\aleph_0, 2)$ and $[D]^k$.

We now return to one of the situations considered by Ahlbrandt and Ziegler in [4] (see Example 6.3.6). So, let $V = V(\aleph_0, 2)$ be a countably infinite dimensional vector space over the field with 2 elements, and $G = GL(\aleph_0, 2)$ its automorphism group. Let $W = V \setminus \{0\}$. We consider finite covers of W where the fibre and binding groups are cyclic of order 2. Let $\pi_0 : C_0 \to W$ be the free cover of W with fibre and binding groups cyclic of order 2 (and each fibre of size 2). Let K_0 be the kernel of this. Note that by Lemma 2.1.4 π_0 splits. In 6.3.6 we described the closed G-invariant subgroups K of K_0: a result which was deduced from the parallel situation of finite-dimensional V.

Ahlbrandt and Ziegler show that

Theorem 6.5.1 *For each possible kernel K we have $H^1_c(G, K_0/K) = \{0\}$.*

From this and Corollary 6.4.5 we obtain immediately.

Corollary 6.5.2 *All finite covers of W with fibres of size 2 split. Any such finite cover is determined (up to isomorphism over W) by its kernel, and the possibilities for the kernels are given in Theorem 6.3.7.* □

Ahlbrandt and Ziegler deduce Theorem 6.5.1 from known results about the vanishing of the first cohomology groups of the finite general linear groups $GL(n,2)$ acting on certain natural modules (duals of exterior powers of $V(n,2)$ (if $n \geq 4$)), together with results on envelopes in totally categorical structures. We shall describe the use of the finite group theoretic results, but avoid mentioning envelopes, substituting instead the following, slightly *ad hoc* result, taken from [26].

Lemma 6.5.3 *Let Γ be a Hausdorff topological group and M a compact topological Γ-module. Suppose there exists $(G_i : i < \omega)$, an increasing chain of subgroups of Γ such that $G = \bigcup_{i<\omega} G_i$ is dense in Γ. Suppose also that for each i we have an open, G_i-invariant subgroup M_i of M, and that $M_{i+1} \leq M_i$ for all $i < \omega$ and $\bigcap_{i<\omega} M_i = \{0\}$. Suppose further that for all i, any continuous derivation from G_i to M/M_i is inner. Then any continuous derivation $d : \Gamma \longrightarrow M$ is inner.*

Proof. Note first that if two continuous derivations $\Gamma \longrightarrow M$ agree on a dense subgroup, then they must be equal. So (as inner derivations are continuous) it will suffice to prove that $\delta = d|G$ is inner. The hypotheses imply that M is metrizable, with a metric θ such that the diameters of the M_i tend to zero.

For every $i < \omega$ there exists $a_i \in M$ such that for all $g \in G_i$ we have

$$\delta(g) + M_i = ga_i - a_i + M_i.$$

By compactness of M we may assume that the a_i converge to some $a \in M$. Let d_a denote the inner derivation obtained from a. Thus, for $g \in G_i$, for every $j > i$ there exists $m_j \in M_j$ such that

$$\theta(\delta(g), d_a(g)) = \theta(ga_j - a_j + m_j, ga - a).$$

Now, the m_j tend to 0 as j tends to infinity, and so (by continuity of the Γ-action) $\theta(\delta(g), d_a(g))$ can be arbitrarily small. So $\delta(g) = d_a(g)$. But this holds for all i, and so we conclude that $d = d_a$, as required. □

Proof of 6.5.1. We use the lemma with $\Gamma = GL(V)$ and $M = K_0/K$. Remember that $K_0 = F_2^W$ and K is a closed, Γ-invariant subgroup of K_0. Let $(V_i : i < \omega)$ be an increasing chain of finite dimensional subspaces of V (of dimension at least 4) with union the whole of V. Let T_i be a complement to V_i in V, and choose these so that $T_i \geq T_{i+1}$ for all i. Let

$$G_i = \{g \in \Gamma : gV_i = V_i \text{ and } gx = x \ \forall x \in T_i\}.$$

Then the G_i form an increasing chain whose union is dense in Γ. Let K_i be those functions in K_0 which are zero on V_i. Thus, K_0/K_i is isomorphic to $F^{V_i\setminus\{0\}}$. Let $M_i = (K + K_i)/K$. Then

$$M/M_i = (K_0/K)/(K + K_i/K) \cong K_0/(K + K_i) \cong (K_0/K_i)/(K + K_i/K_i)$$

and all these isomorphisms hold as isomorphisms of G_i-modules. But now we claim that Theorem 4.1 of [4], and the description of the possibilities for K given in Theorem 3.1 (*ibid.*) show that any derivation from G_i to M/M_i is inner. Put more explicitly, what we want to show in order to apply our lemma, is that $H^1(G_i, M/M_i) = \{0\}$. Fix i, and to ease notation write $X = V_i$. Now, G_i can be identified with the finite general linear group $GL(X)$, and M/M_i is a quotient module of the $GL(X)$-module $F_2^{X\setminus\{0\}}$. The finite-dimensional version of the results in 6.3.6 tell us precisely what are the possibilities for M/M_i. Moreover, the composition factors of M/M_i are well-known $GL(X)$-modules: they are (duals of) exterior powers of X (see lemma 4.3 of [4]). Results of G. B. Bell ([7]) show that the the first cohomology of $GL(X)$ on these modules is trivial, and so (for example, by 6.4.7) $H^1(GL(X), M/M_i) = \{0\}$. The lemma is now applicable, and this finishes the proof of 6.5.1. □

Remark 6.5.4 A completely analogous result for the case of finite covers of the projective space of a countable dimensional vector space over the field with p elements (for a prime p), and fibre groups of order p is given in [26].

The following result is proved using similar methods, but making use of the description of kernels given in 6.3.8 in place of 6.3.6.

Theorem 6.5.5 ([26]) *Let $W = [D]^k$ be the Grassmannian of k-sets from a countably infinite disintegrated set D. Then any finite cover $\pi : C \longrightarrow W$ with fibre groups cyclic of prime order is split. Together with the results in 6.3.8, this gives a complete classification of all such covers.* □

We will not give any details of the proof. Let it suffice to say that the case $p = 2$, $k \geq 2$, is complicated by the fact that some non-trivial cohomology groups are involved: essentially because there are two conjugacy classes of closed full subgroups of $\mathrm{Aut}(C_0)$ which have trivial kernel (cf. Example 3.1.3).

We remark that (with W as above) it follows that any finite cover $\pi : C \longrightarrow W$ with fibre group of odd order is split. By 3.3.5, it suffices to prove this for the case where the kernel is an elementary abelian p-group (for all odd primes p). By 2.1.1, and the fact that any non-trivial finite homomorphic image of $\mathrm{Aut}(W/w)$ has even order, the fibre and binding groups of π are equal. It is then easy to reduce to the case where the binding groups are actually Z_p. This is then handled by Theorem 6.5.5.

6.6 Cohomology and two-graphs

Let X be any set and let V be the vector space over F_2, the field with 2 elements, of all functions from the set of 2-element subsets of X to F_2. We identify the

elements of V with the graphs on X. Then addition in V is the operation of symmetric difference on edge sets. Let V_0 be the subspace of all complete bipartite graphs (including the null and complete graphs). We can look at V_0 as the set of all switching operations: switching a graph corresponds to adding to it a complete bipartite graph. Thus the switching classes are the cosets of V_0 in V.

Let G be the automorphism group of a two-graph $W = (X, T)$, and let (X, R) be any graph giving rise to this two-graph. Recall that any two such graphs are in the same switching class, so we may define a function $d : G \to V_0$ by $g \mapsto (gR - R)$. As noted in [10], this is a derivation. The corresponding element $\gamma \in H^1(G, V_0)$ is called the *first invariant of* T. Theorem 3.1 from [10] asserts that $\gamma = 0$ if and only if there is a graph $R' \in V_0 + R$ such that G is the automorphism group of (X, R'). Recall that in Section 2.3 we constructed a 'double cover' $\pi : M \to W$ where $M = (X^*, R^*)$. It is noticed in [39] (Theorem 3.3) that this is strongly split if and only if G preserves a graph on X in the same switching class as R.

This provides an alternative approach to showing that the double cover in Example 2.3.1 is non-split. Recall that there (X, R) is the 'random graph' and (X, T) the corresponding two-graph. We indicated there that G, its automorphism group, acts doubly-transitively on X. In particular, G cannot preserve any graph on X in the same switching class as R. So by the above remarks, γ here is not zero and the double cover $M \to X$ is not strongly split. As point stabilisers in G are irreducible, this means that $M \to W$ is not split (by Lemma 3.1.2).

The invariant γ has an interpretation in terms of the cohomology groups introduced in Section 6.4. Indeed, let $\pi_0 : M_0 \to W$ be the free reduct of π with fibre and binding groups cyclic of order 2. So the kernel here is $K_0 = F_2^X$ and the kernel K of π is the submodule consisting of the constant functions. The stabiliser S of X^+ is a closed complement to K_0 in $\mathrm{Aut}(M_0)$. In the notation of Section 6.4, let $H_0 = KS$, let $H = \mathrm{Aut}(M)$, and let d_H be the derivation $d_H : G \to K_0/K$ given by Lemma 6.4.3. Now recall that V_0 is identified with the set of switching operations, which was also identified with F_2^X/F_2 (see Section 2.3). It is easy to show that (with these identifications) γ is the element of $H_c^1(G, F_2^X/F_2)$ which d_H gives rise to.

7 Computing cohomology groups

7.1 Dimension shifting and Shapiro's lemma

In this section we give, adapted for our purposes, some results and a technique (known as dimension shifting) which are well-known in the cohomology theory of groups.

Definition 7.1.1 If G is a group and M a G-module then we define the *zero-th cohomology group* $H^0(G, M)$ to be the elements of M fixed by all elements of G. Note that if G is a topological group and M a topological G-module, then this is actually a closed subgroup of M.

Lemma 7.1.2 ('The long exact sequence') *Suppose G is a group and*

$$0 \to K \to M \to N \to 0$$

is an exact sequence of G-modules. Then there is an exact sequence of abelian groups:

$$0 \to H^0(G,K) \to H^0(G,M) \to H^0(G,N) \to H^1(G,K) \to H^1(G,M) \to H^1(G,N).$$

If, moreover, G is a topological group and the short exact sequence is a sequence of topological G-modules in which the homomorphisms are continuous, open maps, then there is a long exact sequence as above in which the H^1 terms are replaced by H^1_c.

Proof. The first part is very well-known and can be found in any book on group cohomology (eg. [8] Proposition III.6.1(ii')). All but one of the maps in the long exact sequence are induced by the maps in the short exact sequence in an obvious way. The map $\phi : H^0(G,N) \to H^1(G,K)$ is known as the connecting map, and is obtained as follows (to ease notation, suppose K is contained in M and identify N with M/K). Let $\bar{a} = a + K \in M/K$ be fixed by all elements of G. Then $ga - a \in K$ for all $g \in G$ and it is easy to check that $d_{\bar{a}} : G \to K$ given by $d_{\bar{a}}(g) = ga - a$ is a derivation, and modulo inner derivations into K this depends only on $\bar{a}$. So we get a well-defined homomorphism $H^0(G,N) \to H^1(G,K)$. The topological version of the result can now be checked using the proof of the group theoretic result (the force of the assumption that the maps in the short exact sequence are continuous, open maps is that K is topologically isomorphic to its image in M and N is topologically isomorphic to M quotiented by this). □

Remarks 7.1.3 1. There are various circumstances in which in a short exact sequence of topological groups, the maps (assumed to be continuous) are automatically open maps. For example, this happens if the groups are compact, or if the groups are Polish groups (see ([42], Proposition 6.3) for the latter).

2. Of course, there are defined, for any group G and G-module M, cohomology groups $H^n(G,M)$ for any $n < \omega$, and the long exact sequence can be continued to involve these. For example, the group $H^2(G,M)$ parametrises equivalence classes of group extensions $1 \to M \to E \to G \to 1$. It is not clear, however, how to make use of these in our context (that of symmetric extensions of a given base structure), or indeed what is the appropriate category in which to develop the general cohomological machinery. It does not seem to be adequate simply to consider continuous 2-cocycles etc. to single out the particular extensions which interest us.

3. Our main use of the long exact sequence will be to effect a trick known as 'dimension shifting': we shift a problem about computing H^1 to one about computing H^0. The procedure, roughly, is this. We want to compute $H^1(G,K)$ for some G-module K. Suppose we can embed K in a module M for which we know $H^1(G,M)$. Then by the long exact sequence, if we can compute H^0 of K, M and M/K, we can read off (at least the size of) $H^1(G,K)$.

All this relies on having a good supply of G-modules whose cohomology we know about. In our context, the appropriate modules are kernels of free finite covers. In the group theoretic terminology (at least if the base W of the cover is transitive) these modules are *coinduced* from a finite module for the stabiliser of a point: the relevant module is the binding group at that point. The next lemma is then seen as a special case of Shapiro's lemma in group cohomology (cf. [8], Proposition III.6.2), and it tells us how to compute the cohomology of the coinduced modules. We give a direct (bare hands) proof of this result which avoids mentioning coinduced modules and any machinery from group cohomology.

Lemma 7.1.4 (Shapiro's lemma) *Let W be a transitive permutation structure and $G = \mathrm{Aut}(W)$. Suppose $\pi : C \longrightarrow W$ is a free finite cover with abelian kernel K. Let $w \in W$, $H = \mathrm{Aut}(W/w)$ and $A = B(w) = \mathrm{Aut}(C(w)/W)$. Then for $i = 0, 1$ we have*

$$H^i(G, K) = H^i(H, A).$$

Moreover $H^1_c(G, K) = H^1_c(H, A)$.

Proof. **Case** $i = 0$. Clearly we get a map $H^0(G, K) \to H^0(H, A)$ by restriction of K to $C(w)$. Conversely, if $a(w) \in A$ is fixed by H, let $k \in K$ restricted to $C(w)$ equal $a(w)$. For each $w' \in W$ let $g \in G$ be such that $gw = w'$ and define $a(w') \in B(w')$ to be the result of conjugating k by g and restricting to $C(w')$. This is independent of the choice of k and g, and the resulting $a \in K$ is fixed by G. The two maps we have described are mutual inverses, and hence the result.

Case $i = 1$. It is clear that any derivation $G \to K$ gives rise to a derivation $H \to A$. We show that any $\delta \in Der(H, A)$ extends to some $d \in Der(G, K)$; any two such extensions differ by an inner derivation; and an inner derivation extends to an inner derivation. The proof is a short computation, but it might help if we rehearse the notation a little further. For $w' \in W$ we denote by $B(w')$ the restriction of K to $C(w')$. Then $K = \prod_{w' \in W} B(w')$, that is, functions f with domain W such that $f(w') \in B(w')$ for all $w' \in W$. We can also regard $B(w')$ as a subgroup of K: identify it with the functions in K which have support w'. With these identifications, for $g \in G$, $f \in K$ and $w' \in W$ we have that the G-action on K is given by $(gf)(w') = g(f(g^{-1}w'))$ (-this is g applied via the module action to the element $f(g^{-1}w')$ of $B(g^{-1}w')$, which we are regarding as an element of K).

So now let $\delta \in Der(H, A)$. Let $\tau : W \longrightarrow G$ be such that $\tau(w')w = w'$ for all $w' \in W$, and $\tau(w) = 1$. For $g \in G$ define $d(g) \in K$ by setting, for $w' \in W$:

$$d(g)(w') = -g\tau(g^{-1}w')\delta(\tau(g^{-1}w')^{-1}g^{-1}\tau(w')).$$

Then, for $g, h \in G$ we have:

$$d(gh)(w') = -gh\tau(h^{-1}g^{-1}w')\delta(\tau(h^{-1}g^{-1}w')^{-1}h^{-1}g^{-1}\tau(w'))$$

$$= -gh\tau(h^{-1}g^{-1}w')\delta([\tau(h^{-1}g^{-1}w')^{-1}h^{-1}\tau(g^{-1}w')][\tau(g^{-1}w')^{-1}g^{-1}\tau(w')])$$

$$= -gh\tau(h^{-1}g^{-1}w')[\tau(h^{-1}g^{-1}w')^{-1}h^{-1}\tau(g^{-1}w')\delta(\tau(g^{-1}w')^{-1}g^{-1}\tau(w'))$$

$$+\delta(\tau(h^{-1}g^{-1}w')^{-1}h^{-1}\tau(g^{-1}w'))]$$
$$= d(g)(w') + g(d(h)(g^{-1}w')).$$

So $d(gh) = d(g) + gd(h)$, whence d is a derivation. To check that d agrees with δ on H note that if $g \in H$ then our formula gives $d(g)(w) = -g\delta(g^{-1})$. As δ is a derivation, $\delta(1) = 0$ and applying the product rule to the equation $gg^{-1} = 1$ gives that $g\delta(g^{-1}) + \delta(g) = 0$. So indeed, $d(g)(w) = \delta(g)$.

Now we show that any two extensions of δ differ by an inner derivation. Equivalently, we show that if $\delta = 0$, then any extension $d \in Der(G, K)$ of δ is inner. Indeed, given such a d we define $f \in K$ as follows. Let $g \in G$ and set

$$f(gw) = g(d(g^{-1})(w)).$$

(This is g applied to an element of $B(w)$.) To check this is well-defined, note that if $h \in H$ then

$$gh(d(h^{-1}g^{-1})(w)) = gh((h^{-1}d(g^{-1}))(w)) = g(d(g^{-1})(w)).$$

Let $g \in G$. For $w' \in W$, let $g' \in G$ be such that $gg'w = w'$. Then

$$(gf - f)(w') = g(f(g^{-1}w')) - f(w')$$
$$= gg'(d(g'^{-1})(w)) - gg'(d(g'^{-1}g^{-1})(w))$$
$$= gg'((d(g'^{-1}) - d(g'^{-1}) - g'^{-1}d(g^{-1}))(w))$$
$$= -gg'(g'^{-1}(d(g^{-1})(g'w))) = -g(d(g^{-1})(g'w))$$
$$= (-gd(g^{-1}))(w') = d(g)(w').$$

So $d(g) = gf - f$, as required.

The fact that an inner derivation δ in $Der(H, A)$ extends to inner derivations in $Der(G, K)$ follows from the above and the observation that δ can be extended to *some* inner derivation. Indeed, suppose that $a \in A$ is such that $\delta(h) = ha - a$ for all $h \in H$. Then a can be regarded as an element of K, and so δ is extended by the inner derivation $d(g) = ga - a$ (for all $g \in G$).

The proof is complete, apart from our assertion that all the above remains true if we consider only *continuous* derivations. So suppose that $\delta : H \longrightarrow A$ is continuous. We must show that $d : G \longrightarrow K$ as defined above is also continuous. So take $g \in G$ and $w_1, \ldots, w_n \in W$. It is enough to find an open subgroup H_1 of G such that if $h \in H_1$ then $d(gh)(w_i) = d(g)(w_i)$. Thus we must ensure that $(gd(h))(w_i) = 0$, that is $d(h)(g^{-1}w_i) = 0$ for $i = 1, \ldots, n$. Clearly we may take $H_1 \leq H$ and H_1 fixing each $g^{-1}w_i$. Then our requirement is that

$$\delta(\tau(g^{-1}w_i)^{-1}h^{-1}\tau(g^{-1}w_i)) = 0.$$

Let $H_0 = \{s \in H : \delta(s) = 0\}$. As δ is continuous and A is finite, this is an open subgroup of H. Thus

$$H_1 = \mathrm{Aut}(W/w, g^{-1}w_1, \ldots, g^{-1}w_n) \cap \bigcap_{i=1}^{n} \tau(g^{-1}w_i)H_0\tau(g^{-1}w_i)^{-1}$$

is an open subgroup of G satisfying our requirements. □

Computation of the groups $H^i_c(H, A)$ is, in the G-finite case, a problem about finite groups:

Lemma 7.1.5 *Suppose H is a topological group and A a finite topological H-module. Suppose further that H has an irreducible closed subgroup T of finite index. Then T acts trivially on A and:*
(i) $H^0(H, A) = H^0(H/T, A)$;
(ii) $H^1_c(H, A) = H^1(H/T, A)$.
In particular, these groups are finite.

Proof. It is clear that T acts trivially on A. We prove (ii). If $d : H \to A$ is a continuous derivation, then d restricted to T is actually a homomorphism, so must be the zero map. Thus d induces a derivation $H/T \to A$, and the result follows easily. □

7.2 Finiteness results

We now use the dimension-shifting technique to prove some results about finite covers. First, for suitable permutation structures W, we describe the minimal finite covers of W which have finite kernels. The result (Theorem 7.2.1) is less elegant than the results of Section 4, but is more general. The key idea in the proof (see Lemma 7.2.4) is an application of dimension shifting. We then show that, given certain chain conditions on covers described in Section 7.2.2, a similar trick can be used to prove that the cohomology groups $H^1_c(\mathrm{Aut}(W), K_0/K)$ from Corollary 6.4.4 have to be finite. All the results here are taken from [25].

7.2.1 Minimal covers with finite kernels

In this section we shall outline the proof of the following result (taken from [25]), which was previously stated as a conjecture in [24]. The key step is Lemma 7.2.4.

Theorem 7.2.1 *Let W be a countable, transitive, irreducible, permutation structure with automorphism group $G = \mathrm{Aut}(W)$. Suppose that G has finitely many orbits on triples from W, and that for all $x, y \in W$, each of $\mathrm{Aut}(W/x)$ and $\mathrm{Aut}(W/x, y)$ has a smallest closed subgroup of finite index. Then there is a natural number r such that if $\pi : C \longrightarrow W$ is a minimal finite cover with finite kernel K, then K can be generated by r elements.*

Henceforth, the hypotheses of the theorem which relate to W will be in force. Recall that if W is irreducible, then a minimal finite cover of W is also irreducible (4.1.1), and if the kernel of an irreducible finite cover is finite, then it is central in the automorphism group (4.1.2) and the cover is minimal. Recall also that by the *rank* of a finite abelian group, we mean the minimum number of elements needed to generate it, and that we call a finite cover with finite kernel *superlinked*. The first two lemmas reduce the proof of the theorem to computation of a cohomology

group. Their proofs are routine, if somewhat technical, and we omit them. See [25] for more details.

Lemma 7.2.2 *Suppose there is a non-trivial, minimal superlinked finite cover π : $C \longrightarrow W$ with kernel K of rank r. Then for some prime p there is a minimal superlinked finite cover of W whose kernel is an elementary abelian p-group of rank r.*□

For a prime p, let F_p denote the field with p elements (and consider this as a trivial G-module). We consider F_p^W as a topological G-module (with the product topology and G-action $(gf)(w) = f(g^{-1}w)$ for $g \in G$, $f \in F_p^W$, $w \in W$). There is a natural G-submodule isomorphic to the trivial module F_p, namely the constant functions $W \longrightarrow F_p$. Denote this by Δ.

Lemma 7.2.3 *There exists a natural number n (depending only on W) with the following property. Suppose that $H_c^1(G, F_p^W/\Delta)$ is finite, of cardinality p^t. Let $\pi : C \longrightarrow W$ be a minimal, superlinked finite cover whose kernel is an elementary abelian p-group of rank r. Then $r \leq t + n$.* □

We omit the proof (the main point is, of course, the machinery in Section 6.4), but we shall say what n is here. By assumption, there is a number l such that any continuous finite image of the stabiliser of a point in W has size at most l. By (2.1) of [17] there exists an integer n such that if T is a finite group of size at most l and $\phi : S \longrightarrow T$ is a Frattini cover with kernel Z, then Z has rank at most n.

The theorem now reduces to the following lemma, whose proof uses the dimension shifting trick.

Lemma 7.2.4 *Let W be a countable, transitive, irreducible, permutation structure with automorphism group $G = \mathrm{Aut}(W)$. Suppose that G has finitely many orbits on triples from W, and that for all $x, y \in W$, each of $\mathrm{Aut}(W/x)$ and $\mathrm{Aut}(W/x, y)$ has a smallest closed subgroup of finite index. Then there is a natural number t such that for every prime p, we have*

$$|H_c^1(G, F_p^W/\Delta)| \leq p^t.$$

Proof. Let $K = F_p^W/\Delta$, and let $W^{(2)}$ denote the set of ordered pairs of distinct elements of W. Clearly, G acts on this. Consider the (topological) G-module $F_p^{W^{(2)}}$ and the continuous map $\alpha : F_p^W \longrightarrow F_p^{W^{(2)}}$ given by $\alpha(f)((x, y)) = f(x) - f(y)$, for $f \in F_p^W$ and $(x, y) \in W^{(2)}$. The kernel of α consists of the constant functions in F_p^W, and so α gives a continuous embedding of K into $F_p^{W^{(2)}}$. We are now in a position to apply the Long Exact Sequence (7.1.2) and the 'dimension-shifting trick' described in 7.1.3(3) to the short exact sequence

$$0 \to K \to F_p^{W^{(2)}} \to \bar{K} \to 0$$

where $\bar{K} = F_p^{W^{(2)}}/im(\alpha)$.

Let $R_1, \ldots, R_s$ be the G-orbits on $W^{(2)}$. Then

$$H_c^1(G, F_p^{W^{(2)}}) = \bigoplus_{i=1}^{s} H_c^1(G, F_p^{R_i}).$$

If $(x, y) \in R_i$, then by Shapiro's lemma (7.1.4)

$$H_c^1(G, F_p^{R_i}) = H_c^1(G_{x,y}, F_p).$$

Now, a continuous derivation into the trivial module is just a continuous homomorphism. So $H_c^1(G_{x,y}, F_p)$ is just the largest elementary abelian p-group which is a continuous image of $G_{x,y}$. By assumption, there is an absolute (finite) bound on this, independent of p. So there exists a natural number m such that for any prime p,

$$|H_c^1(G, F_p^{W^{(2)}})| \leq m.$$

Now, by the long exact sequence (Lemma 7.1.2) applied to the above short exact sequence:

$$|H_c^1(G, K)| \leq |H_c^1(G, F_p^{W^{(2)}})||H^0(G, \bar{K})|.$$

We shall show that there is a bound (independent of p) on the rank of $H^0(G, \bar{K})$.

Apply Pontriagin duality (see Theorem 6.3.2) to the exact sequence of compact G-modules:

$$0 \to F_p \to F_p^W \xrightarrow{\alpha} F_p^{W^{(2)}} \to \bar{K} \to 0.$$

We get an exact sequence of discrete G-modules

$$0 \to (\bar{K})^* \to F_p W^{(2)} \xrightarrow{\beta} F_p W \to F_p \to 0.$$

Here, the map

$$\beta = \alpha^* : F_p W^{(2)} \to F_p W$$

is given by $\beta(x, y) = x - y$.

Under this duality, fixed points in $\bar{K}$ correspond to co-fixed points of $(\bar{K})^*$ (that is, submodules L where $(\bar{K})^*/L$ is the trivial module F_p). So

$$H^0(G, \bar{K}) = H_0(G, ker(\beta))$$

where, for a G-module M,

$$H_0(G, M) = M/\langle gm - m : g \in G, m \in M \rangle$$

is the largest quotient of M on which G acts trivially. Now, it is easy to check that

$$ker(\beta) = \langle (x, y) + (y, z) + (z, x), (x, y) + (y, x) : (x, y, z) \in W^{(3)} \rangle.$$

By assumption, there are only finitely many G-orbits on triples and pairs of distinct elements from W. Let $\{(x_i, y_i, z_i) : i \leq l\}$ and $\{(x^i, y^i) : i \leq s\}$ be representatives

for these orbits. Given any $(x,y,z) \in W^{(3)}$, there exist $g \in G$ and $i \leq l$ such that $g(x,y,z) = (x_i, y_i, z_i)$. Then $(x,y)+(y,z)+(z,x)$ and $(x_i,y_i)+(y_i,z_i)+(z_i,x_i)$ are in the same coset of $\langle gm - m : g \in G, m \in ker(\beta)\rangle$. Similarly, for any $(x,y) \in W^{(2)}$ there exists a $j \leq s$ such that $(x,y)+(y,x)$ and $(x^j,y^j)+(y^j,x^j)$ are in the same coset of $\langle gm - m : g \in G, m \in ker(\beta)\rangle$. This shows that the rank of $H_0(G, ker(\beta))$ is at most $s + l$, which concludes the proof. □

Theorem 7.2.1 now follows immediately from the preceding lemmas.

Remark 7.2.5 Bounds for r in Theorem 7.2.1 can be read off from the above proof. The proof given in [25] avoids the use of Pontriagin duality and gives slightly better bounds. Alternative proofs of Theorem 7.2.1 have recently been given by E. Hrushovski (private communication) and J. Koshan ([45]).

7.2.2 The dcc on covers

The definitions and results in this section are taken from Section 5 of [27]. However, the idea is really due to Ahlbrandt and Ziegler in [2].

Definition 7.2.6 Suppose $\pi : C \longrightarrow W$ is a finite cover of the permutation structure W. Let $\mu : \mathrm{Aut}(C) \to \mathrm{Aut}(W)$ be the restriction map. We say that $\pi : C \longrightarrow W$ has the *descending chain condition* on covers of W (dcc, for short) if any chain

$$\mathrm{Aut}(C) > G_1 > G_2 > \ldots$$

of closed subgroups satisfying $\mu(G_i) = \mathrm{Aut}(W)$ is finite. We say that it has the descending chain condition on quasi-covers of W (qdcc, for short) if any chain

$$\mathrm{Aut}(C/W) > K_1 > K_2 > \ldots$$

of closed normal subgroups of $\mathrm{Aut}(C)$ is finite.

Remarks 7.2.7 **1.** If the kernel of π is abelian then qdcc implies dcc. If also π is split, then dcc implies qdcc.

2. Note that the example of 6.3.6 shows that there are finite covers of $\aleph_0$-categorical structures not having the *ascending* chain condition on covers. We do not know of an $\aleph_0$-categorical W *not* having dcc on all of its finite covers.

Definition 7.2.8 Let W be a countably infinite permutation structure and $<$ a linear ordering on W of type ω. Define a partial ordering $<<$ on W by saying that $a << b$ if and only if there exists $g \in \mathrm{Aut}(W)$ such that $ga = b$ and $gc < b$ for all $c < a$. Say that $<$ is a *nice ordering* of W if $(W, <<)$ has no infinite antichain.

Ahlbrandt and Ziegler [2] show that the Grassmannian of k-sets from a disintegrated set, and the Grassmannian of k-dimensional subspaces from a countably infinite vector space over a finite field have nice orderings. It is easy to show that any enumeration of the rationals (as an ordered set) is a nice ordering. More interestingly, M. Albert and A. Chowdhury ([5]) have shown that the Grassmannian of

ordered k-sets from the rationals has a nice ordering, thereby answering a question raised in the original version of these notes. It is an easy observation that if $\mathrm{Aut}(W)$ has a transitive cyclic subgroup, then W has a nice ordering. So, for example, the random graph has a nice ordering. The following are taken from Section 5 of [27].

Lemma 7.2.9 *If W has a nice ordering and $\pi : C \longrightarrow W$ is a finite cover, then C has a nice ordering.* □

Theorem 7.2.10 *Suppose W has a nice ordering. Then any finite cover $\pi : C \longrightarrow W$ has dcc and qdcc.* □

7.2.3 Finiteness of H^1_c

Given a structure W and a finite cover $\pi : C \to W$ with kernel K, it is natural to ask whether knowing K determines π up to finitely many possibilities. In the context of totally categorical structures (and working also with affine covers as well as finite covers), the statement that this question has an affirmative answer has become known as 'Ziegler's Finiteness Conjecture.' In fact, the conjecture is now known to hold more generally for smoothly approximated structures W. The proof of this is an elegant compactness argument using quasifinite axiomatizability of these structures. Full details will appear in the final version of [16], but a sketch can be found in the notes [14].

In this section we show how dimension-shifting and appropriate chain conditions also give Ziegler's finiteness conjecture for finite covers with abelian kernel (the results are applicable, for example, to the cases where W is a Grassmannian of a disintegrated set, or of a projective space over a finite field). It should be noted however, that neither of these approaches gives effective bounds on the number of covers with a particular kernel, and both rely heavily on the existence of particular enumerations of the base W. The following is the precise formulation of what we shall prove (a similar statement, for arbitrary kernels can be found in [25]).

Theorem 7.2.11 *Suppose W is a countable, transitive, G-finite, permutation structure such that all Grassmannians of W have nice orderings. Let $\pi : C \longrightarrow W$ be a finite cover of W with abelian kernel K_0, and let K be a closed subgroup of K_0 which is normal in* $\mathrm{Aut}(C)$. *Then there are only finitely many* $\mathrm{Aut}(C)$*-conjugacy classes of closed, full subgroups H of* $\mathrm{Aut}(C)$ *with $H \cap K_0 = K$.*

The subgroups H in the above statement are precisely automorphism groups of finite covers of W which are expansions of π and have kernel K. So the theorem implies that there are only finitely many isomorphism types of these. Corollary 6.4.4 shows that in order to prove the theorem, it is enough to prove that the cohomology group $H^1_c(\mathrm{Aut}(W), K_0/K)$ is finite. The dimension-shifting involves embedding the quotient module K_0/K as a submodule of the kernel of some free cover. To do this we need the following lemma.

Lemma 7.2.12 *Assume that W is a countable, transitive permutation structure. Suppose $\pi : C \longrightarrow W$ is a finite cover with abelian kernel K_0, and K is closed*

subgroup of K_0 which is normal in $\mathrm{Aut}(C)$. *Suppose π has qdcc on covers of W. Then there exists a finite subset X of C such that*

$$K = K(X) := \bigcap_{g \in \mathrm{Aut}(C)} K.\mathrm{Aut}(C/gX).$$

Proof. Write C as a union of an increasing chain of finite non-empty subsets $(X_i : i < \omega)$. Then $K(X_1) \geq K(X_2) \geq \ldots$, and we claim that $\bigcap_{i<\omega} K(X_i) = K$. Indeed, for each i, $K \leq K(X_i) \leq K_0$ and each $K(X_i)$ is closed. Let $g \in \bigcap_{i<\omega} K(X_i)$. Write $g = k_i h_i$ where $k_i \in K$ and $h_i \in \mathrm{Aut}(C/X_i)$. As $g \in K_0$ we have that $h_i \in K_0$. Now, the h_i converge to 1, and so the k_i converge to g. Thus, as K is closed, $g \in K$, as required. Each $K(X_i)$ is a closed normal subgroup of $\mathrm{Aut}(C)$, and so the statement now follows from our assumption of qdcc. □

Sketch of Proof of Theorem 7.2.11. Take X given by the above lemma, and let $Y = \pi(X)$. Let $\Gamma = \mathrm{Aut}(C)$, $\Sigma = K.\mathrm{Aut}(C/X)$ and $G = \mathrm{Aut}(W)$. Let C_1 be the cosets of Σ in Γ, considered as a permutation structure with automorphisms those permutations induced by Γ. This can be considered as a finite cover $\pi_1 : C_1 \longrightarrow W_1$, where W_1 is the permutation structure of cosets of G_Y in G (with G acting). Equivalently, W_1 is the G-orbit containing a particular enumeration of Y. There is a natural map $\mathrm{Aut}(C) \longrightarrow \mathrm{Aut}(C_1)$ with kernel K. The point is that the kernel K_1 of π_1 is topologically isomorphic to K_0/K, and what we want to compute is $H^1_c(G, K_1)$.

Now take a reduct $\pi_2 : C_2 \longrightarrow W_1$ of π_1 which is a free cover, as in 2.1.3. Denote its kernel by K_2 and note, of course, that K_1 is a G-submodule of K_2. Shapiro's lemma (7.1.4) tells us how to compute $H^1_c(G, K_2)$, and assuming G-finiteness of W, this will be finite, by 7.1.5. So, by the long exact sequence (7.1.2), we are reduced to computing $H^0(G, K_2/K_1)$. This is just fixed points of G on K_2/K_1. By our assumptions and 7.2.9, W_1 has a nice ordering. So by 7.2.10, π_2 has qdcc. So there can only be finitely many fixed points of G on K_2/K_1. □

8 Problems

This section contains a selection of problems which we came across during the writing of these notes. Presumably some are more tractable than others; indeed, some may be quite straightforward.

Our first problem is open-ended.

Problem 8.1 Take a particular $\aleph_0$-categorical structure W and classify its finite covers.

Here, it might be interesting to take as W a Grassmannian of a structure D for which there is good information about its finite covers (for example, one of the structures in Theorem 4.3.5). But even for cases (iii) and (iv) in 4.3.5, the following seems open.

Problem 8.2 Is the description of the possible kernels of a finite cover of the rationals given in Lemma 3.1.7 true for all primitive W having trivial algebraic closure and with irreducible $\mathrm{Aut}(W)$ and $\mathrm{Aut}(W/w)$ (for $w \in W$)?

Problem 8.3 Suppose W is a countable, G-finite $\aleph_0$-categorical structure and $\pi_0 : C \to W$ is a finite cover. Let K be a closed subgroup of the kernel of π_0. Are there only finitely many isomorphism types of covering expansions of π_0 with kernel K?

As outlined in Section 7, the above has an affirmative answer in the case of π_0 having abelian kernel if either of the following is true in general.

Problem 8.4 Suppose W is a countable, G-finite $\aleph_0$-categorical structure. Does W have a nice enumeration? Does W have qdcc on finite covers?

Problem 8.5 Suppose W is a G-finite $\aleph_0$-categorical structure and $\pi_0 : C_0 \to W$ is a finite cover with abelian kernel K_0. Let π be a covering expansion of π_0 with kernel K. Is the cohomology group $H^1_c(\mathrm{Aut}(W), K_0/K)$ necessarily finite?

Problem 8.6 In cases where Problem 8.5 is known to have an affirmative answer (for example, Grassmannians of vector spaces over finite fields), give explicit bounds on the size of the cohomology groups.

Also related to these is:

Problem 8.7 Is a finite cover of a countable, G-finite $\aleph_0$-categorical structure necessarily G-finite?

The following is taken from [23], where an example is given of a minimal non-trivial free cover (Example 4.6 of [23]).

Problem 8.8 Investigate minimality of free covers. In particular, is the free cover described in Example 4 of 1.2 minimal?

Many of the results we have described have been directed at showing that certain finite covers must split. In fact, all of the non-split examples we have described have involved a non-split cover with finite kernel or have come from a non-split extension of a finite group (via the free cover construction, as in 2.1.5, or a simple amalgamation as in 3.3.6). So we pose the following:

Problem 8.9 Does there exist a transitive, irreducible $\aleph_0$-categorical structure W which has an untwisted, minimal finite cover $\pi : C \longrightarrow W$ with infinite elementary abelian kernel?

There is an example due to Ivanov of a non-free, non-superlinked minimal finite cover of the random two-graph (Example 2.3.1) where the fibre groups are cyclic of order 4.

The following is of course suggested by the work [2] of Ahlbrandt and Ziegler (and others) on quasi finite-axiomatisability of totally categorical structures.

Problem 8.10 Investigate finite axiomatisability of a finite cover $\pi : C \to W$ relative to an axiomatisation of W.

The following is suggested by work in [38] and Example 5 in 1.2.

Problem 8.11 Investigate finite covers which are homogeneous for a finite language.

Affine covers are treated on an equal footing with finite covers in [4] and [34], but have been rather neglected in the present survey. So we pose as our final problem:

Problem 8.12 Develop further the theory of affine covers.

References

[1] E. F. Assmus Jr. and J. Key, *Designs and their Codes*, Cambridge University Press, Cambridge, 1992.

[2] G. Ahlbrandt, M. Ziegler, 'Quasi-finitely axiomatizable totally categorical theories', Ann. Pure Appl. Logic 30 (1986) 63–82.

[3] G. Ahlbrandt, M. Ziegler, 'Invariant subspaces of ${}^{V}V$', J. Algebra 151 (1992) 26–38.

[4] G. Ahlbrandt, M. Ziegler, 'What's so special about $(\mathbf{Z}/4\mathbf{Z})^{\omega}$?', Archive for Mathematical Logic 31 (1991), 115–132.

[5] M. Albert, A. Chowdhury, 'The rationals have an AZ-enumeration', Preprint, July 1996.

[6] J. T. Baldwin, A. H. Lachlan, 'On strongly minimal sets', J. Symbolic Logic 36 (1971), 79–96.

[7] G. B. Bell, 'On the cohomology of the finite special linear groups I, II' , J. Algebra 54 (1978), 216–238, 239–259.

[8] K. S. Brown, *Cohomology of Groups*, Springer GTM 87, Springer Verlag, Berlin 1982.

[9] P. J. Cameron, 'Transitivity of permutation groups on unordered sets', Math.Zeit. 148 (1976), 127 – 139.

[10] P. J.Cameron, 'Cohomological aspects of two-graphs', Math.Zeit. 157 (1977), 101 – 119.

[11] A. R. Camina and D. M. Evans, 'Some infinite permutation modules', Quart. J. Math. Oxford (2) 42 (1991), 15–26.

[12] C. C. Chang, H. J. Keisler, *Model Theory* (3rd edition), North Holland, Amsterdam, 1990.

[13] G. Cherlin, 'Homogeneous directed graphs. The imprimitive case', In: *Logic Colloquium'85*, eds. Paris Logic Group. North-Holland, 1987, 67 – 88.

[14] G. Cherlin, 'Large finite structures with few types', Notes produced for the NATO ASI on Algebraic Model Theory, Toronto, August 1996.

[15] G. Cherlin, L. Harrington, A.H. Lachlan, '$\aleph_0$-categorical, $\aleph_0$-stable structures', Ann. Pure Appl. Logic 28 (1985) 103–135.

[16] G.Cherlin, E. Hrushovski, unpublished manuscript on smoothly approximated structures, Rutgers/MIT, 1991.

[17] J. Cossey, O. H. Kegel, L. G. Kovács, 'Maximal Frattini extensions', Archiv der Math. (Basel) 35 (1980), 210–217.

[18] P. Delsarte, 'On cyclic codes that are invariant under the general linear group', IEEE Trans. Information Theory, Vol. IT-16 (1970), 760–769.

[19] M. M. Erimbetov, 'Categoricity in power and non-two-cardinal formulas of finite rank' (in Russian), Algebra i Logika 13 (1974) 493–500.

[20] D. M. Evans, W.A. Hodges, I.M. Hodkinson, 'Automorphisms of bounded abelian groups', Forum Math. 3 (1991), 523–541.

[21] D. M. Evans, 'Homogeneous geometries', Proc. London Math. Soc. (Series 3), 52 (1986), 305–327.

[22] D. M. Evans, 'A note on automorphism groups of countably infinite structures', Archiv der Math. (Basel) 49 (1987), 479–483.

[23] D. M. Evans, 'Splittings of finite covers of $\aleph_0$-categorical structures', J. London Math. Soc.(2), 54 (1996), 210–226.

[24] D. M. Evans, 'Finite covers with finite kernels', Ann. Pure Appl. Logic, to appear.

[25] D. M. Evans, 'Computation of first cohomology groups for some finite covers', J. Algebra, to appear.

[26] D. M. Evans, D. G. D. Gray, 'Kernels and cohomology groups for some finite covers', Preprint, Norwich, September 1996.

[27] D. M. Evans, E. Hrushovski, 'On the automorphism groups of finite covers', Ann. Pure Appl. Logic 62 (1993), 83–112.

[28] M. D. Fried, M. Jarden, *Field Arithmetic*, Springer-Verlag, Berlin, 1986.

[29] D. G. D. Gray, 'The structure of some permutation modules for the symmetric group of infinite degree', J. Algebra, to appear.

[30] G. Higman, 'On infinite simple groups', Publ. Math. Debrecen 3 (1954), 221–226.

[31] W. A. Hodges, I. M. Hodkinson, D. Lascar and S. Shelah, 'The small index property for ω-stable, ω-categorical structures and for the random graph', J. London Math. Soc. (2) 48 (1993), 204–218.

[32] W. A. Hodges, I. M. Hodkinson, H. D. Macpherson, 'Omega-categoricity, relative categoricity and coordinatisation', Ann. Pure Appl. Logic 46 (1990), 169–199.

[33] W. A. Hodges, 'The structure of totally categorical structures', In: *Automorphisms of First-Order Structures*, eds. R. Kaye, D. Macpherson, Oxford University Press, Oxford, 1994, 111–130.

[34] W. A. Hodges, A. Pillay, 'Cohomology of structures and some problems of Ahlbrandt and Ziegler', J. London Math. Soc. (2) 50 (1994), 1-16.

[35] E. Hrushovski, 'Unidimensional theories: an introduction to geometric stability theory', In: *Logic Colloquium '87, Granada*. North-Holland, Amsterdam, 1989, 73–103.

[36] E. Hrushovski, 'Totally categorical theories', Trans. Amer. Math. Soc. 313 (1989), 131–159.

[37] E. Hrushovski, 'Unimodular minimal theories', J. London Math. Soc. (2) 46 (1992), 385–396.

[38] A. A. Ivanov, 'Some combinatorial aspects of the cover problem for totally categorical theories', In: *Automorphisms of First-Order Structures*, eds. R. Kaye, D. Macpherson, Oxford University Press, Oxford, 1994, 215–231.

[39] A. A. Ivanov, 'Finite covers, cohomology and homogeneous structures', Preprint, Wrocław, 1995.

[40] A. A.Ivanov, D.Macpherson, 'Strongly determined types', Preprint, Wrocław, 1996.

[41] G. D. James, *The Representation Theory of the Finite Symmetric Groups*, Springer LNM 682, Springer, Berlin, 1978.

[42] R. Kaye, D. Macpherson, 'Models and groups', In: *Automorphisms of First-Order Structures*, eds. R.Kaye, D.Macpherson, Oxford University Press, Oxford, 1994, 3 – 31.

[43] H. Kikyo, A. Tsuboi, 'On reduction properties', J. Symb. Logic 59 (1994), 900–911.

[44] J. Koshan, *Superlinked finite covers with central kernels*, M.Sc. Thesis, Simon Fraser University, 1995.

[45] J. Koshan, 'Structure results for transitive, untwisted, superlinked finite covers', Preprint, December 1995.

[46] D. Lascar, 'On the category of models of a complete theory', J. Symb. Logic 47(1982), 249 – 266.

[47] L. Pontriagin, *Topological Groups*, 2nd Edition, Gordon and Breach, New York, 1966.

[48] M. Ziegler, 'Notes on totally categorical theories', unpublished manuscript, Freiburg, 1991.

[49] M. Ziegler, 'Finite covers of disintegrated sets', unpublished notes, Freiburg, 1992.

[50] B. I. Zil'ber, 'Strongly minimal countably categorical theories, II and III', Siberian J. Math., 25(3) (1984), 71–88, and 25(4) (1984), 63–77.

[51] B. I. Zil'ber, *Uncountably categorical theories*, Transl. Math. Monog., Amer. Math. Soc., Providence, 1993.

Authors' addresses:
David M. Evans,
School of Mathematics,
UEA,
Norwich NR4 7TJ,
England.
e-mail: d.evans@uea.ac.uk

Alexandre A. Ivanov,
Institute of Mathematics,
Wrocław University,
pl. Grunwaldzki 2/4,
50-385 Wrocław,
Poland.
e-mail: ivanov@hera.math.uni.wroc.pl

Dugald Macpherson,
Department of Pure Mathematics,
University of Leeds,
Leeds LS2 9JT,
England.
e-mail: pmthdm@amsta.leeds.ac.uk

Definable subgroups of algebraic groups over pseudo-finite fields

Notes by

Zoé Chatzidakis

The aim of these notes is to give an account of a result proved by E. Hrushovski and A. Pillay, describing the definable subgroups of algebraic groups over pseudo-finite fields [8]. This result has numerous applications to algebraic groups defined over finite fields. In section 1, we recall the basic definitions and facts from algebraic geometry. In section 2, we give a brief account of the results needed on pseudo-finite fields, and in section 3 we give a proof of the main result. We conclude with an application to the definability of maximal subgroups of certain algebraic groups defined over the fields $\mathbb{F}_p$.

These notes should be read in conjunction with those of Wilfrid Hodges [6].

Notations and conventions

We denote the algebraic closure of a field F by $\tilde{F}$, and its separable closure by F_s. The field F is perfect if it is of characteristic 0, or if it is of characteristic $p > 0$ and closed under taking p-th roots. The Galois group $\mathcal{G}al(F_s/F)$ acts on $\tilde{F}$, with fixed subfield F^{1/p^∞}, the perfect hull of F. We often identify $Aut(\tilde{F}/F)$ with $\mathcal{G}al(F_s/F)$.

As usual, $\mathbb{Z}$ denotes the ring of integers, $\mathbb{F}_q$ the field with q elements for a prime power q . The language $\mathcal{L}$ is the usual language of rings: $\mathcal{L} = \{+, -, \cdot, 0, 1\}$.

1. Preliminaries in algebraic geometry: definitions and main facts

Let F be a subfield of K, where K is some large algebraically closed field. We refer to [9], Chapter III for proofs. Throughout this section, we work in the theory T_{acf} of algebraically closed fields.

(1.1) Let n be an integer. The set K^n is called the affine space of dimension n (over K); it is also sometimes denoted by $\mathbf{A}^n$, or by $\mathbf{A}^n(K)$. The set $V \subseteq K^n$ is an (affine) algebraic set (also called: Zariski closed set, or closed) if it is the zero-set of a set of polynomials over K (in n indeterminates). If these polynomials have their coefficients in F, we say that V is F-closed.

If V is an algebraic set, we denote by $V(F)$ the set of F-rational points of V (that is, having all their coordinates in F).

(1.2) The topology on K^n whose closed sets are the algebraic sets, is called the Zariski topology. If $S \subseteq K^n$, there is a smallest algebraic set containing S: it is called the Zariski closure of S and denoted by $\bar{S}$.

Suppose that K^* is an algebraically closed field containing K; then the topology induced on K^n by the Zariski topology on K^{*n} coincides with the Zariski topology on K^n.

(1.3) For $S \subseteq K^n$ we define

$$I(S) = \{f \in K[X_1, \ldots, X_n] \mid f(a) = 0 \text{ for all } a \in S\}.$$

Then $\bar{S}$ is precisely the set of zeros of $I(S)$. Observe that because $K[X_1, \ldots, , X_n]$ is noetherian, every descending chain of closed sets is finite.

(1.4) We say that an algebraic set V is defined over F if $I(V)$ is generated by polynomials in $F[X_1, \ldots, X_n]$; we say that F is the field of definition of V if V is defined over F, and F is smallest such. The field of definition of V is unique, and is finitely generated (as a field).

In characteristic 0, an algebraic set V is F-closed if and only if it is defined over F; in characteristic $p > 0$, this is however not the case: the equation $(f(X))^p = 0$ has same zero-set as $f(X) = 0$, and therefore V is F^p-closed whenever it is F-closed. One has: V is F-closed if and only if V is defined over F^{1/p^∞}.

The notion of "defined over" for an algebraic set is therefore stronger than the model-theoretic notion of "definable over".

(1.5) Let V be F-closed; V is F-irreducible if it is not the union of two proper F-closed subsets; V is a variety (or irreducible) if V is not the union of two proper closed subsets.

Clearly, if V is a variety then it is F-irreducible, but the converse is in general false, unless F is separably closed. Using the descending chain condition on closed sets, one shows that any F-closed set V can be written uniquely as $V = V_1 \cup \cdots \cup V_m$, where the V_i's are varieties, and no V_i is contained in the union of the others; the V_i's are called the (irreducible) components of V. If V is F-irreducible and $V = V_1 \cup \cdots \cup V_m$ is the decomposition of V into irreducible components, the varieties V_i are defined over some normal algebraic extension F' of F, and are conjugate under the action of $\mathcal{G}al(F'/F)$. Observe also:

V is F-irreducible if and only if $I(V) \cap F[X_1, \ldots, X_n]$ is a prime ideal.

V is a variety if and only if $I(V)$ is a prime ideal.

If V is a variety and F-closed, then $I(V)$ is the radical of the ideal of $K[X_1, \ldots, X_n]$ generated by $I(V) \cap F[X_1, \ldots, X_n]$.

(1.6) A field F' containing F is a regular extension of F iff it is linearly disjoint from $\tilde{F}$ over F. Equivalent conditions are:

(i) The F-algebra $F' \otimes_F \tilde{F}$ is a domain.

(ii) F' is linearly disjoint from $F^{1/p}$ over F (F' is a separable extension of F) and F is relatively separably closed in F' (or $F' \cap F_s = F$). When F is perfect, the first condition is always satisfied.

(1.7) For an algebraic set V, we define the affine coordinate ring of V to be $K[V] =_{\text{def}} K[X_1, \ldots, X_n]/I(V)$. If V is a variety, then $K[V]$ is an integral domain and its quotient field, $K(V)$, is called the function field of V.

If V is F-closed, we define $F[V] = F[X_1, \ldots, X_n]/(I(V) \cap F[X_1, \ldots, X_n])$. If V is F-irreducible, $F(V)$ is the quotient field of $F[V]$. Then

V is a variety if and only if F is relatively separably closed in $F(V)$.

V is a variety defined over F if and only if $F(V)$ is a regular extension of F.

If V and W are varieties, then so is their cartesian product $V \times W$, and $K[V \times W] = K[V] \otimes_K K[W]$. If V and W are F-closed and F-irreducible, then $V \times W$ may be F-reducible. This happens if $F[V] \cap F_s$ and $F[W] \cap F_s$ are not linearly disjoint over F.

(1.8) For a variety V, we define $dim(V)$ to be the transcendence degree of $K(V)$ over K (or equivalently of $F(V)$ over F if V is defined over F). A point $a \in V$ is called generic over F if the transcendence degree of $F(a)$ over F equals $dim(V)$. Note that if a is generic, then $F(a) \simeq_F F(V)$. If K has infinite transcendence degree, then it contains generic points of any variety (over its field of definition).

For an algebraic set V, we define $dim(V)$ to be the supremum of the dimensions of the irreducible components of V. For $S \subseteq K^n$, we define $dim(S) = dim(\bar{S})$.

(1.9) Let $a \in K^n$; we define $I(a/F)$ to be the ideal consisting of all polynomials $f \in F[X_1, \ldots, X_n]$ such that $f(a) = 0$; then $I(a/F)$ is a prime ideal. If $V \subseteq K^n$ is the associated algebraic set, it is then F-irreducible and $F(a) \simeq_F F(V)$; we call V the locus of a over F. Thus V is a variety if and only if F is relatively separably closed inside $F(a)$. Observe that by definition a is a generic point of V.

(1.10) Let a, V be as in (1.9). The model-theoretic interpretation, in the sense of the theory T_{acf}, is:

The Morley rank and U-rank of $tp(a/F)$ both equal $dim(V)$. The type $tp(a/F)$ is stationary if and only if V is a variety. If V is not a variety, then the multiplicity of $tp(a/F)$ equals the number of irreducible components of V. In terms of field extensions: $tp(a/F)$ is stationary if and only if F is relatively separably closed in $F(a)$; the multiplicity of $tp(a/F)$ equals $[F_s \cap F(a) : F]$.

We define the canonical base of $tp(a/F)$, denoted by $Cb(a/F)$, to be the perfect hull of the field of definition of V; it is definably closed in the sense of T_{acf}, and is contained in the perfect hull of F. It is the smallest definably closed subset of F over which a has same rank and multiplicity as over F. Note that we do not require $tp(a/F)$ to be stationary, for more details on canonical bases of non-stationary types see [2]. By abuse of notation we will write $b = Cb(a/F)$ whenever b is a tuple such that $dcl(b) = Cb(a/F)$.

Observe that, if $tp(a/F)$ is not stationary, then $Cb(a/\tilde{F})$ is contained in the definable closure of $F(a) \cap F_s$, and $tp(a/F(a) \cap F_s) \vdash tp(a/\tilde{F})$.

(1.11) Let $V \subseteq K^n$, $W \subseteq K^m$ be varieties. A morphism from V to W is a map $f = (f_1, \ldots, f_m)$ defined on V and taking its values in W, where each $f_i \in K[V]$. It induces a dual map $f^* : K[W] \to K[V]$, $g \mapsto g \circ f$, which is an inclusion of K-algebras if $f(V)$ is dense in W. A morphism is continuous (for the Zariski topology).

If f is bijective and f^{-1} is also a morphism, then f is called an isomorphism. There are bijective morphisms which are not isomorphisms, for instance in characteristic $p > 0$, the morphism $x \mapsto x^p$. If f is an isomorphism then f^* is an isomorphism, and conversely.

A rational map from V to W is a map $f = (f_1, \ldots, f_m)$ defined on some open subset of V and taking its values in W, and where each $f_i \in K(V)$. It induces a

dual map $f^* : K(W) \to K(V)$, $g \mapsto g \circ f$, which is an inclusion of K-algebras if $f(V)$ is dense in W. A rational map is continuous.

We say that f is birational if there is a rational map $g : W \to V$ such that $f \circ g$ is the identity. If f is birational then f^* is an isomorphism, and conversely. Two varieties are birationally equivalent if there is a birational map between them.

(1.12) A constructible set in K^n is a boolean combination of Zariski closed sets; it can be written as a finite union of sets of the form $V \cap U$, where V is a variety, and U a basic open set, i.e. of the form $\{a \in K^n \mid g(a) \neq 0\}$ for some polynomial g over K.

By quantifier-elimination of the theory T_{acf}, every definable subset of K^n is constructible.

(1.13) Abstract varieties. So far we have talked only of affine algebraic sets and varieties. There is a more general notion of variety, whose definition encompasses both affine varieties and projective varieties. Below, we will list the definitions pertaining to abstract varieties and some of their properties.

(1) An abstract variety $(V, U_i, V_i, \varphi_i)_{i \in I}$, I a finite set of indices, is given by a set $V = \bigcup_{i \in I} U_i$, affine varieties V_i, $i \in I$, and bijections $\varphi_i : U_i \to V_i$ such that for $i \neq j$, $f_{ij} = \varphi_j \varphi_i^{-1} : V_i \to V_j$ is a rational map, defined on the open subset $\varphi_i(U_i \cap U_j)$ of V_i.

(2) The topology on V is then defined in the following manner: a subset W of V is open if and only if $\varphi_i(W \cap U_i)$ is open (for the Zariski topology) in V_i for all $i \in I$. Our assumption on the sets $\varphi_i(U_i \cap U_j)$ implies that each U_i is open in V. This topology is called the Zariski topology.

(3) If all the varieties V_i and rational maps f_{ij} are defined over the field F, we say that V is defined over F. Note that the abstract variety is actually uniquely determined by the data $(V_i, f_{ij})_{i,j \in I}$.

(4) Observe that all varieties V_i have the same dimension, since each map f_{ij} is birational (with inverse f_{ji}) and therefore $F(V_i) \simeq F(V_j)$.

A point $a \in V$ is a generic point of V if $\varphi_i(a)$ is a generic point of V_i for all $i \in I$ (or, equivalently, for some $i \in I$, since one has: if $b = \varphi_i(a)$ is generic, then so is $f_{ij}(b)$).

For $a \in V$, one can define $F(a)$ to be the field $F(\varphi_i(a))$ for some $i \in I$ such that $a \in U_i$. Note that this definition is independent of the choice of i (up to an F-automorphism).

(5) A subvariety of V is an irreducible closed subset W of V. Equivalently, W is a subvariety of V if $\varphi_i(W \cap U_i)$ is a subvariety of V_i for each i. A point $b \in W$ is a generic of W if $\varphi_i(b)$ is a generic of $\varphi_i(W \cap U_i)$ for each i.

(6) Let $S \subseteq V$ be a set. We denote by $\bar{S}$ the closure of S for the topology on V. Then $\bar{S} = \bigcup_{i \in I} \varphi_i^{-1}(\overline{\varphi_i(S \cap U_i)})$.

(7) Let $(V, U_i, V_i, \varphi_i)_{i \in I}$ and $(W, S_j, W_j, \psi_j)_{j \in J}$ be two abstract varieties. A rational map $\theta : V \to W$ is a map with domain an open subset of V, and such that for $i \in I$ and $j \in J$ the maps $\theta_{ij} = \psi_j \theta \varphi_i^{-1} : V_i \to W_j$ are rational maps. Note that θ is continuous for the topology.

(8) If $(V, U_i, V_i, \varphi_i)_{i \in I}$ and $(W, S_j, W_j, \psi_j)_{j \in J}$ are abstract varieties, then their

product is also an abstract variety, given by $(V \times W, U_i \times S_j, V_i \times W_j, \varphi_i \times \psi_j)_{i \in I, j \in J}$.

(1.14) Example. Consider the projective space of dimension n, $\mathbb{P}^n$. It is the set of lines in affine $(n+1)$-space, and can be described as follows: let $S = K^{n+1} \setminus \{0\}$, and define an equivalence relation on S by: $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$ if $\lambda x_0 = y_0, \ldots, \lambda x_n = y_n$ for some $\lambda \in K$. Then $\mathbb{P}^n = S/\sim$. The representative of the equivalence class of $(x_0, \ldots, x_n)$ is often denoted by $(x_0 : \cdots : x_n)$.

One defines (projective) algebraic sets as in the affine case, except that one has to be careful to only consider zero-sets of sets of homogeneous polynomials. We will now show that $\mathbb{P}^n$ has a natural structure of abstract variety, and that the closed sets are precisely finite unions of algebraic sets. For $i = 0, \ldots, m$, consider the hyperplane H_i of $\mathbb{P}^n$ defined by the equation $x_i = 0$, and let $U_i = \mathbb{P}^n \setminus H_i$. There is a natural bijection $\varphi_i : U_i \to K^n = V_i$ given by

$$(x_0 : \cdots : x_n) \mapsto \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i} \ldots, \frac{x_n}{x_i}\right).$$

Then the maps $\varphi_j \varphi_i^{-1} : V_i \to V_j$ are rational maps. Hence $(\mathbb{P}^n, U_i, V_i, \varphi_i)_{0 \leq i \leq n}$ is an abstract variety. One also verifies that algebraic sets are closed, and conversely that an irreducible closed set is an algebraic set (look at the polynomials vanishing at a generic point of the closed set).

(1.15) Recall that a connected algebraic group is a group G, with a structure of abstract variety on G, and such that multiplication: $G \times G \to G$ and inverse: $G \to G$ are rational maps which are everywhere defined (on $G \times G$ and G). If the underlying variety is affine, we say that G is an affine algebraic group.

One can extend this definition to non-connected algebraic groups by defining an "abstract algebraic set": it is a union of sets U_i, each of them in bijection with some algebraic affine set V_i via a map φ_i, with the maps $\varphi_j \varphi_i^{-1}$ defined on an open subset of V_i, and given locally by rational maps.

Then, an algebraic group G is an abstract algebraic set, such that multiplication and inverse are everywhere defined and given locally by rational functions. In particular, closed subgroups of algebraic groups are algebraic groups. See [10] for a precise definition, and [6] for a detailed definition in the affine case and for related results. We conclude this section with two easy results on algebraic groups.

(1.16) Let G be an algebraic group defined over F, let S be a subset of G such that:

(i) S contains all the generics of $\bar{S}$ over F.

(ii) If $a, b \in S$ are generic and independent over F then $ab \in S$ and $a^{-1} \in S$.

Then $\bar{S}$ is a subgroup of G.

Proof. By assumption and since the map $a \mapsto a^{-1}$ is continuous, S^{-1} is a dense subset of the closed set $(\bar{S})^{-1}$ which contains all the generics of S. This implies that $(\bar{S})^{-1} = \overline{S^{-1}} \supseteq \bar{S}$, from which one deduces that $\bar{S} = \bar{S}^{-1}$.

Let $a \in S$ be generic; then the set $S(a) = \{b \in \bar{S} \mid ab \in \bar{S}\}$ is closed and contains all the generic elements of $\bar{S}$ which are independent from a over F; thus

$S(a) = \bar{S}$, and $a\bar{S} = \bar{S}$; similarly, the set $\{a \in \bar{S} \mid a\bar{S} = \bar{S}\}$ is closed, contains all the generic elements of $\bar{S}$ and therefore equals $\bar{S}$; thus $\bar{S}\bar{S} = \bar{S} = \bar{S}^{-1}$, which proves the result.

(1.17) Let G be an algebraic group defined over F, let $a \in G$ and b be a generic of G, independent from a over F. Then ab is also a generic of G, and is independent from a over F.

Proof. This is a simple argument using transcendence degrees. Let $n = dim(G)$. Then $F(a, b) = F(a, ab)$ has transcendence degree n over $F(a)$. Hence

$$n = tr.deg(F(a, ab)/F(a)) \leq tr.deg(F(ab)/F) \leq n,$$

which shows that ab is a generic of G, independent from a over F.

2. Preliminaries on finite and pseudo-finite fields

Definitions. (1) A field F is pseudo-algebraically closed (abbreviated by PAC) if every affine variety defined over F has an F-rational point.

(2) A field F is pseudo-finite if it is PAC, perfect, and has precisely one algebraic extension of degree n for every $n \in \mathbb{N}$.

Let m, n, d be positive integers; there exists an integer e such that, for any field F, if $f_1, \ldots, f_m, f$ are polynomials in n variables over F of total degree $\leq d$ then:

(1) if $f \in I = (f_1, \ldots, f_m)$, then $f = \sum_{i=1}^{m} g_i f_i$ for polynomials g_i of total degree $\leq e$.

(2) if I is not prime, then there are some polynomials g and h of total degree $\leq e$ such that $gh \in I$ but $g, h \notin I$.

The proof can be found e.g. in [4]; from this it follows that "$f_1, \ldots, f_m$ generate a prime ideal in $F[X_1, \ldots, X_n]$" is a first-order property of the coefficients of $f_1, \ldots, f_m$. Moreover, since T_{acf} eliminates quantifiers, there is a quantifier-free formula which defines in all fields F (the coefficients of) the polynomials $f_1, \ldots, f_m$ in n variables and of total degree $\leq d$ whose zero-set is a variety (i.e., such that $f_1, \ldots, f_m$ generate a prime ideal in $\tilde{F}[X_1, \ldots, X_n]$). Thus one can talk about varieties in a first-order way.

Observe also that the statement "F has one extension of degree n" can be formulated by translating in a first-order way: there are $c_1, \ldots, c_n$ such that $f(X) = X^n + c_1 X^{n-1} + \cdots + c_n$ is irreducible, and for all $d_1, \ldots, d_n$ such that $g(X) = X^n + d_1 X^{n-1} + \cdots + d_n$ is irreducible, the field obtained by adjoining to F a root of $f(X)$ contains a root of $g(X)$.

From these we deduce that being pseudo-finite is a first-order property in the language $\mathcal{L}$, and we denote by Psf the theory of all pseudo-finite fields. It is immediate that every finite field is perfect and has one extension of degree n for each $n \in \mathbb{N}$ (the unique extension of $\mathbb{F}_q$ of degree n is $\mathbb{F}_{q^n}$); it also follows from the Lang-Weil theorem that every non-principal ultraproduct of finite fields is PAC. J. Ax [1] showed that pseudo-finite fields are precisely the infinite models of the theory T_f of all finite fields. The fact that every infinite model of the theory of finite fields is a model of Psf follows easily from the Lang-Weil estimates on the number of points in finite fields of varieties; the reverse direction is given by (2.5).

We list below the main properties of the theory Psf; the proofs can be found in [1], [3] and [7]. Let F, F_1 and F_2 be pseudo-finite fields.

(2.1) Let E be a subfield of F_1 and F_2. Then

$$F_1 \equiv_E F_2 \iff (F_1 \cap \tilde{E}) \simeq_E (F_2 \cap \tilde{E}).$$

(2.2) Taking for E the prime field, one obtains invariants for the elementary theories of pseudo-finite fields:

$$F_1 \equiv F_2 \iff Abs(F_1) \simeq Abs(F_2),$$

where $Abs(F_1)$ is the subfield of F_1 of elements algebraic over the prime field.

(2.3) Assume that $F_1 \subseteq F_2$; then, taking $E = F_1$:

$$F_1 \prec F_2 \iff \tilde{F}_1 \cap F_2 = F_1.$$

(2.4) Another application of (2.1) is the following: let E be a subfield of F, and $a, b \in F$; then $tp(a/E) = tp(b/E)$ if and only if there is an E-isomorphism f between $(\widetilde{E(a)} \cap F)$ and $(\widetilde{E(b)} \cap F)$ which sends a to b.

From this one then deduces: let $\varphi(x)$ be a formula (x a tuple of variables); there is a formula $\psi(x)$, boolean combination of sentences of the form $(\exists t\ f(x,t) = 0)$, where $f(x,t) \in \mathbb{Z}[x,t]$, t a single variable, such that

$$\text{Psf} \vdash \varphi(x) \leftrightarrow \psi(x).$$

(2.5) Let E be a perfect field, and assume that E has at most one algebraic extension of each degree. Then there is a field F isomorphic to an ultraproduct of finite fields, such that

$$F \cap \tilde{E} = E.$$

Moreover if E is of characteristic 0, F can be chosen isomorphic to an ultraproduct of prime fields $\mathbb{F}_p$.

This shows that Psf is precisely the theory of all infinite models of T_f, and that the pseudo-finite fields of characteristic 0 are exactly the infinite models of $Th(\mathbb{F}_p \mid p \text{ a prime })$.

(2.6) As an illustration of techniques of proofs, we will show that the algebraic-geometric and model-theoretic notions of algebraic closure coincide:
Let E be a subfield of the pseudo-finite field F, relatively algebraically closed inside F, and let $a \in F$, $a \notin E$; then $tp(a/E)$ is not algebraic.

Proof. Choose a field F' isomorphic to F over E, and linearly disjoint from F over E; because F and F' are linearly disjoint over E, the ring $\tilde{F} \otimes_{\tilde{E}} \tilde{F}'$ is an integral domain; choose (topological) generators σ of $Aut(\tilde{F}/F)$ and σ' of $Aut(\tilde{F}'/F')$ such that σ and σ' have the same restriction to $\tilde{E}$; define τ on the quotient field M of $\tilde{F} \otimes_{\tilde{E}} \tilde{F}'$ by setting $\tau(b \otimes c) = \sigma(b) \otimes \sigma'(c)$ for $b \in \tilde{F}$, $c \in \tilde{F}'$, and extending in

the obvious manner. Lift τ to an automorphism τ_1 of $\tilde{M}$, and let $M_1 \subseteq \tilde{M}$ be the subfield of $\tilde{M}$ fixed by τ_1; then $Aut(\tilde{M}_1/M_1)$ is by definition generated by τ_1, and F and F' are relatively algebraically closed in M_1, since τ_1 extends σ and σ'. By (2.5), there is a pseudo-finite field L containing M_1 and such that M_1 is relatively algebraically closed in L. By (2.3), L is an elementary extension of both F and F', and therefore contains a realisation of $tp(a/E)$ not in F. Hence $tp(a/E)$ is not algebraic.

(2.7) We now give a sharper description of definable sets.

Let $\varphi(x,y)$ be a formula, $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_n)$, let $a \in F^m$ and let $S = \varphi(a,F^n) = \{b \in F^n \mid F \models \varphi(a,b)\}$; there is a positive integer e, an algebraic set V defined over F, and a projection map (on the first n coordinates) π from $V(F)$ onto S, with fibers $\pi^{-1}(y)$ of size $\leq e$ for $y \in S$.

(2.8) Using the Lang-Weil estimates on the number of rational points of varieties in finite fields, (2.7), and some counting arguments, one then obtains similar estimates for definable subsets of finite fields:

Let $\varphi(x,y)$ be a formula in $\mathcal{L}$, with $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_n)$. There is a positive constant C, and a finite set D of pairs (d,μ) with $d \in \{0,1,\ldots,n\}$ and μ a positive rational number, or $(d,\mu) = (0,0)$, such that for each finite field $\mathbb{F}_q$ and tuple $a \in \mathbb{F}_q^m$,

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

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

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

Let a be an m-tuple from the pseudo-finite field F. Then there is a unique pair $(d,\mu) \in D$ such that $F \models \varphi_{(d,\mu)}(a)$; one verifies that $d = dim(\varphi(a,F^n))$. The number μ can be used to define a measure m_S on the definable subsets of $S = \varphi(a,F^n)$: for T a definable subset of S with associated pair (e,ν), define:

$$m_S(T) = \begin{cases} 0 & \text{if } e < d, \\ \nu/\mu & \text{if } e = d. \end{cases}$$

Since the definition of m_S originates from counting points in finite sets, m_S is clearly a finitely additive probability measure, defined on the definable subsets of S, taking only rational values, and invariant under definable bijection.

From these considerations, one obtains easily the following results:

(2.9) (Finiteness of the S_1-rank) Let $S \subseteq F^n$ be definable, with $dim(S) = d$; let $\varphi(x,y)$ be a formula, and $(a_i)_{i\in I}$ a sequence such that $dim(S \cap \varphi(a_i,F^n)) = d$ for all i, and $dim(S \cap \varphi(a_i,F^n) \cap \varphi(a_j,F^n)) < d$ for all $i \neq j$; then I is finite.

Proof. By (2.8) there is a positive rational number $r \leq 1$ such that for any $a \in F^m$, $m_S(\varphi(a,F^n)) > 0$ implies $m_S(\varphi(a,F^n)) > r$; from the additivity of m_S, we obtain that I has less than $(1/r)$ elements.

(2.10) We will denote the second coordinate of the pair associated to a definable set S by $\mu(S)$; then one has:

(a) For a variety V defined over F, $\mu(V(F)) = 1$ (This is the Lang-Weil Theorem).

(b) For disjoint definable subsets S and T of F^n,

$$\mu(S \cup T) = \begin{cases} \mu(S) + \mu(T) & \text{if } dim(S) = dim(T), \\ \mu(S) & \text{if } dim(S) > dim(T), \\ \mu(T) & \text{if } dim(S) < dim(T). \end{cases}$$

(c) If $f : S \to T$ is definable, and $dim(f^{-1}(a)) = d$ for each $a \in T$, then $dim(S) = dim(T) + d$. If moreover $\mu(f^{-1}(a)) = m$ for every $a \in T$, then $\mu(S) = m\mu(T)$.

(2.11) (not the strict order-property) Let $\varphi(x, y)$ be a formula; then every sequence of tuples $a_i \in F$ such that the sets $\varphi(a_i, F^n)$ form a strictly increasing chain, is of bounded length.

We should mention that pseudo-finite fields are unstable: indeed Duret showed they have the independence property [5].

(2.12) Adjoin to the language new constant symbols $c_{i,n}$ for $0 \leq i < n \in \mathbb{N}$ to obtain the language $\mathcal{L}_c$, and consider the extension Psf_c obtained by adding to Psf axioms expressing that the polynomials $X^n + c_{n-1,n}X^{n-1} + \cdots + c_{0,n}$ are irreducible for each n. Every pseudo-finite field then expands (non-uniquely) to a model of Psf_c; also, Psf_c is model-complete, since whenever $(F_1, c) \subseteq (F_2, c)$ are models of Psf_c then F_1 is relatively algebraically closed in F_2.

F admits elimination of imaginaries in the language $\mathcal{L}_c$. Thus every group G interpretable in F is F-definably isomorphic to a group defined in F.

(2.13) Let G be a group definable in F; then there is a connected algebraic group H defined over F, definable subgroups of finite index G_0 of G and H_0 of $H(F)$, and a surjective isomorphism $f : G_0 \to H_0$, defined over F and with finite central kernel.

(2.14) If V is a variety defined over F, then the set $V(F)$ is Zariski dense in V, that is, its Zariski closure equals V (actually, this holds in arbitrary PAC-fields).

Indeed, for $0 \neq g(x) \in F[V]$, the algebraic set $V' = \{(a, b) \mid a \in V,\ bg(a) = 1\}$ is clearly a variety, so it has an F-rational point. This shows that $V(F)$ intersects every open set of K^n defined over F. If W is a proper closed subset of V defined over $\tilde{F}$, then the union of the conjugates of W over F is defined over F, which shows that $V(F)$ is dense in $V(\tilde{F})$. By (1.2) and because $V(\tilde{F})$ is dense in V, we deduce that $V(F)$ is dense in V.

3. Definable subgroups of algebraic groups over finite and pseudo-finite fields

Let F be a pseudo-finite field, contained in the algebraically closed field K. We denote by $qfdcl$ the definable closure in the sense of K; thus $qfdcl(A)$ is the perfect hull of the field generated by A.

When speaking of types, we mean types in the sense of T_{acf}, that is, quantifier-free types. We denote them by $qftp$ to avoid confusion. Independence is also meant in the sense of T_{acf}, that is, two tuples are independent over a set A if they are algebraically independent over A.

Theorem. Let G be a connected algebraic group defined over F. Let G_1 be a proper subgroup of finite index of $G(F)$ which is definable in F. Then there exists a connected algebraic group H defined over F and a surjective F-rational homomorphism $g : H \to G$ with non-trivial finite central kernel, such that $g(H(F))$ is a subgroup of finite index of G_1.

Proof. We may assume that F is uncountable and saturated, and that F_0 is a countable elementary substructure of F containing the parameters necessary to define G and G_1. Observe that F contains generic points of G_1, and that a generic point of G_1 is a generic point of G (since $dim(G) = dim(G_1) =_{\text{def}} n$).

Step 1.

By (2.7), there is a variety V of dimension n defined over F_0, a finite-to-one projection π defined on V, such that $\pi(V(F)) \subseteq G_1$.

Choose b and c in $\pi(V(F))$, independent and generic (over F_0), and let $\hat{b}, \hat{c} \in V(F)$ be such that $\pi(\hat{b}) = b$ and $\pi(\hat{c}) = c$. Let $a = cb^{-1}$; then a is a generic of G_1, but is not necessarily in $\pi(V(F))$.

Let $F_1 = \widetilde{F_0(a)} \cap F$; it is relatively algebraically closed in F; hence $qftp(\hat{b}, \hat{c}/F_1)$ is stationary, because $\hat{b}$ and $\hat{c}$ are in F; let $a^* \in F_1$ be such that $F_0(a^*) = F_0(\hat{b}, \hat{c}) \cap F_1$; by (1.10), $Cb(\hat{b}, \hat{c}/F_1)$ equals (the perfect hull of) $F_0(a^*)$, and $qftp(\hat{b}, \hat{c}/F_0(a^*))$ is stationary. Moreover, $a \in F_0(a^*)$.

Let b^* and c^* be defined by:

$$\begin{aligned} F_0(b^*) &= F_0(\hat{b}) \cap F_0(a^*, \hat{c}) \\ F_0(c^*) &= F_0(\hat{c}) \cap F_0(a^*, b^*). \end{aligned}$$

Then $b \in F_0(b^*)$ and $c \in F_0(c^*)$; furthermore $F_0(b^*) = Cb(a^*, \hat{c}/F_0(\hat{b}))$ and $F_0(c^*) = Cb(a^*, b^*/F_0(\hat{c}))$; from the first equality and the fact that $a^* \in F_0(\hat{b}, \hat{c})$, it follows that $a^* \in F_0(b^*, \hat{c})$; similarly, $a^* \in F_0(b^*, \hat{c})$ and $b^* \in F_0(a^*, \hat{c})$ imply that $a^* \in F_0(b^*, c^*)$ and $b^* \in F_0(a^*, c^*)$.

Let $A = (a^*, a)$, $B = (b^*, b)$ and $C = (c^*, c)$. Summarising the situation, we have:

(i) A, B, C are in F, pairwise independent over F_0, and each of them is quantifier-free definable over F_0 and the other two.

(ii) $a \in F_0(A)$, $b \in F_0(B)$, $c \in F_0(C)$.

(iii) $A \in \widetilde{F_0(a)}$, $B \in \widetilde{F_0(b)}$, $C \in \widetilde{F_0(c)}$.

Step 2. Defining a new group.

We take $A' = (a'^*, a') \in F$ realising $qftp(A/F_0)$ and independent from A, B, C over F_0. We then take $B' = (b'^*, b') \in qfdcl(F_0(A', C))$ such that $qftp(A', B'/F_0(C)) = qftp(A, B/F_0(C))$, and let F_2 be an elementary substructure of F containing $F_0(A')$ and independent from A, B, C over F_0. Then

(1) $B \in qfdcl(F_2(A, B'))$, $B' \in qfdcl(F_2(A, B))$ and $A \in qfdcl(F_2(B, B'))$.

This is because of property (i) above. Note also that A, B and B' are pairwise independent over F_2, and that $ab = a'b' = c$.

We are now ready to apply the machinery of germs of functions. Working in the algebraically closed field K, let P be the set of realisations of $qftp(A/F_2)$ and Q the set of realisations of $qftp(B/F_2)$. Their Zariski closures $\bar{P}$ and $\bar{Q}$ are varieties, because the types of A and B over F_2 are stationary.

Let f, g be two partial definable functions $Q \to Q$; we say that f and g have the same germ if they are defined and equal on a Zariski open subset of Q. By the uniqueness of the non-forking extension of $qftp(B/F_2)$ (because $B \in F$), this happens if and only if: for some $B_1 \in Q$ (or equivalently, for all $B_1 \in Q$), independent over F_2 from the parameters used to define f and g, the tuples $f(B_1)$ and $g(B_1)$ are defined and equal.

Note that we could as well speak of germs of functions $\bar{Q} \to \bar{Q}$, since germs are defined in terms of open subsets of $\bar{Q}$.

To each $A_1 \in P$ we associate a partial function $f_{A_1} : Q \to Q$ as follows: if $B_1 \in Q$ is generic (over $F_2(A_1)$), let $f_{A_1}(B_1)$ be the unique element B_2 such that

$$qftp(A_1, B_1, B_2/F_2) = qftp(A, B, B'/F_2).$$

By (1), the function f_{A_1} is generically invertible. Moreover, given $B_1, B_2 \in Q$ independent over F_2, there is a unique $A_1 \in P$ such that $f_{A_1}(B_1) = B_2$, and this tuple A_1 is independent from B_1 and from B_2 over F_2.

Let $A_1, A_2 \in P$ be independent over F_2; using (1) one checks that for any $B_1 \in Q$ independent from A_1, A_2 over F_2, the tuple $B_2 = f_{A_2} \circ f_{A_1}(B_1)$ is well-defined and in Q. Applying (1) again and using the fact that $tr.deg(A/F_2) = tr.deg(B/F_2) = n$, one sees that each of A_1, A_2, B_1, B_2 is quantifier-freely definable over F_2 and the other three, and therefore that any three tuples from $\{A_1, A_2, B_1, B_2\}$ are independent over F_2.

In particular, B_1 and B_2 are independent over F_2, which implies that there is a unique $A_3 \in P$ such that $f_{A_3}(B_1) = B_2$. Write $B_1 = (b_1^*, b_1)$ and $A_i = (a_i^*, a_i)$ for $i = 1, 2, 3$. From the definition of f_A and the fact that $b' = a'^{-1}ab$ we obtain

$$a'^{-1}a_3b_1 = a'^{-1}a_2a'^{-1}a_1b_1,$$

which shows that $a_3 \in F_2(a_1, a_2)$. Since $A_3 \in \widetilde{F_2(a_3)}$, we deduce that A_3 is algebraic over $F_2(A_1, A_2)$, and therefore that (A_1, A_2, A_3) is independent from B_1

over F_2. Hence $f_{A_3} = f_{A_2} \circ f_{A_1}$, and also $A_3 \in qfdcl(F_2(A_1, A_2))$ (because $A_3 \in qfdcl(A_1, A_2, B_1)$ and $qftp(B_1/F_2)$ is stationary).

Hence we have an F_2-definable partial map $m : P \times P \to P$ which associates to the pair (A_2, A_1) the unique $A_3 \in P$ such that $f_{A_2} \circ f_{A_1} = f_{A_3}$. In the same manner, there is an F_2-definable map $i : P \to P$ which associates to A_1 the unique element A_2 such that $f_{A_2} \circ f_{A_1}$ is the germ of the identity function. We have, for $A_1, A_2 \in P$ independent over F_2:

$$(2) \qquad f_{m(A_2,A_1)} = f_{A_2} \circ f_{A_1} \quad \text{and} \quad f_{i(A_1)} \circ f_{A_1} = f_{A_1} \circ f_{i(A_1)} = id.$$

Observe that each of $A_1, A_2, m(A_1, A_2)$ is definable over F_2 and the other two. From the definition of m and the associativity of composition of germs also follows that for $A_1, A_2, A_3 \in P$ independent over F_2 one has:

$$(3) \qquad m(A_3, m(A_2, A_1)) = m(m(A_3, A_2), A_1).$$

We have a generically defined operation $m : P \times P \to P$, which is generically associative. From this data, a standard argument (detailed below) shows the existence of a definable group.

Let $\psi(x, y, z)$ be a formula in $qftp(A, B, B'/F_2)$ expressing the various relations given by (1), together with $A \in \bar{P}$, and $B, B' \in \bar{Q}$. Since $qftp(B/F_2)$ is definable over F_2, there is a formula $\theta(x) \in qftp(A/F_2)$, such that whenever $K \models \theta(A_1)$ and $B_1 \in Q$ is independent from A_1 over F_2, then:

There is a unique B_2 satisfying $\psi(A_1, B_1, z)$. Furthermore, B_1 is the unique tuple satisfying $\psi(A_1, y, B_2)$ and A_1 is the unique tuple satisfying $\psi(x, B_1, B_2)$.

This is first-order expressible, and precisely says that the formula $\psi(A_1, y, z)$ defines (the graph) of the germ of an invertible function $Q \to Q$, and that A_1 is uniquely defined by this germ. We will denote this function by f_{A_1}. Reasoning in the same manner with equation (2), there is a formula $\varphi(x) \in qftp(A/F_2)$ implying $\theta(x)$, and such that whenever A_1 satisfies $\varphi(x)$ and $A_2 \in P$ is independent from A_1 over F_2 then:

The tuples $m(A_1, A_2)$, $m(A_2, A_1)$ and $i(A_1)$ are defined, satisfy $\theta(x)$ and the identities given by (2). Each of $A_1, A_2, m(A_1, A_2)$ is definable over F_2 and the other two, and similarly for the triple $A_1, A_2, m(A_2, A_1)$.

Observe that this last condition implies in particular that $m(A_1, A_2)$ and $m(A_2, A_1)$ are in P and independent from A_1 over F_2, since $tr.deg(F_2(A_1, m(A_1, A_2))/F_2(A_1)) = tr.deg(F_2(A_1, m(A_2, A_1))/F_2(A_1)) = n$.

Because the set of elements satisfying $\varphi(x)$ contains P, it contains an open subset U_0 of $\bar{P}$, and we may assume that $i(U_0) = U_0$ (if necessary, replace U_0 by $i(U_0) \cap U_0$). For $A_1 \in U_0$, and $A_2 \in P$ independent from A_1 over F_2 we then have:

(a) the formula $\psi(A_1, y, z)$ defines (the graph of) the germ of a generically invertible function f_{A_1} from Q to Q.

(b) $m(A_1, A_2)$, $m(A_2, A_1)$, $i(A_1)$ are defined, are in U_0, and satisfy equation (2).

(c) $m(A_1, A_2), m(A_2, A_1) \in P$, and each of them is equi-definable with A_2 over $F_2(A_1)$ (in particular, they are independent from A_1 over F_2).

Define

$$G' = \{f_{A_1} \circ f_{A_2} \mid A_1, A_2 \in U_0\}.$$

Claim 1. If $A_1 \in P$, then $f_{A_1} \in G'$.

Proof. Let $A_2 \in P$ be independent from A_1 over F_2, and let $A_3 = m(i(A_2), A_1)$. Then $A_3 \in P$, and by (2), $f_{A_1} = f_{A_2} \circ f_{A_3}$.

Claim 2. If $A_3 \in P$ is independent from (A_1, A_2) over F_2, then $m(m(A_3, A_2), A_1)$ is defined and in P.

Proof. By (c), $A_4 = m(A_3, A_2)$ is in P and independent from A_1 over F_2. By (b), $f_{A_4} = f_{A_3} \circ f_{A_2}$. Applying (b) and (c) again, we have:

$$f_{A_3} \circ f_{A_2} \circ f_{A_1} = f_{m(A_4, A_1)} \qquad \text{and} \qquad m(A_4, A_1) \in P.$$

Claim 3. $(G', \circ)$ is a group.

Proof. By (b), G' is closed under inverses. It suffices to show that if A_1, A_2, $A_3 \in U_0$, then $f_{A_1} \circ f_{A_2} \circ f_{A_3} \in G'$. Choose $A_4 \in P$, independent from (A_1, A_2, A_3) over F_2. By (c) and claim 2, the tuples $m(A_1, i(A_4))$ and $m(m(A_4, A_2), A_3)$ are in P, and by (b),

$$\begin{aligned} f_{A_1} \circ f_{A_2} \circ f_{A_3} &= (f_{A_1} \circ f_{A_4}^{-1}) \circ (f_{A_4} \circ f_{A_2} \circ f_{A_3}) \\ &= f_{m(A_1, i(A_4))} \circ f_{m(m(A_4, A_2), A_3)}. \end{aligned}$$

Claim 4. Every element of G' is the product of two elements of P (By abuse of notation we identify P with its image in G' via $(A_1 \mapsto f_{A_1})$).

Proof. Let $A_1, A_2 \in U_0$, and choose $A_3 \in P$ independent from (A_1, A_2) over F_2. Then $m(A_1, i(A_3)) \in P$, $m(A_3, A_2) \in P$ and

$$f_{A_1} \circ f_{A_2} = f_{m(A_1, i(A_3))} \circ f_{m(A_3, A_2)}.$$

Using once more the definability of $qftp(B/F_2)$, one shows that the sets

$$\{(A_1, A_2, A_3, A_4) \in U_0^4 \mid f_{A_1} \circ f_{A_2} = f_{A_3} \circ f_{A_4}\}$$

and

$$\{(A_1, A_2, A_3, A_4, A_5, A_6) \in U_0^6 \mid f_{A_1} \circ f_{A_2} \circ f_{A_3} \circ f_{A_4} = f_{A_5} \circ f_{A_6}\}$$

are definable. Hence the group $(G', \circ)$ is interpretable in K (with parameters from F_2). Since $\bar{P} = \bar{U}_0$ is a variety and contains all the generics of G' (by claim 4, G' is generated by the elements of P), the group G' is connected.

Step 3. Finding the group H.

We now quote a result by Hrushovski, which allows us to find an algebraic group:

Theorem. Let G be a connected group interpretable in an algebraically closed field K. Then there is a connected algebraic group H and a definable isomorphism

$f : G \to H$. Furthermore, if G is definable over the subfield k of K, then so are H and f.

For a proof, see [10] or [11].

Applying this to our situation, and noting that there is a definable injection $P \to G'$, we obtain a connected algebraic group H of dimension n, defined over F_2, and a definable injection $f : P \to H$, such that for $A_1, A_2 \in P$ independent over F_2, $f(m(A_2, A_1)) = f(A_2) \cdot f(A_1)$ and $f(i(A_1)) = f(A_1)^{-1}$.

If $A_1 \in P$, then $A_1 \in F_2(f(A_1)^{1/q})$ for some p-th power q; replacing H by the definably isomorphic algebraic group $\{(x_1^{1/q}, \ldots, x_r^{1/q}) \mid (x_1, \ldots, x_r) \in H\}$, we may assume that

(4) $$\text{if } A_1 \in P, \text{ then } A_1 \in F_2(f(A_1)).$$

Step 4. Definition of the morphism g.

By definition, f_A acts on the second coordinate of $B = (b^*, b)$ like multiplication by $a'^{-1}a$ (inside G). We now look at the set

$$S = \{(f(A_1), a'^{-1}a_1) \mid A_1 \in P\} \subseteq H \times G,$$

where $A_1 = (a_1^*, a_1)$. Then S is irreducible, because $qftp(f(A), a/F_2)$ is stationary. Since it projects onto $f(P) = f(P)^{-1}$, it contains all the generics of $\bar{S}$ and their inverses, and is closed under generic multiplication. By (1.16), $\bar{S}$ is a subgroup of $H \times G$. Since every element of G' is the product of two elements of P, every element of H is the product of two elements of $f(P)$, and therefore $\bar{S}$ projects onto H. From $a_1 \in F_2(f(A_1))$, we deduce that $\bar{S}$ is the graph of a group morphism $g : H \to G$. Because A, and therefore also $f(A)$, is algebraic over $F_2(a)$, the morphism g is finite-to-one. Since A is a generic of G', and g is everywhere defined, g is a morphism of algebraic groups. Also, because a is a generic point of G and G is connected, the morphism g is onto. Since g is finite-to-one, $Ker(g)$ is a finite normal subgroup of H, and therefore central because H is connected.

Step 5. $g(H(F))$ is a subgroup of finite index of G_1.

Since $dim(H) = dim(G)$, (2.9) implies that $[G(F) : g(H(F))]$ is finite. It therefore suffices to show that $g(H(F)) \subseteq G_1$. As every element of $H(F)$ is the product of two generic elements, it is enough to show that the image by g of a generic element of $H(F)$ is in G_1.

Let $e \in H(F)$ be generic. By (4), $e = f(A_1)$ where $A_1 = (a_1^*, a_1) \in F_2(e) \subseteq F$ realises $qftp(A/F_2)$. Since $qftp(B, C/F_2(A))$ is stationary, there are tuples $B_1 = (b_1^*, b_1)$ and $C_1 = (c_1^*, c_1)$ in F such that

$$qftp(A_1, B_1, C_1/F_2) = qftp(A, B, C/F_2).$$

Then b_1^* and c_1^* are in $V(F)$, which implies that b_1 and c_1 are in G_1. Hence $a_1 = c_1 b_1^{-1} \in G_1$, and $g(e) = a'^{-1}a_1 \in G_1$.

Step 6. $Ker(g)$ is non-trivial.

Assume by way of contradiction that g is injective. Then $g^{-1}(a) \in qfdcl(F_2(a))$ for each $a \in G$. Since F is perfect, this implies that $g(H(F)) = G(F)$, which contradicts the fact that G_1 is a proper subgroup of $G(F)$ containing $g(H(F))$.

We can obtain more information on $Ker(g)$ as follows: since G and H are connected, applying (2.10)(a) and (c) we obtain

$$|Ker(g) \cap H(F))|\, \mu(g(H(F))) = \mu(H(F)) = 1 =$$

$$= \mu(G(F)) = [G(F) : g(H(F))]\, \mu(g(H(F))),$$

from which we conclude

$$|Ker(g) \cap (G(F))| = [G(F) : g(H(F))] > 1.$$

Remark. We say that a field is bounded if it has finitely many extensions of degree n for each $n \in \mathbf{N}$. Pseudo-finite fields are bounded. The results stated in sections 2 and 3 are valid in perfect bounded PAC fields, with the exception of those involving the "counting measure" μ. For details, see [7] and [8].

4. Definability of maximal subgroups

Let G be an almost simple algebraic subgroup of GL_n defined over $\mathbf{Z}$. Reducing the equations defining G modulo p gives for almost all prime p, an almost simple algebraic subgroup G_p of GL_n. In this section, we will show that the maximal subgroups of $G_p(\mathbf{F}_p)$ are uniformly definable. For details on reduction modulo p, we refer to section 7 of [6].

Theorem. Let $G \subseteq GL_n$ be an almost simple algebraic group defined over $\mathbf{Z}$. Then there are formulas $\varphi_1(x,y), \ldots, \varphi_s(x,y)$ of $\mathcal{L}$ such that whenever p is a prime and M is a maximal subgroup of $G_p(\mathbf{F}_p)$, then M is definable in $\mathbf{F}_p$ by the formula $\varphi_i(x,b)$ for some $i \leq s$ and tuple b in $\mathbf{F}_p$.

Proof. Let $\varphi_j(x,y_j)$, $j \in \mathbf{N}$, be an enumeration of the $\mathcal{L}$-formulas. If the statement fails, we can find an infinite set $\{p(i) \mid i \in \mathbf{N}\}$ of primes, and for each i a maximal subgroup M_i of $G_{p(i)}(\mathbf{F}_{p(i)})$ such that M_i is not definable by any of the formulas $\varphi_j(x,b_j)$ for $j = 1, \ldots, i$ and b_j a tuple from $\mathbf{F}_{p(i)}$.

Consider a non-principal ultraproduct $(F, +, \cdot, M)$ of the structures $(\mathbf{F}_{p(i)}, +, \cdot, M_i)$; then F is a pseudo-finite field of characteristic 0, and M is a subgroup of $G(F)$, which is contained in no definable proper subgroup of $G(F)$. We will now show that M is definable, which will give a contradiction, and prove our result.

Case 1. M contains a non-trivial unipotent element u.

See section 1 of [6] for definitions and related results. The group $U = \{exp(t \log(u)) \mid t \in F\}$ is a definable, Zariski-irreducible (in F) subgroup of $G(F)$. For almost all i, the i-th coordinate u_i of u is a unipotent element of $M_i \subseteq G_{p(i)}(\mathbf{F}_{p(i)})$. For these i's, $\{exp(t \log(u_i)) \mid t \in \mathbf{F}_{p(i)}\}$ is the subgroup of $G_{p(i)}(\mathbf{F}_{p(i)})$ generated by u_i and is therefore contained in M_i. This implies that U is contained in M.

By the irreducibility theorem (Theorem 24 in [6]) the subgroup H of $G(F)$ generated by the subgroups U^g, $g \in M$, is definable and is normal in M; then the normaliser H_1 of H in $G(F)$ is definable and contains M.

If $H_1 = M$ then M is definable. If $H_1 = G(F)$, then H is normal in $G(F)$ which, since it is infinite and G is almost simple, implies that H has finite index in $G(F)$. As $H \subseteq M \subseteq G(F)$, M is also definable.

Case 2. M contains no non-trivial unipotent element.

Then M_i contains no non-trivial unipotent element for almost all $i \in \mathbb{N}$. Hence each M_i contains an abelian subgroup of index at most d for some d (Facts 38 and 39 of [6]), which implies that M has a normal abelian subgroup A of finite index. Let $B = Z(C_{G(F)}(A))$ be the center of the centralizer in $G(F)$ of A. Since algebraic groups satisfy the descending chain condition on centralizers, $C_G(A) = C_G(A_0)$ for some finite subset A_0 of A. Hence $C_{G(F)}(A) = C_G(A_0) \cap G(F)$ and B are definable in $G(F)$.

Also, $N_{G(F)}(A)$, the normalizer of A in $G(F)$, contains M and is contained in $H = N_{G(F)}(B)$: for $g \in G(F)$, $g^{-1}Bg = Z(C_{G(F)}(g^{-1}Ag))$.

If $H = G(F)$, then B is an infinite normal abelian subgroup of $G(F)$. Because G is almost simple, this implies that $\bar{B} = G$, and therefore that G is abelian, a contradiction. Hence $H = M$ is definable.

References

[1] J. Ax, The elementary theory of finite fields, Annals of Math. 88 (1968), 239 – 271.

[2] E. Bouscaren, The group configuration – after E. Hrushovski, in: The model theory of groups, Nesin-Pillay ed., Notre Dame Math. Lect. 11, Notre Dame (1989).

[3] Z. Chatzidakis, L. van den Dries, A. Macintyre, Definable sets over finite fields, J. reine u. ang. Math. 427 (1992), 107 – 135.

[4] L. van den Dries, K. Schmidt, Bounds in the theory of polynomial rings over fields. A non-standard approach, Invent. Math. 76 (1984), 77 –91.

[5] J.-L. Duret, Les corps pseudo-algébriquement clos non séparablement clos ont la propriété d'indépendance, in: Model theory of algebra and arithmetic, Proc. Karpacz 1979, Springer LN 834 (1980), 136 – 161.

[6] W. Hodges, Groups in pseudo-finite fields, this volume.

[7] E. Hrushovski, A. Pillay, Groups definable in local and pseudo-finite fields, Israel J. of Math 85 (1994), 203 – 262.

[8] E. Hrushovski, A. Pillay, Definable subgroups of algebraic groups over finite fields, J. reine angew. Math 462 (1995), 69 –91.

[9] S. Lang, Introduction to algebraic geometry, Addison-Wesley, Reading 1973.

[10] A. Pillay, Model theory of algebraically closed fields, in: Stability theory and algebraic geometry, an introduction, Bouscaren and Lascar ed., to appear in Springer LN.

[11] A. Pillay, Geometric Stability, Oxford University Press, Oxford 1996.

Author's address:
UFR de Mathématiques, Case 7012,
Université Paris 7,
2 Place Jussieu,
75251 Paris, Cédex 05,
France.
e-mail: zoe@logique.jussieu.fr

Groups in pseudofinite fields

Wilfrid Hodges

These notes should be read with those of Zoé Chatzidakis [2]. We report some results from Hrushovski and Pillay [7]. The main items in this paper are

- an analogue for pseudofinite fields of Zil'ber's Irreducibility Theorem (Theorem 23);
- lemmas relating simplicity properties of an algebraic group G to its restriction $G(F)$ to a pseudofinite field F (§6);
- a fast though non-effective model-theoretic proof of a result of Matthews, Vaserstein and Weispfeiler on reduction at primes (Theorem 33; see also the similar argument in the last section of [2]).

The main difference from [7] is that I avoid the local stability arguments of Hrushovski and Pillay [6]. In fact the proof of the Irreducibility Theorem removes all the stability arguments beyond the '(S1) property', without adding anything in their place. (Later I quote the Theorem of [2] §3, whose proof—at least in its present guise—uses forking, canonical bases and the definability of types.) It took several shoves to remove the parts of the argument that rely on stability; Frank Wagner and John Wilson delivered the final push during the Blaubeuren meeting. I think a fair comment would be that stability theory has powerful methods for showing that things are first-order definable, and this was the role that it played in the original argument. But sometimes, after the event, one sees that other devices may do the job faster. Abraham Robinson was exploring first-order definability in algebraic geometry at the time of his death, before stability methods were widely available (see for example [9]); one can see this work of Hrushovski and Pillay as in some sense a natural continuation of Robinson's.

I thank Zoe Chatzidakis, David Evans and Martin Ziegler for several improvements, and in particular for rescuing me from some errors which would have been a lot more embarrassing in print than they were in front of a friendly audience.

1. An example: one-dimensional matrix groups
2. Algebraic groups
3. Irreducible sets and connected groups
4. Dimension in pseudofinite fields

5. Algebraic groups in pseudofinite fields
6. Almost simple groups
7. Reduction at a prime
8. A theorem on reduction

Throughout, L is the first-order language of rings, $L(X)$ is L with constants added for the elements of a set X, K is an algebraically closed field and F is a pseudofinite field which is a subfield of K.

Except where we say otherwise, 'definable' means definable by a first-order formula with parameters. If X is a subset of K^m which is definable by a first-order formula with parameters in a set Y, we say that X is *definable over* Y. Note that by elimination of quantifiers for algebraically closed fields, the defining formula $\phi(\bar{x})$ can always be chosen to be quantifier-free. When M is a subfield (or even a subring) of K, we write $X(M)$ for the set definable in M by the same quantifier-free formula ϕ; clearly $X(M) = X \cap M^m$. Following [2] (1,4), I have tried to restrict the phrase 'defined over F' to varieties whose ideal of definition is generated by polynomials over F, so that the terminology of algebraic geometers applies.

1 An example: one-dimensional matrix groups

Let n be a positive integer. Then we can write $\mathrm{End}_n(K)$ for the set of $n \times n$ matrices over the algebraically closed field K. This set is in effect the whole of K^{n^2}, so it is a K-closed set, i.e. a closed set in the K-Zariski topology.

Addition of matrices is given by a family of n^2 polynomials s_{ij} with integer coefficients, so that if M and N are matrices in $\mathrm{End}_n(K)$ then $s_{ij}(M, N)$ is the ij-th component of the matrix $M+N$; each s_{ij} is a polynomial in $2n^2$ indeterminates, one for each entry in M or N. Likewise there are polynomials p_{ij} defining the product MN. There is a polynomial 'det' defining the determinant; in fact there are polynomials c_i giving each of the coefficients of the characteristic polynomial, so that for example $\det(M) = (-1)^n c_n(M)$. Also there are polynomials adj_{ij} defining the adjugate matrix $\mathrm{adj}(M)$, so that $M^{-1} = \mathrm{adj}(M)/\det(M)$ when $\det(M) \neq 0$.

Thus we have morphisms

$$s : K^{2n^2} \to K^{n^2}, \quad p : K^{2n^2} \to K^{n^2},$$

$$c_i : K^{n^2} \to K, \quad adj : K^{n^2} \to K^{n^2}.$$

There is not a morphism taking M to M^{-1}, since inverse is not defined on the whole of $\mathrm{End}_n(K)$.

A matrix N is said to be *nilpotent* if $N^k = 0$ for some k. Clearly if this holds then the matrix ranks of $N, N^2, \ldots$ must fall until they reach 0, and so k can be chosen

to be at most n. So assuming that N is nilpotent (and that the characteristic p of K is either 0 or $\geq n$), we can define a further morphism $\exp : \mathrm{End}_n(K) \to \mathrm{End}_n(K)$ by:

$$\exp(N) = I + \frac{N}{1!} + \ldots + \frac{N^{n-1}}{(n-1)!}.$$

One can check that if M, N are nilpotent and commute with each other, then

$$\exp(M+N) = \exp(M).\exp(N).$$

This implies that $\exp(N).\exp(-N) = I$ and hence $\exp(N)$ is invertible. It also implies that if we fix N and define a morphism from K to $\mathrm{End}_n(K)$ by

$$a \mapsto \exp(aN),$$

then *this morphism is a homomorphism from the additive group K_+ of K to the multiplicative group $GL_n(K)$ of invertible matrices in $End_n(K)$.*

As a subset of $\mathrm{End}_n(K)$, $\mathrm{GL}_n(K)$ is an open set, since it is the complement of the set of matrices M with $\det(M) = 0$. But we normally read it another way, namely as a subset of K^{n^2+1}: there are n^2 entries for the matrix and one (say x_d) for $1/\det(M)$. Then $\mathrm{GL}_n(K)$ is a closed subset of K^{n^2+1}, defined by the equation

$$x_d(M).\det(M) = 1.$$

Multiplication is still a morphism on $\mathrm{GL}_n(M)$, since

$$1/\det(MN) = (1/\det(M))(1/\det(N)).$$

Also we can use the extra coordinate to define an inversion morphism $M \mapsto M^{-1}$ on $\mathrm{GL}_n(K)$.

A matrix $U \in \mathrm{GL}_n(K)$ is said to be *unipotent* if $U - I$ is nilpotent. Then since the Jordan normal form of $U - I$ has zeros down the main diagonal, $\det U = 1$. But $\exp(N)$ is unipotent when N is nilpotent, so we have confirmed that the map exp defined above is still a morphism when the extra coordinate $1/\det$ is added.

When U is unipotent, say $U - I = N$, we can define

$$\log(U) = (U-I) - \frac{(U-I)^2}{2} + \ldots + (-1)^n \frac{(U-I)^{n-1}}{n-1}.$$

(with the same assumption as before on the characteristic). Note that because $\log(U)$ is a polynomial in N with no constant term, it is a nilpotent matrix. Hence we can define a morphism α from K_+ to $\mathrm{GL}_n(K)$ by:

$$a \mapsto \exp(a \log U).$$

The image of α contains U, since one can check that

$$\exp \log(U) = U.$$

We shall show later (Lemma 7) that the image is a closed subgroup of $\mathrm{GL}_n(K)$; this fact is far from obvious. Also α is injective (unless $U = 1$), since we can recover a from the equation

$$\log \exp(a \log(U)) = a \log(U).$$

By quantifier elimination there is a quantifier-free formula $\delta(\bar{x})$ which defines in K^{n^2+1} the set of matrices which form the image of α. This image is a subgroup of $\mathrm{GL}_n(K)$ isomorphic to the additive group K_+.

Let M be any subfield of K. Then since exp is defined by polynomials, the restriction of δ to M is exactly the image of the additive group M_+ under α, and it defines a subgroup of $\mathrm{GL}_n(M)$.

For future reference we note an elementary fact.

Proposition 1 *Suppose the field K has prime characteristic p, and U is an element of $GL_n(K)$. Then U is unipotent if and only if it has order p^i for some i.*

PROOF. For left to right, put $U = I + N$ with N nilpotent. Then $N^n = 0$, so $N^{p^n} = 0$ and hence

$$U^{p^n} = (I + N)^{p^n} = I^{p^n} + N^{p^n} = I.$$

Conversely if $U^{p^i} = I$ then

$$(U - I)^{p^i} = U^{p^i} - I^{p^i} = 0.$$

□

2 Algebraic groups

Generalising the example above, we define an *affine algebraic group* (or for short, *affine group*) in the field K to be a (Zariski) closed subset G of some K^m, together with morphisms $\mu : G \times G \to G$ and $\iota : G \to G$ which form the multiplication and inverse operations of a group on G. The group is *defined over* a field $F \subseteq K$ if G, μ, ι are all defined over F.

Since μ is a morphism, translation $g \mapsto fg$ (for a fixed f) is a continuous map from G to G. In fact it is a homeomorphism, because its inverse is a translation $g \mapsto f^{-1}g$. The inversion map ι is a homeomorphism in the Zariski topology, since it is its own inverse.

We note various related notions:

- An *algebraic group* in K is a closed group which is built up by gluing together sets in K^m on which the group operations are defined by polynomials; see

(1.13–15) of [2] for a more precise account. I believe that the theorems below are true for algebraic groups in general (except those explicitly about linear groups), but the reader shouldn't take my word for this. In any case the applications will need only affine groups.

- A group G is said to be *definable in* a field M if the set G is definable in M with parameters, and the group operations are also definable in M with parameters. WARNING: A group *definable over* F and a group *definable in* F are quite different kinds of animal.

- If G is an algebraic group in K which is defined over F, for instance an affine group, then we write $G(F)$ for its restriction to elements lying in F^m. Since G and its operations are defined by polynomials, $G(F)$ is a subgroup of G and it is definable in F.

A subgroup of $\mathrm{GL}_n(K)$ for some n is called a *linear group* in K. We shall not need the following fact, though it may give some reassurance:

Fact 2 *An affine group in K is isomorphic (by a morphism) to a linear group in K. (See Borel [1] Proposition 1.10.)*

Proposition 3 *Let G be an affine group, and H a subgroup of G. Then the closure $\bar{H}$ in the Zariski topology is also an affine group. If H satisfies an identity, for example if H is abelian, then the same holds also for $\bar{H}$.*

PROOF. If $h \in H$ then the set $h^{-1}\bar{H}$ is closed since translations are homeomorphisms. Since $H \subseteq h^{-1}\bar{H}$, we deduce $\bar{H} \subseteq h^{-1}\bar{H}$, and translating back again gives $h\bar{H} \subseteq \bar{H}$. So $H\bar{H} = \bar{H}$. Then if $g \in \bar{H}$, we have $Hg \subseteq \bar{H}$. So $H \subseteq \bar{H}g^{-1}$, and hence $\bar{H} \subseteq \bar{H}g^{-1}$ since $\bar{H}g^{-1}$ is closed. Therefore $\bar{H}g \subseteq \bar{H}$, which proves that $\bar{H}$ is closed under multiplication. A similar argument shows that it is closed under inverse.

The last sentence is trivial, since $\bar{H}$ is precisely the set of all points which satisfy every equation true throughout H. □

Corollary 4 *Let G be an affine group in K and $H \subseteq J$ subgroups of G. If H has finite index k in J, then $\bar{H}$ has index $\leq k$ in $\bar{J}$.*

PROOF. If $j \in J$, then translating H to the coset jH takes $\bar{H}$ to $\overline{jH}$. The union of the closures of these cosets is a closed set containing J, so it contains $\bar{J}$. Thus $\bar{J}$ is contained in at most k cosets of $\bar{H}$. □

Lemma 5 *If X, Y are subsets of an affine group G, then $\bar{X}.\bar{Y} \subseteq \overline{XY}$.*

PROOF. If $x \in X$, then the map $y \mapsto xy$ on G is continuous and takes Y into XY, so that it takes $\bar{Y}$ into $\overline{XY}$. Thus multiplication takes $X \times \bar{Y}$ into $\overline{XY}$. Let y be any element of $\bar{Y}$. Then the map $x \mapsto xy$ on G is continuous and takes X into $\overline{XY}$, so that it also takes $\bar{X}$ into $\overline{XY}$. □

Lemma 6 *If W is a dense open subset of an affine group G, then $G = WW$.*

PROOF. Two dense open subsets of a closed set must meet. Take any $a \in G$, and consider $a.W^{-1}$. This is also a dense open subset of G, so it meets W; say $b = a.w_1^{-1} = w_2$. Then $a = w_2.w_1 \in WW$, so $WW = G$. □

Lemma 7 *If G is an affine group in K and H is a subgroup of G which is definable in K, then H is closed. In particular if F, G are affine groups in K and $\alpha : F \to G$ is a morphism, then the image of α is a closed subgroup of G.*

PROOF. By quantifier elimination H is constructible, and hence it contains a dense open subset of its closure $\bar{H}$. So by the preceding lemma, $\bar{H} = HH = H$. □

We write $Z(G)$ for the centre of the group G.

Proposition 8 *If G is an affine group in K and H is a subgroup which is dense in G, then $Z(H) = Z(G) \cap H$.*

PROOF. The inclusion from right to left is immediate. For left to right, clearly $Z(H) \subseteq H$, so if the inclusion fails, there must be $h \in Z(H) \setminus Z(G)$. Then there is $g \in G$ such that $[h, g] \neq 1$. Since the set $\{g \in G : [h, g] \neq 1\}$ is a nonempty open subset of G and H is dense in G, we find $g' \in H$ such that $[h, g'] \neq 1$, contradiction. □

3 Irreducible sets and connected groups

Recall from [2] (1.5) that a closed set X is *irreducible* if it is not the union of two proper closed subsets. If X is definable but not necessarily closed, we say X is *irreducible* when the Zariski closure of X in K is irreducible. A maximal irreducible closed subset of a closed set X is called a *connected component* of X; X has finitely many connected components, and is their union.

Lemma 9 *If X and Y are irreducible closed subsets of K^m, K^n respectively, then their cartesian product $X \times Y$ is an irreducible closed subset of K^{m+n}.*

PROOF. See for example Hartshorne [5] Exercise I.3.15. □

Lemma 10 *If X is irreducible and f is continuous then fX is irreducible.*

PROOF. Otherwise let Y, Z be two closed proper subsets of the closure $\overline{fX}$ whose union is $\overline{fX}$. Then $f^{-1}X, f^{-1}Y$ are closed proper subsets of $f^{-1}\overline{fX}$ whose union is $f^{-1}\overline{fX}$. Since $X \subseteq f^{-1}\overline{fX}$ and $\bar{X}$ is an irreducible closed set, we infer that $\bar{X}$ lies inside one or other of $f^{-1}Y$ or $f^{-1}Z$, say the former. Then

$$fX \subseteq f\bar{X} \subseteq Y$$

and so $\overline{fX} \subseteq Y$, contradiction. □

An affine group G is said to be *connected* if the set G is irreducible. For example K is irreducible (since K is infinite and all nonempty open subsets of K are cofinite), and thus the additive group K_+ is connected.

Let G be an affine group and X, Y two irreducible components of G which contain the identity element 1. We claim that $X = Y$. For the product $X \times Y$ is irreducible (by Lemma 9), and hence its image XY under multiplication is also irreducible (since multiplication is continuous). Now $X \subseteq XY$, and hence $X = XY$ by maximality. Similarly $Y = XY$, so $X = Y$.

Thus G has a unique connected component containing 1; we write it G°. If g is any element of G°, then gG° is also an irreducible component of G containing 1, so $gG^\circ = G^\circ$. Thus G° is closed under multiplication. A similar argument shows that it is closed under inverse too; so it is a closed subgroup of G. Using conjugation, the same argument shows that it is a normal subgroup of G.

Proposition 11 *G° is a normal subgroup of G. It is also the unique smallest definable subgroup of finite index in G.*

PROOF. We have already proved the first sentence. For the rest, by Lemma 7 the closed subgroups of G coincide with the definable subgroups of G. Now the left cosets of G° are a family of pairwise disjoint irreducible components of G. Since G has only finitely many irreducible components, it follows that the index of G° in G is finite. If H is any closed subgroup of finite index in G, then its left cosets are also closed, hence open, and therefore they partition any subset of G into disjoint open subsets. Since G° is connected, it lies inside one coset of H, and so $G^\circ \subseteq H$. □

Note that by Proposition 11, a connected affine group has no proper definable subgroups of finite index.

Proposition 12 *If G is a connected affine group and W is a nonempty open subset of G, then $W.W = G$.*

PROOF. This follows from Lemma 6, since any nonempty open subset of an irreducible closed set is dense. □

We finish this section with some properties of $G(F)$ when F is pseudofinite and G is a connected group in K. They rest on the following, which holds also for all PAC fields. (See [2] (2.14).)

Fact 13 *Let F be a pseudofinite subfield of the algebraically closed field K, and X an irreducible K-closed set which is definable over F. Then $X(F)$ is Zariski dense in X, i.e. $X = \overline{X(F)}$. Equivalently, $X(F)$ meets every nonempty open subset of X.*

The next proposition lists some typical consequences for affine groups.

Proposition 14 *Let F be a pseudofinite subfield of K, and G a connected affine group definable over F.*

1. *$Z(G(F)) = Z(G) \cap G(F)$.*
2. *If H is a normal subgroup of $G(F)$ which is definable in F, then the closure of H is normal in G.*
3. *If H is a subgroup of finite index in $G(F)$, then $\bar{H} = G$.*
4. *F is an irreducible set.*

PROOF. The set G is irreducible since G is a connected group.

(1) is immediate using Proposition 8.

(2) We claim first that if $g \in G$ then $gHg^{-1} \subseteq \bar{H}$. For consider $h \in H$. The set of $g \in G$ such that $ghg^{-1} \notin \bar{H}$ is an open subset of G, and so if it is not empty it meets $G(F)$, say in g'. Then $g'hg'^{-1} \notin \bar{H}$, contradicting that H is normal in $G(F)$. Now for any $g \in G$ the map $a \mapsto gag^{-1}$ is continuous, and so it takes $\bar{H}$ into $\bar{H}$.

(3) Let $a_1, \ldots, a_n$ be representatives of the left cosets of H in $G(F)$. Then the set

$$X = a_1\bar{H} \cup \ldots \cup a_n\bar{H}$$

is a closed subset of G. Hence its complement in G is open, and meets $G(F)$ if it is not empty. But clearly $G(F) \subseteq X$, and so $X = G$. Since G is connected, $G = \bar{H}$.

(4) The closure of F is K, which is irreducible. □

4 Dimension in pseudofinite fields

We recall the following facts and definitions from Chatzidakis, van den Dries and Macintyre [3], cf. also [2] §2. Throughout, F is a pseudofinite subfield of the algebraically closed field K.

Fact 15 1. *If X is a subset of F^m definable in F, then the* dimension $\dim(X)$ *is the Krull dimension (= Morley rank) of the Zariski closure of X in K^m.*

2. *Let $\phi(\bar{x}, \bar{y})$ be a formula of L and $d < \omega$. Then there is a formula $\theta(\bar{x})$ of L such that if $\bar{b}$ is a tuple from F then*

$$\dim(\phi(F^m, \bar{b})) = d \Leftrightarrow F \models \theta(\bar{b}).$$

3. *If X, Y are definable subsets of F^m, F^n respectively, and f is a definable surjection from X to Y such that the pre-image under f of each element of Y is a set of dimension d, then*

$$\dim(X) = \dim(Y) + d.$$

Given a pseudofinite field F and a formula $\phi(\bar{x})$ of $L(F)$, we define the *dimension* of ϕ to be that of $\phi(F^m)$.

We draw out some immediate consequences.

Proposition 16 *If X, Y are definable subsets of F^m, then*

$$\dim(X \cup Y) = \max(\dim(X), \dim(Y)).$$

PROOF. Taking Zariski closures in K, $\bar{X} \cup \bar{Y} = \overline{X \cup Y}$. But the Morley rank of $\bar{X} \cup \bar{Y}$ is the maximum of the Morley ranks of $\bar{X}, \bar{Y}$, so the result follows by (1) of Fact 15. □

Proposition 17 *A nonempty definable set $X \subseteq F^m$ has dimension 0 if and only if it is finite.*

PROOF. A finite nonempty set is already Zariski closed in K, and so its dimension is 0 by (1) of the Fact. An infinite set has infinite Zariski closure, so its dimension is ≥ 1 by (1) again. □

Proposition 18 *If X, Y are definable subsets of F^m and f is a definable injective map from X to Y, then $\dim(X) \leq \dim(Y)$. If f is a bijection then $\dim(X) = \dim(Y)$.*

PROOF. If f is surjective, this follows from (3) of the Fact together with Proposition 17, since the preimage of each element of Y has dimension 0. If f is not surjective, its image has dimension $\leq \dim(Y)$ by Proposition 16. □

In particular it follows that if X is a definable subset of a group G definable in F, and g is an element of G, then the set $gX = \{ga : a \in X\}$ has the same dimension as X. All the cosets of a definable subgroup have the same dimension.

The next fact from [2] (2.9) is crucial for everything that follows. It is called the *(S1) property*, or for short just *(S1)*.

Fact 19 *Let X be a definable set in F^m, $\phi(\bar{x}, \bar{y})$ a formula of L, $d < \omega$, and suppose there are $\bar{b}_i$ $(i < \omega)$ in F such that each $X \cap \phi(F^m, \bar{b}_i)$ has dimension d and each $X \cap \phi(F^m, \bar{b}_i) \cap \phi(F^m, \bar{b}_j)$ $(i \neq j)$ has dimension $< d$. Then X has dimension $> d$.*

For example:

Proposition 20 *If G is an affine group defined in a pseudofinite field F, and H is a subgroup of G which has the same dimension as G, then H has finite index in G.*

PROOF. By Proposition 18, each coset gH has the same dimension as H. So if H has infinite index in G, we have a contradiction to (S1). □

Corollary 21 *Let G be an affine group which is defined over F. Let H be a subgroup of $G(F)$ which is defined in F. Then H is a subgroup of finite index in $\bar{H}(F)$.*

PROOF. We have $H \leq \bar{H}(F) \leq \bar{H}$. So $\dim(H) = \dim(\bar{H}(F))$, and it follows by the Proposition that H has finite index in $\bar{H}(F)$. □

5 Affine groups in pseudofinite fields

From this point onwards we write x for $\bar{x}$; there is no point in distinguishing elements from tuples.

Theorem 22 *Let F be a pseudofinite subfield of K, and G an affine group definable over F. Let U be a subset of $G(F)$ definable in F which contains the identity 1 and has the same dimension as $G(F)$. Let H be the subgroup of $G(F)$ generated by U. Then H is also definable in F, and has the form*

$$U^{\varepsilon_1} \ldots U^{\varepsilon_k}$$

where $\varepsilon_1, \ldots, \varepsilon_k \in \{1, -1\}$.

PROOF. Put $n = \dim(G(F))$, and let $\delta(x, y)$ be the formula

$$x \in G \wedge y \in G \wedge \dim(xU \cap yU) = n.$$

Put $W = \delta(G(F), 1)$. If $g \in W$, then $gU \cap U$ is not empty, and so there are $u_1, u_2 \in U$ with $gu_1 = u_2$. Then $g = u_2 u_1^{-1} \in U.U^{-1}$, so that $W \subseteq U.U^{-1}$.

Now choose inductively elements a_i of H so that

$$a_i \notin a_0 W \cup \ldots \cup a_{i-1} W,$$

as long as possible. Note that if a_i and a_j are defined with $i < j$, then $a_i^{-1} a_j \notin W$, so

$$\dim(a_j U \cap a_i U) = \dim(a_i^{-1} a_j U \cap U) < n.$$

Since the sets $a_i U$ have dimension n, it follows by (S1) that the choice of the a_i must halt after a finite number of steps, say when a_r has been chosen. Then

$$H = a_0 W \cup \ldots \cup a_r W \subseteq a_0(U.U^{-1}) \cup \ldots \cup a_r(U.U^{-1}) \subseteq H.$$

Choose m so that $a_0, \ldots, a_r \in (UU^{-1})^m$. Then $H = (U.U^{-1})^{m+1}$. □

The following theorem is an analogue of the fundamental Irreducibility Theorem for affine groups over algebraically closed fields. (I use the model theorists' name for Zil'ber's version of it. Algebraic geometers don't seem to have a name for it, apart from 'Proposition 2.2 of Borel [1]'.)

Theorem 23 *Let F be a pseudofinite field, G an affine group defined over F, and for each $i \in I$ let $X(i)$ be an irreducible subset of G which is definable in F and contains the identity element 1. Put $d = \dim G$. Let A be the subgroup of $G(F)$ generated by $\bigcup_{i\in I} X(i)$, and let B be $\bar{A}$. Then there are $i_1, \ldots, i_m \in I$ and $\varepsilon_1, \ldots, \varepsilon_m \in \{1, -1\}$ such that*

1.

$$A = X(i_1)^{\varepsilon_1} \ldots X(i_m)^{\varepsilon_m},$$

2.

$$B = \overline{X(i_1)}^{\varepsilon_1} \ldots \overline{X(i_m)}^{\varepsilon_m},$$

3. B is connected,

4. A has finite index in $B(F)$.

PROOF. Without loss we assume that each $X(i)^{-1}$ is also a set $X(j)$, and that the ε_i to be found are all 1. Choose $(i_1, \ldots, i_k)$ so that the set

$$W = \overline{X(i_1)}.\ldots.\overline{X(i_k)}$$

has maximal dimension in G. Since d is an upper bound on the dimension, there is such a sequence.

Since each $\overline{X(i)}$ is irreducible, the cartesian product

$$\overline{X(i_1)} \times \ldots \times \overline{X(i_m)}$$

is irreducible (by Lemma 9), and so its image under multiplication is also irreducible (by Lemma 10).

Now using Lemma 5,

$$\bar{W} \subseteq \bar{W}.\bar{W} \subseteq \overline{W.W}.$$

The two end terms $\bar{W}$ and $\overline{W.W}$ are both irreducible, and by choice of W they have the same dimension. So they are equal, proving that $\bar{W} = \bar{W}.\bar{W}$. Hence $\bar{W}$ is closed under multiplication by any element of $\bar{W}$. A similar argument shows that $\bar{W}^{-1} \subseteq \bar{W}$ and hence that $\bar{W}$ is closed under inverse. Also the same dimension argument gives that $\overline{X(i)} \subseteq \bar{W}$ for each $i \in I$. Therefore $\bar{W}$ is a closed subgroup of G containing all the sets $\overline{X(i)}$. Since $\bar{W} \subseteq B$, we infer that $\bar{W} = B$. So by Lemma 7, $W = B$. This gives (2) and (3).

Now put $U = X(i_1)\ldots X(i_m)$. Then $U \subseteq A \subseteq B(F) \subseteq B$. But $B \subseteq \bar{U}$ by Lemma 5, and so $\bar{A} = \bar{U} = \overline{B(F)}$. Thus U and $B(F)$ have the same dimension, and Theorem 22 applies. The result is (1) but in general with a larger value of m; this is no loss, since (2), (3) survive when we enlarge W so that its factors are the closures of those of A.

Finally since A and $B(F)$ have the same dimension, Proposition 20 tells us that A has finite index in $B(F)$, which is (4). □

6 Almost simple groups

Henceforth F is a pseudofinite subfield of the algebraically closed field K.

A morphism between affine groups is called an *isogeny* if it is surjective and has finite kernel.

We say that an affine group G over K is *simply connected* if every isogeny $h : H \to G$ with H connected is bijective. (This is an analogue of the notion for topological spaces. But note that a bijective isogeny between affine groups need not be an affine group isomorphism.)

A natural question is whether G connected in K implies that $G(F)$ is connected in F. The next result is a partial positive answer.

Proposition 24 *Let G be a connected and simply connected affine group in K which is definable over F. Then $G(F)$ has no proper subgroups of finite index which are definable in F.*

PROOF. Hrushovski and Pillay [7] prove that if J is a proper subgroup of finite index in $G(F)$ which is definable in F, then there are a connected affine group H defined over F and an isogeny $f : H \to G$ definable over F, such that $f(H(F))$ has finite index in J. (This is the main theorem in [2] §3.) Since G is simply connected, f is bijective. Then for each element g of $G(F) \setminus J$, $f^{-1}(g)$ is in H and is definable over F, and hence is in $H(F)$ since F is pseudofinite and therefore perfect. So $g = ff^{-1}(g)$ is in J; contradiction. □

We say that an affine group G in K is *almost simple* if it is non-abelian and has no closed infinite proper normal subgroups (or equivalently, no closed connected proper nontrivial normal subgroups).

Lemma 25 *Let G be an almost simple infinite affine group in K. Then G is connected, $Z(G)$ is finite and $G/Z(G)$ is simple as an abstract group.*

PROOF. First, G is connected. For otherwise by Proposition 11, G has a proper closed normal subgroup of finite index in G.

Next, the subgroup $Z(G)$ is normal and closed. If it is infinite, it must be the whole of G, which contradicts that G is non-abelian. So $Z(G)$ is finite. If $G/Z(G)$ is not simple, then G has a proper normal subgroup N which properly contains $Z(G)$. Let a be an element of $N \setminus Z(G)$. Then a^G is a subset of N containing more than one element; it is irreducible since G is connected and conjugation is continuous. Hence $a^G(a^{-1})^G$ is irreducible, infinite and contains 1. So by applying the irreducibility theorem (Proposition 2.2 of Borel [1]), the normal subgroup of G generated by a is closed; since it lies inside N, it is proper in G. This contradicts that G is almost simple. □

In the rest of this section, we consider how far almost simplicity is preserved from G to $G(F)$. We begin with two lemmas.

Lemma 26 *Let G be a connected affine group defined over F. Suppose $G(F)$ has finite centre. Then every infinite normal subgroup H of $G(F)$ contains an infinite normal subgroup of $G(F)$ which is definable in F and irreducible in K.*

PROOF. Since H is infinite and $Z(G(F))$ finite, there is $a \in H \setminus Z(G(F))$. Then in F we can define the set $X = a^{G(F)}$. Now $G(F)$ is dense in G by Fact 13, and hence $G(F)$ is irreducible. Therefore X is irreducible by Lemma 10. Since $a \notin Z(G(F))$, X has more than one element, and hence (being irreducible) it must be infinite.

So $X.a^{-1}$ is an irreducible infinite set containing 1. By the Irreducibility Theorem (Theorem 23), the subgroup H_1 of $G(F)$ generated by $X.a^{-1}$ is definable in F and irreducible. Since $X \subseteq H$, we have $H_1 \subseteq H$. Also H_1 is normal in $G(F)$; for this it suffices to test the generators of H_1. Consider $gag^{-1}a^{-1}$ and an element h of $G(F)$; we have

$$hgag^{-1}a^{-1}h^{-1} = (hg)a(hg)^{-1}ha^{-1}h^{-1}$$

$$= (hg)a(hg)^{-1}a^{-1}.(hah^{-1}a^{-1})^{-1} \in H_1.$$

Finally H_1 is infinite. □

The following standard group-theoretic lemma may be worth repeating here:

Lemma 27 *Let G be a group and H a proper subgroup of finite index in G. Then the intersection of all conjugates of H in G is a proper normal subgroup N of finite index in G. If G and H are definable in some surrounding structure, then so is N.*

PROOF. Since H has finite index in G, so does its normaliser $N(H)$. Hence H has only finitely many conjugates in G, and we can define their intersection N by finitely many parameters. The intersection of finitely many subgroups of finite index again has finite index. □

Proposition 28 *If G is an almost simple affine group in K defined over F, then $G(F)$ has a normal subgroup N of finite index which is definable in F, such that $N/Z(N)$ is simple as an abstract group.*

PROOF. By Proposition 14 (1), $Z(G(F)) = Z(G) \cap G(F)$. Hence $G(F)$ has finite centre. By Lemma 25, G is connected.

We claim that $G(F)$ has no infinite normal subgroup of infinite index. For otherwise by Lemma 26 there is an infinite normal subgroup H of infinite index which is irreducible and definable in F. Then H has smaller dimension than $G(F)$ by Proposition 20, and so the closure $\bar{H}$ of H is a closed connected nontrivial proper

subgroup of G (using Proposition 3). But $\bar{H}$ is also normal in G by Proposition 14 (2). This contradicts the assumption that G is almost simple.

We claim next that in F, $G(F)$ has a smallest definable subgroup of finite index. For otherwise there is an infinite descending sequence of definable subgroups of finite index in $G(F)$. Using Lemma 27 we can assume they are all normal, so that their intersection is a normal subgroup of infinite index. By saturating we can assume that this intersection is infinite, thus we get a contradiction to the previous claim. Our claim is proved.

Let N be a smallest subgroup of finite index in $G(F)$ which is definable in F. Then N is normal in $G(F)$ by Lemma 27. By Proposition 14 (3), $\bar{N} = G$, and so $Z(N) = Z(G) \cap N$ by Proposition 8, whence N has finite centre.

It remains to prove that $N/Z(N)$ is simple. For contradiction, let J be a normal subgroup of N which contains a noncentral element a. Recalling the proof of Lemma 26, we form the sets $X = a^{G(F)}$ and $Y = a^N$. We have

$$J \supseteq Ya^{-1} \subseteq Xa^{-1} \subseteq N.$$

As before, Xa^{-1} is irreducible and hence infinite. Also N has finite index in $G(F)$, and hence X is the union of a finite number of conjugates of Y. By Proposition 16 it follows that Ya^{-1} and Xa^{-1} have the same dimension. Hence Ya^{-1} and Xa^{-1} have the same closure in G, since in the Zariski topology on K, a proper closed subset of a closed set W must have strictly smaller dimension. Also as before, the subgroup C of $G(F)$ generated by Xa^{-1} is normal and infinite, so that by our first claim it has finite index in $G(F)$. By the Irreducibility Theorem (Theorem 23), C is definable in F, and hence it must equal N.

Let D be the subgroup of $G(F)$ generated by Ya^{-1}; then $D \subseteq J$. By the Irreducibility Theorem again, D is definable in F and has finite index in $B(F)$ where B is the least closed subgroup B of G containing Ya^{-1}. But B is also a closed subgroup of G containing the closure of Xa^{-1}, and so $N = C \subseteq B(F)$. Thus D has finite index in N, so that D is equal to N by choice of N. But $D \subseteq J \subseteq N$, and thus $J = N$. □

Corollary 29 *Let G be an almost simple and simply connected affine group defined over F. Then $G(F)$ is simple modulo its finite centre.*

PROOF. By Lemma 25, G is connected and its centre is finite. According to Proposition 24, $G(F)$ has no proper subgroups of finite index which are definable in F. It follows at once from the Proposition that $G(F)$ is simple over its centre. □

7 Reduction at a prime

For our remaining two sections, the algebraically closed field K is always the complex numbers $\mathbb{C}$.

Let p be a prime. The well-known reduction homomorphism $r : \mathbb{Z} \to \mathbb{F}_p$ takes each integer n to $n \pmod p$. We can extend r to the ring $\mathbb{Z}_{(p)}$ of all rational numbers whose denominators are prime to p. In fact this ring is local, with maximal ideal $p\mathbb{Z}_{(p)}$, and $r : \mathbb{Z}_{(p)} \to \mathbb{F}_p$ factors out $p\mathbb{Z}_{(p)}$. This extended map r is a ring homomorphism. Also $r(1/a).r(a) = r(1) = 1$, so r takes each element $1/a$ in $\mathbb{Z}_{(p)}$ to $(ra)^{-1}$ in $\mathbb{F}_p$; thus r respects inverses where they exist.

Obviously we can extend r yet again to the product ring $\mathbb{Z}_{(p)}^n$, for any positive integer n, to get a map

$$r : \mathbb{Z}_{(p)}^n \to \mathbb{F}_p^n.$$

If Γ is any subset of $\mathbb{Z}_{(p)}^n$, then we write Γ/p for the image of Γ under r. For example if Γ is a subgroup of $\mathrm{GL}_m(\mathbb{Q})$ generated by a set of matrices whose entries and determinants are all prime to p, then Γ/p is well-defined. (Why the determinants? Because we agreed that in a linear group, one of the coordinates of each element a should be $\det(a)^{-1}$.) The reduction map r restricts to a group homomorphism

$$r : \Gamma \to \Gamma/p$$

because multiplication and inverse are defined by equations.

Algebraic geometers reserve the expression *reduction at p* for a different extension of r. Namely, let $X_0, \ldots, X_{n-1}$ be indeterminates, and extend r to a surjective ring homomorphism

$$\rho : \mathbb{Z}_{(p)}[X_1, \ldots, X_n] \to \mathbb{F}_p[X_1, \ldots, X_n]$$

by applying r to the coefficients. Then for example if I is an ideal in $\mathbb{Z}_{(p)}[X_1, \ldots, X_n]$, ρI is an ideal in $\mathbb{F}_p[X_1, \ldots, X_n]$.

Suppose in particular that V is a $\mathbb{Q}$-closed subset of $\mathbb{C}^n$. This implies (since $\mathbb{Q}$ is perfect, see Borel [1] 12.2) that the ideal $I(V)$ of all polynomials in $\mathbb{C}[X_1, \ldots, X_n]$ which vanish on V is generated by a finite set J of polynomials lying in $\mathbb{Q}[X_1, \ldots, X_n]$; in fact we can choose J so that it lies in $\mathbb{Z}[X_1, \ldots, X_n]$. Then, writing $\bar{\mathbb{F}}$ for the algebraic closure of $\mathbb{F}_p$, the set of points of $\bar{\mathbb{F}}_p^n$ where ρJ vanishes is called the *reduction* of V at p; we shall write it V_p (noting for future reference that it depends on J).

Lemma 30 *If V is as above, then the restriction homomorphism r gives a map*

$$r : V(\mathbb{Z}_{(p)}) \to V_p(\mathbb{F}_p).$$

In general this map is not surjective.

PROOF. If a is a point of $V(\mathbb{Z}_{(p)})$ and F is a polynomial vanishing at a, then $F(a) = 0$, so $(\rho F)(ra) = 0$ when we apply r throughout; hence $ra \in V_p(\mathbb{F}_p)$. For an example where r is not surjective, let V be the set $\{\pm\sqrt{2}\}$ and p the prime 7. The ideal corresponding to V is generated by the polynomial $X^2 - 2$, so $V_7(\mathbb{F}_7)$

consists of the numbers 3 and 4 (mod 7). But there is no rational number a such that $a^2 = 2$. □

There is a problem about this definition of V_p: it depends on the choice of J. For example if we write pJ for the set of polynomials pF with $F \in J$, then pJ generates the same ideal $I(V)$ over the complex numbers, but using pJ instead of J would make V_p the whole of $\tilde{\mathbb{F}}_p^n$. Fortunately the next lemma will save us the trouble of looking for a more canonical definition.

Lemma 31 *If J and J' are finite subsets of $\mathbb{Z}[X_1, \ldots, X_n]$ which generate the same ideal in $\mathbb{C}[X_1, \ldots, X_n]$, and V_p and V'_p are the corresponding sets in $\tilde{\mathbb{F}}_p^n$, then $V_p = V'_p$ for all but finitely many primes p.*

PROOF. Suppose $J = \{F_0, \ldots, F_k\}$ and $J' = \{F'_0, \ldots, F'_m\}$. Since J lies in the ideal generated by J', there are polynomials G_{ij} over the complex numbers, such that each F_i is $\sum_j G_{ij} F'_j$. These equations reduce to a matrix equation $MH = N$ where M and N are integer matrices determined by J' and J respectively, and H is a column matrix consisting of all the coefficients in the polynomials G_{ij}. If such a matrix equation has a solution H in the complex numbers, then it has one in the rationals; so without loss we can assume that the polynomials G_{ij} have rational coefficients. Likewise there are polynomials G'_{ji} defining J' from J, and again they can be chosen with rational coefficients. For any prime p not dividing any of the coefficients of the G_{ij} or the G'_{ji}, ρJ and $\rho J'$ generate the same ideal over $\tilde{\mathbb{F}}_p$. □

If J is a set of polynomials with integer coefficients, then we can write a conjunction $\phi(x_0, \ldots, x_{n-1})$ of equations in the language L of rings, such that

$$\mathbb{C} \models \phi(a) \Leftrightarrow \text{ all polynomials in } J \text{ vanish on } a.$$

Then

$$\mathbb{F}_p \models \phi(ra) \Leftrightarrow \text{ all polynomials in } \rho J \text{ vanish on } ra.$$

This is because we interpret each term n of L in $\mathbb{F}_p$ as a name of n mod p.

Reduction at a prime is a fundamental technique in number theory. Typically one has some property of a $\mathbb{Q}$-closed set X, and one says that there is *good reduction* at the prime p if X/p also has this property, and *bad reduction* at p otherwise. One often finds that there is good reduction at all but finitely many primes, and the task is to calculate the exceptions.

Proposition 32 *Let G be a closed subgroup of $GL_n(\mathbb{C})$ defined over $\mathbb{Q}$, and Γ a finitely generated subgroup of $G(\mathbb{Q})$. Let P be the set of primes p such that $G_p(\tilde{\mathbb{F}}_p)$ is a group and Γ/p is a subgroup of $G_p(\tilde{\mathbb{F}}_p)$. Then P contains all but finitely many primes.*

PROOF. As before, we can suppose that G is defined over $\mathbb{Z}$. Let ψ be a sentence of L expressing that G is a group. Then ψ is true in $\mathbb{C}$, so by compactness, the set

P_1 of all primes p such that ψ is true in $\tilde{\mathbb{F}}_p$ contains all but finitely many primes. For each $p \in P_1$, $G_p(\mathbb{F}_p)$ is a subgroup of $G_p(\tilde{\mathbb{F}}_p)$, since multiplication and inverse are defined by equations. Let P_2 be the set of primes which don't divide either the determinant, or the denominator of any entry, of any generator of Γ. Then Γ/p is a well-defined subgroup of $G_p(\mathbb{F}_p)$ for each $p \in P_1$. The set $P = P_1 \cap P_2$ will serve. □

8 A theorem on reduction

The following theorem is from Matthews, Vaserstein and Weispfeiler [8]. The bad news is that their proof uses the structure theory of Lie algebras and the classification of finite simple groups. The good news is that they have an explicit bound on the rogue primes. Our proof (from Hrushovski and Pillay [7]) is non-effective but really rather easy.

Theorem 33 *Let G be an almost simple, simply connected closed subgroup of $GL_n(\mathbb{C})$ which is defined over $\mathbb{Q}$, and let Γ be a finitely generated subgroup of $G(\mathbb{Q})$ which is Zariski-dense in G. Then for all but finitely many primes p, $\Gamma/p = G_p(\mathbb{F}_p)$.*

Before we prove this, here is an example. Let G be the group $\mathrm{SL}_2(\mathbb{C})$ of 2×2 matrices with determinant 1. Let Γ be the subgroup of $G(\mathbb{Q})$ generated by the matrices

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then G is simple, simply connected and defined over $\mathbb{Q}$, and it has dimension 3. To show that Γ is Zariski-dense in G, one method is to check first that the group

$$\bar{\Gamma} \cap \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : a, d \in \mathbb{C} \right\}$$

is infinite and hence has Morley rank at least 1. Multiplication by powers of

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

gives infinitely many cosets of this subgroup, so that

$$\bar{\Gamma} \cap \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, b, d \in \mathbb{C} \right\}$$

has Morley rank at least 2; then a similar argument with the third generator of Γ shows that $\bar{\Gamma}$ has Morley rank at least 3. Since $\bar{\Gamma} \subseteq \mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SL}_2(\mathbb{C})$ is irreducible, this shows that $\bar{\Gamma} = \mathrm{SL}_2(\mathbb{C})$.

We try two primes. At $p = 2$, $\Gamma/2$ is undefined because of the $1/2$ in the first generator. At $p = 3$, $\Gamma/3$ is defined, and it is generated by the matrices

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

One can verify that $\Gamma/3$ is the whole of $\mathrm{SL}_2(\mathbb{F}_3)$. The theorem says that all but finitely many primes behave like 3.

PROOF of Theorem 33. The proof rests on the following construction. Let S be the set of primes $\geq n$ which are in the set P of Proposition 32 (here we use the fact that Γ is finitely generated). Let $\mathcal{U}$ be a nonprincipal ultrafilter on S. We form the ultraproduct $G^{\mathcal{U}} = \Pi_{p\in S} G_p(\mathbb{F}_p)/\mathcal{U}$. As always with ultraproducts, we can add other structure. For example Γ/p is a subgroup of $G_p(\mathbb{F}_p)$ for each prime p. We add a symbol Γ' which picks out, for each prime $p \in S$, the set Γ/p in the p-th factor. Then the interpretation $\Gamma^{\mathcal{U}}$ of Γ' in the ultraproduct is a subgroup of $G^{\mathcal{U}}$. Likewise we can add the field $\mathbb{F}_p$ to the p-th factor, getting an ultraproduct field $F^{\mathcal{U}}$.

Since $\mathcal{U}$ is a nonprincipal ultrafilter and the primes S are arbitrarily large, the field $F^{\mathcal{U}}$ is pseudofinite and of characteristic 0. Since $G_p(\mathbb{F}_p)$ is a subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$ for each $p \in S$, the group $G^{\mathcal{U}}$ is a subgroup of $\mathrm{GL}_n(F^{\mathcal{U}})$.

Lemma 34 $G^{\mathcal{U}} \subseteq G$.

PROOF. We have to show that if $a \in G^{\mathcal{U}}$ and $E(x)$ is one of the defining equations of G, then a satisfies E. Multiplying out denominators, we can write E in the language of rings. Each element $a(p)$ satisfies E in $G_p(\mathbb{F}_p)$, and so a satisfies E by Loś's theorem. □ Lemma

Lemma 35 $G(\mathbb{Q}) \subseteq G^{\mathcal{U}}$.

PROOF. Every element of $G(\mathbb{Q})$ is a matrix whose entries are explicitly definable in $\mathbb{C}$ by formulas of the language L. □ Lemma

In particular $\Gamma \subseteq \Gamma^{\mathcal{U}}$. Since Γ is Zariski-dense in G, it follows that $\Gamma^{\mathcal{U}}$ is also Zariski-dense in G.

Lemma 36 *$\Gamma^{\mathcal{U}}$ contains an infinite subgroup which is definable in $F^{\mathcal{U}}$ and irreducible.*

PROOF. It suffices to show that $\Gamma^{\mathcal{U}}$ contains a unipotent element U. For then we can apply the map $\exp(a\log(U))$ of Section 1, which maps the additive group $F_+^{\mathcal{U}}$ to an infinite subgroup J of $G^{\mathcal{U}}$. (We chose S so that the restriction on characteristic is met.) As noted in section 1, this subgroup J is definable in $F^{\mathcal{U}}$ by the definition of $\exp(a\log(U))$, which is a polynomial. So it is irreducible by Proposition 14 (4). In the ultraproduct, $U(p)$ is a unipotent matrix in Γ/p for all but finitely many primes p; hence the matrices $\exp(a\log(U(p)))$, as a runs through $\mathbb{F}_p$, are powers of $U(p)$ and therefore lie in Γ/p too. It follows that $J \subseteq \Gamma^{\mathcal{U}}$.

To find U we go by contradiction. Suppose no element of $\Gamma^{\mathcal{U}}$ is unipotent. Since there is a formula uniformly defining the unipotent elements in $\mathrm{GL}_n(A)$ for

any field A (namely $(u-I)^n = 0$), it follows that only finitely many of the groups Γ/p contain a unipotent element. Then Proposition 1 implies that for all but finitely many primes p in S, Γ/p contains no elements of order p.

Now we quote a fact:

Fact 37 *Let p be a prime and G a finite group with no p-power elements. If $\theta : G \to GL_n(\mathbb{F}_p)$ is a representation, it can be lifted to a p-adic representation $\theta' : G \to GL_n(\hat{\mathbb{Z}}_p)$, so that θ is recovered by factoring out the maximal ideal of $\hat{\mathbb{Z}}_p$.*

PROOF. See Proposition 43 of Serre [10].

Since the field $\hat{\mathbb{Z}}_p$ of p-adic numbers is embeddable in $\mathbb{C}$, this fact tells us that for all but finitely many primes p the group Γ/p has a faithful representation in $\mathrm{GL}_n(\mathbb{C})$. We need another fact:

Fact 38 (Jordan) *Let G be a finite subgroup of $GL_n(\mathbb{C})$. Then G has a normal abelian subgroup of index at most $d(n)$, where $d(n)$ is an integer depending only on n.*

PROOF. See Curtis and Reiner [4] p. 258ff.

So for all but finitely many primes p, Γ/p has a normal abelian subgroup Δ_p of index at most $d(n)$. Adding a symbol for the groups Δ_p to the ultraproduct, we recover a normal abelian subgroup $\Delta^{\mathcal{U}}$ of finite index in $\Gamma^{\mathcal{U}}$. Then by Corollary 4 the closure of $\Delta^{\mathcal{U}}$ has finite index in the closure of $\Gamma^{\mathcal{U}}$, which is G. Since G was assumed connected, it must be the closure of $\Delta^{\mathcal{U}}$, and hence by Proposition 3 must be abelian, contradicting that G is almost simple. This contradiction proves the lemma. □ Lemma

Consider the irreducible infinite subgroup J of $\Gamma^{\mathcal{U}}$ in Lemma 36. By the Irreducibility Theorem (Theorem 23), the normal closure H of J (generated by the groups J^g with $g \in \Gamma^{\mathcal{U}}$) is a subgroup of $\Gamma^{\mathcal{U}}$ which is definable in $F^{\mathcal{U}}$. Then the normaliser H_1 of H in $G(F^{\mathcal{U}})$ is definable in $F^{\mathcal{U}}$ and contains $\Gamma^{\mathcal{U}}$. Thus we have $H \subseteq \Gamma^{\mathcal{U}} \subseteq H_1 \subseteq G(F^{\mathcal{U}})$; we shall show that the end terms of this chain are equal.

Since G is the closure of $\Gamma^{\mathcal{U}}$, it is also the closure of H_1. Hence H_1 and $G(F^{\mathcal{U}})$ have the same dimension, so that by Proposition 20 it follows that H_1 has finite index in $G(F^{\mathcal{U}})$, besides being definable in $F^{\mathcal{U}}$. This is where at last we invoke Proposition 24 (proved in [2]), to deduce that $H_1 = G(F^{\mathcal{U}})$. So H is an infinite normal subgroup of $G(F^{\mathcal{U}})$ definable in $F^{\mathcal{U}}$. Then $H.Z(G(F^{\mathcal{U}}))$ is a normal subgroup of $G(F^{\mathcal{U}})$ containing the finite centre of $G(F^{\mathcal{U}})$. So by Corollary 29 (another application of Proposition 24), $H.Z(G(F^{\mathcal{U}})) = G(F^{\mathcal{U}})$, and hence H has finite index in $G(F^{\mathcal{U}})$. Then by Proposition 24 a third time, $H = G(F^{\mathcal{U}})$, trapping $\Gamma^{\mathcal{U}}$ so that $\Gamma^{\mathcal{U}} = G(F^{\mathcal{U}})$.

Since we showed in Lemma 34 that $G^{\mathcal{U}} \subseteq G$ and hence $\subseteq G(F^{\mathcal{U}})$, it follows that $G^{\mathcal{U}} \subseteq \Gamma^{\mathcal{U}}$, and hence $G_p(\mathbb{F}_p) = \Gamma/p$ for all p in some set in $\mathcal{U}$. But the ultrafilter

was chosen arbitrarily, and hence $G_p(\mathbb{F}_p) = \Gamma/p$ for all but finitely many primes in S, and therefore for all but finitely many primes. □

References

[1] A. Borel, Linear algebraic groups, Springer-Verlag, New York 1991.

[2] Z. Chatzidakis, Definable subgroups of algebraic groups over pseudo-finite fields (this volume).

[3] Z. Chatzidakis, L. van den Dries and A. J. Macintyre, Definable sets over finite fields. J. Reine Angew. Math. 427 (1992) 107-135.

[4] C. W. Curtis and I. Reiner, Representation theory of finite groups and associative algebras, Wiley, New York 1962.

[5] R. Hartshorne, Algebraic geometry, Springer-Verlag, New York 1977.

[6] E. Hrushovski and A. Pillay, Groups definable in local fields and pseudo-finite fields, Israel J. Math. 85 (1994) 203–262.

[7] E. Hrushovski and A. Pillay, Definable subgroups of algebraic groups over finite fields, J. Reine Angew. Math. 462 (1995) 69–91.

[8] C. R. Matthews, L. N. Vaserstein and B. Weisfeiler, Congruence properties of Zariski-dense subgroups I, Proc. London Math. Soc. 48 (1984) 514–532.

[9] A. Robinson and P. Roquette, On the finiteness theorem of Siegel and Mahler concerning diophantine equations, J. Number Theory 7 (1975) 121–176.

[10] J.-P. Serre, Linear representations of finite groups, Springer-Verlag, New York 1977.

Author's address:
School of Mathematical Sciences,
Queen Mary and Westfield College,
Mile End Road,
London E1 4NS,
England.
e-mail: w.hodges@qmw.ac.uk

The group of automorphisms of the field of complex numbers leaving fixed the algebraic numbers is simple

Daniel Lascar

In this short article, we will prove the following theorem:

Theorem 1 *The group of automorphisms of* $\mathbb{C}$ *which leave fixed every algebraic number is simple.*

(Here, $\mathbb{C}$ denotes the fields of complex numbers.)

This result is already mentioned in [1], where it is proved with the help of the continuum hypothesis, which will not be used here. (In fact it appeared as a particular case of a much more general result). A preliminary version of the present paper is to be found in french in [2].

Lets us first introduce some notation:

- G will denote the group of automorphisms of $\mathbb{C}$ which leave fixed the algebraic numbers;
- Ω will denote the set of algebraically closed subfields of $\mathbb{C}$ which are of cardinality strictly less than $2^{\aleph_0}$;
- if $K \in \Omega$, G_K will denote the group of automorphisms of K which leave the algebraic numbers fixed and $Aut_K(\mathbb{C})$ the group of automorphims of $\mathbb{C}$ which leaves fixed the elements of K.

We will need the following lemma.

Lemma 2 *Assume that* $g \in G$ *and* $K \in \Omega$ *are such that, for all* $a \in \mathbb{C}$, $g(a)$ *is algebraic over* $K(a)$. *Then* g *is the identity on* $\mathbb{C}$.

A proof of Lemma 2 is given in [1]. M. Ziegler (private communication) has given another proof which works in any characteristic. We give here a third proof which is completely elementary and which can be easily generalized to non zero characteristic.

Proof. Let g and K be as in the lemma.

Assume for a while that there exists an element $a \in \mathbb{C} - K$ such that $g(a) = a$. We show that, in these conditions g is the identity map. Let b be another element of $\mathbb{C}$ which is not algebraic over $K(a)$. Then $g(b)$ is algebraic over $K(b)$, and so is $g(b) - b$. On the other hand, $g(a + b) = a + g(b)$, and so, $a + g(b)$ is algebraic over $K(a + b)$, and so is $a + g(b) - (a + b) = g(b) - b$. Thus, we see that $g(b) - b$ is algebraic over $K(b)$ and over $K(a+b)$, and must belong to K, since b and $a+b$ are algebraically independent over K. Set $c = g(b) - b$. For the same reason, $g(ab) - ab$

belongs to K. But $g(ab) = g(a)g(b) = ab + ac$, so ac belongs to K, and this implies $c = 0$. This means that g is the identity on $\mathbb{C} - K$, thus on $\mathbb{C}$ itself.

We now return to the general case. Let a an element which is not in K. Set $a' = g(a)$ and let $P(X, X')$ be a polynomial with coefficients in K such that $P(a, a') = 0$ and of minimal degree. It is easily seen that $P(a, X')$ is a polynomial with coefficient in $K(a)$ of minimal degree such that $P(a') = 0$. We also notice that, if c does not belong to K and if c' is such that $P(c, c') = 0$, then there exist a K-automorphism of $\mathbb{C}$ which sends a to c and a' to c'. We show that $P(a, X') = 0$ has only one solution in $\mathbb{C}$.

Indeed, let b be another element of $\mathbb{C}$ which is not algebraic over $K(a)$ and K_1 be the algebraic closure of $K(b)$. The polynomial $P(a, X')$ is again of minimal degree among all polynomials with coefficients in K_1 for which a' is a zero. So, if $a'' \in \mathbb{C}$, and $P(a, a'') = 0$, then there exists a K_1-automorphism f of $\mathbb{C}$ which leaves a fixed and sent a' to a''. Then, the automorphism $h = g \circ f \circ g^{-1} \circ f^{-1}$ leaves b fixed and sends a'' to a' and is such that, for all $c \in \mathbb{C}$, $h(c)$ is algebraic over $K(c)$. We just saw that this implies that $a' = a''$.

So, we infer that the degree of P in the variable X' is equal to 1. For the same reason, its degree in the variable X is also equal to 1. In other words, $g(a) = \alpha a + \beta$, for some elements α and β of K. If we apply the same argument to a^2, we see that there exist γ and δ in K such that $g(a^2) = \gamma a^2 + \delta = \alpha^2 a^2 + 2\alpha a\beta + \beta^2$, which implies that $\beta = 0$, and, using this fact for $a+1$, we get $g(a+1) = \alpha a + 1 = \varepsilon(a+1)$ for some $\varepsilon \in K$; this implies that $\alpha = 1$ and that g is the identity map.
♡

The Theorem 1 is an immediate consequence of the following proposition:

Proposition 3 *Let g be an element of G, g not the identity. Then :*

$$G = g^G \circ (g^{-1})^G \circ (g^{-1})^G \circ g^G.$$

(g^G denotes the set $\{h^{-1}gh \ ; \ h \in G\}$).

Proof. For this proof, we fix an element g of G, $g \neq 1$. We will need some more notation.

• χ denotes the map from $G \times G$ to G defined by $\chi(h, k) = h^{-1} \circ g \circ k^{-1} \circ g^{-1} \circ k \circ h$. We see that $\chi(h, g) \in g^G \circ (g^{-1})^G$.

• If $K \in \Omega$ and $g[K] = K$, χ_K denotes application from $G_K \times G_K$ to G_K defined by $\chi_K(h, k) = h^{-1} \circ (g \mid K) \circ k^{-1} \circ (g^{-1} \mid K) \circ k \circ h$.

The combinatorial part of the proof is contained in the following lemma:

Lemma 4 *Let $K \in \Omega$, $g[K] = K$, h, k in G_K , $K' \in \Omega$ and $f \in G_{K'}$. Suppose that $K \subset K'$ and that f extends $\chi_K(h, k)$. Then, there exist h' and k' in G such that $\chi(h', k')$ extends f.*

We will need the two following sublemmas. Here and later, independent means algebraically independent and the dimension of a field is the cardinality of a transcendence basis of this field.

Sub-lemma 5 *Assume that K_0 , K_1 and K_2 are in Ω, that $K_0 \subset K_1$, that $K_0 \subset K_2$ and that K_1 and K_2 are independent over K_0. Moreover, assume that $h_1 \in G_{K_1}$, that $h_2 \in G_{K_2}$, and that $h_1 \mid K_0 = h_2 \mid K_0 = h$. Then there exists an automorphism h in G which extends both h_1 and h_2.*

Proof. Let K be the field generated by $K_1 \cup K_2$. We may define a map h of K to K in the following way: every element b of K can be written $R(\overline{b_0}, \overline{b_1}, \overline{b_2})$, where R is a rational function with coefficients in $\mathbf{Z}$, $\overline{b_0}$ is a sequence of elements in K_0, $\overline{b_1}$ is a sequence of elements in $K_1 - K_0$, and $\overline{b_2}$ is a sequence of elements in $K_2 - K_0$. We set: $h(b) = R(h_0(\overline{b_0}), h_1(\overline{b_1}), h_2(\overline{b_2}))$. It is straightforward to check that the map h is well defined ($h(b)$ does not depend on the representation $R(\overline{b_0}, \overline{b_1}, \overline{b_2})$ which has been chosen) and that h is an automorphism of K which can be extended to $\mathbf{C}$.
♡

Sub-lemma 6 *Let $K \in \Omega$ be such that $g[K] = K$ and λ a cardinal not bigger than $2^{\aleph_0}$. Then there exists $K_1 \in \Omega$, $K \subset K_1$ such that K_1 and $g[K_1]$ are independent over K and $\dim(K_1/K) = \lambda$.*

Proof. By induction we construct a sequence $(a_i \ ; \ i \in \lambda)$ of points in $\mathbf{C}$ such that: for all $i \in \lambda$, a_i is not algebraic over $K(\{a_j; j < i\} \cup \{g(a_j) \ ; \ j < i\})$ and $g(a_i)$ is not algebraic over $K(\{g(a_j) \ ; \ j < i\} \cup \{a_j \ ; \ j \leq i\})$. This is possible by lemma 2. Then, set K_1 to be the algebraic closure of $K(a_i; i < \lambda)$.
♡

We can now start the proof of Lemma 4.

Proof. We can find $h_1 \in G$ extending h which leaves K' setwise fixed. Set $h_2 = h_1 \mid K'$. We have to construct h' and k' in G extending respectively h and k such that:

(1) $\quad h'^{-1} \circ g \circ k'^{-1} \circ g^{-1} \circ k' \circ h'$ extends f

but we will rather construct $a \in Aut_K(\mathbf{C})$ and k' extending k such that:

(2) $\quad a^{-1} \circ g \circ k'^{-1} \circ g^{-1} \circ k' \circ a$ extends $h_2 \circ f \circ h_2^{-1}$.

(Then it will suffice to take $h' = a \circ h_1$). From Sub-lemma 6, we know that there exists $K_0 \in \Omega$ such that $K \subsetneq K_0$, $\dim(K_0/K) = \dim(K'/K)$ and that K_0 and $K_1 = g[K_0]$ are independent over K. Choose $a \in Aut_K(\mathbf{C})$ in such a way that $a[K'] = K_1$. Let $f_1 \in G_{K_1}$ be the map $a \circ h_2 \circ f \circ h_2^{-1} \circ a^{-1}$ (in fact, we should write $f_1 = (a \mid K') \circ h_2 \circ f \circ h_2^{-1} \circ (a \mid K')^{-1}$).

Now, we have to find $k' \in G$ extending k such that:

(3) $\quad g \circ k'^{-1} \circ g^{-1} \circ k'$ extends f_1.

Let $k_1 \in G_{K_1}$ extending k and $k_0 \in G_{K_0}$ extending $g^{-1} \circ k_1^{-1} \circ f_1 \circ g$ (or, more exactly, extending $(g \mid K_0)^{-1} \circ f_1 \circ k_1^{-1} \circ (g \mid K_0)$). Then $g \circ k_0^{-1} \circ g^{-1}$ equals $f_1 \circ k_1^{-1}$ and we see that $k_0 \mid K = k$ (because f_1 extends $g \circ k^{-1} \circ g^{-1} \circ k$). By Sublemma 5, there exists $k' \in G$ extending both k_0 and k_1. Then $g \circ k'^{-1} \circ g^{-1} \circ k'$ extends $g \circ k_0^{-1} \circ g^{-1} \circ k_1 = f_1$.
♡

Here is a slight modification of Lemma 4

Lemma 7 *Let* $K \in \Omega$, $g[K] = K$, h *and* k *in* G_K, $K' \in \Omega$, $f \in G$ *and* $b \in G_{K'}$. *Assume that* $K \subset K'$ *and that* b *extends* $\chi_K(h,k)$. *Then, there exists* $K_1 \in \Omega$, h_1 *and* k_1 *in* G_{K_1} *such that* $K' \subset K_1$, $g[K_1] = K_1 = f[K_1]$, $card(K_1) = card(K')$ *and* $\chi_{K_1}(h_1, k_1)$ *extends* b.

Proof. We first construct automorphisms h' and k' as in the preceding lemma, and then, by an argument of Löwenheim-Skolem type, we find a subfield say $K_1 \in \Omega$ such that $K' \subset K_1$, $g[K_1] = h'[K_1] = k'[K_1] = f[K_1] = K_1$ and $card(K_1) = card(K')$. Then, it suffices to set $h_1 = h' \mid K_1$ and $k_1 = k' \mid K_1$.
♡

Let $f \in G$. We are going to show that there exist h_*, k_*, h'_*, and k'_* such that $\chi(h_*, k_*) \circ (\chi(h'_*, k'_*))^{-1} = f$. The following lemma is a first step in this direction.

Lemma 8 *Let* $K \in \Omega$, $g[K] = K$, h , k , h' , k' *in* G_K , $a \in \mathbb{C}$ *and assume that* f *extends* $\chi_K(h,k) \circ (\chi(h',k'))^{-1}$. *Then there exist* $K_1 \in \Omega$ $a \in K_1$ *and* h_1, k_1, h'_1, k'_1 *in* G_{K_1} *extending* h, k, h', k' *respectively such that* $g[K_1] = h_1[K_1] = k_1[K_1] = h'_1[K_1] = k'_1[K_1] = K_1$, $card(K_1) = card(K)$ *and* f *extends* $\chi_{K_1}(h_1, k_1) \circ (\chi_{K_1}(h'_1, k'_1))^{-1}$.

Proof. We first choose $K^1 \in \Omega$ such that $card(K^1) = card(K)$, $a \in K^1$, $g[K^1] = K^1$ and $f[K^1] = K^1$. Let h'^1 and k'^1 be in G_{K^1} extending respectively h' and k'. We see that $(f \mid K^1) \circ \chi_{K^1}(h'^1, k'^1)$ extends $\chi_K(h,k)$. By Lemma 7, we find $K^2 \in \Omega$ such that $card(K^2) = card(K)$, $g[K^2] = K^2$, $f[K^2] = K^2$ and h^2 and k^2 in G_{K^2} extending respectively h and k such that $\chi_{K^2}(h^2, k^2)$ extends $(f \mid K^1) \circ \chi_{K^1}(h'^1, k'^1)$. In other words, $(f \mid K^2)^{-1} \circ \chi_{K^2}(h^2, k^2)$ extends $\chi_{K^1}(h'^1, k'^1)$. Again by Lemma 7, we find $K^3 \in \Omega$ such that $card(K^3) = card(K)$, $g[K^3] = K^3$, $f[K^3] = K^3$ and h'^3 and k'^3 in G_{K^3} extending respectively h'^1 and k'^1 such that $\chi_{K^3}(h'^3, k'^3)$ extends $(f \mid K^2) \circ \chi_{K^2}(h^2, k^2)$. Now, $(f \mid K^3)^{-1} \circ \chi_{K^3}(h'^3, k'^3)$ extends $\chi_{K^2}(h^2, k^2)$ and this way, we build a increasing sequence $(K^i \ ; \ i \in \omega)$ of elements of Ω and increasing sequences $(h^i \ ; \ i \in \omega, i \text{ even})$, $(k^i \ ; \ i \in \omega, i \text{ even})$, $(h^i \ ; \ i \in \omega, i \text{ odd})$, $(k'^i \ ; \ i \in \omega, i \text{ odd})$ in such a way that, setting $K_1 = \bigcup_{i \in \omega} K^i$, $h_1 = \bigcup_{i \in \omega} h^{2i}$, $k_1 = \bigcup_{i \in \omega} k^{2i}$, $h'_1 = \bigcup_{i \in \omega} h'^{2i+1}$, $k'_1 = \bigcup_{i \in \omega} k'^{2i+1}$, we get what was needed.
♡

We can now complete the proof of Proposition 3. Let $\left\{a_{\alpha+1} \ ; \ \alpha \in 2^{\aleph_0}\right\}$ be an enumeration of $\mathbb{C}$; we build by induction increasing sequences $(K_\alpha \ ; \ \alpha \in 2^{\aleph_0})$ of elements of Ω, $(h_\alpha \ ; \ \alpha \in 2^{\aleph_0})$, $(k_\alpha \ ; \ \alpha \in 2^{\aleph_0})$, $(h'_\alpha \ ; \ \alpha \in 2^{\aleph_0})$, $(k'_\alpha \ ; \ \alpha \in 2^{\aleph_0})$ of elements of G_{K_α}, such that for all $\alpha \in 2^{\aleph_0}$, $f[K_\alpha] = K_\alpha$, $g[K_\alpha] = K_\alpha$, $a_\alpha \in K_{\alpha+1}$ and f extends $\chi_{K_\alpha}(h_\alpha, k_\alpha) \circ (\chi(h'_\alpha, k'_\alpha))^{-1}$. We start with K_0 equal to the field of algebraic numbers and h_0, k_0, h'_0, k'_0 equal to the identity on K_0; afterwards, we use Lemma 8 at non limit stages, and take the union of what we have got so far at limit stages. It suffices to set $h_* = \bigcup_{\alpha \in 2^{\aleph_0}} h_\alpha$, $k_* = \bigcup_{\alpha \in 2^{\aleph_0}} k_\alpha$, $h'_* = \bigcup_{\alpha \in 2^{\aleph_0}} h'_\alpha$, $k'_* = \bigcup_{\alpha \in 2^{\aleph_0}} k'_\alpha$,
♡

References

[1] DANIEL LASCAR, *Les automorphismes d'un ensemble fortement minimal,* **Journal of Symbolic Logic** Vol.57, Number 1, March 1992, pp 238-251.

[2] DANIEL LASCAR, *Le groupe des automorphismes de* $\mathbb{C}$ *laissant les nombres algébriques fixes est simple*, séminaire de structures algébriques ordonnées 1993-1994, prépublication de l'équipe de logique n 45.

Author's address:
UFR de Mathématiques, Case 7012,
Université Paris 7,
2 place Jussieu,
75251 Paris, Cédex 05,
France.
e-mail: lascar@logique.jussieu.fr

The automorphism group of the field of complex numbers is complete

David M. Evans and Daniel Lascar

1 Introduction and notation

A group is complete if all its automorphisms are inner. The aim of this article is to prove the theorem stated in the title, assuming the continuum hypothesis. We will prove the following more general theorem:

Theorem 1 *Assume that L is an algebraically closed field of characteristic zero, either countable of infinite transcendence degree, or of cardinality $2^\lambda = \lambda^+$ for some infinite cardinal λ. Then its automorphism group is complete.*

See the final section for various generalisations, including to non-zero characteristic.

Some notation

In all the paper except the final section, L will be a field satisfying the hypotheses of the theorem; $Aut(L)$ will denote its automorphism group, and $Aut(Aut(L))$ the automorphism group of $Aut(L)$. If $X \subseteq L$, then

$$Aut(L/X) = \{f \in Aut(L); f \text{ is the identity on } X\}.$$

We will denote by Ω the set of algebraically closed subfields of L whose transcendence degree is strictly less than the one of L.

If α is a map from a set X to a set Y and Z is a subset of X, then $\alpha[Z]$ will denote the image of Z under α. If G is a group and X a subset of G, then $\langle X \rangle$ will denote the subgroup of G generated by X.

The proof will rest essentially on two ingredients. The first one is the small index property. If H is a subgroup of $Aut(L)$, we will say that H has small index if the index of H in $Aut(L)$ is not bigger that the cardinality of L. We will use the following result (see [3] or [4]):

Theorem 2 *Let H be a subgroup of small index in $Aut(L)$. Then there exists $k \in \Omega$ such that $Aut(L/k) \subseteq H$.*

The second ingredient is a theorem of Evans and Hrushovski and needs some explanation. If X is a subset of L, we will denote the algebraic closure of the field generated by X by $acl(X)$. The map acl from $\wp(L)$, the power set of L, to itself satisfies the following properties for any subsets X, Y of L and any x, y in L:

1. $X \subseteq acl(X)$ and if $X \subseteq Y$, then $acl(X) \subseteq acl(Y)$;
2. $acl(acl(X)) = acl(X)$;
3. If $x \notin acl(X)$ and $x \in acl(X \cup \{y\})$, then $y \in acl(X \cup \{x\})$;
4. $acl(X) = \bigcup_{X_0 \subseteq X, X_0 \text{ finite}} acl(X_0)$.

We have here what is called a pregeometry. We want a geometry, that is a map which, in addition satisfies the property that the closure of a singleton is the singleton itself. For that purpose, we consider the set

$$\mathbb{G} = \{k \subseteq L\,;\, k \text{ is algebraically closed and of transcendence degree } 1\}\,.$$

The map acl naturally induces a map cl from $\wp(\mathbb{G})$ into itself: if $X \subseteq \mathbb{G}$

$$cl(X) = \left\{k \in \mathbb{G}\,;\, k \subseteq acl(\bigcup X)\right\},$$

and now we have a geometry, that is the following properties are satisfied:

1. $X \subseteq cl(X)$ and if $X \subseteq Y$, then $cl(X) \subseteq cl(Y)$;
2. $cl(cl(X)) = cl(X)$;
3. If $x \notin cl(X)$ and $x \in cl(X \cup \{y\})$, then $y \in cl(X \cup \{x\})$;
4. $cl(X) = \bigcup_{X_0 \subseteq X, X_0 \text{ finite}} cl(X_0)$;
5. $cl(\{x\}) = \{x\}$.

We will denote by $Aut(\mathbb{G})$ the automorphism group of the geometry $\mathbb{G}$, that is the group of permutations β of $\mathbb{G}$ such that, for every $x \in \mathbb{G}$ and $X \subseteq \mathbb{G}$, $x \in cl(X)$ if and only if $\beta(x) \in \beta[X]$.

Now let $\alpha \in Aut(L)$; α naturally induces a map φ_α from $\mathbb{G}$ onto itself by: $\varphi_\alpha(k) = \alpha[k]$, and it is clear that $\varphi_\alpha \in Aut(\mathbb{G})$. Moreover, the map φ from $Aut(L)$ to $Aut(\mathbb{G})$ defined by: $\varphi(\alpha) = \varphi_\alpha$ is a group homomorphism. The theorem of Evans and Hrushovski (Theorem A of [2]) states:

Theorem 3 *The map φ is a surjective homomorphism from $Aut(L)$ onto $Aut(\mathbb{G})$.*

On the other hand,

Theorem 4 *The map φ is injective.*

Proof. This is an easy consequence of Theorem 1.1 of [2]. See also the remark after Lemma 2.5 of [2].
♡

The intuition behind the proof of Theorem 1 is to 'interpret' the field L into the group $Aut(L)$. Given this interpretation, any automorphism $\alpha \in Aut(Aut(L))$ induces naturally an automorphism γ of L. It will only remain to prove that α is nothing else than conjugation by γ.

The interpretation is done in two steps. First, in the second section, we interpret the geometry $\mathbb{G}$ into $Aut(L)$. It is there that the small index property is used. The precise result that we will get is:

Proposition 5 *Let* $\alpha \in Aut(Aut(L))$. *Then for any algebraically closed field* k *of finite transcendence degree, there exists a unique algebraically closed field* k' *of finite transcendence degree such that* $\alpha[Aut(L/k)] = Aut(L/k')$.

The second step is the interpretation of L into $\mathbb{G}$. This is contained in the paper of Evans and Hrushovski ([2]), and gives Theorem 3 above. We will prove Theorem 1 from Proposition 5 in Section 3. In the final section, we will indicate how this result can be generalised.

2 Proof of Proposition 5

We first state some facts that will be needed.

Lemma 6 *Assume that* K_0, K_1 *and* K_2 *are elements of* Ω *and that* K_1 *and* K_2 *are independent over* K_0. *Then*

$$Aut(L/K_0) \subseteq \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle .$$

Proof. Let $G = \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle$. Without loss, we can assume that $K_0 \subseteq K_1 \cap K_2$. Then independence of K_1 and K_2 over K_0 means that $K_0 = K_1 \cap K_2$ and there exists a transcendence basis X of L with $X_i = X \cap K_i$ a transcendence basis for K_i, for $i = 0, 1, 2$. Let $g \in Aut(L/K_0)$. There exists $Y \subseteq X$ of cardinality less than that of X such that $K_1, K_2, g[K_1] \subseteq acl(Y)$. Find $Z \subseteq X \setminus Y$ with the same cardinality as $X_1 \setminus X_0$. Then there exists $h \in Aut(L/K_2)$ such that $h[X_1] = X_0 \cup Z$. Let $K_3 = acl(h[X_1])$. Thus $Aut(L/K_3) \leq G$. There exists $k \in Aut(L/K_3)$ such that g and k have the same restriction to X_1. Moreover, as any automorphism of K_1 which fixes K_0 can be extended to an automorphism of L fixing K_3, we can choose k so that it has the same restriction to K_1 as g. Thus $k^{-1}g \in Aut(L/K_1)$, so $g \in G$.
♡

Lemma 7 *Assume that* K_0, K_1 *and* K_2 *are elements of* Ω, *that* $K_0 = K_1 \cap K_2$ *and that* K_1 *and* K_2 *are of finite transcendence degree over* K_0. *Then*

$$Aut(L/K_0) = \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle .$$

Proof. Let n_1, n_2, n_3 be the transcendence degrees of K_1, K_2, $acl(K_1 \cup K_2)$ over K_0 respectively. The proof will proceed by induction on $m = n_1 + n_2 - n_3$. If $m = 0$, then K_1 and K_2 are independent over K_0, and the result follows from Lemma 6. We will need the following fact

Fact : *Let L_0, L_1, L_2 be algebraically closed fields, included in some big field H and assume that $L_0 = L_1 \cap L_2$. Let a be a finite sequence of elements of H which is algebraically independent over $L_1 \cup L_2$. Then $acl(L_1 \cup \{a\}) \cap acl(L_2 \cup \{a\}) = acl(L_0 \cup \{a\})$.*

It is sufficient to prove this fact in the case where a is a single element (by induction). Let $\bar{c}_1$ and $\bar{c}_2$ be transcendence basis of L_1 and L_2 respectively over L_0. Assume, toward a contradiction that $\alpha \in acl(L_1 \cup \{a\}) \cap acl(L_2 \cup \{a\})$, $\alpha \notin acl(L_0 \cup \{a\})$. Let $P_1(x, a, \bar{z}_1)$ and $P_2(x, a, \bar{z}_2)$ be minimal polynomials with coefficients in L_0 such that $P_1(\alpha, a, \bar{c}_1) = 0$ and $P_2(\alpha, a, \bar{c}_2) = 0$. Consider the set of elements y such that:

- the degree in x of both $P_1(x, y, \bar{c}_1)$ and $P_2(x, y, \bar{c}_2)$ is positive;
- $P_1(x, y, \bar{c}_1) = 0$ and $P_2(x, y, \bar{c}_2) = 0$ have a common root.

This set is definable (or constructible, if you prefer), and since it contains every point which is not algebraic over $L_1 \cup L_2$, it contains all but a finite number of points. So, it contains a point $b \in L_0$ such that, in addition $P_1(x, b, \bar{z}_1)$ is of positive degree in $\bar{z}_1$. So, if β is such that $P_1(\beta, b, \bar{c}_1) = 0$ and $P_2(\beta, b, \bar{c}_2) = 0$, then $\beta \in L_1 \cap L_2$, but $\beta \notin L_0$, otherwise $\bar{c}_1$ would be algebraically dependent over L_0.

We now prove the lemma from the fact. Let $(a_1, a_2, \ldots, a_{n_1})$ and $(b_1, b_2, \ldots, b_{n_2})$ be transcendence bases of K_1 and K_2 respectively over K_0. We may assume that $b_1, b_2, \cdots, b_m$ are algebraic over $K_0 \cup \{a_1, a_2, \ldots, a_{n_1}, b_{m+1}, b_{m+2}, \ldots, b_{n_2}\}$. Now let $(c_1, c_2, \ldots, c_{n_2})$ be such that $c_{m+1}, c_{m+2}, \ldots, c_{n_2}$ are algebraically independent over $acl(K_1 \cup K_2)$ and such that there exists an automorphism of L leaving K_1 pointwise fixed and mapping $(b_1, b_2, \ldots, b_{n_2})$ onto $(c_1, c_2, \ldots, c_{n_2})$. Set $K_2' = acl(K_0 \cup \{c_1, c_2, \ldots, c_{n_2}\})$. Clearly $Aut(L/K_2') \subseteq \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle$ and $K_0 = K_2 \cap K_2'$. So, we just have to prove that

$$\langle Aut(L/K_2) \cup Aut(L/K_2') \rangle = Aut(L/K_0).$$

Since K_2 and K_2' are independent over K_1, $K_2 \cap K_2' \subseteq K_1$, so $K_2 \cap K_2' = K_0$. Now, $c_1 \in acl(K_1 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\})$, $c_1 \notin acl(K_0 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\})$ and by the fact,

$$acl(K_1 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\}) \cap acl(K_2 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\}) = acl(K_0 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\}).$$

Thus $c_1 \notin acl(K_2 \cup \{c_{m+1}, c_{m+2}, \ldots, c_{n_2}\})$, and this proves that the transcendence degree of $acl(K_2 \cup K_2')$ is at least $2n_2 - m + 1$. Thus, we may apply the induction hypothesis.

♡

Lemma 8 *Assume that K_0, K_1 and K_2 are elements of Ω, that $K_0 = K_1 \cap K_2$ and that K_1 is of finite transcendence degree over K_0. Then*

$$Aut(L/K_0) = \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle .$$

Proof. We know that there exists K_2' such that $K_0 \subseteq K_2' \subseteq K_2$, K_2' of finite transcendence degree over K_0 and K_1 and K_2 independent over K_2'. By Lemma 6, $Aut(L/K_2') \subseteq \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle$ and by Lemma 7 we have that $Aut(L/K_0) \subseteq \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle$.
♡

Remark: If L is of uncountable transcendence degree, it is not true in general that

$$\langle Aut(L/K_1) \cup Aut(L/K_2) \rangle = Aut(L/K_1 \cap K_2)$$

for arbitrary $K_1, K_2 \in \Omega$. Indeed, an easy inductive argument shows that if each of K_1, K_2 is of finite transcendence degree over the other then we have that $g[K_1]$ is of finite transcendence degree over K_1 for $g \in \langle Aut(L/K_1) \cup Aut(L/K_2) \rangle$. Now, by Example 3.3 of [1], there exist $K_1, K_2 \in \Omega$ of infinite transcendence degree, each of transcendence degree 1 over the other, and such that $K_1 \cap K_2$ is algebraic. So for these K_1, K_2 the above equality does not hold.

Lemma 9 *Assume that $K \in \Omega$, $f \in Aut(L)$ and that, for every $a \in L$, $f(a)$ is algebraic over $K(a)$. Then f is the identity on L.*

Proof. This is an easy consequence of Lemma 2.5 of [2]. See also Lemma 2 of [5].
♡

Let $\mathcal{H} = \{H \leq Aut(L)$; the index of H in $Aut(L)$ is small$\}$ and for $k \in \Omega$, $Aut(L/\{k\})$ the set of automorphisms which leave k fixed setwise. It is clear that, for every $k \in \Omega$, $Aut(L/\{k\}) \in \mathcal{H}$.

The proof of Proposition 5 will be split into two cases, according whether the transcendence degree of L is countable or uncountable.

The countable case.

In this case, Ω is just the set of subfields of L of finite transcendence degree. Let K_0 be the algebraic closure in L of the prime subfield. Let

$$\mathcal{H}_1 = \{Aut(L/\{k\})\,;\, k \in \Omega\} .$$

If $H \in \mathcal{H}$ we know by Theorem 2 and Lemma 7 that there exists a unique minimal $k \in \Omega$ such that $Aut(L/k) \subseteq H$. In fact, $k = \bigcap\{k' \in \Omega\,;\, Aut(L/k') \subseteq H\}$. We will call this subfield k the support of H and denote it by $\sigma(H)$.

We claim that for every $H \in \mathcal{H}$ which does not contain $Aut(L/K_0)$, there exists $H_1 \in \mathcal{H}_1 \setminus Aut(L)$ such that $H \subseteq H_1$: it suffices to take $H_1 = Aut(L/\{\sigma(H)\})$. Obviously, if k and k' are distinct elements of Ω and $k' \neq K_0$ then $Aut(L/\{k\}) \not\subseteq Aut(L/\{k'\})$.

Lemma 10 *Let $H \in \mathcal{H}$, $k \in \Omega$, $k_1 = \sigma(H)$ and assume that $H \subseteq Aut(L/\{k\})$. Then*

$$k_1 = k \text{ if and only if } \bigcap_{g \in Aut(L/\{k\})} g^{-1}Hg \neq \{e\}.$$

Proof. If we assume that $\sigma(H) = k$, then $Aut(L/k) \subseteq H$, and, for every $g \in Aut(L/\{k\})$, $Aut(L/k) = g^{-1}.Aut(L/k).g \subseteq g^{-1}Hg$, so

$$Aut(L/k) \subseteq \bigcap_{g \in Aut(L/\{k\})} g^{-1}Hg.$$

Conversely, suppose that $k_1 \neq k$. Since $Aut(L/k_1) \subseteq H \subseteq Aut(L/\{k\})$, we see that $k \subseteq k_1$. We conclude the proof of the lemma by showing that, under these conditions,

$$\bigcap_{g \in Aut(L/\{k\})} g^{-1}.Aut(L/\{k_1\}).g = \{e\}$$

or, equivalently

$$\bigcap_{g \in Aut(L/\{k\})} Aut(L/\{g[k_1]\}) = \{e\}.$$

So, let $f \in \bigcap_{g \in Aut(L/\{k\})} Aut(L/\{g[k_1]\})$. Let $(a_1, a_2, \ldots, a_n)$ be a transcendence basis of k_1 over k and k_2 be the algebraic closure of $k(a_1, a_2, \ldots, a_{n-1})$. We claim that for every $b \in L$, $f(b)$ is algebraic over $k_2(b)$. This is trivial if $b \in k_2$, for there exists $g \in Aut(L/\{k\})$ such that $g[k_1] \cap k_1 = k_2$ so f stabilises k_2. If $b \notin k_2$ then there exists $g \in Aut(L/\{k\})$ such that $g(a_i) = a_i$ for $i = 1, 2, \ldots, n-1$ and $g(a_n) = b$. Thus $g[k_1]$ is equal to k_3, the algebraic closure of $k(a_1 a_2 \ldots a_{n-1} b)$, and since $f[k_3] = k_3$, $f(b)$ is algebraic over $k_2(b)$. By Lemma 9, this implies that f is the identity.
♡

Note in particular this implies that $H \in \mathcal{H}$ does not contain $Aut(L/K_0)$ if and only if $\bigcap_{g \in Aut(L)} g^{-1}Hg = \{e\}$, that is, H is core-free. Moreover, if $H \in \mathcal{H}$ does not contain $Aut(L/K_0)$ then $H \leq Aut(L/\{\sigma(H)\}) < Aut(L)$. So $\mathcal{H}_1 \setminus Aut(L)$ is the set of core-free maximal elements of $\mathcal{H}$. Consequently, it is left fixed by any $\alpha \in Aut(Aut(L))$.

We are now done: Let

$$\mathcal{F} = \{(H, Aut(L/\{k\}) ; H \in \mathcal{H}, k \in \Omega, H \subseteq Aut(L/\{k\}), \sigma(H) = k\}.$$

Then (by Lemma 10) this family $\mathcal{F}$ is left fixed by any α in $Aut(Aut(L))$, and the same is true for the class

$$\left\{ \bigcap_{(H,H_1) \in \mathcal{F}} H ; H_1 \in \mathcal{H}_1 \right\}$$

and since, for $H_1 = Aut(L/\{k\}) \in \mathcal{H}_1$, $\bigcap_{(H,H_1) \in \mathcal{F}} H = Aut(L/k)$, we see that this last class is exactly equal to $\{Aut(L/k) ; k \in \Omega\}$.

The uncountable case

We recall that we have assumed that $card(L) = \lambda^+ = 2^\lambda$. Let

$$\mathcal{H}_2 = \{Aut(L/K)\,;\, K \text{ is an algebraically closed subfield of } L \text{ and } card(K) = \lambda\}.$$

Lemma 11 *$\mathcal{H}_2$ is the unique subset of $\mathcal{H}$ (the class of subgroups of $Aut(L)$ of small index) satisfying the 3 following conditions:*

1. *$\mathcal{H}_2$ is the conjugacy class of one of its elements;*
2. *$\mathcal{H}_2$ is closed under countable decreasing intersections;*
3. *If $H \in \mathcal{H}$, then there exists $H' \in \mathcal{H}_2$ such that $H' \subseteq H$.*

Proof. First, we see that $\mathcal{H}_2$ satisfies the 3 properties: the first one comes from the fact that, for every two algebraically closed subfields K_1 and K_2 of L of cardinality λ, there exist an automorphism of L mapping K_1 onto K_2. The second follows from the fact that, if K^i is an algebraically closed subfield of L of cardinality λ for $i \in \omega$ and $Aut(L/K^0) \geq Aut(L/K^1) \geq \ldots$, then $K^0 \subseteq K^1 \subseteq \ldots$ and

$$\bigcap_{i\in\omega} Aut(L/K^i) = Aut(L/\bigcup_{i\in\omega} K^i).$$

The third comes from the small index property (Theorem 2).

Assume now that $\mathcal{K} \subseteq \mathcal{H}$ satisfies the three conditions above. We want to prove that $\mathcal{K} = \mathcal{H}_2$. Because of condition 1, its suffices to prove that $\mathcal{K}$ and $\mathcal{H}_2$ intersect. Using condition 3, one can construct inductively subgroups H_i and K^i for $i \in \omega$ such that the H_i belong to $\mathcal{H}_2$ and the K^i belong to $\mathcal{K}$ and $H_0 \supseteq K^0 \supseteq H_1 \supseteq K^1 \supseteq \ldots \supseteq H_i \supseteq K^i \supseteq \ldots$. Then, by condition 2

$$\bigcap_{i\in\omega} H_i = \bigcap_{i\in\omega} K^i \in \mathcal{H} \cap \mathcal{K}.$$

♡

It follows that $\mathcal{H}_2$ is left fixed by any $\alpha \in Aut(Aut(L))$. To conclude the proof of Proposition 5, we have to show the same property for the class

$$\mathcal{H}_3 = \{Aut(L/k)\ ;\ k \text{ is an algebraically closed subfield of } L \text{ of finite transcendence degree}\}.$$

This will be an immediate consequence of the above lemma and of the following lemma:

Lemma 12 *Let $H \in \mathcal{H}$. Then $H \in \mathcal{H}_3$ if and only if the following two conditions are satisfied:*

1. *If $\mathcal{X}$ is a subset of $\mathcal{H}_2$ which is downward directed, of cardinality at most λ and such that $\bigcap \mathcal{X} \subseteq H$, then there exists $H_1 \in \mathcal{X}$ such that $H_1 \subseteq H$.*

2. $H = \langle \bigcup \{H_1 ; H_1 \in \mathcal{H}_2 \text{ and } H_1 \subseteq H\} \rangle$.

(**Remark:** we say that $\mathcal{X}$, a subset of $\mathcal{H}_2$, is downard directed if the intersection of any two elements of $\mathcal{X}$ contains an element of $\mathcal{X}$. Since any element of $\mathcal{X}$ is of the form $Aut(L/K)$ for some algebraically closed subfied of L, this means that the set $\mathcal{K} = \{K ; Aut(L/K) \in \mathcal{X}\}$ is upward directed, that is, the union of two elements of $\mathcal{K}$ is contained in an element of $\mathcal{K}$.)

Proof. Assume that $H \in \mathcal{H}_3$. The second condition follows from Lemma 6. Suppose now that $\mathcal{X}$ is as in condition (1). Let k be the algebraically closed field of finite transcendence degree such that $H = Aut(L/k)$. The hypothesis of condition (1) tells us that $\bigcap \mathcal{X} \subseteq Aut(L/k)$, and this implies that $k \subseteq acl(\bigcup \mathcal{K})$. So, because the family $\mathcal{K}$ is upward directed, this implies that k is included in one of the elements of $\mathcal{K}$, thus there exists an element H_1 of $\mathcal{X}$ included in H.

Conversely, assume that $H \in \mathcal{H}$ and that the two conditions are satisfied. By the small index property, we know that there exists $K^0 \in \Omega$ such that $Aut(L/K^0) \subseteq H$. Let $\{a_i ; i \in \mu\}$ (μ a cardinal not bigger than λ) be a transcendence basis of K^0 and let K^1 be an algebraically closed subfield of L of transcendence degree λ and independent from K^0. Set

$$\mathcal{X} = \left\{ Aut(L/acl(K^1 \cup \{a_i ; i \in s\})) ; s \text{ is a finite subset of } \mu \right\}.$$

Then the hypotheses of condition (1) are satisfied, so, for some finite subset s of μ, $Aut(L/acl(K^1 \cup \{a_i ; i \in s\})) \subseteq H$. But we know that $Aut(L/K^0) \subseteq H$, and that K^0 and $acl(K^1 \cup \{a_i ; i \in s\})$ are independent over $acl(a_i ; i \in s)$. Thus, by Lemma 6, $Aut(L/acl(a_i ; i \in s)) \subseteq H$. By Lemma 8, there exists an algebraically closed subfield k of L of finite transcendence degree, such that, for every $K \in \Omega$, $Aut(L/K) \subseteq H$ if and only if $k \subseteq K$. But in this case $Aut(L/K) \subseteq Aut(L/k)$ so by condition (2) we get $H = Aut(L/k)$. $\heartsuit$

3 Proof of Theorem 1

Let $\mathcal{T}$ be the lattice of algebraically closed subfields of L of finite transcendence degree. Let $\alpha \in Aut(Aut(L))$. Then Proposition 5 allows us to define a map ψ_α from $\mathcal{T}$ into itself by: for every $k \in \mathcal{T}$, $\alpha[Aut(L/k)] = Aut(L/\psi_\alpha(k))$. Obviously, this map is bijective (because $\psi_{\alpha^{-1}}$ is the inverse mapping), and is a lattice isomorphism. So, it maps elements of $\mathbb{G}$ (which are minimal elements of $\mathcal{T}$) into elements of $\mathbb{G}$. We will also denote by ψ_α the restriction of ψ_α to $\mathbb{G}$. It is easy to see that $\psi_\alpha \in Aut(\mathbb{G})$ and that the map ψ from $Aut(Aut(L))$ into $Aut(\mathbb{G})$ defined by $\psi(\alpha) = \psi_\alpha$ is a group homomorphism.

Lemma 13 *The map ψ is injective.*

Proof. Let $\alpha \in Aut(Aut(L))$, and assume that ψ_α is the identity on $\mathbb{G}$ (that is, for all $k \in \mathbb{G}$, $\alpha[Aut(L/k)] = Aut(L/k)$). We have to prove that α is the identity on $Aut(L)$.

For all $g \in Aut(L)$ and $k \in \mathbb{G}$, we have: $\alpha[Aut(L/g[k])] = Aut(L/g[k])$. But $Aut(L/g[k]) = g \cdot Aut(L/k) \cdot g^{-1}$, so that the above identity yields:

$$Aut(L/g[k]) = \alpha[gAut(L/k)g^{-1}] = \alpha(g) \cdot \alpha[Aut(L/k)] \cdot \alpha(g^{-1}) = \alpha(g) \cdot Aut(L/k) \cdot \alpha(g)^{-1} = Aut(L/\alpha(g)[k])$$

and it follows that $g[k] = \alpha(g)[k]$ and that $g^{-1} \cdot \alpha(g)[k] = k$. This means exactly that $\varphi(g^{-1} \cdot \alpha(g))$ is the identity on $\mathbb{G}$, and by Theorem 4, $g^{-1} \cdot \alpha(g)$ is the identity on L and so α is the identity on $Aut(L)$.
♡

We have already a homomorphism φ from $Aut(L)$ into $Aut(\mathbb{G})$ and a homomorphism ψ from $Aut(Aut(L))$ into $Aut(\mathbb{G})$. Call γ the homomorphism from $Aut(L)$ into $Aut(Aut(L))$, defined by: for all $g \in Aut(L)$, $\gamma(g)(f) = g \cdot f \cdot g^{-1}$ (that is, $\gamma(g)$ is conjugation by g).

Lemma 14 *The diagram*

$$\begin{array}{ccc} Aut(L) & \xrightarrow{\varphi} & Aut(\mathbb{G}) \\ & \overset{\psi}{\nearrow} & \\ \downarrow \gamma & & \\ Aut(Aut(L)) & & \end{array}$$

commutes.

Proof. This is more or less evident: let $g \in Aut(L)$. Then for all $k \in \mathbb{G}$ $\psi(\gamma(g))(k)$ is the element k' of $\mathbb{G}$ such that:

$$Aut(L/k') = \gamma(g)[Aut(L/k)] = g \cdot Aut(L/k) \cdot g^{-1} = Aut(L/g[k]),$$

so that $\psi(\gamma(g))(k) = g[k] = \varphi(g)(k)$.
♡

We are now ready to finish the proof of Theorem 1. The map φ is surjective, so ψ is also surjective. Thus by Lemma 13 ψ is a bijection. Then by Lemma 14, $\gamma = \psi^{-1} \cdot \varphi$ is surjective, which is what we need.

4 Generalisations

A more general version of Theorem 1 is:

Theorem 15 *Assume that L is an algebraically closed field, either countable of infinite transcendence degree, or of cardinality $2^\lambda = \lambda^+$ for some infinite cardinal λ. Let K_0 be an algebraically closed subfield of L such that the transcendence rank of L over K_0 is equal to the cardinality of L. Then the group $Aut(L/\{K_0\})$ is complete.*

We indicate how to modify the proof of Theorem 1 to give this. First, assume that the characteristic of L is $p \neq 0$ and K_0 is the algebraic closure of the prime

subfield. Theorem 3 remains true, as do the results on the small index property which we used and the basic algebraic lemmas 6, 7 and 8. The main difference lies in Theorem 4. In this case, φ is a surjective homomorphism, whose kernel is the set of Frobenius automorphisms,

$$\Phi = Ker(\varphi) = \{f_n\,;\, n \in \mathbb{Z}\}$$

where f_n is the automorphism of L which maps any $x \in L$ to x^{p^n}. In Lemma 10 the conclusion should be modified to

$$k_1 = k \text{ if and only if } \bigcap_{g \in Aut(L/\{k\})} g^{-1}Hg \not\subseteq \Phi.$$

Thus the only difficulty is encountered in the proof of Lemma 13: if $\alpha \in Ker(\psi)$, then for every $g \in Aut(L)$, there exists $n \in \mathbb{Z}$ (call it $n(g)$) such that $\alpha(g) = g \cdot f_{n(g)}$, and it is easy to see that the map $g \mapsto n(g)$ (call it n) is a group homomorphism from $Aut(L)$ into $\mathbb{Z}$. The kernel of n being of countable index, there exists a subfield k of L of small transcendence degree such that $Aut(L/k) \subseteq Ker(n)$, and since $Ker(n)$ is normal in $Aut(L)$, it is easy to see that $Aut(L/K_0) \subseteq Ker(n)$.

We can then factorise n

$$\begin{array}{ccc} Aut(L) & \xrightarrow{n} & \mathbb{Z} \\ \downarrow r & \nearrow n_0 & \\ Aut(K_0) & & \end{array}$$

where r is the restriction map: $r(g) = g \mid K_0$.

So, to prove that ψ is injective, it suffices to prove that n is trivial (that is that $Ker(n) = Aut(L)$), or equivalently that n_0 is trivial. But we know exactly what is $Aut(K_0)$: it is (isomorphic to) $\hat{\mathbb{Z}}$, the profinite completion of $\mathbb{Z}$ and we will finish the proof by proving that there is no non-trivial homomorphism from $\hat{\mathbb{Z}}$ to $\mathbb{Z}$.

Indeed, $\hat{\mathbb{Z}}$ is isomorphic to $\prod_{q \text{ prime}} \mathbb{Z}_q$ where $\mathbb{Z}_q$ denotes the group of q-adic integers. Let P_1 and P_2 be a partition of the set of prime numbers into two infinite sets. Then $\hat{\mathbb{Z}} = \prod_{q \in P_1} \mathbb{Z}_q \times \prod_{q \in P_2} \mathbb{Z}_q$. Let $a \in \prod_{q \in P_1} \mathbb{Z}_q$. Then a is divisible by any prime number in P_2, and so is $h(a)$, if h is any homomorphism from $\hat{\mathbb{Z}}$ to $\mathbb{Z}$, and thus $h(a) = 0$. Similarly if $a \in \prod_{q \in P_2} \mathbb{Z}_q$.

Now for the generalisation to working over an arbitrary subfield K_0 as in Theorem 15. The proof is about the same, we just have to use the full strength of Theorem A of [2] in place of Theorem 3, and results in [3] and [4] give us the small index properties we require. We always work with subfields of L containing K_0 and work with transcendence rank over K_0.

As a final remark, we sketch an alternative approach to the uncountable case of Theorem 1 (and Theorem 15) which uses the full strength of Theorem A of [2], but avoids Lemmas 7 and 8. Let $\alpha \in Aut(Aut(L))$. Then α preserves $\mathcal{H}_2$ as before, and without loss we may assume that α stabilises some $Aut(L/K^0) \in \mathcal{H}_2$. So α acts on $\Omega_0 = \{K \in \Omega; K \geq K^0\}$ preserving the geometry. By Theorem A of [2] there exists $g \in Aut(L/\{K^0\})$ such that for all $K \in \Omega_0$, $\beta = \gamma(g)^{-1}\alpha$

stabilises $Aut(L/K)$. An argument as in Lemma 13 (supplemented by the argument from above in the non-zero characteristic case) shows that β fixes every element of $Aut(L/\{K^0\})$. Then one notes that if $H, H' \in \mathcal{H}_2$ are not contained in $Aut(L/K^0)$ then $H \cap Aut(L/\{K^0\}) = H' \cap Aut(L/\{K^0\})$ if and only if $H = H'$. Thus β stabilises every $H \in \mathcal{H}_2$ and so (by Lemma 6) stabilises every $Aut(L/\{K\})$, for $K \in \Omega$. It follows that β is the identity, that is, $\alpha = \gamma(g)$.

References

[1] C. J. Ash and John W. Rosenthal: Intersections of algebraically closed fields, *Annals of Pure and Applied Logic* 30 (1986), pp. 103-119.

[2] David Evans and Ehud Hrushovski: The automorphism group of the combinatorial geometry of an algebraically closed field, *J. London Math. Soc.*, (2), 52 (1995), pp. 209-225.

[3] Daniel Lascar: Les automorphismes d'un ensemble fortement minimal, *Journal of Symbolic Logic*, vol. 57 (March 1992) pp. 238-251.

[4] Daniel Lascar and Saharon Shelah: Uncountable saturated structures have the small index property, *Bull. London Math. Soc.*, 25 (1993), pp.125-131.

[5] Daniel Lascar: The group of automorphisms of the complex numbers leaving fixed the algebraic numbers is simple, *This volume.*

Authors' addresses:
David M. Evans,
School of Mathematics,
UEA,
Norwich NR4 7TJ,
England.
e-mail: d.evans@uea.ac.uk

Daniel Lascar,
UFR de Mathématiques, Case 7012,
Université Paris 7,
2 place Jussieu,
75251 Paris, cédex 05,
France.
e-mail: lascar@logique.jussieu.fr

The algebra of an age

Peter J. Cameron

Abstract

Associated with any oligomorphic permutation group G, there is a graded algebra $\mathcal{A}^G$ such that the dimension of its nth homogeneous component is equal to the number of G-orbits on n-sets. I show that the algebra is a polynomial algebra (free commutative associative algebra) in some cases, and pose some questions about transitive extensions.

1 The algebra

Let Ω be an infinite set. Let $\binom{\Omega}{n}$ denote the set of n-element subsets of Ω, V_n the vector space of functions from $\binom{\Omega}{n}$ to $\mathbb{Q}$. Set $\mathcal{A} = \bigoplus_{n\geq 0} V_n$, with multiplication defined as follows: for $f \in V_n$, $g \in V_m$, and $X \in \binom{\Omega}{n+m}$,

$$(fg)(X) = \sum_{Y \in \binom{X}{n}} f(Y)g(X \setminus Y).$$

This is the *reduced incidence algebra* of the poset of finite subsets of Ω (Rota [13]). It is a commutative and associative algebra with identity, but is far from an integral domain: any function with finite support is nilpotent.

Now, if G is any permutation group on Ω, let $\mathcal{A}^G = \bigoplus_{n\geq 0} V_n^G$, where V_n^G consists of the functions in V_n which are G-invariant (where G acts on V_n in the natural way: $f^g(X) = f(Xg^{-1})$). Now a function in V_n is fixed by G if and only if it is constant on the G-orbits. So, if G is *oligomorphic* (that is, G has only finitely many orbits on n-sets for all n), then $\dim(V_n^G) = f_n(G)$ is the number of orbits of G on $\binom{\Omega}{n}$.

If G has a finite orbit, then $\mathcal{A}^G$ contains non-zero nilpotents. I *conjecture* that conversely, if G has no finite orbits, then $\mathcal{A}^G$ is an integral domain. This question arose originally in studying the rate of growth of the numbers $f_n(G)$ for oligomorphic groups. The only evidence for it, apart from the fact that no counterexamples are known, is the following observation. Let $f \in V_n$ and $g \in V_m$ be such that $fg \neq 0$. Let X and Y be sets in the support of f and g respectively. By the Separation Lemma (Neumann [10], Lemma 2.3), if G has no finite orbits, then there is a translate Y' of Y such that $X \cap Y' = \emptyset$. Now we have a non-zero contribution to $(fg)(X \cup Y')$, though this may be cancelled out by other terms in the sum.

There is a stronger form of the conjecture, as follows. Let e be the constant function in V_1 with value 1. It is known that e is a non-zero-divisor in $\mathcal{A}$, and lies

in $\mathcal{A}^G$ for any group G. (This implies that multiplication by e is a monomorphism from V_n^G to V_{n+1}^G, and hence that $f_{n+1}(G) \geq f_n(G)$ for any n: see Cameron [1].) I *conjecture* that, if G has no finite orbits, then e is prime in $\mathcal{A}^G$, in the sense that if $e|fg$ then $e|f$ or $e|g$. This would imply that $\mathcal{A}^G$ is an integral domain.

There is a combinatorial version of this algebra, defined as follows. Let $\mathcal{C}$ be a class of finite relational structures closed under isomorphism and under taking induced substructures. Let $V_n(\mathcal{C})$ be the vector space of functions from the isomorphism types of n-element structures in $\mathcal{C}$ to $\mathbb{Q}$, and $\mathcal{A}(\mathcal{C}) = \bigoplus_{n\geq 0} V_n(\mathcal{C})$, with multiplication defined just as before.

The *age* of a relational structure M on Ω is the class of all finite structures embeddable in M as induced substructures. M is *homogeneous* if every isomorphism between finite induced substructures of M extends to an automorphism of M. Now we have:

- If $\mathcal{C}$ is the age of a relational structure M on Ω, then $\mathcal{A}(\mathcal{C})$ is a subalgebra of the reduced incidence algebra $\mathcal{A}$ on Ω (and this is equivalent to $\mathcal{C}$ having the *joint embedding property*, that is, any two members of $\mathcal{C}$ can be simultaneously embedded in a member of $\mathcal{C}$).

- If $\mathcal{C}$ is the age of a homogeneous relational structure M on Ω, then $\mathcal{A}(\mathcal{C}) = \mathcal{A}^G$, where $G = \mathrm{Aut}(M)$ (and this is equivalent to $\mathcal{C}$ having the *amalgamation property*, that is, any amalgam of two members of $\mathcal{C}$ with a common substructure can be embedded in a member of $\mathcal{C}$).

See, for example, Cameron [3] for discussion.

2 Polynomial algebras

There are only two techniques I know for determining the structure of the algebras $\mathcal{A}^G$ or $\mathcal{A}(\mathcal{C})$. The first is based on the simple observation that, regarding $G \times H$ as a permutation group on the disjoint union of the sets on which G and H act, we have

$$\mathcal{A}^{G\times H} = \mathcal{A}^G \otimes_{\mathbb{Q}} \mathcal{A}^H.$$

Let S denote the symmetric group on an infinite set. Then $\mathcal{A}^S$ is a polynomial ring in one variable (generated by the element e). Hence $\mathcal{A}^{S^n}$ is a polynomial algebra in n variables.

Now let H be a finite permutation group of degree n. Then the *wreath product* $S\mathrm{Wr}H$ is the semidirect product of S^n by H, and so $\mathcal{A}^{S\mathrm{Wr}H}$ consists of the invariants of H in the polynomial algebra (in the classical sense, where H acts as a linear group by permutation matrices). For example, if H is the symmetric group S_n, then $\mathcal{A}^{S\mathrm{Wr}S_n}$ is the polynomial algebra generated by the n elementary symmetric functions, by Newton's Theorem. (Note that $\mathcal{A}^{S\mathrm{Wr}H}$ is always an integral domain, but almost never a polynomial algebra.)

In this case, the numbers $f_n(S \operatorname{Wr} H)$ can be calculated by Molien's Theorem, which turns out to be a special case of a "cycle index theory" for oligomorphic permutation groups (see [3]).

The second approach requires that the class $\mathcal{C}$ has a "good notion of connectedness", as follows. I will give an axiomatic treatment, since in one of the examples below, words like "connected" and "involvement" have meanings quite different from their usual ones. We require

- a distinguished subclass of $\mathcal{C}$ consisting of "connected" structures;
- a partial order $\leq$ called "involvement" on the class of n-element structures for each n;
- a binary, commutative and associative "composition" $\circ$ such that, if X and Y are structures with n and m points respectively, then $X \circ Y$ is a structure with $n + m$ points.

Assume that the following conditions hold:

A1 Let S be a structure which is partitioned into disjoint induced substructures $S_1, S_2, \ldots$. Then $S_1 \circ S_2 \circ \ldots \leq S$.

A2 Any structure has a unique representation as a composition of connected structures.

Theorem 2.1 *If all the above conditions hold, then $\mathcal{A}(\mathcal{C})$ is a polynomial algebra, generated by the characteristic functions of the connected structures.*

Proof. If $|S| = n$, then S is a disjoint union $S_1 \cup S_2 \cup \ldots$ of connected structures; so we have a bijection between characteristic functions χ_S (the basis elements of $V_n(\mathcal{C})$) and monomials $\phi_S = \chi_{S_1}\chi_{S_2}\ldots$ of total weight n. Consider the matrix expressing the monomials ϕ_S in terms of the basis elements $\chi_{S'}$. The coefficient of χ_S in the row corresponding to ϕ_S is non-zero. Suppose that $\chi_{S'}$ also has non-zero coefficient. Then S' can be partitioned into induced substructures isomorphic to $S_1, S_2, \ldots$; so $S = S_1 \circ S_2 \circ \ldots \leq S'$. Thus the matrix is upper triangular with non-zero diagonal, and hence invertible. So the monomials of weight n form a basis for $V_n(\mathcal{C})$, and the theorem is proved.

Example 1. Let M be the countable "random graph" [4], whose age $\mathcal{C}$ is the class of all finite graphs. Let "connected" have its usual meaning, "involvement" mean "spanning subgraph", and "composition" be disjoint union (with no edges between the parts). Then A1 and A2 hold, and so $\mathcal{A}(\mathcal{C}) = \mathcal{A}^{\mathrm{Aut}(M)}$ is a polynomial algebra, whose generators correspond to the finite connected graphs.

This method works for many other ages, both of homogeneous structures (for example, the class of K_n-free graphs for fixed n [8]), and not (for example, bipartite graphs, N-free graphs [5]).

Example 2. Let $\mathcal{C}$ be the age of a homogeneous structure M, and let $G = \mathrm{Aut}(M)$. Let $\mathcal{C}'$ be the class of structures over a language with the relation symbols for $\mathcal{C}$ and one new binary symbol E, in which E is an equivalence relation each of whose classes carries a $\mathcal{C}$-structure (with no instances of relations holding between points in different E-classes). Then $\mathcal{C}'$ is the age of a homogeneous structure consisting of the disjoint union of countably many copies of M, with automorphism group $G \mathrm{Wr} S$, where S is the symmetric group of countable degree. Now let "connected" mean "only one E-class", "involvement" mean "inclusion of all relations", and "composition" mean "disjoint union". Then A1 and A2 hold.

The conclusion is that $\mathcal{A}^{G\mathrm{Wr}S}$ is always a polynomial algebra; the number of generators of degree n is equal to the number of orbits of G on n-sets.

Example 3. Let A be a fixed alphabet of finite size q, and let $\mathcal{C} = A^*$ be the set of words in A. (Here a word of length n is regarded as an n-set carrying a total order and q unary relations $R_1, \ldots, R_q$, where each element of the set satisfies exactly one of the unary relations; the word $a_1 a_2 \ldots a_q$ corresponds to the n-set $\{x_1, \ldots, x_n\}$, with $x_1 < x_2 < \ldots < x_n$ and in which x_i satisfies R_{a_i}.) The algebra $\mathcal{A}(A^*)$ is the *shuffle algebra* which arises in the theory of free Lie algebras [12]. The name comes from the fact that the product of two words is the sum of all words which can be obtained by "shuffling" them together, with appropriate multiplicities. For example,

$$(aab) \cdot (ab) = abaab + 3aabab + 6aaabb.$$

Also, A^* is the age of a homogeneous relational structure $M(q)$ which is order-isomorphic to $\mathbb{Q}$ and in which the set of elements satisfying each relation R_i is dense; in other words, a partition of $\mathbb{Q}$ into q dense subsets. Such a partition is unique up to order-isomorphism of $\mathbb{Q}$. Let $G(q) = \mathrm{Aut}(M(q))$.

Take a total order on A, and define the *lexicographic order* on A^* in the usual way: that is, $a_1 \ldots a_m < b_1 \ldots b_n$ if and only if *either*

- $m < n$, and $a_i = b_i$ for $i = 1, \ldots, m$; *or*
- for some $l < \min\{m, n\}$, we have $a_i = b_i$ for $i = 1, \ldots, l$, and $a_{l+1} < b_{l+1}$.

A non-empty word $w \in A^*$ is a *Lyndon word* if, whenever $w = xy$ with x, y non-empty, we have $w < y$; that is, w is less than any proper cyclic shift of itself. The number of Lyndon words of length n is $(1/n)\sum_{d|n} \mu(d) q^{n/d}$, where μ is the Möbius function. (This well-known number counts several other things, for example, irreducible polynomials over $\mathbb{F}_q$ if q is a prime power; see [12].) The following combinatorial properties hold for Lyndon words:

Lemma 2.2 *(i) Any word w has a unique expression in the form $w = w_1 w_2 \ldots$, where $w_1, w_2, \ldots$ are Lyndon words with $w_1 \geq w_2 \geq \ldots$.*

(ii) Given Lyndon words $w_1, w_2, \ldots$ with $w_1 \geq w_2 \geq \ldots$, the lexicographically greatest shuffle of these words is the concatenation $w_1 w_2 \ldots$.

Hence, if we let "connected" mean "Lyndon word", "involvement" mean "lexicographic order reversed", and "composition" mean "concatenation in decreasing lexicographic order", then A1 and A2 hold, and we conclude that $\mathcal{A}(A^*) = \mathcal{A}^{G(q)}$ is a polynomial algebra generated by the Lyndon words (a result of Radford [11]).

3 Transitive extensions

Not much is known in general about how the algebra $\mathcal{A}^G$ is affected by group-theoretic or model-theoretic constructions (direct products with product action, wreath products, covers and quotients, etc.). This section contains some comments about transitive extensions.

The permutation group H on Ω is a *transitive extension* of G if H is transitive and the stabiliser H_α of the point α, acting on $\Omega \setminus \{\alpha\}$, is isomorphic to G as permutation group. Note that, in this situation, H is closed if and only if G is closed.

A general question: *Let H be a transitive extension of G. What is the relation between $\mathcal{A}^H$ and $\mathcal{A}^G$?*

We can regard the group induced on Ω by G as the direct product of G (in its given action) with the trivial group of degree 1. For the latter group (K, say), the algebra $\mathcal{A}^K$ is generated by an element k of degree 1 with $k^2 = 0$. In other words, $\mathcal{A}^K \cong T(\mathbb{Q})$, the algebra of 2×2 upper triangular matrices with constant diagonal over $\mathbb{Q}$. Hence, using G^+ for the group induced on Ω by G, we have

$$\mathcal{A}^{G^+} \cong \mathcal{A}^G \otimes_{\mathbb{Q}} T(\mathbb{Q}) \cong T(\mathcal{A}^G).$$

However, we can only say that, since $G^+ \leq H$, the algebra $\mathcal{A}^H$ is a subalgebra of $T(\mathcal{A}^G)$. This does not seem to help to decide, for example, whether $\mathcal{A}^H$ is an integral domain.

There is a special class of transitive extensions for which a bit more can be said. We say that the transitive extension H of G is *curious* if H has a transitive subgroup (on the whole of Ω) which is isomorphic to G. In the case where G and H are closed, this means that H is a reduct of G. If H is a curious transitive extension of G, then $\mathcal{A}^H$ is a subalgebra of $\mathcal{A}^G$; in particular, $\mathcal{A}^H$ is an integral domain if $\mathcal{A}^G$ is. Perhaps it is possible to weave together the embeddings of $\mathcal{A}^H$ in $\mathcal{A}^G$ and in $T(\mathcal{A}^G)$ to get better information.

Example 1 (continued). A *two-graph* on Ω is a set T of 3-element subsets of Ω such that any 4-subset contains an even number of members of T (Seidel [14]).

Given a graph Γ on Ω, let $T(\Gamma)$ be the set of *odd triples* of Γ (those containing an odd number of edges). Then $T(\Gamma)$ is a two-graph on Ω. Every two-graph arises in this way.

Let R be the random graph on Ω_0. Take a new point ∞, and define T to be the two-graph on $\Omega = \Omega_0 \cup \{\infty\}$ derived from R (with ∞ as an isolated vertex). Then $\mathrm{Aut}(T)$ is a transitive extension of $\mathrm{Aut}(R)$. Moreover, it is curious; for the two-graph derived from R without an isolated vertex is clearly a reduct of R, and

is isomorphic to T. (In fact, T is the unique countable universal homogeneous two-graph.) See Thomas [16].

Problem. Is $\mathcal{A}^{\mathrm{Aut}(T)}$ a polynomial algebra?

Remark. Mallows and Sloane [9] showed that the numbers of two-graphs and *even graphs* (graphs with all valencies even) on n points are equal. Hence, if $\mathcal{A}^{\mathrm{Aut}(T)}$ is a polynomial algebra, then its generators are in one-to-one correspondence (preserving degree) with the finite *Eulerian graphs* (the connected even graphs).

Example 3 (continued). Let $G(q)$ be as in Example 3 in the preceding section. Then $G(q)$ has a transitive extension $H(q)$ defined as follows.

On the set of complex roots of unity, put $z_1 \equiv z_2$ if $z_2 z_1^{-1}$ is a qth root of unity. Let Ω be a dense subset containing exactly one member of each equivalence class of this relation. (Such a set is unique up to permutation preserving the cyclic order. If we choose a random member of each class, the resulting set almost surely has this property.) Now define binary relations $R_1, R_2, \ldots, R_q$ by $(z_1, z_2) \in R_j$ if and only if

$$\frac{2\pi(j-1)}{q} < \arg(z_2 z_1^{-1}) < \frac{2\pi j}{q}.$$

The structure $N(q)$ consists of the circular order and the relations $R_1, R_2, \ldots, R_q$. It is $\aleph_0$-categorical. Note that, if $z_1 \neq z_2$, then $(z_1, z_2) \in R_j$ for a unique value of j; and the converse of R_j is R_{q+1-j}. Let $H(q) = \mathrm{Aut}(N(q))$.

Now take $z \in \Omega$. Define a map $\phi : \Omega \setminus \{z\} \to (0,1)$ by letting $\phi(w)$ be the fractional part of $\frac{q}{2\pi} \arg(zw^{-1})$. Then $\phi(\Omega \setminus \{z\}) = (0,1) \cap \mathbb{Q}$. If we give $\phi(w)$ the colour j if $(z,w) \in R_j$, then each colour class is dense. Moreover, the structure $N(q)$ can be recovered uniquely from this information. So $H(q)$ is a transitive extension of $G(q)$.

This extension is also curious. If we repeat the above construction, but with z a point on the unit circle which is not a root of unity, we obtain a bijection from all of Ω to a countable dense subset of $(0,1)$ partitioned into q dense subsets.

Problem. Is $\mathcal{A}^{H(q)}$ a polynomial algebra?

Remark. For $q = 2$, the relations R_1 and R_2 are a converse pair of tournaments, each of which is isomorphic to the countable universal homogeneous *local order* [2], *locally transitive tournament* [7], or *vortex-free tournament* [6]: these are three alternative names for a tournament having no subtournament consisting of a directed 3-cycle dominating or dominated by a vertex. This structure is further discussed in the lectures of Evans, Ivanov and Macpherson.

Orbits of $H(q)$ on n-sets are parametrised by two-way infinite "shift register sequences" (x_i) with elements in $\{1, \ldots, q\}$ satisfying $x_i + n \equiv x_i + 1 \pmod{q}$ for all i. For $q = 2$, the sequences counting these orbits is listed as M0324 in the *Encyclopedia of Integer Sequences* [15], where further references can be found.

On the assumption that $\mathcal{A}^{H(2)}$ is a polynomial algebra, it is possible to compute the numbers of generators of each degree. The resulting sequence appears to be "unknown"; in particular, it is not in the Encyclopedia [15].

The group $H(2)$ does not have a transitive extension. Nevertheless, the following occurrence is suggestive.

Knuth [6] defines a *CC-structure* to be a set with a ternary relation satisfying five universal axioms, of which the first three assert that the induced structure on any 3-set is a circular order. The letters CC stand for "counter-clockwise"; and, given a set Ω of points in the Euclidean plane with no three collinear, the relation R such that $R\alpha\beta\gamma$ holds if and only if the points α, β, γ occur in the counter-clockwise sense, is a CC-structure. Such a CC-structure is called *representable*. There is a countable universal representable CC-structure, defined by choosing a countable dense set of points in the Euclidean plane with no three collinear. It is not homogeneous; indeed, the class of CC-structures (or of representable CC-structures) does not have the amalgamation property.

Given a ternary relation R on Ω whose restriction to any 3-set is a circular order, there is a derived tournament R_α on $\Omega \setminus \{\alpha\}$ defined by $R_\alpha\beta\gamma \Leftrightarrow R\alpha\beta\gamma$. Knuth's fifth axiom for CC-structures implies that R_α is a local order for any point α. Indeed, if we take the universal representable CC-structure above, and project $\Omega \setminus \{\alpha\}$ radially onto the unit circle with centre α, we obtain the homogeneous local order $N(2)$.

Problem. Do there exist countable CC-structures (or representable ones) with large automorphism groups, or with other nice model-theoretic properties?

Acknowledgment. I am grateful to R. A. Bailey, R. M. Bryant and D. G. Fon-Der-Flaass for their help with the contents of this paper.

References

[1] P. J. Cameron, Transitivity of permutation groups on unordered sets, *Math. Z.* **48** (1976), 127–139.

[2] P. J. Cameron, Orbits of permutation groups on unordered sets, II, *J. London Math. Soc.* (2) **23** (1981), 249–265.

[3] P. J. Cameron, *Oligomorphic Permutation Groups*, London Math. Soc. Lecture Notes **152**, Cambridge University Press, Cambridge, 1990.

[4] P. J. Cameron, The random graph, pp. 333–351 in *The Mathematics of Paul Erdős, II* (ed. R. L. Graham and J. Nešetřil), Algorithms and Combinatorics **14**, Springer, Berlin, 1997.

[5] J. Covington, A universal structure for N-free graphs, *Proc. London Math. Soc.* (3), **58** (1989), 1–16.

[6] D. E. Knuth, *Axioms and Hulls*, Lecture Notes in Computer Science **606**, Springer, Berlin, 1992.

[7] A. H. Lachlan, Countable homogeneous tournaments, *Trans. Amer. Math. Soc.* **284**, 431–461.

[8] A. H. Lachlan and R. E. Woodrow, Countable ultrahomogeneous undirected graphs, *Trans. Amer. Math. Soc.* **262** (1980), 51–94.

[9] C. L. Mallows and N. J. A. Sloane, Two-graphs, switching classes, and Euler graphs are equal in number, *SIAM J. Appl. Math.* **28** (1975), 876–880.

[10] P. M. Neumann, The lawlessness of finitary permutation groups, *Arch. Math.* **26** (1975), 561–566.

[11] D. E. Radford, A natural ring basis for the shuffle algebra and an application to group schemes, *J. Algebra* **58** (1979), 432–454.

[12] C. Reutenauer, *Free Lie Algebras*, London Math. Soc. Monographs (New Series) **7**, Oxford University Press, 1993.

[13] G.-C. Rota, On the foundations of combinatorial theory, I: Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie* **2** (1964), 340–368.

[14] J. J. Seidel, A survey of two-graphs, pp. 481–511 in *Proc. Int. Colloq. Teorie Combinatorie*, Accad. Naz. Lincei, Roma, 1977.

[15] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, New York, 1995.

[16] S. R. Thomas, Reducts of the random graph, *J. Symbolic Logic* **56** (1991) 176–181.

Author's address:
School of Mathematical Sciences,
Queen Mary and Westfield College,
London E1 4NS,
England.
e-mail: P.J.Cameron@qmw.ac.uk

Elimination of inverses in groups

Maurice Boffa

1 Terminology

I say that a group G *satisfies a formula* $\varphi(\overline{x})$ (with free variables $\overline{x}$) if it satisfies the sentence $\forall\overline{x}\varphi(\overline{x})$. In this case, I simply write $G \vDash \varphi(\overline{x})$ instead of $G \vDash \forall\overline{x}\varphi(\overline{x})$.
An *identity* is a *non-trivial* atomic formula of the language of groups $L_g = \{\cdot, ^{-1}, 1\}$, i.e. one which can be put (using group axioms) in the form $t(\overline{x}) = 1$ where $t(\overline{x})$ is a *non-trivial* element of the free group on $\overline{x}$.
A *monoidal identity* is a *non-trivial* atomic formula of the language of monoids $L_m = \{\cdot, 1\}$, i.e. one of the form $\alpha(\overline{x}) = \beta(\overline{x})$ where $\alpha(\overline{x})$ and $\beta(\overline{x})$ are *distinct* elements of the free monoid on $\overline{x}$.

2 Two facts and an open question

Fact 1. *If a group satisfies an identity (resp. a monoidal identity), then it satisfies an identity (resp. a monoidal identity) in 2 variables.*

It suffices to replace each variable x_i by $y^i x y^{-i}$ (resp. xy^i).

Fact 2. *If a group satisfies a finite disjunction of identities, then it satisfies an identity.*

It suffices to show that if $t_1(\overline{x}) = 1$ and $t_2(\overline{x}) = 1$ are identities, then there is an identity $t_3(\overline{x}) = 1$ such that any group satisfies

$$(t_1(\overline{x}) = 1 \vee t_2(\overline{x}) = 1) \rightarrow t_3(\overline{x}) = 1.$$

If $t_1(\overline{x})$ and $t_2(\overline{x})$ don't commute in the free group on $\overline{x}$, then we take $t_3(\overline{x}) = [t_1(\overline{x}), t_2(\overline{x})]$; if they commute, then they are powers of a common element, i.e. $t_1(\overline{x}) = u(\overline{x})^{n_1}$ and $t_2(\overline{x}) = u(\overline{x})^{n_2}$, and we take $t_3(\overline{x}) = u(\overline{x})^{n_1 n_2}$.

As we shall see later, the following question (formulated in [1], [11]) has a positive answer for several classes of groups, but remains open in its full generality :

(Q) *If a group satisfies a finite disjunction of monoidal identities, does it satisfy a monoidal identity ?*

3 Model theoretic characterization of groups satisfying an identity

Let * denote the ultrapower construction modulo a (fixed) nonprincipal ultrafilter over ω. I refer the reader not versed in logic to the *Handbook of Mathematical Logic* (edited by J. Barwise, North-Holland, 1977) for the model-theoretic notions which I require, in particular for ultrapowers.

Theorem 1 ([8]). *For a group G, the following are equivalent :*

(1) *G satisfies an identity.*

(2) *$G^\star$ has no free subgroup of rank 2.*

(1) $\rightarrow$ (2) is obvious; the converse holds since (2) means that $G^\star$ satisfies the (countably infinite) disjunction of all identities in 2 variables, thus (since $G^\star$ is ω_1-saturated) that G satisfies a finite disjunction of such identities, and (by Fact 2) this implies (1).

Corollary 1. *For a commutative field K, the following are equivalent :*

(1) *$SL_2(K)$ satisfies an identity.*

(2) *K is finite.*

This holds since $(SL_2(K))^\star$ is canonically isomorphic to $SL_2(K^\star)$ and since, for K infinite, $K^\star$ contains two algebraically independent elements (over the prime subfield), so that (by exercise 2.2 of [14]) $SL_2(K^\star)$ has a free subgroup of rank 2.

Example. Let K be the algebraic closure of a finite field. Then $SL_2(K)$ is an example of group which satisfies no identity and which has no free subgroup of rank 2. This shows that $G^\star$ cannot be replaced by G in condition (2) of Theorem 1.

Corollary 2. *For a group G, the following are equivalent :*

(1) *G satisfies an identity.*

(2) *G^ω has no free subgroup of rank 2.*

(1) $\rightarrow$ (2) is obvious; the converse holds since (2) implies that $G^\star$ (which is a quotient of G^ω) has no free subgroup of rank 2.

Let $t_i(x, y) = 1$ $(i \in \omega)$ enumerate all identities in 2 variables. The previous result has a direct proof : if G satisfies no identity, then choose $x_i, y_i \in G$ such that $t_i(x_i, y_i) \neq 1$, so that (x_i), (y_i) generate a free subgroup of G^ω. A similar argument for monoidal identities gives

Theorem 2. *For a group G, the following are equivalent :*

(1) *G satisfies a monoidal identity.*

(2) *G^ω has no free submonoid of rank 2.*

For finite disjunctions of monoidal identities, the line of proof of Theorem 1 gives

Theorem 3. *For a group G, the following are equivalent :*

(1) *G satisfies a finite disjunction of monoidal identities.*

(2) *G^* has no free submonoid of rank 2.*

From Theorems 2 and 3, we get a new formulation of question (Q) :

If G^ω has a free submonoid of rank 2, does G^* also have a free submonoid of rank 2 ?

4 Elimination of inverses

Here is my terminology, where e.i. (resp. s.e.i.) means "elimination of inverses" (resp. "strong elimination of inverses"). Given a group G and an *atomic* formula $\varphi(\overline{x})$ of L_g, I will say that G has *e.i.* (resp. *s.e.i.*) *for* $\varphi(\overline{x})$ if there is an *open* (resp. *atomic*) formula $\theta(\overline{x})$ of L_m such that $G \vDash (\varphi(\overline{x}) \leftrightarrow \theta(\overline{x}))$. Here, *open* means *quantifier-free.*
Given a group G, I will say that G has *e.i.* (resp. *s.e.i.*) if it has e.i. (resp. s.e.i.) for all atomic formulas of L_g.

In other words :

G has e.i. iff	each *open* formula of L_g is equivalent in G to an *open* formula of L_m;
G has s.e.i. iff	each *atomic* formula of L_g is equivalent in G to an *atomic* formula of L_m.

These notions are closely linked to monoidal identities, as shown by the following basic results :

Theorem 4 ([1]). *For a group G, the following are equivalent :*

(1) *G has e.i. for the formula $x^{-1}y = y^{-1}z$.*

(2) *G satisfies a finite disjunction of monoidal identities (in 2 variables).*

(3) *G has e.i.*

Theorem 5. *For a group G, the following are equivalent :*

(1) *G has s.e.i. for the formula $x^{-1}y = y^{-1}z$.*

(2) *G satisfies a monoidal identity (in 2 variables).*

(3) *G has s.e.i.*

We first consider Theorem 5.
(1) $\rightarrow$ (2), since if $x^{-1}y = y^{-1}z$ is equivalent in G to an atomic formula $\theta(x,y,z)$ of L_m then :

either θ is trivial and then G satisfies the monoidal identity $x = 1$,

or θ is non-trivial and then G satisfies the monoidal identity $\theta(x, xy, xy^2)$.

(2) $\rightarrow$ (3), since in a group any monoidal identity in 2 variables can be put in the form $\gamma(x,y)x = \delta(x,y)y$, that is $xy^{-1} = \gamma(x,y)^{-1}\delta(x,y)$, which allows to put each term $t(\overline{x})$ of L_g in the form $\gamma(\overline{x})^{-1}\delta(\overline{x})$ where $\gamma(\overline{x})$ and $\delta(\overline{x})$ are terms of L_m.

Theorem 4 has a more complicated proof. Here is a sketch (see [1] for details).
(1) $\rightarrow$ (2), since if $x^{-1}y = y^{-1}z$ is equivalent in G to a boolean combination of monoidal identities $\alpha_i(x,y,z) = \beta_i(x,y,z)$ then G satisfies the disjunction of the following monoidal identities :

$$y = 1, \alpha_i(x, xy, xy^2) = \beta_i(x, xy, xy^2), \alpha_i(x, xy, xy^3) = \beta_i(x, xy, xy^3).$$

(2) $\rightarrow$ (3), since (2) implies that in G each term $t(\overline{x})$ of L_g can be in some sense locally written in the form $\gamma(\overline{x})^{-1}\delta(\overline{x})$ with $\gamma(\overline{x})$ and $\delta(\overline{x})$ in L_m.

Corollary. *e.i. and s.e.i. are preserved with respect to forming subgroups, homomorphic images, finite extensions, and finite products. Moreover, s.e.i. is preserved with respect to forming arbitrary powers.*

This is obvious except for the case of finite products, which can be obtained from the fact that a finite product of groups which have no free submonoid of rank 2 has itself no free submonoid of rank 2 (see [9], proposition 4.21).

Our question (Q) has now two new formulations :

If a group has e.i., does it have s.e.i. ?
Is e.i. preserved with respect to forming arbitrary powers ?

Remark. Since two elements of a free group either are free or commute, we see that a free group has s.e.i. for any atomic formula of L_g in 2 variables. This explains why we need more than 2 variables in the formula of condition (1) of Theorems 4 and 5.

For $n \in \omega$, let $I_n(x,y)$ denote the monoidal identity defined inductively as follows : $I_0(x,y)$ is $x = y$ and $I_{n+1}(x,y)$ is $I_n(xy, yx)$. These monoidal identities were first considered by A.I. Shirshov [12] under the name of ν-identities, and rediscovered in [3] under the name of Thue-Morse identities.

Theorem 6. *Any nilpotent group of class $\leqslant n$ satisfies $I_n(x,y)$. Consequently, any nilpotent group has s.e.i..*

The proof (by induction on n) results from the following facts (where Z denotes the centre of G) :

(i) if I_n is $\alpha = \beta$, then I_{n+1} is $\alpha\beta = \beta\alpha$,

(ii) if $G/Z \vDash (\alpha = \beta)$, then $G \vDash (\alpha\beta = \beta\alpha)$.

Example. Let us show how a nilpotent group G of class $\leqslant 2$ has s.e.i. for the formula $x^{-1}y = y^{-1}z$. Since G satisfies $I_2(x, y)$, i.e. $xyyx = yxxy$, i.e. $x^{-1}y = yyx(xxy)^{-1}$, we get (in G) :

$$x^{-1}y = y^{-1}z \leftrightarrow yyyx = zxxy.$$

5 Linear groups with elimination of inverses

By definition a *linear group* (over a commutative field K) is a subgroup of $GL_n(K)$ for some positive integer n. Each linear group G carries the topology induced by the *Zariski topology* of K^{n^2}, which is *noetherian* (i.e. satisfies the descending chain condition for closed sets), and the following holds (for details see [14], ch.5 and 14) :

(i) G has only finitely many connected components and these are also its irreducible components;

(ii) the *identity component* G^0 is a closed normal subgroup of finite index in G;

(iii) the connected components of G are exactly the cosets of G^0 in G.

As mentioned in [1], J. Tits has pointed out to me that the irreducibility of $G^0 \times G^0$ $(= (G \times G)^0)$ entails the following result :

Theorem 7. *Question (Q) has a positive answer for linear groups.*

Indeed, if a linear group G satisfies a finite disjunction of monoidal identities $\alpha_i(x, y) = \beta_i(x, y)$, then $G^0 \times G^0$ is the (finite) union of the closed subsets

$$\left\{(x, y) \in G^0 \times G^0 \;\middle|\; \alpha_i(x, y) = \beta_i(x, y)\right\}$$

and its irreducibility shows that $G^0 \vDash (\alpha_i(x, y) = \beta_i(x, y))$ for some i, so that $G \vDash (\alpha_i(x^m, y^m) = \beta_i(x^m, y^m))$ where m is the index of G^0 in G.

A stronger result is the following definite characterization of linear groups with e.i. :

Theorem 8 ([2]). *For a linear group G, the following are equivalent :*

(1) *G has e.i..*

(2) *G is nilpotent-by-finite (equivalently : G^0 is nilpotent).*

Here is the line of proof :

(a) A linear group with e.i. is soluble-by-finite.
This follows from a result of Platonov (see 10.15 of [14]) : a linear group which satisfies an identity is soluble-by-finite.

(b) A finitely generated linear group with e.i. is nilpotent-by-finite.
This follows from (a) and a result of Rosenblatt [9] : a finitely generated soluble group which has no free submonoid of rank 2 is nilpotent-by-finite.

(c) A linear group G with e.i. is nilpotent-by-periodic.
Indeed, G is a subgroup of some $GL_n(K)$ and (b) implies that H^0 is nilpotent (of class necessarily $\leqslant n$) for any finitely generated subgroup H of G. It follows that $\varinjlim(H^0)$ is a nilpotent normal subgroup N of G such that G/N is periodic.

(d) A connected linear group G with e.i. is nilpotent.
Indeed, G is a subgroup of some $GL_n(K)$ and we may assume that K is an algebraically closed field. Its Zariski closure $\overline{G}$ in $GL_n(K)$ is then a connected linear algebraic group. Since $\overline{G}$ satisfies the same identities as G (see 10.7 of [14]), it follows from (a) and (c) that $\overline{G}$ is soluble and nilpotent-by-periodic. From the structure of the connected soluble linear algebraic groups (see 14.22 of [14]) it follows finally that $\overline{G}$ is nilpotent-by-a periodic torus.
But we may assume that K contains a transcendental element, in which case a periodic torus is necessarily trivial.

Corollary 1 (for $n \geqslant 2$). *For a subgroup G of $GL_n(K)$ the following are equivalent :*

(1) *G has e.i..*

(2) *G satisfies the monoidal identity $I_{n-1}(x^m, y^m)$ where m is the index of G^0 in G.*

This follows from Theorem 6 and the fact that (for $n \geqslant 2$) a connected nilpotent subgroup of $GL_n(K)$ is of class $\leqslant n-1$.

Corollary 2. *For a simple linear group G, the following are equivalent :*

(1) *G has e.i..*

(2) *G is finite.*

Application. In [6] it is shown that if a cancellative linear semigroup satisfies an identity then it has a group of fractions which is linear and satisfies the same identity, so that Theorem 8 can be generalized as follows : a cancellative linear semigroup satisfies an identity iff it has a group of fractions which is nilpotent-by-finite.

6 Soluble groups with elimination of inverses

Rosenblatt's result used in the proof of Theorem 8 immediately gives :

Theorem 9. *For a finitely generated soluble group G, the following are equivalent :*

(1) *G has e.i..*

(2) *G is nilpotent-by-finite.*

We cannot expect the same result for all soluble groups, since the soluble group $(S_3)^\omega$ (where S_3 is the symmetric group of degree 3) has e.i. (it satisfies $x^6 = 1$) but is not nilpotent-by-finite.

But, according to [4], there is for each n, m an open formula $\theta_{n,m}(x, y)$ of L_g which expresses that the group generated by x, y has a nilpotent normal subgroup of class $\leqslant n$ and index $\leqslant m$. And (by Theorem 9) a soluble group with e.i. (as well as its elementary extensions) satisfies the (countably infinite) disjunction of all these formulas and so (by compactness) it satisfies one of them. So we get :

Theorem 10 ([7]). *For a soluble group G, the following are equivalent :*

(1) *G has e.i..*

(2) *For some n, m : G satisfies $\theta_{n,m}(x, y)$, i.e. each 2-generated subgroup of G has a nilpotent normal subgroup of class $\leqslant n$ and index $\leqslant m$.*

Note that $\theta_{n,m}(x, y)$ implies $I_n(x^k, y^k)$ where k is the lowest common multiple of $1, 2, \ldots, m$.

Corollary 1. *For a soluble group G, the following are equivalent :*

(1) *G has e.i..*

(2) *G satisfies the monoidal identity $I_n(x^k, y^k)$ for some n, k.*

Corollary 2. *Question (Q) has a positive answer for soluble groups.*

7 Bounded elimination of inverses

This notion was first investigated by Point [7] and later by Shalev [11] (in terms of collapsing groups introduced in [10]).

Definitions ([7]).

(i) A monoidal identity (in 2 variables) $\alpha(x, y) = \beta(x, y)$ is in *normal form* of *length* ℓ if α, β have the same length ℓ and if they differ in their first letter as well as in their last letter.
For example, $I_n(x, y)$ is in normal form of length 2^n.

(ii) A group has *e.i. of complexity* $\leqslant \ell$ if it satisfies a finite disjunction of monoidal identities in normal form of length $\leqslant \ell$.

(iii) A group G has the *ℓ-Milnor property* if for all x, y in G the subgroup generated by $\left\{ y^i x y^{-i} \;\middle|\; i \in \mathbb{Z} \right\}$ is already generated by $\left\{ y^i x y^{-i} \;\middle|\; 1 \leqslant i \leqslant \ell \right\}$ (this will be called the ℓ-Milnor condition on x, y).

Remarks.

(i) From the cancellation law of groups and the fact that $\alpha = \beta \rightarrow \alpha\beta = \beta\alpha$ it follows that any monoidal identity *entails* one which is in normal form. This implies that a group has e.i. iff it has e.i. of complexity $\leqslant \ell$ for some ℓ.

(ii) The following fact has its origins in [5] and [9] : if $\alpha(x, y) = \beta(x, y)$ is in normal form of length ℓ, then $\alpha(xy, y) = \beta(xy, y)$ implies the $(\ell - 1)$-Milnor condition on x, y. The proof consists in expressing $\alpha(xy, y) = \beta(xy, y)$ in terms of elements of $\{ y^i x y^{-i} \mid i \in \mathbb{Z} \}$.

Example. $xxyx = yxxy$ is in normal form of length 4. If we replace x by xy we get $xyxyyxy = yxyxyy$, i.e. $x(yxy^{-1})(y^3xy^{-3}) = (yxy^{-1})(y^2xy^{-2})$, which implies the 3-Milnor condition on x, y.

From Remark (ii) we get

Theorem 11 ([7]). *A group with e.i. of complexity* $\leqslant \ell$ *has the* $(\ell - 1)$*-Milnor property.*

This result (together with some ingredients of the theory of finite groups, including the classification of finite simple groups) has led to

Theorem 12 ([7]). *There is a function* $e(\ell)$ *such that every finite group* G *with e.i. of complexity* $\leqslant \ell$ *has a nilpotent normal subgroup* N *such that* G/N *has exponent* $\leqslant e(\ell)$.

And Zelmanov's solution to the restricted Burnside problem (see [13] for details) immediately gives :

Corollary. *There is a function* $m(r, \ell)$ *such that every* r*-generated finite group with e.i. of complexity* $\leqslant \ell$ *has a nilpotent normal subgroup of index* $\leqslant m(r, \ell)$.

But the strongest result in this direction is the following reformulation of theorem A' of [11] :

Theorem 13 ([11]). *There are functions* $n(r, \ell)$ *and* $m(r, \ell)$ *such that every* r*-generated residually finite group with e.i. of complexity* $\leqslant \ell$ *has a nilpotent normal subgroup of class* $\leqslant n(r, \ell)$ *and index* $\leqslant m(r, \ell)$.

For $n(\ell) = n(2,\ell)$ and $k(\ell)$ = the least common multiple of $1, 2, \ldots, m(2,\ell)$ we get

Corollary 1. *Every residually finite group with e.i. of complexity* $\leqslant \ell$ *satisfies* $I_{n(\ell)}(x^{k(\ell)}, y^{k(\ell)})$.

Corollary 2. *Question (Q) has a positive answer for residually finite groups.*

Final remarks.

(i) Since finitely generated nilpotent-by-finite groups are residually finite, we can replace in Theorem 13 "residually finite" by any property X such that every finitely generated X-group with e.i. is nilpotent-by-finite (for example, X = linear or soluble). If moreover X is inherited by subgroups, then this can also be done in Corollaries 1 and 2.

(ii) There are finitely generated groups with e.i. (and even s.e.i.) which are not nilpotent-by-finite (for example, the infinite Burnside groups).

References

[1] M. Boffa, L'élimination des inverses dans les groupes, *C.R. Acad. Sc. Paris* 303 (1986), série I, p. 587-589.

[2] M. Boffa & R.M. Bryant, Les groupes linéaires vérifiant une identité monoïdale, *C.R. Acad. Sc. Paris* 308 (1989), série I, p. 127-128.

[3] M. Boffa & F. Point, Identités de Thue-Morse dans les groupes, *C.R. Acad. Sc. Paris* 312 (1991), série I, p. 667-670.

[4] L. van den Dries & A.J. Wilkie, Gromov's theorem on groups of polynomial growth and elementary logic, *J. Algebra* 89 (1984), p. 349-374.

[5] J. Milnor, Growth of finitely generated solvable groups, *J. Differential Geometry* 2 (1968), p. 447-449.

[6] J. Okninski, Linear semigroups with identities, in : *Semigroups-Algebraic theory and applications to formal languages and codes*, World Sci., 1993, p. 201-211.

[7] F. Point, Groups with identities, *Annals of Pure and Applied Logic* 45 (1989), p. 171-188.

[8] G. Revesz, Universal properties of generators of a variety : groups and skew fields (abstract), *J. Symbolic Logic* 52 (1987), p. 340.

[9] J.M. Rosenblatt, Invariant measures and growth conditions, *Trans. Amer. Math. Soc.* 193 (1974), p. 33-53.

[10] J.F. Semple & A. Shalev, Combinatorial conditions in residually finite groups, I, *J. Algebra* 157 (1993), p. 43-50.

[11] A. Shalev, Combinatorial conditions in residually finite groups, II, *J. Algebra* 157 (1993), p. 51-62.

[12] A.I. Shirshov, On certain near-Engel groups (in russian), *Algebra i Logika Sem.* 2 (1963), n^0 5, p. 5-18.

[13] M. Vaughan-Lee, *The Restricted Burnside Problem*, 2nd edition, Oxford Univ. Press, 1993.

[14] B.A.F. Wehrfritz, *Infinite Linear Groups*, Springer, 1973.

Author's address:
Université de Mons-Hainaut
Institut de Mathématique et d'Informatique
Avenue Maistriau, 15
B-7000 Mons
Belgium.
e-mail: boffa@sun1.umh.ac.be

Model theoretic properties of polycyclic-by-finite groups

Francis Oger

If R is a (possibly noncommutative) ring, then, by [P, Corollary 2.18, p. 37], any R-module M is characterized up to elementary equivalence by the invariants $|\varphi(M)/(\varphi(M) \cap \psi(M))| \in \{1, 2, ..., \infty\}$, where φ and ψ are positive primitive formulas with one free variable. Some algebraic invariants which characterize abelian groups up to elementary equivalence, and which can be written in the form $|\varphi(M)/(\varphi(M) \cap \psi(M))|$, had been previously given by W. Szmielew and by P.C. Eklof and E.R. Fisher (see [EF]). The two following consequences are easily proved:
1) Two abelian groups, or two modules, M, N, are elementarily equivalent if and only if they satisfy the same sentences with one alternation of quantifiers.
2) For each integer $n \geq 2$, two abelian groups, or two modules, M, N, are elementarily equivalent if and only if the direct product of n copies of M and the direct product of n copies of N are elementarily equivalent.

For nonabelian groups in general, it is not possible to obtain such a characterization of elementary equivalence, since S. Burris proved in [Bu] that, for each integer n, there exist two soluble groups which satisfy the same sentences with n alternations of quantifiers without being elementarily equivalent. Concerning 2), L. Manevitz proposes the following problem in [Mn, p. 9]:

Conjecture. *For each integer $n \geq 2$, two groups M, N are elementarily equivalent if and only if the direct product of n copies of M and the direct product of n copies of N are elementarily equivalent.*

This problem may be considered for any sort of structure. The "only if" part is always true. In [Mn], L. Manevitz mentions that the two following structures M, N are not elementarily equivalent, though $M \times M$ and $N \times N$ are isomorphic: M is the set $\mathbf{N}$ with the map $n \to n + 1$, and N is the disjoint union of two copies of M. The conjecture is also believed to be false for groups in general.

We are going to see that, in some classes of groups, it is possible to obtain algebraic characterizations of elementary equivalence which can be expressed with one or two alternations of quantifiers. We shall use these characterizations in order to prove that the conjecture is true for the groups that we consider.

First, we give a few definitions and notations. The finite images of a group G are the finite groups H such that there exists a surjective homomorphism from G to H. For each group G and for each integer n, we denote by G^n the subgroup which is generated by the n-th powers of elements of G, and $\times^n G$ the direct product of

n copies of G. For any properties P, Q defined in the class of groups, a group G is P-by-Q if there exists a normal subgroup H of G which satisfies P and such that G/H satisfies Q. The definitions and results of group theory which are used here can be found in [S]. Concerning model theory, the reader is referred to [CK].

In our proofs, and especially for the conjecture, we use some results concerning the cancellation properties and the decompositions of a group in direct products of indecomposable groups. The first results were proved by R. Hirshon in [H] and other papers. Later on, some generalizations and other results were obtained in [O8].

Any group can be decomposed into a finite direct product of indecomposable groups if and only if it satisfies the maximal condition on direct factors. This condition is satisfied, for instance, by polycyclic-by-finite groups, and in particular by finitely generated finite-by-nilpotent groups, since they satisfy the maximal condition on subgroups. On the other hand, J.M.T. Jones proved in [J] that, for each integer $n \geq 3$, there exists a nontrivial finitely generated group G which satisfies $G \cong \times^n G$ and $G \ncong \times^k G$ for $2 \leq k \leq n-1$.

The decomposition is unique for finite groups; this is the Remak-Krull-Schmidt property. On the other hand, R. Hirshon and other authors gave examples of finitely generated abelian-by-finite groups or finitely generated nilpotent groups which satisfy $\mathbf{Z} \times G \cong \mathbf{Z} \times H$, or $\times^n G \cong \times^n H$ for an integer $n \geq 2$, without being isomorphic. Moreover, in [Ba], G. Baumslag constructed, for any integers $m, n \geq 2$, a finitely generated nilpotent group which has a decomposition with m factors and a decomposition with n factors.

In [O8], we introduce a slightly different notion of decomposition. We say that a group G is **Z**-*indecomposable* if it is not isomorphic to $\times^k \mathbf{Z}$ for an integer $k \geq 1$ and if, for each integer $n \geq 1$, $(\times^n \mathbf{Z}) \times G \cong A \times B$ implies that A or B is isomorphic to $\times^k \mathbf{Z}$ for some integer $k \geq 1$. We say that two groups G, H are **Z**-*equivalent*, and we write $G \approx_{\mathbf{Z}} H$, if there exist two integers m, n such that $(\times^m \mathbf{Z}) \times G \cong (\times^n \mathbf{Z}) \times H$. In that case, we necessarily have $\mathbf{Z} \times G \cong \mathbf{Z} \times H$, or $(\times^k \mathbf{Z}) \times G \cong H$ for an integer $k \geq 1$, or $G \cong (\times^k \mathbf{Z}) \times H$ for an integer $k \geq 1$.

The **Z**-*decompositions* of a group G are the relations $G \approx_{\mathbf{Z}} A_1 \times ... \times A_m$ with $A_1, ..., A_m$ **Z**-indecomposable. We identify two **Z**-decompositions $G \approx_{\mathbf{Z}} A_1 \times ... \times A_m$ and $G \approx_{\mathbf{Z}} B_1 \times ... \times B_n$ if $m = n$ and if there exists a permutation σ of $\{1, ..., n\}$ such that $A_i \approx_{\mathbf{Z}} B_{\sigma(i)}$ for each $i \in \{1, ..., n\}$. The following result generalizes the Remak-Krull-Schmidt property for finite groups:

Theorem 1 [O8, Prop. 1, p. 1999 and Th., p. 2001]. *Any group G which satisfies the property P below has one and only one* **Z**-*decomposition:*
(P) $G/[G, G]$ is finitely generated and G satisfies the maximal condition on direct factors.

The two first propositions below, as well as the second part of the third one, are consequences of this result:

Proposition 1 [O8, Cor. 3, p. 2005]. *Any P-group only has finitely many decompositions in direct products of indecomposable groups.*

Proposition 2 [O8, Cor. 1, p. 2002]. *For each P-group U and for any groups G, H, $U \times G \cong U \times H$ implies $\mathbf{Z} \times G \cong \mathbf{Z} \times H$.*

Proposition 3. *Let G and H be groups. Then:*
1) [H, Th. 1, p. 135] *If $\mathbf{Z} \times G \cong \mathbf{Z} \times H$, then there exists an integer $n \geq 1$ such that $\times^n G \cong \times^n H$.*
2) [O8, Cor. 2, p. 2003] *The converse is true if G satisfies P.*

By considering ultrapowers, we see that, for each integer $n \geq 2$, the statement of the conjecture is equivalent to the following one:
For any groups G, H, if $\times^n G$ and $\times^n H$ are isomorphic, then G and H are elementarily equivalent.
According to Proposition 3 above and Proposition 4 below, the last statement is true if G satisfies P.

Proposition 4 [O5]. *Any groups G, H such that $\mathbf{Z} \times G \cong \mathbf{Z} \times H$ are elementarily equivalent.*

Many reasons make it natural to search for an algebraic characterization of elementary equivalence for polycyclic-by-finite groups. If G is such a group, then we have $\cap_{k \geq 1} G^k = 1$, and G/G^k is finite for each integer $k \geq 1$. Two polycyclic-by-finite groups G, H have the same finite images if and only if they satisfy $G/G^k \cong H/H^k$ for each integer $k \geq 1$.

Two finitely generated abelian groups which have the same finite images are isomorphic. On the other hand, various authors gave examples of nonisomorphic polycyclic-by-finite groups, in particular finitely generated abelian-by-finite groups and finitely generated nilpotent groups, which have the same finite images. However, F.J. Grünewald, P.F. Pickel and D. Segal proved that any class of polycyclic-by-finite groups which have the same finite images is a finite union of isomorphism classes (see [S, Chap. 10]).

For each polycyclic-by-finite group G and for each integer n, G^n is definable in G, since there exists an integer $r(n)$ such that each element of G^n is a product of $r(n)$ n-th powers. It follows that two elementarily equivalent polycyclic-by-finite groups necessarily have the same finite images (see [O2, pp. 470, 475]).

In [R], D. Raphael obtains a stronger result:
If G and H are polycyclic-by-finite groups which satisfy the same sentences with one alternation of quantifiers, then, for each integer $m \geq 1$, there exists a subgroup H_m of G with $H_m \cong H$ and $|G : H_m|$ prime to m, and a subgroup G_m of H with $G_m \cong G$ and $|H : G_m|$ prime to m.
The conclusion implies that G and H have the same finite images.

The following result implies that two finitely generated abelian-by-finite groups are elementarily equivalent if they satisfy the same sentences with one alternation of quantifiers:

Theorem 2 [O4, Corollary, p. 1042]. *Two finitely generated abelian-by-finite groups are elementarily equivalent if and only if they have the same finite images.*

It follows that the conjecture is true for finitely generated abelian-by-finite groups. For any two such groups G, H, and for each integer $n \geq 1$, if $\times^n G$ and

$\times^n H$ are elementarily equivalent, then, for each integer $k \geq 1$, $(\times^n G)/(\times^n G)^k \cong \times^n(G/G^k)$ and $(\times^n H)/(\times^n H)^k \cong \times^n(H/H^k)$ are isomorphic; the last property implies $G/G^k \cong H/H^k$, because the finite group $\times^n(G/G^k) \cong \times^n(H/H^k)$ has a unique decomposition in direct product of indecomposable groups. Consequently, G and H are elementarily equivalent.

Theorem 2 cannot be generalized to polycyclic-by-finite groups. For instance, by [O1] and [R], there exist examples G, H of finitely generated torsion-free nilpotent groups of class 2 such that:
1) G and H do not satisfy the same sentences with one alternation of quantifiers.
2) For each integer $m \geq 1$, there exists a subgroup H_m of G with $H_m \cong H$ and $|G : H_m|$ prime to m, and a subgroup G_m of H with $G_m \cong G$ and $|H : G_m|$ prime to m.

Anyhow, the following result gives a characterization of elementary equivalence for finitely generated nilpotent groups; a sketch of the proof will be given after some remarks.

Theorem 3 [O6], [O9]. *If G and H are finitely generated finite-by-nilpotent groups, then the following properties are equivalent:*
1) G and H are elementarily equivalent;
2) G and H satisfy the same sentences with two alternations of quantifiers;
3) $\mathbf{Z} \times G \cong \mathbf{Z} \times H$.

Remark. In [O7] and [O9], we obtain similar results for the following classes of structures, where $n \geq 2$ is an integer:
a) the $(n+2)$-tuples $(A_1, ..., A_{n+1}, f)$, with $A_1, ..., A_{n+1}$ finitely generated abelian groups and $f : A_1 \times ... \times A_n \to A_{n+1}$ n-linear;
b) the triples (A, B, f), with A, B finitely generated abelian and $f : \times^n A \to B$ n-linear (in particular integral quadratic forms);
c) the pairs (A, f), with A finitely generated abelian and $f : \times^n A \to A$ n-linear (in particular finitely generated Lie rings).

Theorem 3, in conjunction with Proposition 3, implies the conjecture for finitely generated finite-by-nilpotent groups: If two finitely generated finite-by-nilpotent groups G, H satisfy $\times^n G \equiv \times^n H$ for an integer $n \geq 1$, then we have $\mathbf{Z} \times (\times^n G) \cong \mathbf{Z} \times (\times^n H)$, and there exists an integer $k \geq 1$ such that $\times^k(\times^n G) \cong \times^{nk} G$ and $\times^k(\times^n H) \cong \times^{nk} H$ are isomorphic. Consequently, we have $\mathbf{Z} \times G \cong \mathbf{Z} \times H$ and $G \equiv H$.

Theorem 3, as well as Theorem 2, cannot be generalized to polycyclic-by-finite groups. We can see it by considering the semi-direct products $G = I \rtimes \langle \xi_p \rangle$ and $H = J \rtimes \langle \xi_p \rangle$, where p is a prime number, $\langle \xi_p \rangle$ is the cyclic group of order p generated by a primitive p-th root ξ_p of 1 in $\mathbf{C}$, and I, J are nontrivial ideals of $\mathbf{Z}[\xi_p]$, with ξ_p acting on I and J by multiplication.

The groups G, H were introduced by D.S. Warhurst in [W, pp. 33-34]. They are extensions of a finitely generated torsion-free abelian group by a cyclic group of order p. For each integer $k \geq 1$, we have $G/G^k \cong H/H^k$, because I/kI and J/kJ are isomorphic as $\mathbf{Z}[\xi_p]$-modules. Consequently, G and H are elementarily equivalent by Theorem 1.

Now, let us suppose that I is a principal ideal and J is a nonprincipal ideal (J exists for p large enough). Then, G can be generated by 2 elements, while H can only be generated by 3 elements. G and H cannot satisfy $\mathbf{Z} \times G \cong \mathbf{Z} \times H$ since they are nonisomorphic and $Z(G) = Z(H) = 1$.

Proof of Theorem 3 (sketch). As 3) implies 1) by Proposition 4, we just have to show that 2) implies 3). We prove that the following property is a consequence of 2):

4) *For each integer* $m \geq 1$, *there exists a subgroup* H_m *of* G *with* $H_m \cong H$, $[G,G] \subset H_m$ *and* $|G : H_m|$ *prime to* m, *and a subgroup* G_m *of* H *with* $G_m \cong G$, $[H,H] \subset G_m$ *and* $|H : G_m|$ *prime to* m.

For a suitable choice of m, 4) implies $A \times G \cong A \times H$ for a finitely generated abelian group A, and therefore $\mathbf{Z} \times G \cong \mathbf{Z} \times H$ according to Proposition 2.

In order to prove that 2) implies 4), we use arguments which are essentially similar to those of [O3, pp. 63-67] and [R]. Here, the key point is to show that, for each integer $m \geq 1$ and for each finite sequence $\overline{x}$ which generates G, there exists a $\forall\exists$ formula $\varphi(\overline{u})$ such that: 1) G satisfies $\varphi(\overline{x})$; 2) for each finitely generated finite-by-nilpotent group H and for each finite sequence $\overline{y} \subset H$, if H satisfies $\varphi(\overline{y})$, then we have $[H,H] \subset \langle \overline{y} \rangle$.

For each group M, we consider the subgroups $\Gamma_i(M)$ with $\Gamma_1(M) = M$ and $\Gamma_{i+1}(M) = [M, \Gamma_i(M)]$ for $i \geq 1$. For each integer $i \geq 1$, the map $M \times M \to M : (x,y) \to [x,y]$ induces a bilinear map from $(M/[M,M]) \times (\Gamma_i(M)/\Gamma_{i+1}(M))$ to $\Gamma_{i+1}(M)/\Gamma_{i+2}(M)$. Moreover, if M is finitely generated finite-by-nilpotent, then, for each integer $i \geq 1$, there exists an integer $r(i) \geq 1$ such that each element of $\Gamma_{i+1}(M)$ can be written as $[x_1,y_1]...[x_{r(i)},y_{r(i)}]$ with $x_1, ..., x_{r(i)} \in M$ and $y_1, ..., y_{r(i)} \in \Gamma_i(M)$. Consequently, the subgroups $\Gamma_i(M)$ for $i \geq 1$ are defined by existential formulas, and this fact can be expressed by a unique $\forall\exists$ sentence since only finitely many of them are distinct.

There exist an integer $c \geq 1$ such that $\Gamma_{c+1}(G)$ is finite and, for each $i \in \{1, ..., c-1\}$, some sequences of terms $\sigma_i(\overline{u}), \tau_i(\overline{u})$ such that the following properties are satisfied by $\overline{x}$ in G:

$M = \langle \sigma_i(\overline{u}), \{z \in M \mid [z, \Gamma_i(M)] \subset \Gamma_{i+2}(M)\} \rangle$ and $\Gamma_i(M) = \langle \tau_i(\overline{u}), \{z \in M \mid [M,z] \subset \Gamma_{i+2}(M)\} \rangle$.

It follows from the lemma below that these two properties can be expressed by a $\forall\exists$ formula. If this formula is satisfied by $\overline{y}$ in H, then we have $\Gamma_{i+1}(H) = \langle [\sigma_i(\overline{y}), \tau_i(\overline{y})], \Gamma_{i+2}(H) \rangle$.

Moreover, there exists a finite sequence of terms $\tau_c(\overline{u})$ such that $\Gamma_{c+1}(G) = \tau_c(\overline{x})$. The property $\Gamma_{c+1}(M) = \tau_c(\overline{u})$ can also be expressed by a $\forall\exists$ formula. But the property $\Gamma_{c+1}(H) = \tau_c(\overline{y})$ and the properties $\Gamma_{i+1}(H) = \langle [\sigma_i(\overline{y}), \tau_i(\overline{y})], \Gamma_{i+2}(H) \rangle$ for $1 \leq i \leq c-1$ imply $[H,H] \subset \langle \overline{y} \rangle$.

Lemma. *For each quadruple* $A = (A_1, A_2, A_3, f)$ *with* A_1, A_2, A_3 *finitely generated abelian groups and* $f : A_1 \times A_2 \to A_3$ *bilinear, there exist a* $\forall\exists$ *formula* $\varphi(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ *and some sequences* $\bar{x}_1 \subset A_1$, $\bar{x}_2 \subset A_2$, $\bar{x}_3 \subset A_3$ *such that: 1)* A *satisfies* $\varphi(\bar{x}_1, \bar{x}_2, \bar{x}_3)$; *2) for each quadruple* $B = (B_1, B_2, B_3, g)$ *and for any sequences* $\bar{y}_1 \subset B_1$, $\bar{y}_2 \subset B_2$, $\bar{y}_3 \subset B_3$, *if* B *satisfies* $\varphi(\bar{y}_1, \bar{y}_2, \bar{y}_3)$, *then* B_1 *is generated*

by $\bar{y}_1$ *and* $\ker_1(g) = \{y_1 \in B_1 \mid g(y_1, B_2) = 0\}$, *while* B_2 *is generated by* $\bar{y}_2$ *and* $\ker_2(g) = \{y_2 \in B_2 \mid g(B_1, y_2) = 0\}$.

Remark. The lemma can be generalized to the $(n+2)$-tuples $(A_1, ..., A_{n+1}, f)$ with $n \geq 2$, $A_1, ..., A_{n+1}$ finitely generated abelian groups and $f : A_1 \times ... \times A_n \to A_{n+1}$ n-linear.

Remark. The property 2) of the lemma implies that $g(B_1, B_2)$ is generated by $g(\bar{y}_1, \bar{y}_2)$. When we prove that 2) implies 4) in Theorem 3, we use this fact, and also a careful analysis of the complexity of the formula φ which is constructed in the proof of the lemma.

Proof of the lemma (sketch). We construct the formula φ in two steps:

First, we consider an integer $m \geq 2$ and, for each $i \in \{1,2,3\}$, a sequence $\bar{x}_i = (x_{i,1}, ..., x_{i,m(i)})$ which generates A_i and a sequence of variables $\bar{u}_i = (u_{i,1}, ..., u_{i,m(i)})$. We construct a $\forall\exists$ formula $\chi(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ such that: 1) A satisfies $\chi(\bar{x}_1, \bar{x}_2, \bar{x}_3)$; 2) for each quadruple $B = (B_1, B_2, B_3, g)$ and for any sequences $\bar{y}_1 \subset B_1$, $\bar{y}_2 \subset B_2$, $\bar{y}_3 \subset B_3$, if B satisfies $\chi(\bar{y}_1, \bar{y}_2, \bar{y}_3)$, then there exists an injective homomorphism $\theta = (\theta_1, \theta_2, \theta_3) : A \to B$ such that, for each $i \in \{1,2,3\}$, $\theta_i(\bar{x}_i) = \bar{y}_i$ and $|B_i/\theta_i(A_i)|$ is prime to m.

Then, we show that, for a suitable choice of m, there exists a $\forall\exists$ formula $\psi(\bar{u}_1, \bar{u}_2)$ such that:
1) A satisfies $\psi(\bar{x}_1, \bar{x}_2)$;
2) for each quadruple $B = (B_1, B_2, B_3, g)$ with $A \subset B$ and B_1/A_1, B_2/A_2 finite, if B satisfies $\psi(\bar{x}_1, \bar{x}_2)$, then, for each $i \in \{1,2\}$, $|B_i/\langle A_i, \ker_i(g)\rangle|$ divides m.

For this integer m, we consider the formula

$$\varphi(\bar{u}_1, \bar{u}_2, \bar{u}_3) = \chi(\bar{u}_1, \bar{u}_2, \bar{u}_3) \wedge \psi(\bar{u}_1, \bar{u}_2),$$

which is satisfied by $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ in A. For each quadruple $B = (B_1, B_2, B_3, g)$ and for any sequences $\bar{y}_1 \subset B_1$, $\bar{y}_2 \subset B_2$, $\bar{y}_3 \subset B_3$, if B satisfies $\varphi(\bar{y}_1, \bar{y}_2, \bar{y}_3)$, then there exists an injective homomorphism $\theta = (\theta_1, \theta_2, \theta_3) : A \to B$ such that, for each $i \in \{1,2,3\}$, $\theta_i(\bar{x}_i) = \bar{y}_i$ and $|B_i/\theta_i(A_i)|$ is prime to m. For each $i \in \{1,2\}$, we have $B_i = \langle \theta_i(A_i), \ker_i(g)\rangle = \langle \bar{y}_i, \ker_i(g)\rangle$ since $|B_i/\langle \theta_i(A_i), \ker_i(g)\rangle|$ is prime to m and divides m.

If χ exists for an integer m, then it exists for each integer which divides m. So, we can suppose that m is divisible by the cardinals of the torsion subgroups $t(A_1), t(A_2), t(A_3)$. The formula χ says that $(\bar{y}_1, \bar{y}_2, \bar{y}_3)$ satisfies an appropriate finite set of relations, which define a "presentation" of A on $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$, and that $B_i = \langle \bar{y}_i, mB_i\rangle$ and $|B_i/mB_i| = |A_i/mA_i|$ for each $i \in \{1,2,3\}$. The first part implies that there exists a homomorphism $\theta = (\theta_1, \theta_2, \theta_3) : A \to B$ such that $\theta_i(\bar{x}_i) = \bar{y}_i$ for each $i \in \{1,2,3\}$. It follows from the second part that, for each $i \in \{1,2,3\}$, $|B_i/\theta_i(A_i)|$ is prime to m; then, θ_i is injective since $mt(A_i) = 0$.

Now, we come to the construction of ψ. We can suppose f nondegenerate, since there are some universal formulas which define $\ker_1(g)$ and $\ker_2(g)$ for each quadruple $B = (B_1, B_2, B_3, g)$.

For the remainder of the proof, we proceed as follows: we consider some properties of $(A, \bar{x}_1, \bar{x}_2)$ which can be expressed by a unique $\forall\exists$ formula, and we restrict

ourselves to quadruples B containing A, with B_1/A_1 and B_2/A_2 finite, such that $(B, \bar{x}_1, \bar{x}_2)$ satisfies this formula. At the end, we obtain a bound on $|B_1/A_1|$ and $|B_2/A_2|$ which only depends on A, and we denote by $\psi(\bar{u}_1, \bar{u}_2)$ the conjunction of the formulas which have been considered.

As A_1 and A_2 are generated by $\bar{x}_1$ and $\bar{x}_2$, A satisfies $(\forall v_2)(f(\bar{x}_1, v_2) = 0 \rightarrow v_2 = 0) \wedge (\forall v_1)(f(v_1, \bar{x}_2) = 0 \rightarrow v_1 = 0)$. We can suppose that B also satisfies this formula.

We denote by S_B the set of all pairs $(\theta_1, \theta_2) \in \mathrm{End}(B_1) \times \mathrm{End}(B_2)$ such that $g(\theta_1(y_1), y_2) = g(y_1, \theta_2(y_2))$ for any elements $y_1 \in B_1$ and $y_2 \in B_2$, with the product $(\theta_1, \theta_2)(\theta_1', \theta_2') = (\theta_1'\theta_1, \theta_2\theta_2')$. We have $(a\mathrm{Id}_{B_1}, a\mathrm{Id}_{B_2}) \in S_B$ for each $a \in \mathbf{Z}$. $(S_B, +)$ is a finitely generated abelian group and $(S_B, +, .)$ is a not necessarily commutative ring.

Any element $(\theta_1, \theta_2) \in S_B$ is completely determined by $\theta_1(\bar{x}_1)$ and $\theta_2(\bar{x}_2)$: if $\theta_2(\bar{x}_2) = 0$, then, for each $z_1 \in B_1$, we have $g(\theta_1(z_1), \bar{x}_2) = g(z_1, \theta_2(\bar{x}_2)) = g(z_1, 0) = 0$, and therefore $\theta_1(z_1) = 0$; similarly, we have $\theta_2(z_2) = 0$ for $z_2 \in B_2$ if $\theta_1(\bar{x}_1) = 0$.

Now write $\bar{x} = (\bar{x}_1, \bar{x}_2)$ and identify each $\theta \in S_B$ with $\theta(\bar{x}) = (\theta_1(\bar{x}_1), \theta_2(\bar{x}_2))$. There exists a quantifier-free formula which defines the pairs $\bar{y} = (\bar{y}_1, \bar{y}_2) \in S_A$ in $(A, \bar{x})$. We can suppose that the same formula defines the pairs $\bar{y} \in S_B$ in $(B, \bar{x})$.

As $(S_A, +)$ is finitely generated, there exist sequences of terms $\rho^1(\bar{u}), ..., \rho^p(\bar{u})$ such that any element of S_A which commutes with $\rho^1(\bar{x}), ..., \rho^p(\bar{x})$ is in the center R_A of S_A. Consequently, there exists a quantifier-free formula which defines the center R_A of S_A, and we can suppose that the same formula defines the center R_B of S_B. As A_1 and A_2 are generated by $\bar{x}_1$ and $\bar{x}_2$ as R_A-modules, we can also suppose that B_1 and B_2 are generated by $\bar{x}_1$ and $\bar{x}_2$ as R_B-modules.

Moreover, there exist some prime ideals $P_1, ..., P_s$ of R_A such that $P_1...P_s = 0$, and, for each $i \in \{1, ..., s\}$, some sequences of terms $\sigma^{i,1}(\bar{u}), ..., \sigma^{i,t(i)}(\bar{u})$ such that $P_i = \sigma^{i,1}(\bar{x})R_A + ... + \sigma^{i,t(i)}(\bar{x})R_A$. We assume $R_A/P_1, ..., R_A/P_t$ infinite and $R_A/P_{t+1}, ..., R_A/P_s$ finite for an integer $t \leq s$. We can suppose that the ideals $Q_i = \sigma^{i,1}(\bar{x})R_B + ... + \sigma^{i,t(i)}(\bar{x})R_B$ are prime, and satisfy $Q_1...Q_s = 0$ and $|R_B/Q_i| = |R_A/P_i|$ for $t+1 \leq i \leq s$.

For each $i \in \{1, ..., t\}$, we write $R_i = R_A/P_i$ and $S_i = R_B/Q_i$. We have a canonical injection $\Phi_i : R_i = R_A/(Q_i \cap R_A) = (R_A + Q_i)/Q_i \subset S_i$. As B_1/A_1 and B_2/A_2 are finite, R_B/R_A is also finite, and the same property is true for S_i/R_i. Moreover, R_i and S_i are subrings of the ring of integers of an extension of finite degree of $\mathbf{Q}$, since they are integral domains, and $(R_i, +)$ and $(S_i, +)$ are finitely generated torsion-free abelian groups. Consequently, S_i is contained in the integral closure $\bar{R}_i$ of R_i, and we have $|S_i/R_i| \leq |\bar{R}_i/R_i|$.

From this, we deduce some bounds which only depend on A, for $|R_B/R_A|$, and also for $|B_1/A_1|$ and $|B_2/A_2|$ since we have $A_i = R_A\bar{x}_i$ and $B_i = R_B\bar{x}_i$ for $i = 1, 2$.

Remark. As early as 1959, A.I. Mal'cev used a correspondance between rings and groups in order to prove some results concerning the model-theoretic properties of a class of nilpotent groups (see [Ml, Chap. 15, pp. 124-137]). In [GS, p. 172], F.J. Grünewald and R. Scharlau also constructed rings from the nilpotent

groups that they considered; the idea was suggested to them by J. Tits. Later on, A.I. Myasnikov also considered rings associated to multilinear maps and nilpotent groups.

Problem 1. Can two finitely generated nilpotent groups satisfy the same sentences with one alternation of quantifiers without being elementarily equivalent?

Problem 2. Is there an integer n such that two polycyclic-by-finite groups (respectively two finitely generated soluble groups, two finitely generated groups) which satisfy the same sentences with n alternations of quantifiers are elementarily equivalent?

References

[Ba] G. Baumslag, Direct decompositions of finitely generated torsion-free nilpotent groups, Math. Z. 145 (1975), 1-10.
[Bu] S. Burris, Bounded boolean powers and $\equiv_n$, Algebra Universalis 8 (1978), 137-138.
[CK] C.C. Chang and H.J. Keisler, Model Theory, Studies in Logic 73, North-Holland, Amsterdam, 1973.
[EF] P.C. Eklof and E.R. Fisher, The elementary theory of abelian groups, Ann. Math. Logic 4 (1972), 115-171.
[GS] F.J. Grünewald and R. Scharlau, A note on finitely generated torsion-free nilpotent groups of class 2, J. Algebra 58 (1979), 162-175.
[H] R. Hirshon, The cancellation of an infinite cyclic group in direct products, Arch. Math. 26 (1975), 134-138.
[J] J.M.T. Jones, On isomorphisms of direct powers, Word Problems II, Studies in Logic 95, North-Holland, Amsterdam, 1980, pp 215-245.
[Ml] A.I. Mal'cev, The Metamathematics of Algebraic Systems, Studies in Logic 66, North-Holland, Amsterdam, 1971.
[Mn] L.M. Manevitz, Applied model theory and metamathematics. An Abraham Robinson memorial problem list, Israel J. Math. 49 (1984), 3-14.
[O1] F. Oger, Des groupes nilpotents de classe 2 sans torsion de type fini ayant les mêmes images finies peuvent ne pas être élémentairement équivalents, C.R. Acad. Sci. Paris 294 (1982), 1-4.
[O2] F. Oger, Equivalence élémentaire entre groupes finis-par-abéliens de type fini, Comment. Math. Helvetici 57 (1982), 469-480.
[O3] F. Oger, Elementary equivalence and genus of finitely generated nilpotent groups, Bull. Austral. Math. Soc. 37 (1988), 61-68.
[O4] F. Oger, Elementary equivalence and profinite completions: a characterization of finitely generated abelian-by-finite groups, Proc. Amer. Math. Soc. 103 (1988), 1041-1048.
[O5] F. Oger, Cancellation of abelian groups of finite rank modulo elementary equivalence, Math. Scand. 67 (1990), 5-14.
[O6] F. Oger, Cancellation and elementary equivalence of finitely generated finite-by-nilpotent groups, J. London Math. Soc. (2) 44 (1991), 173-183.

[O7] F. Oger, Isomorphism and elementary equivalence of multilinear maps, Lin. Multilin. Algebra 36 (1994), 151-174.
[O8] F. Oger, The direct decompositions of a group G with G/G' finitely generated, Trans. Amer. Math. Soc. 347 (1995), 1997-2010.
[O9] F. Oger, Elementary equivalence for nilpotent groups and multilinear maps: a characterization with two alternations of quantifiers, to appear.
[P] M. Prest, Model Theory and Modules, London Math. Soc. Lecture Notes Series 130, Univ. Press, Cambridge, 1988.
[R] D. Raphael, Commensurability and elementary equivalence of polycyclic groups, Bull. Australian Math. Soc. 53 (1996), 425-439.
[S] D. Segal, Polycyclic Groups, Cambridge Tracts in Math. 82, Univ. Press, Cambridge, 1983.
[W] D.S. Warhurst, Topics in Group Rings, Thesis, University of Manchester, 1981.

Author's address:
Equipe de Logique Mathématique
Université Paris VII, C.N.R.S.
2 place Jussieu, case 7012
75251 Paris Cédex 05
France.
e-mail: oger@logique.jussieu.fr

Non-standard Free Groups

I. M. Chiswell

1. Denote by L_0 the first-order language of groups, $\{\cdot, ^{-1}, 1\}$, and for every cardinal r let F_r denote the free group of rank r. It is well-known that, for any cardinals r, s greater than 1, F_r and F_s have the same universal theory. This is because the free group of countably infinite rank embeds in the free group of rank 2, so if r, $s \leq \omega$, F_r embeds in F_s, while if $\omega \leq r \leq s$, then F_r is an elementary substructure of F_s, by a theorem of Vaught (see Theorem 4, §38 in [16]). It follows that F_r, F_s have the same universal theory for all r, $s > 1$ using Lemma 3.7 in [5]. We denote by Φ the set of all universal and existential sentences true in the non-abelian free groups, so that a group G has the same universal theory as the non-abelian free groups if and only if $G \models \Phi$. We shall also use the following simple remark.

Remark. If A and B are structures for a first-order language L and B has the same universal theory as A, then there is an index set I and an ultrafilter $\mathcal{D}$ on I such that B is a substructure of the ultraproduct $A^I/\mathcal{D}$. If A is a substructure of B, then A and B have the same universal theory in L if and only if B is embeddable in some ultraproduct $A^I/\mathcal{D}$.

For the first part, see [5; Ch. 9, Lemma 3.8]. If $A \subseteq B \subseteq A^I/\mathcal{D}$, it follows from Lemma 3.7 in [5], and the fact that A and $A^I/\mathcal{D}$ are elementarily equivalent (see [5; Ch. 5, Lemma 2.3]), that A and B have the same universal theory.

During an investigation of the model theory of the non-abelian free groups, Gaglione and Spellman [12] noticed a connection between the universal theory of these groups and the fully residually free groups studied by B. Baumslag [3]. The general use of the word "residually" in this context can be described as follows. Let $\mathcal{X}$ be a class of structures for a first-order language and let n be a positive integer. A structure A for the language is said to be n-residually $\mathcal{X}$ if, given elements $a_1, \ldots, a_n$ and $b_1, \ldots b_n$ of A with $a_i \neq b_i$ for $1 \leq i \leq n$, there exists $B \in \mathcal{X}$ and an epimorphism $\phi : A \to B$ such that $\phi(a_i) \neq \phi(b_i)$ for $1 \leq i \leq n$. We abbreviate 1-residually $\mathcal{X}$ to residually $\mathcal{X}$, and A is said to be fully residually $\mathcal{X}$ if it is n-residually $\mathcal{X}$ for all $n \geq 1$. We shall only be using two cases. One is where the language is L_0, and A and all members of $\mathcal{X}$ are groups. In this case we may take all the b_i in the definition to be equal to 1. The other is where the language is the first-order language of rings, $\{+, -, \cdot, 0, 1\}$, which we denote by L_1. Again if A and all members of $\mathcal{X}$ are rings, we may take all b_i to be zero in the definition.

The following theorem lists some results on residually free groups which are relevant to what follows. Before stating it, a definition is needed. We call a group G commutative transitive if given a, b, $c \in G$ such that $[a, b] = 1$, $[b, c] = 1$ and $b \neq 1$, then $[a, c] = 1$. Equivalently, centralisers of non-identity elements of G are

abelian. Note that the property of being commutative transitive can be expressed by a universal sentence in L_0, and since free groups are commutative transitive, models of Φ are commutative transitive. Also, the property of being non-abelian can be expressed by an existential sentence in L_0, so models of Φ are non-abelian.

Theorem 1.1.

(i) *A residually free group is fully residually free if and only if it is commutative transitive.*

(ii) *A group is fully residually free if and only if it is 2-residually free.*

(iii) *A two generator subgroup of a residually free group is either free of rank 2 or abelian.*

(iv) *If $G = A * B$ is the free product of two non-trivial groups A and B, then G is residually free if and only if A and B are both fully residually free.*

Proof. For the proof of (i) see [3; Theorem 1]. It is easy to see that a 2-residually free group is commutative transitive, and (ii) follows (this was noted by Remeslennikov [21; Theorem 1]). Part (iii) is Lemma 1 of §4 in [4], and for (iv) see [3; Theorem 6]. □

It was shown by Gaglione and Spellman [12] that, if G is a non-abelian residually free group, then G is fully residually free if and only if it is a model of Φ. (The main point is that if G is non-abelian and fully residually free then $G \models \Phi$. The converse follows from Theorem 1.1(i) and the remarks preceding Theorem 1.1). Remeslennikov [21] showed that a finitely generated group is a model of Φ if and only if it is non-abelian and fully residually free. It was observed by the author [8] that this can be easily improved, to show that a group is a model of Φ if and only if it is non-abelian and locally fully residually free. We present a detailed and slightly simplified version of this in §2. (The class of fully residually free groups is clearly subgroup closed, so locally fully residually free means that every finitely generated subgroup is residually free).

Note that free abelian groups are fully residually infinite cyclic.To see this, it suffices to observe that finitely generated free abelian groups are fully residually infinite cyclic, by induction on the rank-it is easy to see directly that the free abelian group of rank 2 is fully residually infinite cyclic. It follows that an abelian group is fully residually free if and only if it is fully residually free abelian. Also, it follows that an abelian group is locally fully residually free if and only if it is torsion-free.

Models of Φ are of interest because of another fact noted by Gaglione and Spellman [13], [14] (see also Remeslennikov [22]). They are examples of groups which act freely on a Λ-tree, a generalisation of an ordinary tree introduced by Morgan and Shalen [19], and extensively studied by Alperin and Bass [1]. In fact the argument shows that any locally fully residually free group, abelian or not, acts freely on a Λ-tree. This will be discussed in §3.

Remeslennikov's result made use of rings which are fully residually $\mathbf{Z}$ as rings. There are ring-theoretic analogues of some of the group theoretic results, and these and other properties of residually $\mathbf{Z}$ rings have been considered by the author [9]. We shall give a brief account of some of these results in §4.

2. We shall prove the characterisation given in the introduction of groups with the same universal theory as the non-abelian free groups. We do this by characterising subgroups of ultrapowers of F_2.

Lemma 2.1. *Let G be a locally fully residually free group. Then G embeds in some ultrapower of F_2.*

Proof. Let I be the set of all finite subsets of G. For $E \in I$, let $a_E = \{E' \in I \mid E \subseteq E'\}$. Then $a_E \neq \emptyset$ since $E \in a_E$, and if $E, E' \in I$, then $a_{E \cup E'} \subseteq a_E \cap a_{E'}$. Hence, there is an ultrafilter $\mathcal{D}$ on I such that $a_E \in \mathcal{D}$ for all $E \in I$.

For $E \in I$, let G_E denote the subgroup of G generated by E. Since all countable free groups embed in F_2, there is a group homomorphism $\phi_E : G_E \to F_2$ such that, for all $x \in E \setminus \{1\}$, $\phi_E(x) \neq 1$. Extend ϕ_E to G by putting $\phi_E(x) = 1$ for $x \notin G_E$. Now define $\phi : G \to F_2^I/\mathcal{D}$ by $\phi(x) = \langle \phi_E(x) \rangle_{E \in I}$, where $\langle \phi_E(x) \rangle_{E \in I}$ means the equivalence class of $(\phi_E(x))_{E \in I}$ in $F_2^I/\mathcal{D}$. If $x, y \in G$ then ϕ_E restricted to G_E is a group homomorphism for all $E \in a_{\{x,y\}}$, and $a_{\{x,y\}} \in \mathcal{D}$. It follows that ϕ is a group homomorphism.

Suppose $\phi(x) = 1$; then $\phi_E(x) = 1$ for almost all E, that is, for all $E \in A$, where A is some element of $\mathcal{D}$. Then $a_{\{x\}} \cap A \in \mathcal{D}$, so is non-empty. Take $E \in a_{\{x\}} \cap A$; then $\phi_E(x) = 1$ and $x \in E$, so $x = 1$ by the definition of ϕ_E, hence ϕ is an embedding. □

The main point is that there is a converse to Lemma 2.1, that is, finitely generated subgroups of ultrapowers of F_2 are fully residually free. This is due to Remeslennikov [21], and we shall give his argument with trivial modifications. As we indicated earlier, this depends on a result about fully residually $\mathbf{Z}$ rings. Fix a set I and an ultrafilter $\mathcal{D}$ on I, and denote the ultrapower $X^I/\mathcal{D}$ by ${}^*X$. We shall view ${}^*\mathbf{Z}$ as an extension of the ring $\mathbf{Z}$, and there is an obvious identification of ${}^*\mathrm{SL}_2(\mathbf{Z})$ with $\mathrm{SL}_2({}^*\mathbf{Z})$.

Lemma 2.2. *Let R be a finitely generated subring of ${}^*\mathbf{Z}$. Then as a ring, R is fully residually $\mathbf{Z}$. That is, given $n \geq 1$ and $x_1, \ldots, x_n \in R \setminus \{0\}$, there is a ring homomorphism $\phi : R \to \mathbf{Z}$ such that $\phi(x_i) \neq 0$ for $1 \leq i \leq n$.*

Proof. There is a short exact sequence

$$J \rightarrowtail \mathbf{Z}[y_1, \ldots, y_k] \xrightarrow{\theta} R$$

where the y_i are commuting indeterminates, and J is a finitely generated ideal. Choose generators $f_1, \ldots, f_m$ for J and $g_j \in \mathbf{Z}[y_1, \ldots, y_k]$ such that $\theta(g_j) = x_j$ for $1 \leq j \leq n$. Denoting $z_1, \ldots, z_k$ by $\mathbf{z}$, the sentence

$$\exists z_1 \cdots \exists z_k (f_1(\mathbf{z}) = 0 \wedge \ldots \wedge f_m(\mathbf{z}) = 0 \wedge \neg(g_1(\mathbf{z}) = 0) \wedge \ldots \wedge \neg(g_n(\mathbf{z}) = 0))$$

in L_1 is then valid in ${}^*\mathbf{Z}$ (assign $\theta(y_j)$ to the variable z_j), hence is valid in $\mathbf{Z}$. Thus there are integers $u_1, \ldots, u_k$ in $\mathbf{Z}$ such that $f_1(\mathbf{u}) = \cdots = f_m(\mathbf{u}) = 0$ and $g_j(\mathbf{u}) \neq 0$ for $1 \leq j \leq n$. The mapping $\mathbf{Z}[y_1, \ldots, y_k] \to \mathbf{Z}$ which sends y_j to u_j and is the identity on $\mathbf{Z}$ induces a ring homomorphism $\phi : R \to \mathbf{Z}$, with $\phi(x_j) = g_j(\mathbf{u}) \neq 0$, as required. □

Let p be an odd positive prime in $\mathbf{Z}$ and let $K = \mathrm{Ker}(\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/p\mathbf{Z}))$. It is well-known that K is a free group. One way to see this is to note that $\mathrm{SL}_2(\mathbf{Z})$ is a free product of two cyclic groups of orders 4 and 6 amalgamating their subgroups of order 2 (see [23; Ch.I, 4.2]). One can obtain explicit generators for the cyclic free factors from the proof, and it is easily checked that K intersects these free factors, and so all their conjugates, trivially. Hence K is free by the subgroup theorem for amalgamated free products (see Ch. I, §4.3, Prop. 18 in [23]). It is easy to see that K is non-abelian, so there is an embedding of F_2 into K, and this has an extension to an embedding of ${}^*F_2$ into ${}^*K$.

Lemma 2.3. *Let G be a finitely generated subgroup of ${}^*F_2$. Then G is fully residually free.*

Proof. By the above we may assume $G \subseteq {}^*K \subseteq \mathrm{SL}_2({}^*\mathbf{Z})$. Let $\{g_1, \ldots, g_n\}$ be a set of generators for G, and for simplicity assume this set is closed under taking inverses. We can write

$$g_j = \begin{pmatrix} 1 + pa_j & pb_j \\ pc_j & 1 + pd_j \end{pmatrix}$$

where a_j, b_j, c_j and d_j are all in ${}^*\mathbf{Z}$. Let R be the subring of ${}^*\mathbf{Z}$ generated by $\{a_j, b_j, c_j, d_j \mid 1 \leq j \leq n\}$. If $g \in G$, an easy induction on the length of a word in the g_j representing g shows that we can write

$$g = \begin{pmatrix} 1 + pa & pb \\ pc & 1 + pd \end{pmatrix}$$

where $a, b, c, d \in R$. In particular, $G \subseteq \mathrm{SL}_2(R)$.

Suppose $h_1, \ldots, h_k$ are non-identity elements of G. If

$$h_j = \begin{pmatrix} a_{11}^j & a_{12}^j \\ a_{21}^j & a_{22}^j \end{pmatrix}$$

then by Lemma 2.2 there is a ring homomorphism $\phi : R \to \mathbf{Z}$ such that $\phi(a_{kl}^j) \neq 0$ whenever $a_{kl}^j \neq 0$ and $\phi(a_{kl}^j) \neq 1$ whenever $a_{kl}^j \neq 1$. Let $\psi : \mathrm{SL}_2(R) \to \mathrm{SL}_2(\mathbf{Z})$ be the group homomorphism induced by ϕ. Then $\psi(h_j) \neq 1$ for $1 \leq j \leq k$ and for $g \in G$ written as above, we have

$$\psi(g) = \psi \begin{pmatrix} 1 + pa & pb \\ pc & 1 + pd \end{pmatrix} = \begin{pmatrix} 1 + p\phi(a) & p\phi(b) \\ p\phi(c) & 1 + p\phi(d) \end{pmatrix} \in K.$$

Thus $\psi|_G$ maps G into K which is free. □

Lemmas 2.2 and 2.3 establish our characterisation of subgroups of ultrapowers of F_2.

Theorem 2.4. *Let G be any group. Then G embeds in some ultrapower ${}^*F_2$ of the free group of rank 2 if and only if G is locally fully residually free.*

□

In the case of a finitely presented subgroup of ${}^*F_2$, there is a simpler proof of Lemma 2.3, similar to the proof of Lemma 2.2. (See the proof of Theorem 6 in [11]). This suggested the following question, which was recorded in [15].

Question. Is every finitely generated fully residually free group finitely presented?

We shall return to this question later. It is now easy to give a characterisation of models of Φ.

Theorem 2.5. *Let G be any group. Then the following are equivalent.*

(1) *G is non-abelian and locally fully residually free.*

(2) *G is a model of the set of sentences Φ.*

Proof. Assume (1). By Theorem 1.1(iii), G contains a subgroup isomorphic to F_2. Thus by Theorem 2.4, $F_2 \subseteq G \subseteq {F_2}^I/\mathcal{D}$ for some ultrapower ${F_2}^I/\mathcal{D}$. It follows from the remark at the beginning of §1 that G has the same universal theory as F_2, so (2) holds. Conversely if (2) holds, then by the same remark, G embeds in some ultrapower of F_2, so by Theorem 2.4, G is locally fully residually free. As we observed before Theorem 1.1, G is non-abelian. □

Theorem 2.5 generalises the results that non-abelian fully residually free groups are models of Φ [12], and that non-abelian locally free groups are models of Φ ([11], after Question 3).

3. We begin by giving the definition of a Λ-tree. If Λ is a (totally) ordered abelian group, written additively, a Λ-metric on a set X is a mapping $d : X \times X \to \Lambda$ satisfying the usual axioms for a metric with values in $\mathbb{R}$, and given such a metric the pair (X, d) is called a Λ-metric space. The mapping $\Lambda \times \Lambda \to \Lambda$ given by $(a, b) \mapsto |a - b|$, where $|x| = \max\{x, -x\}$, makes Λ itself into a Λ-metric space. A segment in an arbitrary Λ-metric space (X, d) is the image of an isometry $\alpha : [a, b] \to X$, where $[a, b] = \{x \in \Lambda; a \leq x \leq b\}$ (and $a \leq b$). The endpoints of the segment are $\alpha(a)$ and $\alpha(b)$. A Λ-metric space (X, d) is *geodesic* if, for all $x, y \in X$, there is a segment in X with endpoints x and y.

Definition. A Λ-metric space (X, d) is a Λ-tree if

(a) it is geodesic

(b) the intersection of two segments with a common endpoint is a segment

(c) if two segments intersect in a single point, which is an endpoint of both, then their union is a segment.

If X is the set of vertices of an ordinary tree and d is the path metric on X ($d(x, y)$ is the number of edges in the reduced path joining x and y), it is not difficult to

see that (X, d) is a $\mathbb{Z}$-tree, and all $\mathbb{Z}$-trees arise this way (see Lemmas 1.8 and 1.9 in [8]). Thus Λ-trees can be viewed as generalisations of ordinary trees.

It is an easy consequence of Axiom (b) that given x, y in X, there is a unique segment in X whose set of endpoints is $\{x, y\}$, which we denote by $[x, y]$.

Suppose G is a group acting on a Λ-tree (X, d) as isometries. To each $x \in X$ there is associated a "based length function" $L_x : G \to \Lambda$ given by $L_x(g) = d(x, gx)$. This satisfies the following axioms of Lyndon for a "length function" $L : G \to \Lambda$.

(1) $L(1) = 0$

(2) For all $g \in G$, $L(g) = L(g^{-1})$.

(3) For all $g, h, k \in G$, $(c(g,h) > c(h,k)$ implies $c(h,k) = c(k,g))$, where $c(g,h)$ is defined to be $\frac{1}{2}(L(g) + L(h) - L(g^{-1}h))$.

(4) For all $g, h \in G$, $c(g,h) \in \Lambda$.

Axiom (3) implies that, for all $g, h, k \in G$, at least two of $c(g,h)$, $c(h,k)$, $c(k,g)$ are equal, with the third no smaller. In the case of the function L_x above, it follows from Axiom (b) for a Λ-tree that $[x, gx] \cap [x, hx] = [x, w]$ for some $w \in X$, and it is easily seen, using 2.11 in [1], that $c(g,h) = d(x, w)$, which is why L_x satisfies (3) and (4).

There is a classification of isometries of Λ-trees. If G is acting on the Λ-tree (X, d) as isometries and $g \in G$, there are three possibilities for g. Either g is *elliptic* (i.e. has a fixed point), an *inversion* (g has no fixed point but g^2 does have a fixed point), or else it is *hyperbolic*. In the last case, g has an *axis*, i.e. there is a subtree A_g of X which is isomorphic as a metric space to a subtree of Λ, on which the action of g is equivalent to the restriction of a translation on Λ. We put $\ell(g)$ equal to the amplitude of the translation. If g is not hyperbolic we put $\ell(g) = 0$, to obtain a function $\ell : G \to \Lambda$. If g is not an inversion, we have $\ell(g) = \min_{x \in X} L_x(g)$. This is obvious if g is elliptic (the minimum is attained at any fixed point x of g), and when g is hyperbolic, it can be shown that the minimum is attained precisely at the points on the axis A_g. (See [1], §6, for proofs of all these assertions). An example of an inversion is an automorphism of an ordinary tree which interchanges the endpoints of an edge. There is, so to speak, a "phantom" fixed point in the middle of the edge, which is not part of the corresponding $\mathbb{Z}$-tree. This partly explains why we put $\ell(g) = 0$ when g is an inversion.

The function ℓ is variously called the unbased length function, translation length function or hyperbolic length function for the action of G on X. We have the following connection with the based length functions.

Lemma 3.1. *For any isometry g from X onto X and any $x \in X$,*

$$\ell(g) = \max\{L_x(g^2) - L_x(g), 0\}.$$

Proof. See [1; 7.1(c)] □

We call an action of a group G on a Λ-tree free if g acts as a hyperbolic isometry, for all $1 \neq g \in G$. If Λ is an ordered abelian group, we call a group Λ-free if it

has a free action on some Λ-tree, and a group is called tree-free if it is Λ-free for some Λ. We note that Gaglione and Spellman ([13]) have shown that the class of non-abelian tree-free groups is the model class $\mathbf{M}(\Theta)$ of some set of sentences Θ in the first-order language L_0 of groups, and that Θ may be taken to be a set of Π_2 sentences.

We shall make use of the following result.

Theorem 3.2. *Let G be a group and $L : G \to \Lambda$ a function satisfying Lyndon's Axioms (1)-(4). Then there are a Λ-tree (X, d), an action of G on X and a point $x \in X$ such that $L = L_x$.*

Proof. See [1; 5.4]. □

If F is a free group, F acts on its Cayley graph, relative to some fixed basis, so acts freely on the corresponding $\mathbb{Z}$-tree (The action is just the regular action of F on itself by left translation, so no non-identity element is elliptic or an inversion, since F is torsion-free). Using the vertex 1 as basepoint, it is easy to see that the corresponding Lyndon length function $L = L_1$ is given by: $L(u)$ is the length of the reduced word on $X^{\pm 1}$ representing u, and $c(u, v)$ is the length of the longest common initial segment of the reduced words representing u and v. Since the action is free, $L(u^2) > L(u)$ for all $u \in F \setminus \{1\}$ by Lemma 3.1.

Let I be an index set and let $\mathcal{D}$ be an ultrafilter on I. As before, if $\{X_i; i \in I\}$ is a family of sets, we denote the equivalence class of an element $(x_i)_{i \in I}$ of $\prod_{i \in I} X_i$ in the ultraproduct $\prod_{i \in I} X_i/\mathcal{D}$ by $\langle x_i \rangle_{i \in I}$. Also, if X is a set, we denote the ultrapower $X^I/\mathcal{D}$ by ${}^*X$.

There is an extension of the length function L on the free group F to ${}^*L : {}^*F \to {}^*\mathbb{Z}$, given by ${}^*L(\langle u_i \rangle) = \langle L(u_i) \rangle$, and ${}^*L$ satisfies Lyndon's Axioms (1)-(4) (either by a direct check, or by invoking Loš's Theorem). Further, ${}^*L(u^2) > {}^*L(u)$ for all $u \in {}^*F \setminus \{1\}$. Therefore there is an action of ${}^*F$ on a ${}^*\mathbb{Z}$-tree by Theorem 3.2, and it is a free action by Lemma 3.1. Thus ${}^*F$, and so all its subgroups, are tree-free. Hence, using Theorem 2.4, we obtain the following result.

Theorem 3.3. *Any locally fully residually free group is tree-free.*

□

This is the argument used in [14]; there is also a more direct way of constructing a tree on which ${}^*F$ acts freely, using the next theorem. Again let I be an index set and let $\mathcal{D}$ be an ultrafilter on I. For every $i \in I$ let G_i be a group acting as isometries on a Λ_i-tree (X_i, d_i), where Λ_i is an ordered abelian group. Put

$$G = \prod_{i \in I} G_i/\mathcal{D},\ \Lambda = \prod_{i \in I} \Lambda_i/\mathcal{D},\ X = \prod_{i \in I} X_i/\mathcal{D}$$

so that G is a group, Λ is an ordered abelian group and X is a Λ-metric space with metric $d = \prod_{i \in I} d_i/\mathcal{D}$, that is,

$$d(\langle x_i \rangle_{i \in I}, \langle y_i \rangle_{i \in I}) = \langle d_i(x_i, y_i) \rangle_{i \in I}.$$

Further, the actions of G_i on X_i induce an action of G on X as isometries. Let ℓ_i denote the hyperbolic length function for the action of G_i on X_i.

Theorem 3.4. *In this situation, (X, d) is a Λ-tree, and if ℓ is the hyperbolic length function for the action of G on X, then $\ell = \prod_{i \in I} \ell_i / \mathcal{D}$. In particular, if G_i acts freely on X_i for almost all i (i.e. the set of i for which G_i acts freely belongs to $\mathcal{D}$), then G acts freely on X.*

Proof. See [7; Lemma 5]. □

In another paper, Remeslennikov [22] has shown that if G is a finitely generated fully residually free group, then G is Λ-free for some finitely generated ordered abelian group Λ. This makes use of a construction in [19] of a Λ-tree on which the group $\mathrm{SL}_2(F)$ acts, where F is a field equipped with a valuation with value group Λ. (See also Appendices A and B in [1]).

If Λ is a finitely generated ordered abelian group, it has a decomposition $\Lambda = \Lambda_1 \oplus \cdots \oplus \Lambda_k$ for some $k \geq 0$, where the ordering is the lexicographic ordering and $\Lambda_1, \ldots, \Lambda_k$ are of rank 1 (i.e. archimedean). We believe this observation is due to Zaitseva [25]. Thus Λ_i is isomorphic to $\mathbb{Z}^{m_i}$ for some m_i, and has the ordering induced from some embedding of $\mathbb{Z}^{m_i}$ into the additive group of $\mathbb{R}$.

This brings us to some recent developments, involving cases when all the m_i can be taken equal to 1. Let R be a ring; an R-group is a group G with a mapping $G \times R \to G$, $(g, r) \mapsto g^r$, satisfying the following, for all $g, h \in G$ and $r, s \in R$.

(i) $g^1 = g$

(ii) $g^{(r+s)} = g^r g^s$

(iii) $g^{rs} = (g^r)^s$

(iv) $g(hg)^r = (gh)^r g$.

This definition is due to Lyndon [17], who established a normal form for the free R-group on a given set of generators in the case that R is a polynomial ring $\mathbb{Z}[t_1, \ldots, t_n]$ over $\mathbb{Z}$. (The existence of free R-groups follows from general results in universal algebra. See Cor. 2, §25 in [16]). We denote the free R-group on r generators by $(F_r)^R$. Let $\mathcal{D}$ be a non-principal ultrafilter on ω and let ${}^*F_2$ denote the ultrapower $F_2^\omega / \mathcal{D}$. In his M.Sc. thesis, Pfander has shown that $(F_2)^{\mathbb{Z}[t]}$ embeds in ${}^*F_2$, where $\mathbb{Z}[t]$ is the polynomial ring in one variable (see [20]; this article contains a discussion of Lyndon's normal form for elements of $(F_2)^{\mathbb{Z}[t]}$, which gives some insight into the structure of this group). In fact, for every non-standard integer η in ${}^*\mathbb{Z}$, there is an embedding $\rho_\eta : \mathbb{Z}[t] \to {}^*\mathbb{Z}$ and, using Lyndon's normal form, a corresponding embedding of groups $\bar{\rho}_\eta : (F_2)^{\mathbb{Z}[t]} \to {}^*F_2$. Further, the Lyndon length function ${}^*L$ corresponding to the length function $L : F_2 \to \mathbb{Z}$ given by a basis of F_2 (see the discussion after Theorem 3.2), when restricted to $\bar{\rho}_\eta((F_2)^{\mathbb{Z}[t]})$, takes values in $\rho_\eta(\mathbb{Z}[t])$, an isomorphic copy of $\mathbb{Z}[t]$.

Now if G is a finitely generated subgroup of $\bar{\rho}_\eta((F_2)^{\mathbb{Z}[t]})$, Pfander further shows that there is some integer m such that, for all $g \in G$, ${}^*L(g)$ is a polynomial of degree at most $m-1$, so that ${}^*L$ maps G into some subgroup of ${}^*\mathbb{Z}$ which is isomorphic to $\mathbb{Z}^m = \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}$ with the lexicographic ordering. (This follows because $\rho_\eta(t)$ is a non-standard integer). Using results of Bass [2; 3.5] concerning actions on Λ-trees,

together with an inductive argument, one can deduce that G is finitely presented. Further details are contained in [20].

This is of course relevant to the question raised in the last section, asking whether or not finitely generated subgroups of ${}^*F_2$ are finitely presented, and appears to be the only progress which has been made on it. If one could improve Remeslennikov's argument to show that if G is a finitely generated fully residually free group, then G is Λ-free for $\Lambda = \mathbb{Z}^m$ with the lexicographic ordering, for some m, the last part of Pfander's argument would work to show G is finitely presented, giving an affirmative answer to the question. However, attempting this seems to raise basic questions about extending a valuation on a field to an extension field.

A very recent development is an interesting paper by Fine et al. [10], in which it is observed that $(F_r)^R$ is fully residually free, merely assuming that the additive group R^+ of R is torsion-free and $\mathbb{Z}$ is a pure subgroup of R^+. Examples of such rings are the rings in which every finitely generated subring is fully residually $\mathbb{Z}$, which are considered in §3 below. A particular example is the polynomial ring $\mathbb{Z}[t_1, \ldots, t_n]$. Further, under these assumptions on R, finitely generated subgroups of $(F_r)^R$ embed in $(F_\omega)^{\mathbb{Z}[t]}$, and finitely generated subgroups of $(F_\omega)^{\mathbb{Z}[t]}$ are Λ-free for $\Lambda = \mathbb{Z}^m$ with the lexicographic ordering. The argument of Pfander then shows that finitely generated subgroups of $(F_\omega)^{\mathbb{Z}[t]}$ are finitely presented. The main results in [10] concern residually free groups. It is shown that every 3-generator fully residually free group is embeddable in $(F_\omega)^{\mathbb{Z}[t]}$, and every 2-free residually free group is 3-free. (A group is r-free if every subgroup generated by r or fewer elements is free). Also, there is a classification of fully residually free groups of rank at most 3 (rank meaning minimal number of generators). A fully residually free group of rank 1 is infinite cyclic (residually free groups are clearly torsion-free), and the rank 2 case is covered by Theorem 1.1(iii). In [10], the authors show that if G is a fully residually free group of rank 3, then G is either free of rank 3, free abelian of rank 3, or else G has a one-relator presentation

$$G = \langle x, y, t \mid tvt^{-1} = v \rangle$$

where v is a word on $\{x^{\pm 1}, y^{\pm 1}\}$ representing a non-trivial element on the free group on $\{x, y\}$ which is not a proper power.

It should be pointed out that not all tree-free groups are locally fully residually free. For example, it is shown after Theorem 3 in [15] that $G = \langle x, y, z \mid x^2y^2z^2 = 1 \rangle$ is tree-free, but not a model of the set of sentences Φ, so is not fully residually free by Theorem 2.5. The argument shows that this group is not residually free.

One can ask what properties of (locally) fully residually free groups hold for all tree-free groups. For example, if G is a finitely generated tree-free group, is G Λ-free for some finitely generated Λ? In connection with parts (i) and (iii) of Theorem 1.1, it is known that tree-free groups are commutative transitive, and that a two-generator subgroup is either free of rank 2 or free abelian (see [6] and [24]). Other open questions are: is a tree-free group locally indicable, and is a tree-free group orderable, or at least right orderable? (A group is locally indicable if every non-trivial finitely generated subgroup admits a homomorphism onto the infinite cyclic group. It is known that locally indicable implies right orderable).

4. In Lemma 2.2 the idea of a fully residually $\mathbf{Z}$ ring was used. We now consider residually and fully residually $\mathbf{Z}$ rings in slightly more detail, discussing some of the results in [9]. By considering the additive commutator $(xy - yx)$ of two elements x, y, we see that a ring which is residually $\mathbf{Z}$ is commutative. We also have the following simple result, which was noted as part of Theorem 2 in [21]. The equivalence of (1) and (3) is, of course, an analogue of Theorem 1.1(ii).

Lemma 4.1. *Let R be a ring. The following are equivalent.*

(1) *R is 2-residually $\mathbf{Z}$*

(2) *R is residually $\mathbf{Z}$ and has no zero-divisors*

(3) *R is fully residually $\mathbf{Z}$.*

Proof. Left as an exercise. □

Note also that a non-zero residually $\mathbf{Z}$ ring has its prime ring isomorphic to $\mathbf{Z}$, since it admits a ring homomorphism to $\mathbf{Z}$. The following is an analogue of Lemma 2.1, and the proof is similar.

Lemma 4.2. *Let R be a non-zero ring such that all finitely generated subrings are fully residually $\mathbf{Z}$. Then R embeds in some ultrapower of $\mathbf{Z}$.*

□

This enables us to prove an analogue of Theorem 2.4, which is a slight generalisation of Theorem 2 in [21].

Theorem 4.3. *Let R be a non-zero ring. The following are equivalent.*

(1) *All finitely generated subrings of R are fully residually $\mathbf{Z}$*

(2) *R embeds as a ring in some ultrapower of $\mathbf{Z}$*

(3) *R has the same universal theory as $\mathbf{Z}$ in the first-order language L_1 of rings.*

Proof. (1) implies (2) by Lemma 4.2, and (2) implies (1) by Lemma 2.2. Assume (2); then $\mathbf{Z} \subseteq R \subset \mathbf{Z}^I/\mathcal{D}$ for some index set I and ultrafilter $\mathcal{D}$ on I. It follows from the remark at the beginning of §1 that (3) holds. Conversely (3) implies (2) by the same remark. □

There is a characterisation of finitely generated residually $\mathbf{Z}$ rings, and of finitely generated fully residually $\mathbf{Z}$ rings. Let n be a positive integer, and let $\mathbf{Z}[x_1, \ldots, x_n]$ be the polynomial ring in n commuting indeterminates. If S is a subset of $\mathbf{Z}^n$ we denote by $I(S)$ the set

$$\{f \in \mathbf{Z}[x_1, \ldots, x_n] \mid f(\mathbf{a}) = 0 \text{ for all } \mathbf{a} \in S\}$$

where $\mathbf{a}$ means $a_1, \ldots, a_n$. Thus $I(S)$ is an ideal in $\mathbf{Z}[x_1, \ldots, x_n]$.

Lemma 4.4. *Let R be a finitely generated ring. The following are equivalent.*

(1) *R is residually $\mathbb{Z}$*

(2) *R is isomorphic to $\mathbb{Z}[x_1, \ldots, x_n]/I(S)$ for some positive integer n and subset S of $\mathbb{Z}^n$.*

Proof. Assume (1). Since R is finitely generated, it is isomorphic to $\mathbb{Z}[x_1, \ldots, x_n]/I$ for some ideal I and positive integer n. Let $S = \{\mathbf{a} \in \mathbb{Z}^n \mid f(\mathbf{a}) = 0 \text{ for all } f \in I\}$. Then $I \subseteq I(S)$. Suppose $I \neq I(S)$, and let $f \in I(S) \setminus I$. Then $f + I \neq 0$ in $\mathbb{Z}[x_1, \ldots, x_n]/I$, so there is a ring homomorphism $\phi : \mathbb{Z}[x_1, \ldots, x_n]/I \to \mathbb{Z}$ such that $\phi(f + I) \neq 0$. Let ψ be the lift of ϕ to a homomorphism $\mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{Z}$; thus if $\psi(x_i) = a_i$, then $\psi(g) = g(\mathbf{a})$. If $g \in I$, then $\psi(g) = \phi(g + I) = 0$, hence $\mathbf{a} \in S$. But $f(\mathbf{a}) = \phi(f + I) \neq 0$, contradicting $f \in I(S)$. Thus $I = I(S)$.

Conversely, assume (2), and suppose $f + I(S) \neq 0$ in $\mathbb{Z}[x_1, \ldots, x_n]/I(S)$, that is, $f \notin I(S)$. Choose $\mathbf{a} \in S$ such that $f(\mathbf{a}) \neq 0$. Let $\psi : \mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{Z}$ be the ring homomorphism sending x_i to a_i for $1 \leq i \leq n$. Then if $g \in I(S)$, $\psi(g) = g(\mathbf{a}) = 0$, so ψ induces a ring homomorphism $\phi : \mathbb{Z}[x_1, \ldots, x_n]/I(S) \to \mathbb{Z}$, and $\phi(f + I(S)) = f(\mathbf{a}) \neq 0$, showing that the ring $\mathbb{Z}[x_1, \ldots, x_n]/I(S)$ is residually $\mathbb{Z}$. □

Corollary. *Let R be a non-zero finitely generated ring. The following are equivalent.*

(1) *R is fully residually $\mathbb{Z}$*

(2) *R is isomorphic to $\mathbb{Z}[x_1, \ldots, x_n]/I(S)$ for some positive integer n and subset S of $\mathbb{Z}^n$ such that $I(S)$ is prime.* □

Examples of fully residually $\mathbb{Z}$ rings are $\mathbb{Z}[x_1, x_2, x_3]/(x_1^2 + x_2^2 - x_3^2)$, which is studied in [9], and $\mathbb{Z}[x_1, x_2]/(x_1^2 - nx_2^2 - 1)$, where n is a positive integer which is not a square, which is given in [10]. The paper [9] contains other results on these classes of rings, such as the following.

Proposition 4.5.

(1) *If A and B are residually $\mathbb{Z}$ rings, then so is $A \otimes_{\mathbb{Z}} B$.*

(2) *If A and B are fully residually $\mathbb{Z}$ rings, then so is $A \otimes_{\mathbb{Z}} B$.*

(3) *If all finitely generated subrings of A and B are fully residually $\mathbb{Z}$ rings, then the same is true of $A \otimes_{\mathbb{Z}} B$.* □

In view of the fact that tensor product is the coproduct in the category of commutative rings, Proposition 4.5 may be viewed as a (not very precise) analogue of Theorem 1.1(iv).

References

[1] R.C. Alperin and H. Bass, "Length functions of group actions on Λ-trees". In: Combinatorial group theory and topology (ed. S.M.Gersten and J.R.Stallings), Annals of Mathematics Studies **111**, pp 265-378. Princeton: University Press 1987.

[2] H. Bass, "Group actions on non-archimedean trees". In: Arboreal group theory, (ed. R.C. Alperin), MSRI Publications **19**, pp 69-131. New York: Springer-Verlag 1991.

[3] B. Baumslag, "Residually free groups", *Proc. London Math. Soc.* (3) **17** (1967), 402–418.

[4] G. Baumslag, "On generalised free products", *Math. Z.* **78** (1962), 423–438.

[5] J.L. Bell and A.B. Slomson, *Models and ultraproducts.* Amsterdam: North-Holland 1971.

[6] I.M. Chiswell, "Harrison's theorem for Λ-trees", *Quart. J. Math Oxford (2)* **45** (1994), 1-12.

[7] I.M. Chiswell, "Generalised trees and Λ-trees". In: Combinatorial and Geometric Group Theory, Edinburgh 1993, London Mathematical Society Lecture Notes 204, pp 43-55. Cambridge: University Press 1994.

[8] I.M. Chiswell, "Introduction to Λ-trees". In: Semigroups, formal languages and groups, (ed. J. Fountain), pp 255-293. Dordrecht, Boston, London: Kluwer 1995.

[9] I.M. Chiswell, "Rings which are residually $\mathbb{Z}$", preprint.

[10] B. Fine, A.M. Gaglione, A. Myasnikov, G. Rosenberger and D. Spellman, "A classification of fully residually free groups", preprint.

[11] A.M. Gaglione and D. Spellman, "More model theory of free groups", *Houston J. Math.* **21** (1995), 225-245.

[12] A.M. Gaglione and D. Spellman, "Even more model theory of free groups". In: Infinite groups and rings, 37-40. River Edge, NJ: World Sci. Publishing 1993.

[13] A.M. Gaglione and D. Spellman, "Does Lyndon's length function imply the universal theory of free groups?". In: The mathematical legacy of Wilhelm Magnus: groups, geometry and special functions (Brooklyn, NY, 1992). Contemp. Math. **169**, 277-281. Amer. Math. Soc., Providence, RI, 1994.

[14] A.M. Gaglione and D. Spellman, "Every 'universally free' group is tree free". In: Group theory (Granville, Ohio 1992), 149-154. River Edge, NJ: World Sci. Publishing 1993.

[15] A.M. Gaglione and D. Spellman, "Generalisations of free groups: some questions", *Commun. Algebra* **22(8)** (1994), 3159-3169.

[16] G. Grätzer, *Universal algebra*. Princeton: van Nostrand 1968.

[17] R.C. Lyndon, "Groups with parametric exponents", *Trans. Amer. Math. Soc.* **96** (1960), 518-533.

[18] R.C. Lyndon, "The equation $a^2b^2 = c^2$ in free groups", *Michigan Math. J.* **6** (1959), 155-164.

[19] J.W. Morgan and P.B. Shalen, "Valuations, trees and degenerations of hyperbolic structures: I", *Annals of Math. (2)* **122** (1985), 398-476.

[20] P.H. Pfander, "Finitely generated subgroups of the free $\mathbf{Z}[t]$ group on two generators", this volume.

[21] V.N. Remeslennikov, "∃-free groups", *Siberian Math. J.* **30** (1989), 193-197.

[22] V.N. Remeslennikov, "∃-free groups as groups with a length function", *Ukrainian Math. J.* **44** (1992), 813-818.

[23] J.-P. Serre, *Trees*. New York: Springer 1980.

[24] M. Urbański and L. Zamboni, "On free actions on Λ-trees", *Math. Proc. Cambridge Philos. Soc.* **113** (1993), 535-542.

[25] M.I. Zaitseva, Uspekhi Mat. Nauk 8(**53**) (1953), 135-137.

Author's address:
School of Mathematical Sciences
Queen Mary and Westfield College
University of London
Mile End Road
London E1 4NS.
e-mail: i.m.chiswell@qmw.ac.uk

Finitely generated subgroups of the free $\mathbf{Z}[t]$-group on two generators

Patrick H. Pfander

Contents

Abstract

A theorem of Schreier states that subgroups of free groups are free. From the perspective of nonstandard analysis it is natural to ask whether an analogous result holds for an ultrapower of a free group. This question leads to the investigation of presentations of finitely generated subgroups of ultrapowers of free groups. In this paper we prove that finitely generated subgroups of the free $\mathbf{Z}[t]$-group, which we will denote by $\mathcal{F}_2$, are finitely presented. Indeed these groups are isomorphic to groups constructed from finitely generated free groups by finitely many applications of the operations of taking amalgamated products and HNN-extensions. The bearing of this result upon the original question is that in the course of the proof we shall embed $\mathcal{F}_2$ into ${}^*F_2$ and so the result gives us some information about finitely generated subgroups of ${}^*F_2$.

1 Introduction

Because of the interest in the elementary theory of free groups, the structure of subgroups of an ultrapower of a free group has received some attention. For example Gaglione and Spellmann conjecture in [Ga 94] that all finitely generated

subgroups of ${}^*F_2$ are finitely presented. The question as to whether or not finitely generated subgroups of ${}^*F_2$ are finitely presented arises naturally as a generalization of Schreier's theorem, which says that subgroups of free groups are free (and hence finitely generated subgroups are finitely presented). That Schreier's theorem does not itself extend to ${}^*F_2$ can be seen from the following example. Let x and y be generators for F_2. Consider the two sequences $(x^n)_{n\in\omega}$ and $(x^{n^2})_{n\in\omega}$ of elements of F_2. We may regard ${}^*F_2$ as an ultraproduct $\prod F_2/\mathcal{D}$, where $\mathcal{D}$ is an non principal ultrafilter on ω and so these sequences give rise to two elements of ${}^*F_2$, which we shall denote respectively as $\langle x^n\rangle$ and $\left\langle x^{n^2}\right\rangle$. These elements commute although there is no single element of ${}^*F_2$ of which they are both a power. Consequently $\langle x^n\rangle$ and $\left\langle x^{n^2}\right\rangle$ generate a subgroup of ${}^*F_2$ which is isomorphic to $\mathbb{Z}\times\mathbb{Z}$. This shows that *freeness* is not the same as $*$*-freeness.*

The result of this paper is that finitely generated subgroups of the so-called free $\mathbb{Z}[t]$-group on two generators, (which we will denote by $\mathcal{F}_2$) are finitely presented. The bearing of this result upon the original question is that in the course of the proof we shall embed $\mathcal{F}_2$ into ${}^*F_2$ and so the result gives us some information about finitely generated subgroups of ${}^*F_2$.

In fact the main tool we shall use, Bass-Serre theory, actually gives rise to concrete presentations of such groups. It will turn out that every finitely generated subgroup of $\mathcal{F}_2$ is isomorphic to a group constructed from finitely generated free groups by finitely many applications of the operations of taking amalgamated products and HNN-extensions.

We now give a brief outline of the content of the paper. After setting up the universe of discourse we spend section 3 on introducing the free $\mathbb{Z}[t]$-group on two generators $\mathcal{F}_2$ and on embedding this group into ${}^*F_2$. This proof relies mainly on Lyndon's normal form for $\mathcal{F}_2$ and is rather technical. The proof of the main result does not require details from this section. Now let G be a finitely generated subgroup of $\mathcal{F}_2$. In section 5 we will show, in Lemma 5.2, that there is a Lyndon length function $\mathcal{L} : G \to (\mathbb{Z}^m, <)$, where $<$ is the lexicographic ordering. By a construction of Chiswell in [Ch 76] this gives rise to a free action of G on a $\mathbb{Z}^m$-tree without inversions. Now a theorem by Bass, which generalizes Serres structure theorem (Theorem 13, [Se]) applies. This will tell us that G is isomorphic to a certain fundamental group of a graph of groups. This group will be finitely presented, having the aforementioned structure.

Finally I want to express my gratitude to Professor Angus Macintyre and Professor Knut Radbruch for their intellectual and personal support, which enabled me to write this paper. I also want to thank Simon Chatterjee for his help on the texnical side.

2 Notations and the mathematical set up

Let $S = \{\cdot, ^{-1}, e\}$ be the language of groups. We will denote the free group on two generators, given as a S-structure, by F_2. The set $\{x, y\}$ will be a set of generators for F_2. Let W^0 be the set of words in the letters x, x', y, y' for F_2 and let $\mathbf{e}$ be the

empty word. We denote the natural map $W^0 \to F_2$ by π.

Let $\mathbf{Z}$ be the integers considered as an ordered abelian group and $L : F_2 \to \mathbf{Z}$ the length function for F_2. So $g \in F_2$ is mapped by L to the length of the shortest word $w_g \in W^0$ representing g. Let min be the arithmetic function $\min : \mathbf{Z} \times \mathbf{Z} \to \mathbf{Z}$.

In order to apply the theorem of Łoš to objects like π or L, we need a mathematical set up consisting of a superstructure containing F_2, W^0 and $\mathbf{Z}$. Thus we define $M = F_2 \sqcup W^0 \sqcup \mathbf{Z}$, the disjoint union of F_2, W^0 and $\mathbf{Z}$. We denote the language of the superstructure $V(M)$ by $S_{V(M)}$.

Let $\mathcal{D}$ be a non-principal ultrafilter on ω. Now let ${}^*M = \prod M/\mathcal{D}$ and $V({}^* M)$ be the corresponding superstructure. By the theorem of Łoš there is a natural elementary embedding

$$^* : V(M) \to V(^*M).$$

So every constant symbol a in $S_{V(M)}$ has an interpretation ${}^*a$ in $V({}^*M)$ and every function symbol $f \in S_{V(M)}$ has an interpretation ${}^*f$ in $V({}^*M)$. We have ${}^*F_2 \cong \prod F_2/\mathcal{D}$ and ${}^*(W^0) \cong \prod W^0/\mathcal{D}$. We will denote elements $g \in {}^*F_2$ as $g = \langle g_n \rangle$.

We will refer to ${}^* \leq \in V({}^*M)$ as $\leq$ and to ${}^*\min \in V({}^* M)$ as min, since no ambiguity will arise.

3 $\mathbf{Z}[t]$-groups, their normal form and the embedding $\mathcal{F}_2 \to {}^*F_2$

3.1 $\mathbf{Z}[t]$-groups

Definition 3.1 *[Ly 60] Let R be an unital ring. Define S' to be the language $\{\cdot, ^{-1}, e, \bar{f}_\alpha\}_{\alpha \in R}$, where $\bar{f}_\alpha$ is an unary function symbol. An R-group G is a S'-structure satisfying the following axioms (we write g^α for $f_\alpha(g)$.):*

1. *the usual group axioms, formulated in the language $S = \{\cdot, ^{-1}, e\}$.*
2. $g^1 = g$
3. $g^{(\alpha+\beta)} = g^\alpha \, g^\beta$
4. $g^{\alpha\beta} = (g^\alpha)^\beta$
5. $g(hg)^\alpha = (gh)^\alpha g$, *for all $g, h \in G$ and for all $\alpha, \beta \in R$.*

In the class of 'n-generator R-groups' free objects exist, since this class is equationally defined : the number of generators is fixed so it follows from a theorem of Birkhoff (see for example Corollary 2 ,§25, chapter 4 in [Gr]) that there is a free algebra in the class of n-generator R-groups.

We will be concerned with the free $\mathbf{Z}[t]$-group on two generators, which we will denote by $\mathcal{F}_2$. Then, as in [Ly 60], let $F^0 = F_2$ and define recursively,

$$F^i = \left\langle g^\alpha \, : \, g \in F^{i-1} \, ; \alpha \in \mathbf{Z}[t] \right\rangle_S ,$$

where $\langle \ldots \rangle_S$ denotes 'generation as a group in the language S'. Then $\mathcal{F}_2$ equals $\bigcup_{i\in\omega} F^i$ (see [Ly 60]). Let $H : \mathcal{F}_2 \to \omega$ be the function

$$H(g) = \min\{i : g \in F^i\},$$

and call $H(g)$ the *height of* g. Using this notion we have

$$F^i = \{g \in \mathcal{F}_2 : g \text{ has height } H(g) \leq i\}.$$

In order to distinguish between *words* and *group elements* we need the following definitions, parallelling the construction of $\mathcal{F}_2$ and F^i above. Let W^0 be the set of all words in the letters x, x', y, y', i.e. W^0 is recursively defined as follows:

1. $x, y, x', y' \in W^0$

2. if $w_1, w_2 \in W^0$ then $w_1 w_2 \in W^0$.

Further more define recursively W^i,

1. If $w \in W^{i-1}$, then $w \in W^i$.

2. If $w_1, w_2 \in W^i$ then $w_1 w_2 \in W^i$.

3. If $w \in W^{i-1}$, $\alpha \in \mathbb{Z}[t]$ then $(w, \alpha) \in W^i$. We will write w^α instead of (w, α).

Let $W = \bigcup_{i\in\omega} W^i$ be the set of words. We can define a height function on W too. Then $H(w) = 0$ if $w \in W^0$ and w is a *word of height 0.* If $w \in W^K \backslash W^{K-1}$ we define $H(w) = K$ and say that w is a *word of height* K. The general form of $w \in W$ is $w = w_1^{\alpha_1} \ldots w_r^{\alpha_r}$, where $H(w_i) \leq H(w)$. Now we define recursively:

$$\begin{aligned} \pi^0 : W^0 &\to F^0 = F_2 \\ \pi^i : W^i &\to F^i, \quad i \geq 1 \end{aligned}$$

where $\pi^0 = \pi$ is the natural map and π^i, $i \geq 1$, is defined as

$$\begin{aligned} \pi^i(w) &= \pi^{i-1}(w), && \text{iff } H(w) \leq i-1, \\ \pi^i(w) &= \pi^{i-1}(v)^\alpha && \text{iff } H(w) = i, \text{ and } w = v^\alpha, \text{ and } H(v) = i-1. \\ \pi^i(w_1 w_2) &= \pi^i(w_1)\pi^i(w_2) && \text{for all } w_1, w_2 \in W. \end{aligned}$$

Note that $\pi^i | W^{i-1} = \pi^{i-1}$. So we can define

$$\begin{aligned} \pi' : W &\to \mathcal{F}_2 \\ \pi'(w) &= \pi^i(w) \quad \text{iff } w \text{ of height } i. \end{aligned}$$

We call π' the *natural map.*

3.2 Nonstandard retraction maps

We will identify polynomials $\alpha(t) \in \mathbf{Z}[t]$ with the induced polynomial function, $\alpha : \mathbf{Z} \to \mathbf{Z}$, $n \mapsto \alpha(n)$. Let for any $n \in \omega$

$$\begin{array}{rcl} \rho_n : \mathbf{Z}[t] & \to & \mathbf{Z} \\ \alpha(t) & \mapsto & \alpha(n) \end{array}$$

be the evaluation map. These maps are ring homomorphisms. We use these maps to define nonstandard evaluation maps for all $\eta = \langle \eta_n \rangle \in {}^*\mathbf{Z}$, i.e.

$$\begin{array}{rcl} \rho_\eta : \mathbf{Z}[t] & \to & {}^*\mathbf{Z} \\ \alpha & \mapsto & \langle \rho_{\eta_n}(\alpha) \rangle \; = \; \langle \alpha(\eta_n) \rangle . \end{array}$$

Then ρ_η is a ring homomorphism. Moreover

$$\begin{array}{rcll} c & \mapsto & \langle c \rangle & \text{if } c \text{ constant} \\ t & \mapsto & \langle \eta_n \rangle \; = \eta & \end{array} \tag{1}$$

Lemma 3.2 *Let $\eta \in {}^*\mathbf{Z} \setminus \mathbf{Z}$, $\eta > 0$ and let $\mathbf{Z}[t]$ be equipped with the lexicographic ordering. Then $\rho_\eta : \mathbf{Z}[t] \to {}^*\mathbf{Z}$ is injective and order preserving.*

Proof Let $\alpha \in \mathbf{Z}[t]$ be in the kernel of ρ_η. Then $\rho_\eta(\alpha) = 0$, what is equivalent to $\{n : \rho_{\eta_n}(\alpha) = 0\} \in \mathcal{D}$. The ultrafilter $\mathcal{D}$ is non-principal over ω, whence $\rho_{\eta_n}(\alpha) = 0$ for infinitely many n. It follows that $\alpha \equiv 0$ and that ρ_η is injective. The map ρ_η is order preserving by (1) and the fact that $\eta > c$ for all $c \in \mathbf{Z}$. □

Every map ρ_n induces a retraction map $\bar{\rho}_n : W \to W^0$ by the following definition

$$\begin{array}{rcl} \bar{\rho}_n(\mathsf{e}) & = & \mathsf{e} \\ \bar{\rho}_n(x) & = & x \\ \bar{\rho}_n(x') & = & x' \\ \bar{\rho}_n(y) & = & y \\ \bar{\rho}_n(y') & = & y' \\ \bar{\rho}_n(w) & = & \bar{\rho}_n(w_1)\bar{\rho}_n(w_2) \text{ if } w = w_1 w_2 \\ \bar{\rho}_n(w^\alpha) & = & \begin{cases} \underbrace{(\ldots(\bar{\rho}_n(w)\bar{\rho}_n(w))\bar{\rho}_n(w) \ldots\ldots \bar{\rho}_n(w))}_{\rho_n(\alpha)\text{-times}} & \text{if } \rho_n(\alpha) > 0 \\ \underbrace{(\ldots(\bar{\rho}_n(w')\bar{\rho}_n(w'))\bar{\rho}_n(w') \ldots\ldots \bar{\rho}_n(w'))}_{|\rho_n(\alpha)|\text{-times}} & \text{if } \rho_n(\alpha) < 0 \\ \mathsf{e} & \text{if } \rho_n(\alpha) = 0, \end{cases} \end{array}$$

where, if $w = w_1^{\varepsilon_1} \ldots w_n^{\varepsilon_n}$, and $\varepsilon_i \in \{-1, 1\}$, then $w' = w_n^{-\varepsilon_n} \ldots w_1^{-\varepsilon_1}$.

Now pick for any $g \in \mathcal{F}_2$ a representative $w \in W$ and define $\bar{\rho}_n : \mathcal{F}_2 \to F_2$ by

$$\bar{\rho}_n(g) = \pi(\bar{\rho}_n(w)).$$

This is well defined and the following diagram commutes:

$$\begin{array}{ccc} W & \xrightarrow{\bar{\bar{\rho}}_n} & W^0 \\ \downarrow{\scriptstyle \pi'} & & \downarrow{\scriptstyle \pi} \\ \mathcal{F}_2 & \xrightarrow{\bar{\rho}_n} & F_2 \end{array}$$

Since $\bar{\rho}_n \circ \pi'(w) = \pi \circ \bar{\bar{\rho}}_n(w)$ holds by definition of $\bar{\rho}_n$, the maps $\bar{\rho}_n$ have the following properties:

$$\begin{array}{lcll} \bar{\rho}_n(g^\alpha) & = & g^{\rho_n(\alpha)} & \text{iff } H(g) = 0 \\ \bar{\rho}_n(g^\alpha) & = & \bar{\rho}_n(g)^{\rho_n(\alpha)} & \text{iff } H(g) > 0 \\ \bar{\rho}_n(g_1 g_2) & = & \bar{\rho}_n(g_1)\bar{\rho}_n(g_2) & \text{for all } g_1, g_2 \in \mathcal{F}_2. \end{array}$$

Hence $\bar{\rho}_n$ is a group homomorphism for any $n \in \omega$.

Now we use $\bar{\bar{\rho}}_n$ and $\bar{\rho}_n$ in order to define nonstandard retraction maps. For $\eta = \langle \eta_n \rangle \in {}^*\mathbf{Z}$, we define $\bar{\bar{\rho}}_\eta : W \to {}^*(W^0)$, by

$$\bar{\bar{\rho}}_\eta(w) = \langle \bar{\bar{\rho}}_{\eta_n}(w) \rangle,$$

and analogously $\bar{\rho}_\eta : \mathcal{F}_2 \to {}^*F_2$.

$$\bar{\rho}_\eta(g) = \langle \bar{\rho}_{\eta_n}(g) \rangle. \tag{2}$$

Hence $\bar{\rho}_\eta$ is a group homomorphism. It follows by an application of the theorem of Łoš that the following diagram commutes.

$$\begin{array}{ccc} W & \xrightarrow{\bar{\bar{\rho}}_\eta} & {}^*(W^0) \\ \downarrow{\scriptstyle \pi'} & & \downarrow{\scriptstyle {}^*\pi} \\ \mathcal{F}_2 & \xrightarrow{\bar{\rho}_\eta} & {}^*F_2 \end{array}$$

3.3 Lyndon's normal form for $\mathcal{F}_2$

By [Ly 60] there is a normal form for elements of free $\mathbf{Z}[t_1, \ldots, t_m]$-groups available. We will outline this normal form in the special case where we deal with $\mathcal{F}_2$, the free $\mathbf{Z}[t]$-group in two generators.

In a free group the normal form of a word, representing a group element g, is the unique reduced word representing g. Lyndon apes this idea in the sense that he associates with any word $w \in W$, a word w_g which is in some sense reduced. Roughly speaking this means that , if $w \in W$, $w = w_1^{\alpha_1} \ldots w_r^{\alpha_r}$, then $\bar{\bar{\rho}}_n(w_i^{\alpha_i}) \in W^0$ does not cancel with any adjacent factor. This is generally not the case for all retractions ρ_n, as the following example shows: let $w = x(yx)^{t-5}$. Then, with $w_1 = x$ and $w_2 = yx$ we see that $\bar{\bar{\rho}}_n(w_1)$ does not cancel with $\bar{\bar{\rho}}_n(w_2^{t-5})$ if $n > 5$. On the other hand let $n = 2$. Then $\bar{\bar{\rho}}_2(w_1)$ and $\bar{\bar{\rho}}_2(w_2^{t-5}) = x^{-1}y^{-1}x^{-1}y^{-1}x^{-1}y^{-1}$, cancel on the left. It depends on the sign of $\rho_n(\alpha_{i-1}), \rho_n(\alpha_i)$ and $\rho_n(\alpha_{i+1})$ whether or not cancellation between adjacent factors happens. This leads to the assignment of a collection of normal words to a given word w. For each exponent α_i in w we distinguish between the 'normal form of w if $\rho(\alpha_i) > 0$' , the 'normal form of w if $\rho(\alpha_i) < 0$' and the 'normal form of w if $\rho(\alpha_i) = 0$'. We will call formulae of the type $\rho(\alpha_i) > 0, \rho(\alpha_i) < 0$ and $\rho(\alpha_i) = 0$, *conditions*. Formally this gives rise to the

need to regard a normal form of $w \in W$ as a tuple (w', C), where $w' \in W$ and C is a set of conditions. Then we say a given retraction ρ_n *satisfies* C, if the conditions in C hold for ρ_n.

Moreover we can associate arbitrary finite sets of conditions C with any word $w \in W$ and write (w, C). We consider also the empty set as a set of conditions. Then we define which of these words with conditions are in normal form. A set of conditions C for a word w need not be satisfiable. Then we say C is *inconsistent*. If there is at least one $n \in \omega$ such that ρ_n satisfies C, we say C is *consistent*. Lyndon defines two sets of words with conditions W_1, W_2, to be equivalent, if they have the same sets of values, i.e. if for $W_1 = \{(w_1, C_1), \ldots, (w_m, C_m)\}$ and $W_2 = \{(w_1', D_1), \ldots, (w_l', D_l)\}$ and for all $n \in \omega$,

$$\{\bar{\rho_n}(w_i) : \rho_n \text{ satisfies } C_i, 1 \leq i \leq m\} = \{\bar{\rho_n}(w_j') : \rho_n \text{ satisfies } D_j, 1 \leq j \leq l\}.$$

Equivalently we define two sets of words with conditions W_1, W_2 to be equivalent, if it is possible to pass from one to the other using the following steps:

1. replace $\{(w, C_1), \ldots, (w, C_m)\} \subseteq W_1$ by $\{(w, D_1), \ldots, (w, D_l)\} \subseteq W_2$ if the same retractions satisfy one of the C_i as satisfy one of the D_j.

2. replace $(w, C) \in W_1$ by $(w', C) \in W_2$, where the set C contains a condition $\rho(\alpha - \beta) = 0$ for some $\alpha, \beta \in \mathbb{Z}[t]$ and w and w' only differ in that one contains α at a certain place, where w' contains β.

3. replace $(w, C) \in W_1$ by (w', C), if $\pi(w) = \pi(w')$ follows from the axiom of R-groups.

Lyndon defines recursively when a word with conditions (w, C) is in normal form. Let w be a word of height 0. Then, for any consistent C, (w, C) is in normal form if w is a reduced word.

Let w be a word of height $H(w) \geq 1$, Then (w, C) is in normal form, if

1. $w = u_1 v_1^{\alpha_1} u_2 v_2^{\alpha_2} \ldots v_r^{\alpha_r} u_{r+1}$, where α_i is non-constant for $i \in \{1, \ldots, r\}$

2. and the properties **N1** - **N6**, as stated below, hold for (w, C).

N 1 *For all i, $1 \leq i \leq r$, (u_i, C) is a normal word of height less or equal to $H(w) - 1$ and (v_i, C) in (1) is a normal word of height $H(w) - 1$.*

N 2 *C contains all the conditions $\rho(\alpha_i) > 0$ and does not imply any of the conditions $\rho(\alpha_i) < k$, for all exponents α_i in w, and any $k \in \mathbb{Z}$. For all i, $1 \leq i \leq n$ there is no $z \in W$ and $\beta \in \mathbb{Z}[t] \setminus \{\pm 1\}$, such that (v_i, C) is equivalent to (z^β, C). And if $H(w) \geq 2$, then (v_i, C) is not equivalent to a word (s, C), where s is of height $H(w) - 2$.*

N 3 *Let ρ_n be a retraction satisfying C. Then $\bar{\rho_n}(v_i) \in W^0$ is a non trivial reduced word beginning with a unique $L(v_i)$ as left letter and ending with a unique $R(v_i)$ as right letter. $L(v_i)$ and $R(v_i)$ do not depend on $n \in \omega$. Each $u_i \neq 1$ satisfies the analogous condition.*

N 4 *1. If $u_i \neq \mathbf{1}$ then $R(u_i)L(v_i) \neq \mathbf{1}$.*

2. If $u_{i+1} \neq \mathbf{1}$ then $R(v_i)L(u_{i+1}) \neq \mathbf{1}$.

3. $R(v_i)L(v_i) \neq \mathbf{1}$.

4. $R(v_i)R(v_{i+1}) \neq \mathbf{1}$, if $u_i = \mathbf{1}$.

N 5 *1. If $u_i \neq \mathbf{1}$ then $R(u_i) \neq R(v_i)$.*

2. If $u_i = \mathbf{1}$ then $R(v_i) \neq R(v_{i+1})$.

N 6 *For $1 \leq i \leq r$, there is no word z of height less than or equal to $H(w) - 1$, such that (u_{i+1}, C) is equivalent to $(v_i z, C)$, and such that for all ρ_n satisfying C, $R(v_i)$ does not cancel with $\bar{\rho}_n(z)$.*

We will make use of the following theorem , proved in [Ly 60]):

Lyndon's Theorem I *There is an effective procedure to associate with each word $w \in W$ a finite set of normal words $\{(w_1, C_1), \ldots, (w_m, C_m)\}$, such that if C is the empty set of conditions (w, C) is equivalent to $\{(w_1, C_1), \ldots, (w_m, C_m)\}$ (in the sense explained above).*

The different sets C_i cover the various trichotomies $\rho(\alpha_i) > 0$, $\rho(\alpha_i) = 0$ and $\rho(\alpha_i) < 0$.

For our purpose we single out one particular (w_i, C_i). By the properties of the ultrafilter $\mathcal{D}$ we know that there is an i, $1 \leq i \leq m$, such that

$$A_i = \{n : \rho_n \text{ satisfies } C_i\} \cap \omega \in \mathcal{D}. \tag{3}$$

We assume without loss of generality that (w_1, C_1) is such a word and define it to be a *good normal word for w*.

Lemma 3.3 *There exists $N \in \omega$ such that ρ_n satisfies C_1 if $n > N$. Hence A_1 is cofinite in ω.*

Proof Since (w_1, C_1) is a normal word we know by property **N2** , that C_1 contains all conditions $\rho_n(\alpha_i) > 0$, where α_i is an exponent occurring in w_1. By (3) we know that infinitely many retractions ρ_n, $n \in \omega$, satisfy C_1. Since $\mathcal{D}$ is a non principal ultrafilter it follows that for all exponents α_i in w_1, $\lim_{n \to \infty} \alpha_i(n) = +\infty$ must hold. For every α_i there is a $N_{\alpha_i} \in \omega$ such that $\alpha_i(n) > 0$ for all $n > N_{\alpha_i}$. If we define $N = \max\{N_{\alpha_i} : \alpha_i \text{ an exponent in } w_1\}$, we get

$$\alpha_i(n) > 0 \text{ for all } \alpha_i, \text{ exponent in } w_1, \text{ and all } n > N.$$

This is equivalent to saying

$$\rho_n \text{ satisfies } C_1 \text{ for all } n > N.$$

It follows that A_1 is cofinite in ω. □

It will be of use to add another condition into C_1. We define

$$C_1^N = C_1 \cup \{\rho(t - N) > 0\}.$$

Then ρ_n satisfies C_1^N if and only if $n > N$. Moreover (w_1, C_1^N) is still a normal word.

Now we embark on assigning a normal word w_g to a group element $g \in \mathcal{F}_2$. We do this by choosing an arbitrary word $w \in W$ representing $g \in \mathcal{F}_2$. Let (w_1, C_1) be a good normal word for w. Then we associate to g the normal word w_1 . We have to show that this assignment is well defined and does not depend on any set of conditions. The tool to show this is Lyndon's uniqueness Lemma (Lemma D, [Ly 60]).

Lyndon's Uniqueness Lemma *Let (w, C) and (v, C) be normal words. If we have $\pi'(w) = \pi'(v)$ then $w = v$ holds.*

So let $g \in \mathcal{F}_2$ and let $w, w' \in W$ be two words representing g. Let (w_1, C_1) be a good normal word for w and (w_1', C_1') be a good normal word for w'. Furthermore let $N, N' \in \omega$ be such that ρ_n satisfies C_1 if $n > N$ and ρ_n satisfies C_1' if $n > N'$. These N and N' do exist by the previous Lemma. Define $M = \max\{N, N'\}$. So we get that ρ_n satisfies C_1 and C_1' if $n > M$. Moreover (w_1, C_1^M) and $(w_1', C_1'^M)$ are normal words. Now ρ_n satisfies C_1^M if and only if ρ_n satisfies $C_1'^M$ if and only if $n > M$. Thus we can replace $C_1'^M$ by C_1^M, whence (w_1', C_1^M) is a normal word too. This implies by Lyndon's uniqueness lemma that $w_1 = w_1'$. This argument shows as well that the good normal word w_1, associated to g is independent of a concrete set of conditions.

We fix some notation: We will denote the unique normal word for $g \in \mathcal{F}_2$ by w_g. Say $w_g = u_1 v_1^{\alpha_1} u_2 v_2^{\alpha_2} \dots v_r^{\alpha_r} u_{r+1}$. Let $\bar{\bar{\rho}}_\eta(w_g) = \langle w_{gn} \rangle$, i.e. $w_{gn} = \bar{\bar{\rho}}_{\eta_n}(w_g)$ and we can write as well:

$$\begin{array}{lcll} w_{gn} & = & u_1 \in W^0 & \text{iff } H(g) = 0 \\ w_{gn} & = & u_{1n} v_{1n}{}^{\rho_{\eta_n}(\alpha_1)} \dots v_{rn}{}^{\rho_{\eta_n}(\alpha_r)} u_{r+1n} \in W^0 & \text{iff } H(g) \geq 1 \end{array} \tag{4}$$

Note that if the height of g is 1, we can write

$$w_{gn} = u_1 v_1^{\rho_{\eta_n}(\alpha_1)} \dots v_r^{\rho_{\eta_n}(\alpha_r)} u_{r+1} \in W^0,$$

since u_i, v_i are of height 0, and so $u_i, v_i \in W^0$.

Lemma 3.4 *Let $g \in \mathcal{F}_2$ be such that $g \neq 1$ and let w_g be the normal form of g, i.e. $\bar{\rho}_\eta(w_g) = \langle w_{gn} \rangle$ as defined above. Then*

1. $\{n : u_{in}, v_{in}$ *do not cancel* $\} \in \mathcal{D}$
2. $\{n : v_{in}, u_{i+1n}$ *do not cancel* $\} \in \mathcal{D}$
3. $\{n : v_{in}$ *are cyclically reduced* $\} \in \mathcal{D}$
4. $\{n : w_{gn} \in F_2$ *is non trivial and reduced* $\} \in \mathcal{D}$, *i.e. g is '$* - reduced$'.

Proof First note that 1. - 3. are clear if g is of height 0. So let g be of height $H(g) \geq 1$. By the definition of the normal word w_g, we know that all ρ_n, such that $n > N$, satisfy C_1^N. Since $\eta \in {}^*\mathbb{Z}\backslash\mathbb{Z}$, we have $\{n : \eta_n > N\} \in \mathcal{D}$. If $\eta_n > N$ then ρ_{η_n} satisfies C_1^N. By the filter properties of $\mathcal{D}$ we see that $\{n : \rho_{\eta_n} \text{ satisfies } C_1^N\} \in \mathcal{D}$. Furthermore, if ρ_{η_n} satisfies C_1^N, we know from the property **N3** that $\bar{\bar{\rho}}_{\eta_n}(u_i)$ is a nontrivial reduced word which has the right letter $R(u_i)$ and $\bar{\bar{\rho}}_{\eta_n}(v_i)$ is a nontrivial reduced word which has the left letter $L(u_i)$. By **N4.1** we know that $R(u_i)L(u_i) \neq e$. This means that, if ρ_{η_n} satisfies C_1^N then u_{in} and v_{in} do not cancel, i.e.

$$\{n : u_{in}, v_{in} \text{ do not cancel } \} \supseteq \{n : \rho_{\eta_n} \text{ satisfies } C_1^N\} \in \mathcal{D},$$

and it follows by the filter properties of $\mathcal{D}$ that

$$\{n : u_{in}, v_{in} \text{ do not cancel } \} \in \mathcal{D}.$$

Part 2. and 3. follow equally using **N4.2** , respectively **N4.3** , instead of **N4.1** .

In order to prove 4. note that $w_{gn} = \bar{\bar{\rho}}_{\eta_n}(w_g)$ is a non trivial reduced word if $\bar{\bar{\rho}}_{\eta_n}$ satisfies C_1^N. So

$$\{n : w_{gn} \in F_2 \text{ is non trivial and reduced } \} \supseteq \{n : \rho_{\eta_n} \text{ satisfies } C_1^N\} \in \mathcal{D},$$

and it follows that $\{n : w_{gn} \in F_2 \text{ is non trivial and reduced } \} \in \mathcal{D}$. This finishes the proof. □

Lemma 3.5 *For any $\eta \in {}^*\mathbb{Z} \setminus \mathbb{Z}$, $\eta > 0$, the map*

$$\bar{\rho}_\eta : \mathcal{F}_2 \longrightarrow {}^*F_2$$

is an embedding of groups.

Proof The map $\bar{\rho}_\eta$ is a is a group homomorphism since ρ_{η_n}s is a is a group homomorphism for all $n \in \omega$, as already pointed out. Let $g \in \mathcal{F}_2$ such that $g \neq 1$. By the previous Lemma we know that $\{n : w_{gn} \in F_2 \text{ is non trivial and reduced } \} \in \mathcal{D}$. It follows that

$$\bar{\rho}_\eta(g) = \langle \pi(\bar{\bar{\rho}}_\eta(w_g)) \rangle = \langle \pi(w_{gn}) \rangle \neq 1 \in {}^*F_2.$$

This shows that $\bar{\rho}_\eta$ is injective. □

Note that $\rho_\eta(\mathbb{Z}[t])$ is a subring of ${}^*\mathbb{Z}$, isomorphic to $\mathbb{Z}[t]$. We will identify $\mathcal{F}_2$ with its isomorphic image $\bar{\rho}_\eta(\mathcal{F}_2)$ and W with $\bar{\bar{\rho}}_\eta(W)$. Moreover we identify π^{ρ_η} with π, where π^{ρ_η} is the map which makes the following diagram commute.

$$\begin{array}{ccc} W & \xrightarrow{\bar{\bar{\rho}}_\eta} & \bar{\bar{\rho}}_\eta(W) \\ \downarrow \pi & & \downarrow \pi^{\rho_\eta} \\ \mathcal{F}_2 & \xrightarrow{\bar{\rho}_\eta} & \bar{\rho}_\eta(\mathcal{F}_2) \end{array}$$

The injectivity of $\bar{\rho}_\eta$ enables us to associate with every $\bar{\rho}_\eta(g) \in {}^*F_2$ a normal word $\bar{\bar{\rho}}_\eta(w_g) \in {}^*(W^{\mathrm{O}})$. Then we will identify $\bar{\rho}_\eta(g)$ and g in ${}^*F_2$ and identify $\bar{\bar{\rho}}_\eta(w_g)$ with w_g in ${}^*(W^{\mathrm{O}})$. We will write

$$w_g = u_1 v_1^{\rho_\eta(\alpha_1)} u_2 v_2^{\rho_\eta(\alpha_2)} \dots v_r^{\rho_\eta(\alpha_r)} u_{r+1}.$$

By this convention we can write $w_g = \langle w_{gn} \rangle$ and so we have as in (4):

$$w_{gn} = u_{1n} v_{1n}^{\rho_{\eta_n}(\alpha_1)} \dots v_{rn}^{\rho_{\eta_n}(\alpha_r)} u_{r+1n} \in W^{\mathrm{O}}.$$

4 Λ-trees and Bass-Serre theory

The reader should be familar with the notion of Λ-metric-spaces, actions of groups on graphs and Bass-Serre theory. Here we restrict ourselves to setting some notation which will be used later on.

Let Λ be a totally ordered group, written additively. For every such group we can define a map $|\cdot|_\Lambda : \Lambda \to \Lambda$, by mapping

$$\lambda \mapsto \begin{cases} \lambda & \text{if } \lambda \geq 0 \\ -\lambda & \text{if } \lambda < 0. \end{cases}$$

If there is no need to specify Λ or Λ is clear, we write $|\lambda|$ instead of $|\lambda|_\Lambda$, where $\lambda \in \Lambda$.

Let (X, d) be a Λ-metric space and let $x, y, v \in X$. As in [Al 87] we then define

$$x \wedge_v y := \frac{1}{2}(d(x,v) + d(y,v) - d(x,y)) \in \frac{1}{2}\Lambda.$$

Bass-Serre Theory for $\Lambda = \mathbb{Z}$

Graph of groups ([Se], Definition 8). Let Y be a connected nonempty graph. A graph of groups $\mathcal{G}$ consists of Y and

1. a group $\mathcal{G}_x$ for every $x \in V(Y)$,
2. a group $\mathcal{G}_e$ for every $e \in E(Y)$, such that $\mathcal{G}_e = \mathcal{G}_{\bar{e}}$,
3. and an embedding $(\cdot)^e : \mathcal{G}_e \to \mathcal{G}_{t(e)}$, $g \mapsto g^e$, for all $e \in E(Y)$.

Consequently we denote the image of $\mathcal{G}_e$ in $\mathcal{G}_{t(e)}$ by $\mathcal{G}_e^e$. Note that by 2. and 3. there is an embedding $\mathcal{G}_e \to \mathcal{G}_{o(e)}$, defined by $g \mapsto g^{\bar{e}}$. Then the image of $\mathcal{G}_e$ in $\mathcal{G}_{o(e)}$ is denoted by $\mathcal{G}_e^{\bar{e}}$.

The fundamental group of a graph of groups Next we define, as in [Se], chapter I, §5.1

$$F(\mathcal{G}, Y) = \Big(* \underset{x \in Y}{} \mathcal{G}_x * E(Y) \Big) / \left\langle \mathcal{R}'^G \right\rangle, \tag{5}$$

the free product of the vertex groups $\mathcal{G}_x$ and the free group generated by the elements of $E(Y)$, modulo the normal subgroup generated by the following relations, which constitute $\mathcal{R}'$:

1. $e^{-1} = \bar{e}$ for all $e \in E(Y)$
2. $eg^e e^{-1} = g^{\bar{e}}$ for all $e \in E(Y)$ and all $g \in \mathcal{G}_e$.

Let T be a maximal subtree of Y. We define

$$\mathcal{R} = \mathcal{R}' \cup \{e = 1 : e \in E(T)\}$$

and

$$\pi_1(\mathcal{G}, Y, T) = \Big(* \underset{x \in Y}{} \mathcal{G}_x * E(Y) \Big) / \left\langle \mathcal{R}^G \right\rangle. \tag{6}$$

We call $\pi_1(\mathcal{G}, Y, T)$ the *fundamental group of $\mathcal{G}$ relative to T.*

Serre's structure theorem *Let G be a group acting without inversions on a nonempty connected graph X. Let $Y = G \setminus X$ and T be a maximal subtree of Y. Let $\mathcal{G}$ be the graph of groups constructed in [Se] Then $G \cong \pi_1(\mathcal{G}, Y, T)$ if and only if X is a tree.*

Base change

We will need to apply the theorem about base change (Proposition 4.4 in [Al 87]) in a special case, presented in the following Lemma.

Lemma 4.1 *Let $\Lambda = (\mathbf{Z}^m, <)$, where $m \in \omega$ and $<$ is the lexicographic ordering and $\Lambda^* = \mathbf{Z}$ is equipped with the usual ordering. Let*

$$\begin{aligned} pr : \mathbf{Z}^m &\to \mathbf{Z} \\ (a_1, \ldots, a_m) &\mapsto a_m, \end{aligned}$$

be the projection map where a_m is the biggest component. Let (X, d) be a $\mathbf{Z}^m$-tree. Then there is a tree (X^*, d^*) and a mapping $\psi : X \to X^*$ such that for all $x, y \in X$

$$d^*(\psi(x), \psi(y)) = pr(d(x, y)). \tag{7}$$

Moreover, if (X', d') is another tree and $\psi' : X \to X'$ a map satisfying

$$d'(\psi'(x), \psi'(y)) = pr(d(x, y)) \text{ for all } x, y \in X,$$

then there is a unique metric morphism of trees $\mu : X^* \to X'$ such that $\mu \circ \psi = \psi'$.

Notation (Analogous to the notation in [Ba 91]).
From now on we will write x^*, for $\psi(x)$ and for any subset $Z \subseteq X$ we will write Z^* instead of $\psi(Z)$. We let $X^* = \mathbf{Z} \otimes_{\mathbf{Z}^m} X$. Furthermore we define

$$X(x^*) = \{y \in X : d(x, y) \in ker(pr)\} = \psi^{-1}(\psi(x)).$$

Since $ker(pr) = \mathbf{Z}^{m-1}$ we have that $X(x^*)$ is a $\mathbf{Z}^{m-1}$-tree.

Proofs of the following Lemma can be found in [Ba 91]

Lemma 4.2 *With the assumptions above the following is true*

1. *If G acts on X as isometries , then G acts on X^* as isometries.*
2. *If G acts without inversions on X, then G acts without inversions on X^*.*

The following Lemma is a synthesis of Propositions 1.6.a. and 1.7.d. in [Ba 91]. We are giving it here, since it introduces the map τ_ε, which is then important for Bass' generalized structure theorem ([Ba 91]). It combines base change with Serre's structure theorem.

Lemma 4.3 *Let (X,d) be a Λ-tree and let G be a group which acts on X as isometries. Denote the end stabilizers by $G_\varepsilon = \{g \in G : g\varepsilon = \varepsilon\}$. Then for every end ε there is a homomorphism*

$$\tau_\varepsilon : G_\varepsilon \to \Lambda.$$

If the action of G on X is free then τ_ε is injective.

Bass' generalized structure theorem

(Theorem 3.5., [Ba 91]).
Let $\Lambda = (\mathbb{Z}^m, <)$, where $m \in \omega$ and $<$ is the lexicographic ordering. Let G be a group acting on a $\mathbb{Z}^m$-tree (X,d) as isometries and without inversions. Let $X^* = \mathbb{Z} \otimes_{\mathbb{Z}^m} X$ and $Y = G \setminus X^*$. Moreover let T be a maximal subtree of X^* and $p : X^* \to Y$ be the projection map. We denote points of X^* by x^*, the points of Y by $\tilde{x} = p(x^*)$.

Then G acts on X^* without inversions and there is a graph of groups $\mathcal{G}$ such that $G \cong \pi_1(\mathcal{G}, Y, T)$. Furthermore if $\mathcal{G}_{\tilde{x}}$ denote the vertex groups and $\mathcal{G}_e$ denote the edge groups, the following holds

1. *for each $\tilde{x} \in Y$, $\mathcal{G}_{\tilde{x}}$ acts on the $\mathbb{Z}^{m-1}$-tree $X(x^*)$.*
2. *for each edge $e \in E(Y)$ there is an end ε_e in $X(o(e))$ of full $\mathbb{Z}^{m-1}$-type, satisfying*

$$\mathcal{G}_e^{\bar{e}} = (\mathcal{G}_{o(e)})_{\varepsilon_e}, \tag{8}$$

i.e. the image of $\mathcal{G}_e$ in $\mathcal{G}_{o(e)}$ is the end stabilizer of ε_e [of the action of $\mathcal{G}_{o(e)}$ on $X(o(e))$]. That means that the elements $g \in \mathcal{G}_{o(e)}$ stabilizing e are exactly the elements $g \in \mathcal{G}_{o(e)}$ stabilizing ε_e. Then we also have $\mathcal{G}_e^e = (\mathcal{G}_{t(e)})_{\varepsilon_{\bar{e}}}$. Furthermore

$$\tau_{\varepsilon_e}(g^{\bar{e}}) = -\tau_{\varepsilon_{\bar{e}}}(g^e) \text{ for } g \in \mathcal{G}_e \ . \tag{9}$$

If $f \neq e$ are edges of Y and $o(f) = o(e) = \tilde{x}$, then ε_f and ε_e are in distinct orbits of the action of $\mathcal{G}_{\tilde{x}}$ on $Ends(X(x^*))$.

3. *If G acts freely on (X,d) then $\mathcal{G}_{\tilde{x}}$ acts freely on $X(x^*)$ for all $\tilde{x} \in V(Y)$. Moreover $\tau_{\varepsilon_e} \circ {}^e : \mathcal{G}_e \to \mathbb{Z}^m$, $g \mapsto \tau_{\varepsilon_e}(g^e)$ is injective and $\mathcal{G}_e$ are finitely generated free abelian groups.*

5 Lyndon length functions

Definition 5.1 Cf. [Ly 63]. *Let Λ be an ordered abelian group. Let G be a group. A mapping $L : G \to \Lambda$ is called a Lyndon length function if it has the following properties:*

1. $L(1) = 0$
2. $L(g) = L(g^{-1})$ *for all* $g \in G$
3. *Let* $\delta : G \times G \to \frac{1}{2}\Lambda$ *be the map*

$$\delta(f,g) = \frac{1}{2}(L(f) + L(g) - L(f^{-1}g)).$$

   Then $\delta(f,g) \in \Lambda$ *for all* $f, g \in G$.

4. $\delta(f,g) \geq \min\{\delta(h,g), \delta(f,h)\}$ *for all* $f, g, h \in G$.

Let $L : F_2 \longrightarrow \mathbf{Z}$ be the usual length function for F_2. So L maps $g \in F_2$ to the length of the shortest word $w \in W^0$ which represents g. This is obviously a Lyndon length function.

Now in $V(^*M)$ there is a map

$$^*L : {}^*F_2 \longrightarrow {}^*\mathbf{Z}.$$

This map is defined as

$$^*L(\langle g_n \rangle) = \langle L(g_n) \rangle$$

for $\langle g_n \rangle \in {}^*F_2$. Since we can express 1.-4. in Definition 5.1 as first order formulae in the language of $S_{V(M)}$, it follows by the theorem of Loš that these properties hold for $^*L$ as well. Whence $^*L$ is a Lyndon length function too. Let $\mathcal{L}' = {}^*L|\mathcal{F}_2$ be the restricted Lyndon length function $^*L$ for $\mathcal{F}_2$,

$$\mathcal{L}' : \mathcal{F}_2 \longrightarrow {}^*\mathbf{Z}.$$

The length of $g = \langle g_n \rangle \in {}^*F_2$ will be calculated via the length of the normal word w_g, representing g. This is admissible since $\{n : w_{gn} \in F_2 \text{ is reduced }\} \in \mathcal{D}$ by Lemma 3.4. By convention we then write $\mathcal{L}(g) = \mathcal{L}(w_g)$.

Lemma 5.2 *Let G be a finitely generated subgroup of $^*F_2$ with generators from $\mathcal{F}_2$ and let $\{g_1, \ldots g_n\}$ be a generating system. Then there is an $m \in \omega$ and a Lyndon length function*

$$\mathcal{L} : G \longrightarrow (\mathbf{Z}^m, <),$$

where $<$ is the lexicographic ordering on $\mathbf{Z}^m$.

Proof First we want to show that that the image of $\mathcal{L}'$ is contained in $\rho_\eta(\mathbb{Z}[t])$. We do this by defining inductively maps $\mathcal{L}_K : F^K \to \mathbb{Z}[t]$, for $K \in \omega$, where here $F^K = \{g \in \mathcal{F}_2 : H(g) \leq K\}$ such that the following equation holds

$$\rho_\eta(\mathcal{L}_K(g)) = \mathcal{L}'(g) \text{ for all } g \in \mathcal{F}_2 \text{ such that } H(g) = K. \tag{10}$$

The start of the induction is trivial, since then $g \in F_2$ and $\mathcal{L}'|F_2 = L$ holds.

Assume we have shown that there is a map $\mathcal{L}_K : F^K \to \mathbb{Z}[t]$ such that (10) holds for $K \geq 1$. Let $g \in \mathcal{F}_2$ be of height $H(g) = K + 1$, given in normal representation as follows:

$$w_g = u_1 v_1^{\rho_\eta(\alpha_1)} u_2 v_2^{\rho_\eta(\alpha_2)} \dots v_r^{\rho_\eta(\alpha_r)} u_{r+1},$$

where $u_i, v_i \in \mathcal{F}_2$ of height $H(g) \leq K$. We can calculate $\mathcal{L}'(g) = \mathcal{L}'(w_g)$, since by Lemma 3.4.4 $\{n : w_{gn} \text{ is reduced } \} \in \mathcal{D}$. We get

$$\mathcal{L}'(g) = \mathcal{L}'(u_1) + \mathcal{L}'(v_1^{\rho_\eta(\alpha_1)}) + \dots + \mathcal{L}'(v_r^{\rho_\eta(\alpha_r)}) + \mathcal{L}'(u_{r+1}), \tag{11}$$

since by Lemma 3.4.1 and 3.4.2. we know that $\{n : u_{in} \text{ and } v_{in} \text{ do not cancel } \} \in \mathcal{D}$ and $\{n : v_{in} \text{ and } u_{i+1n} \text{ do not cancel } \} \in \mathcal{D}$, where $u_i = < u_{in} > \in {}^*F_2$ and $v_i = < v_{in} > \in {}^* F_2$. Furthermore

$$\mathcal{L}'(v_i^{\rho_\eta(\alpha_i)}) = \rho_\eta(\alpha_i)\mathcal{L}'(v_i), \tag{12}$$

since $\{n : v_{in} \text{ are cyclically reduced } \} \in \mathcal{D}$ by Lemma 3.4.3. Inserting (12) into (11) we get

$$\begin{aligned} \mathcal{L}'(g) = \quad & \mathcal{L}'(u_1) + \rho_\eta(\alpha_1)\mathcal{L}'(v_1) + \mathcal{L}'(u_2) + \dots \\ & \dots + \rho_\eta(\alpha_r)\mathcal{L}'(v_r) + \mathcal{L}'(u_{r+1}) \in {}^* \mathbb{Z}. \end{aligned} \tag{13}$$

Now we will apply the induction hypothesis on u_i, v_i, since they have height $H \leq K$. This tells us that there is the map $\mathcal{L}_K : F^K \to \mathbb{Z}[t]$, satisfying (10):

$$\rho_\eta(\mathcal{L}_K(u_i)) = \mathcal{L}'(u_i), \quad \rho_\eta(\mathcal{L}_K(v_i)) = \mathcal{L}'(v_i).$$

We substitute this into (13) and get

$$\begin{aligned} \mathcal{L}'(g) = \quad & \rho_\eta(\mathcal{L}_K(u_1)) + \rho_\eta(\alpha_1)\rho_\eta(\mathcal{L}_K(v_1)) + \rho_\eta(\mathcal{L}_K(u_2)) + \dots \\ & \dots + \rho_\eta(\alpha_r)\rho_\eta(\mathcal{L}_K(v_r)) + \rho_\eta(\mathcal{L}_K(u_{r+1})), \end{aligned}$$

and since ρ_η is a ring homomorphism this is equivalent to

$$\mathcal{L}'(g) = \rho_\eta(\mathcal{L}_K(u_1) + \alpha_1\mathcal{L}_K(v_1) + \mathcal{L}_K(u_2) + \dots + \alpha_r\mathcal{L}_K(v_r) + \mathcal{L}_K(u_{r+1})) \in \mathbb{Z}[t].$$

Now define

$$\mathcal{L}_{K+1} : F^{K+1} \to \mathbb{Z}[t]$$

by

$$\begin{aligned} g \mapsto \quad & \mathcal{L}_K(u_1) + \alpha_1\mathcal{L}_K(v_1) + \dots + \alpha_r\mathcal{L}_K(v_r) + \mathcal{L}_K(u_{r+1}) && \text{iff } H(g) = K + 1 \\ g \mapsto \quad & \mathcal{L}_H(g) && \text{iff } H(g) \leq K. \end{aligned} \tag{14}$$

$\mathcal{L}_{K+1}$ is well-defined. Note that $\mathcal{L}_{K+1}|F^K = \mathcal{L}_K$. Furthermore (10) holds by definition. Now we define

$$\begin{array}{rcl} \mathcal{L} : \mathcal{F}_2 & \to & \mathbf{Z}[t] \\ g & \mapsto & \mathcal{L}_H(g), \quad \text{where } H \text{ is the height of } g. \end{array}$$

$\mathcal{L}$ defines a Lyndon length function by (10) and Lemma 3.2.1 which said that ρ_η is an order preserving isomorphism $\mathbf{Z}[t] \xrightarrow{\rho_\eta} \rho_\eta(\mathbf{Z}[t])$. This implies that $<$ is the lexicographic ordering on $\mathbf{Z}[t]$. Now we let G be a finitely generated subgroup of ${}^*F_2$ where the generators $g_1, \ldots, g_n$ are chosen from $\mathcal{F}_2$. Assume without loss of generality that g_1 is the longest generator, i.e. $\mathcal{L}(g_1) \geq \mathcal{L}(g_i)$ for all $i \in \{1, \ldots, n\}$. Say $\mathcal{L}(g_1)$ is a polynomial of degree m'. Let $g \in G$ be such that $g = g_{i_1} \ldots g_{i_k}$, where $g_{i_l} \in \{g_i, g_i^{-1} : \ g_i \text{ generator of G }\}$ for $1 \leq l \leq k$. Thus $\mathcal{L}(g) \leq \mathcal{L}(g_{i_1}) + \ldots + \mathcal{L}(g_{i_k})$, where $\mathcal{L}(g_{i_l})$ are polynomials of degree less or equal to m'. Hence for all $g \in G$ the length $\mathcal{L}(g)$ is a polynomial of degree less or equal to m'. This is equivalent to saying

$$\mathcal{L} : G \to \mathbf{Z}^m,$$

where $m = m' + 1$. □

We will also use the notation $|g|$ $(= \mathcal{L}(g))$ for the length of $g \in G$.

6 Chiswell's Λ -tree (X, d)

The next Lemma is essentially Theorem 3.10 in [ChPre].

Lemma 6.1 *Let G be a finitely generated subgroup of ${}^*F_2$ with generators from $\mathcal{F}_2$. Then there exists an $m \in \omega$ such that G acts freely without inversions on a Λ-tree (X, d), where $\Lambda = (\mathbf{Z}^m, <)$, and $<$ is the lexicographic ordering.*

Construction Before we describe (X, d), note first that G itself gives rise to a Λ-metric space (G, d'), where the metric $d' : G \times G \to \Lambda$ is defined as $d'(g, h) = \mathcal{L}(g^{-1}h)$ for all $g, h \in G$.

Now we construct X as $X = Z/\sim$, where

$$Z := \{(x, n) \in G \times \Lambda : 0 \leq n \leq |x|\}$$

and

$$(x, n) \sim (y, m) \text{ iff } \begin{cases} \text{both} & n = m \\ \text{and} & x \wedge_{\mathbf{1}} y \geq n \text{ are satisfied} \end{cases} \tag{15}$$

Denote the equivalence class of (x, n) by $\langle x, n \rangle$. The point $\langle \mathbf{1}, 0 \rangle$ is the root of the Λ-tree. The metric $d : X \times X \to \Lambda$ is defined by

$$d(\langle x, n \rangle, \langle y, m \rangle) = m + n - 2\min\{n, m, x \wedge_{\mathbf{1}} y\} \in \Lambda. \tag{16}$$

This map is well defined. Then (X, d) constitutes a Λ-tree by Theorem 3.8 in [ChPre], respectively Lemma 3 in [Ch 76]. Moreover, again by [Ch 76] the following

formula holds for $g, x \in G$, $0 \leq n \leq |x|$:

$$g \cdot \langle x, n\rangle = \begin{cases} \langle g, |g| - n\rangle & \text{iff} \quad n \leq g^{-1} \wedge_{\mathbf{1}} x \\ \langle gx, |g| + n - 2(g^{-1} \wedge_{\mathbf{1}} x)\rangle & \text{iff} \quad n \geq g^{-1} \wedge_{\mathbf{1}} x. \end{cases} \tag{17}$$

□

Basechange in Chiswell's Λ-tree (X, d)

Now we apply the base change Lemma 4 to X. So the homomorphism of ordered abelian groups is again the projection map $pr : \mathbb{Z}^m \to \mathbb{Z}$, where $(a_1, \ldots, a_m) \mapsto a_m$. So points of $X^* = \mathbb{Z} \otimes_{\mathbb{Z}^m} X$ will be denoted by $\langle x, n\rangle^*$. As indicated in section 4 on base change, and Lemma 4.2 the isometric action of G on X without inversions, induces an isometric action of G on X^* without inversions. By the definition of the action of G on X^* ($gx^* = (gx)^*$), the following formula holds

$$g \cdot \langle x, n\rangle^* = \begin{cases} \langle g, |g| - n\rangle^* & \text{iff} \quad n \leq g^{-1} \wedge_{\mathbf{1}} x \\ \langle gx, |g| + n - 2(g^{-1} \wedge_{\mathbf{1}} x)\rangle^* & \text{iff} \quad n \geq g^{-1} \wedge_{\mathbf{1}} x. \end{cases} \tag{18}$$

Lemma 6.2 *Let G be a finitely generated group, such that $\{g_1, \ldots, g_n\} \subset \mathcal{F}_2$ is a generating system. Let X be the Λ-tree from Lemma 6.1 and $X^* = \mathbb{Z} \otimes_\Lambda X$. Then $Y = G\backslash X^*$ is a finite graph.*

Proof First note that every point $\langle g, |g|\rangle^*$ lies in the orbit of the basepoint $\langle \mathbf{1}, 0\rangle^*$ because $g \cdot \langle \mathbf{1}, 0\rangle^* = \langle g, |g|\rangle^*$.

We define

$$S := \{\langle g_i, k\rangle^* : g_i \text{ generator or the inverse of a generator of } G,\ 0 \leq k \leq |g_i|^*\}.$$

Since X^* is a $\mathbb{Z}$-tree, we know that $|g_i|^* \in \mathbb{Z}$. If $0 \leq k_1, k_2 \leq |g_i|^*$ and $k_1 - k_2 \in ker(pr) = \mathbb{Z}^{m-1}$ then by 7

$$d^*(\langle g_i, k_1\rangle^*, \langle g_i, k_2\rangle^*) = pr(d(\langle g_i, k\rangle, \langle g_i, k\rangle)) = 0$$

holds. Thus $\langle g_i, k_1\rangle^* = \langle g_i, k_2\rangle^*$ holds and it follows, that S is a finite set.

In order to prove the Lemma, it is enough to show that every point $\langle x, n\rangle^* \in X^*$, where $1 \leq n \leq |x|$, lies in the orbit of a point in S. We will proceed by induction on the minimum length of $x \in G$, written in the generators of G.

Note that $\langle \mathbf{1}, 0\rangle^* \in S$ since we have $\langle \mathbf{1}, 0\rangle \sim \langle g_i, 0\rangle$.

1. First let $\langle x, n\rangle^*$ be such that x be a generator of G. Then for all n, $0 \leq n \leq |x|$, $\langle x, n\rangle^*$ is defined and obviously $\langle x, n\rangle^* \in S$. If x is the inverse of a generator of G, then x^{-1} is a generator of G and we calculate

$$x^{-1} \langle x, n\rangle^* = \left\langle x^{-1}, |x| - n\right\rangle^*,$$

since $n \leq x \wedge_e x = |x|$. So $\langle x, n\rangle^*$ is in the orbit of element of $\langle x^{-1}, |x| - n\rangle^* \in S$.

2. Assume that for all elements x of complexity at most j the points $\langle x, n\rangle^*$, $0 \leq n \leq |x|$, are in the orbit of an element of S. Let u have complexity $j+1$ and suppose $u = hu'$, where $h \in \{g_i, g_i^{-1} : g_i \text{ generator of } G\}$ and $u' \in G$ is of complexity j. Let $n \in \Lambda$ be such that $\langle u, n\rangle^* \in X^*$. Choosing $g = h^{-1}$ in (18), we get

$$h^{-1} \cdot \langle hu', n\rangle^* = \begin{cases} \langle h^{-1}, |h^{-1}| - n\rangle^* & \text{iff } n \leq h \wedge_{\mathbf{1}} u \\ \langle u', |h| + n - 2(h \wedge_{\mathbf{1}} u)\rangle^* & \text{iff } n \geq h \wedge_{\mathbf{1}} u. \end{cases}$$

First assume that $n \leq h \wedge_{\mathbf{1}} u$. Now, if $h = g_j^{-1}$, we succeed immediately in the first case, since then $\langle x, n\rangle^*$ is in the orbit of $\langle g_j, |g_j| - n\rangle^* \in S$. If $h = g_k$, then an argument analogous to the first step of the induction shows that $\langle u, n\rangle^*$ is in the orbit of $\left\langle g_k^{-1}, |g_k^{-1}| - n\right\rangle^* \in S$.

If $n \geq h \wedge_e u$, then we know that $\langle u, n\rangle^*$ is in the orbit of an element of the form $\langle u', n'\rangle^* \in X^*$, which by the induction hypothesis is in the orbit of an element of S.

Let $e \in E(X^*)$ such that $o(e) = \langle x, n\rangle^*$ and $t(e) = \langle x, n+1\rangle^*$. Let $\langle x, n\rangle^*$ be in the orbit of $s_i \in S$ and $\langle x, n+1\rangle^*$ be in the orbit of $s_j \in S$. Then e is in the orbit of an edge $f \in E(X^*)$ such that $o(f) = s_i$ and $t(f) = s_j$. Since S is finite there are only finitely many such edges and this implies that $E(Y)$ is finite. □

7 The presentation of G

Proposition 1 *Let G be a finitely generated subgroup of ${}^*F_2$ with generators from $\mathcal{F}_2$. Then G is finitely presented.*

Proof Let G be a finitely generated subgroup of ${}^*F_2$, such that $\{g_1, \ldots, g_n\} \subset \mathcal{F}_2$ is a generating set for G. By Lemma 5.2 we know that there is a Lyndon length function $\mathcal{L} : G \to (\mathbf{Z}^m, <)$, where $<$ is the lexicographic ordering on $\mathbf{Z}^m$. We proceed by induction on the rank m of Λ :

If $m = 1$ then Lemma 6.1 tells us that G acts freely on a $\mathbf{Z}$-tree. It follows by by Theorem 4 in [Se], that G is a free group. Since G is finitely generated it follows that G is finitely presented.

Assume we have shown that G is finitely presented if $k \leq m-1$. Let G be generated by $g_1, \ldots, g_n \in \mathcal{F}_2$ such that $m = k$. Lemma 6.1 tells us that G acts freely and without inversions on the $\mathbf{Z}^m$-tree (X, d), where $\mathbf{Z}^m$ is equipped with the lexicographic ordering. Now we can apply Bass's generalized structure theorem, introduced in section 4. Then

$$G \cong \pi_1(\mathcal{G}, Y, T)$$

where

1. $X^* = \mathbf{Z} \otimes_{\mathbf{Z}^m} X$, as defined in section 4.
2. $Y = G \backslash X^*$.

3. T is a maximal subtree of Y.

4. $\mathcal{G}$ is the graph of groups constructed in section 4 for the action of G on X^*, as in section 4.

Let $p : X^* \to Y$ be the projection map. We shall denote points of Y by $\tilde{x}, \tilde{y}, \tilde{z}, \ldots$.

We know the presentation of $\pi_1(\mathcal{G}, Y, T)$ from the definition in section 4, i.e. by equation (6):

$$\pi_1(\mathcal{G}, Y, T) = \Big(* \underset{\tilde{x} \in Y}{} \mathcal{G}_{\tilde{x}} * E(Y) \Big) / \left\langle \mathcal{R}^G \right\rangle .$$

If $e \in E(Y)$ is an edge such that $t(e) = \tilde{x}$ then we denote the image of $g \in \mathcal{G}_e$ in $\mathcal{G}_{\tilde{x}}$ by g^e. The set $\mathcal{R}$ consists of the following relations

1. $e^{-1} = \bar{e}$ for all $e \in E(Y)$

2. $eg^e e^{-1} = g^{\bar{e}}$ for all $e \in E(Y)$ and all $g \in \mathcal{G}_e$

3. $e = 1$ if $e \in E(T)$.

By Bass' generalized structure theorem $\mathcal{G}_{\tilde{x}}$ is $\mathbb{Z}^{m-1}$-free. Thus $\mathcal{G}_{\tilde{x}}$ acts freely on $X(\langle x, n \rangle^*)$, where $p(\langle x, n \rangle) = \tilde{x}$.

Now we can apply the induction hypothesis on the vertex stabilizers. It follows that all the $\mathcal{G}_{\tilde{x}}$ are finitely presented. Then $* \underset{\tilde{x} \in Y}{} \mathcal{G}_{\tilde{x}}$ is finitely presented. Further, by Lemma 6.2, $E(Y)$ is a finite set and so contributes finitely many generators to $\pi_1(\mathcal{G}, Y, T)$.

Finally we have to show that the set $\mathcal{R}$ of relations is equivalent to a finite set $\mathcal{R}'$ of relations. There are only finitely many relations given by 1. and 3. above because Y is a finite graph by Lemma 6.2. Since the action of G on X is free, Bass's generalized structure theorem tells us that $\mathcal{G}_e$ is a finitely generated free abelian group and finitely presented.

It is enough to consider the relations in 2. above for the generators of the edge groups $\mathcal{G}_e$. Suppose $g = g_{i_1} \ldots g_{i_k} \in \mathcal{G}_e$, where g_{i_l} is a generator or the inverse of a generator of $\mathcal{G}_e$, then the relation $eg^e e^{-1} = g^{\bar{e}}$ can be derived from the relations

$$eg_i^e e^{-1} = g_i^{\bar{e}}, \quad 1 \leq i \leq n.$$

First we see that by $eg_i^e e^{-1} = g_i^{\bar{e}}$ we have also $(eg_i^e e^{-1})^{-1} = (g_i^{\bar{e}})^{-1}$ and it follows that

$$eg_i^{-1^e} e^{-1} = g_i^{-1^{\bar{e}}},$$

since $(\cdot)^e$ and $(\cdot)^{\bar{e}}$ are group homomorphisms. Thus we can derive the relation

$$(eg_{i_1}^e e^{-1})(eg_{i_2}^e e^{-1}) \ldots (eg_{i_k}^e e^{-1}) = g_{i_1}^{\bar{e}} g_{i_2}^{\bar{e}} \ldots g_{i_k}^{\bar{e}}.$$

This is equivalent to

$$eg_{i_1}^e g_{i_2}^e \ldots g_{i_k}^e e^{-1} = g_{i_1}^{\bar{e}} g_{i_2}^{\bar{e}} \ldots g_{i_k}^{\bar{e}}$$

and since $(\cdot)^e$ and $(\cdot)^{\bar{e}}$ are group homomorphisms we get

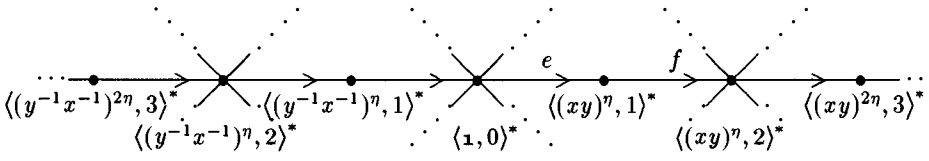
$$eg^e e^{-1} = g^{\bar{e}}.$$

Thus we get the finite set $\mathcal{R}'$ of relations by deleting from $\mathcal{R}$ all of the relations of the second type except the ones for the generators of the $\mathcal{G}_e$. This proves the proposition. □

Example Let $g_1 = x$, $g_2 = y$ and $g_3 = (xy)^\eta$, all given in normal form. Let $G = \langle g_1, g_2, g_3 \rangle \subset {}^*F_2$. Then $\mathcal{L}(g_1) = \mathcal{L}(g_2) = 1$ and $\mathcal{L}(g_3) = (0,2) \in \mathbf{Z}[t]$. So the upper bound for the degree of polynomials in $\mathcal{L}(G)$ is 1 and

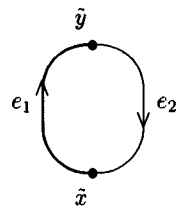
$$\mathcal{L} : G \longrightarrow \mathbf{Z} \times \mathbf{Z}.$$

This implies by Lemma 6.1, that G acts freely without inversions on a $\mathbf{Z}^2$-tree X. As above let $X^* = \mathbf{Z} \otimes_{\mathbf{Z}^2} X$. Then we denote, for example $\langle (xy)^\eta, (0,1) \rangle^*$ by $\langle (xy)^\eta, 1 \rangle^* \in X^*$. We can draw X^* schematically as the following graph:



The big crosses in the points $\left\langle (xy)^{l \cdot \eta}, m \right\rangle^*$ denote branches in the tree X^*. In fact in any such point there is one branch (i.e. edges of length one in X^*) for any $g \in F_2$. Two elements $g, g' \in F_2$ have the same branch, if and only if the reduced word representing $g^{-1}g'$ is of the form $(xy)^n$, for some $n \in \mathbf{Z}$.

By Lemma 6.2 we know that $Y = G \setminus X^*$ is a finite graph. In our case Y is the following graph:



Here $\tilde{x} = G\langle 1, 0 \rangle^*$ and $\tilde{y} = G\langle (xy)^\eta, 1 \rangle^*$. The maximal subtree T consists of the points $\tilde{x}$, $\tilde{y}$ and the edge e_1. In order to find the presentation of G we first need to know the edge and vertex groups $\mathcal{G}_{e_1}, \mathcal{G}_{e_2}, \mathcal{G}_{\tilde{x}}, \mathcal{G}_{\tilde{y}}$. To do this we need to fix the lift $j : Y \to X^*$. j is naturally defined on $T \subset Y$, we lift $\tilde{x} \in Y$ to $\langle 1, 0 \rangle^* \in X^*$, $\tilde{y} \in Y$ to $\langle (xy)^\eta, 1 \rangle^* \in X^*$ and $e_1 \in Y$ to $e \in X^*$. We decide to lift e_2 to the edge $f \in X^*$, and this choice determines $\gamma_{e_2} = (y^{-1}x^{-1})^\eta$. Now we get

1. $\mathcal{G}_{\tilde{x}}$ is the stabilizer of $\langle 1, 0 \rangle^*$ in X^*, $G_{\langle 1,0 \rangle^*}$ as defined in section 4 on the construction of the graph $\mathcal{G}$. In our case this is the subgroup of G generated by x and y, i.e. the free group in two generators.

2. $\mathcal{G}_{\bar{y}}$ is the stabilizer of $j\langle\langle(xy)^\eta, 1\rangle^*$ in X^*. This is in our case the subgroup of G generated by xy.

3. $\mathcal{G}_{e_1}$ is the stabilizer of e in X^* and in this example equals the cyclic subgroup of G generated by xy.

4. $\mathcal{G}_{e_2}$ is also $\langle xy\rangle$, since f is stabilized by xy.

Note that all vertex groups are indeed free groups and so act freely on a $\mathbb{Z}$-tree by theorem 4 in [Se].

Since we know that $G \cong \pi_1(\mathcal{G}, Y, T)$, we know that G is generated by the vertex groups $\mathcal{G}_{\bar{x}}$, $\mathcal{G}_{\bar{y}}$ and γ_{e_1}, $\gamma_{\bar{e_1}}$, γ_{e_2} and $\gamma_{\bar{e_2}}$, subject to the following relations :

1. $\gamma_{e_1} = 1 \in G$, since $e_1 \in E(T)$.

2. $\gamma_{\bar{e_i}} = \gamma_{e_i}^{-1}$ for $i \in \{1, 2\}$.

3. $\gamma_{e_1} xy \gamma_{e_1}^{-1} = xy$ and $\gamma_{e_2} xy \gamma_{e_2}^{-1} = xy$.

Since we know that $\gamma_{e_2} = (y^{-1}x^{-1})^\eta$, and using the symbol a for γ_{e_2}, we see that G has a presentation as

$$G = \left\langle x, y, a \mid axya^{-1} = xy \right\rangle.$$

References

[Al 87] Roger Alperin and Hyman Bass,"Length functions of group actions on Λ-trees". In: Combinatorial group theory and topology (ed. by S.M.Gersten and John R.Stallings), Annals of Mathematics Studies **111**, pp 265-378. Princeton University Press 1987.

[Ba 91] Hyman Bass, "Group Actions on Non-Archimedean Trees". In: Arboreal Group Theory (ed. by Roger C. Alperin), MSRI Publications **19**, pp 69-131. Springer Verlag, New York, 1991.

[Ch 76] I.M.Chiswell, "Abstract length functions in groups". In: *Mathematical Proceedings of the Cambridge Philosophical Society* **80** (1976), pp 451-463.

[ChPre] I.M.Chiswell, "Introduction to Λ-trees". Preprint.

[Co] Daniel E. Cohen,"Combinatorial Group Theory: a topological approach", LMS Student Texts 14, Cambridge University Press, 1989.

[Ga 94] A.M.Gaglione and D.Spellmann, "Generalisations of Free Groups: Some Questions". In: *Communications in Algebra*, **22(8)**, (1994), pp 3159 - 3169.

[Gr] G.Graetzer,"Universal Algebra". D.van Nostand Company, 1968.

[La] S. Lang, "Algebra", third edition, Addison Wesley, 1993.

[Li 88] T. Lindsrøm, "An invitation to Nonstandard Analysis". In: Nonstandard Analysis and its Applications (ed. by N. Cutland), LMS Student Texts **10**, Cambridge University Press, 1988.

[Ly 60] Roger C.Lyndon, "Groups with parametric exponents", *Transactions American Mathematical Society* **96** (1960) ,pp 518-533.

[Ly 63] Roger C.Lyndon, "Length functions in Groups", *Mathematica Scandinavia* **12** (1963), pp 209-234.

[Pf] Patrick Pfander, "Finitely generated subgroups of ${}^*F_2$", Masters thesis, 1994

[Re 89] V.N. Remeslennikov, "$\exists$-free Groups", *Siberian Mathematical Journal* **30** (1989), pp 193-197.

[Se] Jean-Pierre Serre, "Trees", Springer Verlag, New York, 1980.

Author's address:
Am Pfingstborn 17,
D-55262 Heidesheim,
Germany.

Rings of definable scalars and biendomorphism rings

Kevin Burke, Mike Prest

Abstract

The definable additive endomorphisms of a module form a ring which we call the ring of definable scalars of the module. One is lead, by various routes - model theoretic and algebraic - to consider these endomorphisms and the rings they form. In this paper we show that these rings may be realised as biendomorphism rings of suitably saturated modules and also as endomorphism rings of certain functors. We also consider rings of type-definable scalars and the context of arbitrary sorts.

1 Introduction

Let us consider a (right) module M over a ring R. The elements of R act as scalars on M but, on this particular module, other scalars may act. For instance on any torsionfree divisible $\mathbb{Z}$-module the ring, $\mathbb{Q}$, of rationals has an action extending the action of $\mathbb{Z}$ via the natural embedding of rings $\mathbb{Z} \hookrightarrow \mathbb{Q}$. We require that such "scalars" commute with the R-endomorphisms of the module and hence that they should belong to the biendomorphism ring of the module. But we shall also require that our scalars be definable from the R-action, thus excluding some biendomorphisms. For instance, the biendomorphism ring of the prüfer group $\mathbb{Z}_{p^\infty}$, regarded as a $\mathbb{Z}$-module, is the ring of p-adic integers. This ring is uncountable, so is too large to consist entirely of definable scalars. In fact, it is easily seen (it also follows from Theorem 2.5) that the only actions which are definable from the $\mathbb{Z}$-action are the scalar multiplications by elements of the localisation $\mathbb{Z}_{(p)}$ of $\mathbb{Z}$ at p.

The sense of the term "definable" is such that if two modules are elementarily equivalent then they have the same ring of definable scalars. We will require, furthermore, that if a scalar is defined on a module then it should also be defined (by the same formula) on every power of the module. We begin by showing that these conditions already force our scalars to be pp-definable (for pp-formulas, pp-types and other background from the model theory of modules see [Zie84], [Pre88b] or, for a brief account [Pre93]).

In Section 4 we include a sort for (the logical equivalence class of) each pp-formula $\psi(\bar{v})$ and a relation symbol ϕ/ψ for each pp-pair $\psi \to \phi$ so that we are working in the language L^{eq+} (see [KuPr92]). Given a model of a theory T, a corresponding L^{eq+}-structure contains elements of sort $\psi(\bar{v})$ which correspond to the cosets of $M^{l(\bar{v})}$ modulo $\psi(M)$, and we denote these elements by $\bar{a}_\psi$ and the group they form by M_ψ. The relation symbol ϕ/ψ is interpreted so that $M_\psi \models \phi/\psi(\bar{a}_\psi)$ iff $M \models \phi(\bar{a})$. For a pp-type p with $\psi \notin p$ a pp-formula in the same finite set of free variables, we define the type of p modulo the sort ψ, p_ψ, to be the set $\{\theta + \psi/\psi : \theta \in p\}$.

2 Rings of definable scalars

Lemma 2.1 *Suppose that the formula $\sigma(x,y)$ defines an additive function f on M and also defines, for each n, the function f^n on M^n. Then there is a pp-formula $\rho(x,y)$ such that for all n, $M^n \models \sigma(x,y) \leftrightarrow \rho(x,y)$.*

Proof Let L' denote the language of R-modules enriched by a unary function symbol F. For each n, let $M_n = (M^n, f^n)$ be the L'-structure with F interpreted as f^n. Then $M_n = M_1^n$ and $M_n \models F(x) = y \leftrightarrow \sigma(x,y)$ by hypothesis. Thus, for every n, $M_1^n \models F(x) = y \leftrightarrow \sigma(x,y)$ and hence (see e.g. [ChKe73, 6.3.14]) $M_1^{\aleph_0} \models F(x) = y \leftrightarrow \sigma(x,y)$. That is, in $M^{\aleph_0}$ the formula $\sigma(x,y)$ defines the function $f^{\aleph_0}$. So it will be enough to show that $f^{\aleph_0}$ is pp-definable. But it is in general the case that if N is a module with $N \equiv N^{\aleph_0}$ then any subgroup of any finite power of N which is definable in N is actually pp-definable (apply the method of proof of [Pre88b, 16.5]). □

Since the elementary classes of modules which are closed under formation of direct products and pure submodules are in bijective correspondence with the closed subsets of the Ziegler spectrum, Zg_R, of R it is natural to associate rings of definable scalars to closed subsets of this space. Notice that the set of points of Zg_R on which ρ defines a scalar is the Ziegler-closed subset: $[x = x/\exists y \rho(x,y)] \cap [\rho(0,y)/y = 0]$.

Recall [Zie84] that the points of Zg_R are the (isomorphism classes of) indecomposable pure-injective R-modules and the basic open sets are those of the form $(\phi/\psi) = \{N : \phi(N)/\psi(N) \neq 0\}$ where $\psi \leq \phi$ are pp-formulas. We use the notation $[\phi/\psi]$ for the complement of (ϕ/ψ). By the **support**, $supp(M)$, of a module M we mean the set $\{N \in Zg_R : N$ is a direct summand of some (pure-injective) module elementarily equivalent to $M\}$. By [Zie84, 4.10] this is a closed subset of Zg_R and every closed subset is the support of some module.

Fix a closed subset $C = supp(M)$ of Zg_R. Consider the set of all those pp-formulas $\rho(x,y)$ which define a total function from the first argument to the second argument on M (so that $M \models \forall x \exists y \rho(x,y)$ and $M \models \rho(0,y) \to y = 0$). Hence $\rho(x,y)$ defines a total function on any module with support contained in C. Let $\sim$ denote the equivalence relation on this set given by $\rho \sim \rho'$ iff $M \models \forall x \forall y(\rho(x,y) \leftrightarrow \rho'(x,y))$. Let R_C denote the set of $\sim$-equivalence classes: R_C is the **ring of definable scalars** for C (the operations are addition and composition, as in R). If C is the closed set of some theory $T = Th(M)$ then we also

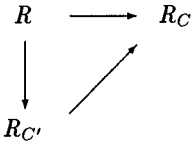
use the notations R_T and R_M for R_C.

We make some immediate observations. The first points out that the set we have defined, equipped with the natural operations, is a ring and that there is a canonical morphism from R to this ring (so to be precise we should mean by the "ring of definable scalars" the R-algebra $R \to R_C$).

Proposition 2.2 *1. There is a natural ring structure on R_C with multiplication given by composition and where addition is pointwise addition. There is a natural ring morphism $R \to R_C$.*

2. Every module M with $supp(M) \subseteq C$ has a natural structure as an R_C-module. The restriction of this structure along the natural map $R \to R_C$ yields the initial R-module structure on M.

3. Corresponding to each inclusion $C \subseteq C'$ of closed subsets of Zg_R we have a commutative diagram as shown.



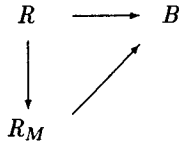
Proof (1) The operations are given as follows. $\rho + \rho'$ is defined by the pp-formula $\exists u, v(\rho(x,u) \wedge \rho'(x,v) \wedge y = u + v)$. $\rho\rho'$ is defined by the pp-formula $\exists z(\rho(x,z) \wedge \rho'(z,y))$ (recall that we're dealing with right modules). The map $R \to R_C$ is that which sends $r \in R$ to the class of the formula $y = xr$. Parts (2) and (3) are immediate from what has been said above. □

Observe that since pp-formulas define our scalars, if $M = M' \oplus M''$ is a decomposition of R-modules then it is also a decomposition of R_M-modules.

Lemma 2.3 *The ring of definable scalars for Zg_R is precisely R.*

Proof Suppose that $\rho(x,y)$ is a pp-formula defining a function on every right R-module; say $\rho(x,y)$ is $\exists \bar{v}(x\ y\ \bar{v})H = 0$. In particular, there is some $s \in R$ with $R \models \rho(1,s)$; say $R \models (1\ s\ \bar{r})H = 0$. Then, for every R-module M and every $m \in M$ we have $m.(1\ s\ \bar{r})H = 0$ that is $(m\ ms\ m\bar{r})H = 0$, and so $M \models \rho(m, ms)$. Hence ρ defines multiplication by s in every R-module, as required. □

Lemma 2.4 *Let M_R be any module, let $S = End(M_R)$ be its endomorphism ring and let $B = End(_SM) = Biend(M_R)$ be its biendomorphism ring. Then the ring R_M of definable scalars of M is a subring of B (by an embedding extending the canonical maps of R to R_M and to B).*



Proof Suppose that $r \in R_M$ is defined by the formula $\rho_r(x, y)$. Let $f \in S$. Then $M \models \rho_r(a, b)$ implies $M \models \rho_r(fa, fb)$: that is, $(fa)r = fb = f(ar)$. Thus every element of R_M may be regarded (canonically) as a member of B, as required. □

Later we will see some conditions under which this inclusion is an equality.

One source of rings of definable scalars is localisation (see [Ste75] for undefined terms).

Theorem 2.5 *[Pre95] Let τ be a hereditary torsion theory of finite type on the category of R-modules, cogenerated by the injective R-module E. Let $T = Th(E^{\aleph_0})$. Then the ring R_T of definable scalars for T coincides with the localisation R_τ of R at τ.*

The above result covers Ore localisation for instance (and so justifies the remarks relating to the prüfer group $\mathbb{Z}_{p^\infty}$ in the introduction). The result follows from (but also inspires) a more general result, Proposition 4.6, which we will prove below. We also recall [Ste75] at this point that the localisation R_τ above can be obtained as the biendomorphism ring of a suitably large power of the injective cogenerator E. This result also follows from and inspires one below (Theorem 4.7). The next result covers Ore localisation but also, for instance, universal localisation (see, e.g., [Sch85]).

Theorem 2.6 *[Pre96] Let $R \to S$ be a ring epimorphism. Then the ring of definable scalars of the module S_R is exactly S. More precisely, if C is the closed subset of the Ziegler spectrum associated to $Th(S_R)$ then $R \to S$ is isomorphic to $R \to R_C$.*

It is easily checked that if R is $\mathbb{Z}$ or $K[X]$ (K a field) then one obtains, as rings of definable scalars, exactly the epimorphisms from R.

Example 2.7 *$R \to R_C$ need not be an epimorphism of rings. We will prove below that if M is a module of finite length over its endomorphism ring then all its biendomorphisms are pp-definable and hence that its ring of definable scalars equals its biendomorphism ring. So all we have to do is exhibit a finite-dimensional module M over an algebra R with the natural map from R to $Biend(M)$ not an epimorphism of rings.*

We may take R to be the algebra $K[a, b, c : ab = ac = bc = 0 = a^2 = b^2 = c^2]$ where K is any field. For the module take the four-dimensional indecomposable string module M with K-basis $x, xa = yb, y, yc$.

Since M is indecomposable its endomorphism ring S is a local ring, with $S/J(S) \cong K$. In the radical of S we have the radical elements a, b, c of R (which, note, is commutative) and we also have the element - let us call it d - which sends x to yc and sends y to 0. A one-line computation shows that the image of x under any element of S must be contained in $xK \oplus xaK \oplus ycK$ and similarly for y. So we have found a K-basis for S.

So we have $S \cong R[d : d^2 = 0 = ad = bd = cd]$. Since this ring is commutative and contains R we conclude that it is also the biendomorphism ring of N. Consider

the canonical embedding of R into S followed by composition with either of the morphisms $S \to S$ the first being the identity and the second the map which sends d to 0 and fixes R. Since the compositions are equal we conclude that $R \to S = R_N$ is not an epimorphism.

It is reasonable to regard rings of definable scalars as localisations in some sense. For instance, the proof of [Zie84, 5.4] applies to show that, although the ring of definable scalars at a point (i.e. of an idecomposable pure-injective) need not be local, at least its centre will be a local ring.

3 Rings of Type-Definable Scalars and Biendomorphism rings

Throughout this section we suppose that N_R is a right pure-injective module and let $S = End(N_R)$. Our pp-types refer only to sets of pp-formulas, not, as they are sometimes defined, to sets of pp-formulas together with the negations of some other pp-formulas.

Definition 3.1 *We say that an element $c \in N^I$ for some set I is a* **generic element** *for N if for each $d \in N$ we have $pp^{N^I}(c) \subseteq pp^N(d)$.*

So the pp-type of a generic element $c \in N^I$ is equivalent to $v = v$ modulo $Th(N)$. The most important property for us is that if N^I contains a generic element then N^I is a cyclic module over its endomorphism ring with each generic a generator. This holds because N^I is pure-injective and the pp-type of any element of this module is an intersection of pp-types realised in N and hence contains the pp-type of the generic c. We note that the set of pp-formulas $\{\phi : Th(N) \models \phi(v) \leftrightarrow v = v\}$ is the pp-type of any generic of N. Finding a generic for a given pure-injective is easy.

Proposition 3.2 *Suppose $\{a_i : i \in I\}$ is a generating set for N considered as a left S-module. Then the element $(a_i)_{i\in I} \in N^I$ is a generic for N.*

Proof Let $J \subseteq I$ be finite, $d \in N$ and let $s_j \in S$ for each $j \in J$ so that $d = \sum_{j\in J} s_j a_j$. Then we can define a map $f : N^I \to N$ with $f : (a_i)_{i\in I} \mapsto d$ as follows. Let $s \in End(N_R^I)$ have the action of s_j on the $j-$component of N^I for $j \in J$ and the action of the identity map on components not indexed by J. Next define $\pi_J : N^I \to N^{|J|}$, which projects N^I to the components indexed by J. Finally let σ be the summation map $N^{|J|} \to N$. We then define f to be $\sigma\pi_J s$. □

We now define the ring R_N^∞ of **type-definable scalars** for the module N. We take as the elements of R_N^∞ those pp-types $p(x, y)$ such that for each $a \in N$ we have $N \models p(a, b)$ for some unique $b \in N$, factored by the equivalence relation $\sim_N$ where $p_1 \sim_N p_2$ iff $N \models p_1(a, b)$ holds precisely when $N \models p_2(a, b)$ does. We frequently confuse pp-types with the equivalence classes that they lie in. We define addition and multiplication by using generics for N as follows. Let $(a_i)_{i\in I} \in N^I$ be a generic and let $p, q \in R_N^\infty$. Suppose that the elements d_i and e_i are defined

so that $N \models p(a_i, d_i)$ and $N \models q(a_i, e_i)$ which we denote respectively as $a_i p = d_i$ and $a_i q = e_i$. Then we define $p + q$ to be $pp^{N^I}((a_i)_{i\in I}, (d_i + e_i)_{i\in I})$. Now suppose that $f : N^I \to N$ has $f : (a_i)_{i\in I} \mapsto 0$. Then $f : (d_i)_{i\in I} \mapsto 0$ since $0p = 0$ and likewise $f : (e_i)_{i\in I} \mapsto 0$. Thus $f : (d_i + e_i)_{i\in I} \mapsto 0$ and so if $N \models (p+q)(0, d)$ then $d = 0$. Hence $p + q \in R_N^\infty$. Multiplication is defined similarly. We define pq to be $pp^{N^I}((a_i)_{i\in I}, (b_i)_{i\in I})$ where $d_i q = b_i$. Now clearly we have $N^I \models p((a_i)_{i\in I}, (d_i)_{i\in I})$ and $N^I \models q((d_i)_{i\in I}, (b_i)_{i\in I})$. If $N \models pq(a', b')$ then we have some map $f : N^I \to N$ with $f : (a_i)_{i\in I} \mapsto a', (b_i)_{i\in I} \mapsto b'$. Suppose d' is the image of $(d_i)_{i\in I}$ under f. Thus $a'p = d'$ and $d'q = b'$ and since p and q are in R_N^∞, $a'pq = b'$ is uniquely defined and so $pq \in R_N^\infty$.

We note that since any two generics for a given pure-injective have the same pp-type and live in pure-injective modules there are homomorphisms between these modules which map these elements to each other. This shows that our definitions of addition and multiplication are independent of the choice of generic.

We next give some simple properties of the ring R_N^∞.

Proposition 3.3 *Let R_N be the ring of definable scalars and I be any set. Then*

$$R_N \subseteq R_N^\infty = R_{N^I}^\infty \subseteq Biend(N_R^I) \subseteq Biend(N_R^{(I)}) = Biend(N_R)$$

where the identifications and embeddings are the obvious ones.

Proof We prove the inclusions from left to right. First, $R_N \subseteq R_N^\infty$ since any pp-formula ρ corresponding to a definable scalar clearly defines an element of R_N^∞ (given by the pp-type generated by ρ, the set of all pp-formulas lying above ρ in the pp-lattice of $Th(N)$).

For the first equality, suppose that $p \in R_N^\infty$. Then clearly $p \in R_{N^I}^\infty$ by letting p act componentwise on N^I. We want to show that these are the only possible type-definable scalars on N^I, i.e. we will show that if $p \in R_{N^I}^\infty$ then p acts componentwise on N^I and that the action of p on each component is the same. If $N^I \models p((a_i)_i, (b_i)_{i\in I})$ then clearly $N \models p(a_i, b_i)$ for each $i \in I$. Also if $N \models p(c, d)$ and $N \models p(c, d')$ then $d = d'$ for otherwise we would have $N^I \models p(0, (d - d')_i)$ with $d - d'$ non-zero, contradicting $p \in R_{N^I}^\infty$. This gives the first equality.

The second inclusion is just a type-definable scalar version of Lemma 2.4. We have for each $a \in N^I$ and $p \in R_{N^I}^\infty$ that $ap = b$ iff $N^I \models p(a, b)$ and so for any $s \in S = End(N_R^I)$ we have $N^I \models p(sa, sb)$ and so $(sa)p = sb = s(ap)$. The final equality and inclusion are text-book results. □

If N^I contains a generic then the second inequality in Proposition 3.3 becomes an equality.

Proposition 3.4 *Suppose that N^I contains a generic for N. Then*

$$R_N^\infty = Biend(N^I).$$

Proof Let $k \in Biend(N^I)$, $c \in N^I$ be a generic for N and $p = pp^{N^I}(c, ck)$. Then if $N^I \models p(a, b)$ we have a map $f \in End(N^I)$ with $f : c, ck \mapsto a, b$ and

$b = f(ck) = (fc)k = ak$. Since c is a generic, p defines a total function. So $p \in R_{N^I}^{\infty} = R_N^{\infty}$ and this association of biendomorphisms with elements of $R_{N^I}^{\infty}$ is clearly the inverse of the embedding $R_{N^I}^{\infty} \subseteq Biend(N^I)$ defined earlier. □

Corollary 3.5 *For any pure-injective module N_R there exists a set I with $R_N^{\infty} = Biend(N^I)$.*

Proposition 3.6 *If N_R is finitely generated over its endomorphism ring then $R_N^{\infty} = Biend(N_R)$. If $T = Th(N_R)$ is totally transcendental then $R_N = R_N^{\infty}$. Hence if both these conditions hold then $R_N = Biend(N_R)$.*

Proof If $\{a_i : i \leq n\}$ is a generating set for ${}_SN$ then $(a_i)_{i \leq n} \in N^{n+1}$ is a generic for N and $Biend(N^{n+1}) = Biend(N_R)$. The second assertion follows from the fact that every pp-type consistent with a tt theory T is finitely generated modulo T so that $R_N = R_N^{\infty}$ in this case. □

In fact when both conditions of Proposition 3.6 hold one may write down explicitly the formula which defines a biendomorphism. For, if $\bar{c} \in N^k$ is a generating (so generic) sequence for the tt module N_R over its endomorphism ring, if g is a biendomorphism of N and if $\theta(\bar{z}, \bar{w})$ is a pp-formula which generates the type of $(\bar{c}, \bar{c}g)$ modulo the theory of N then one may check that the formula $\rho(u, v)$ which is

$$\exists x_1, ..., x_k, y_1, ..., y_k(u = \Sigma x_i \wedge v = \Sigma y_i \wedge ...$$

$$\bigwedge_i \exists z_{i1}, ..., z_{ik}, w_{i1}, ..., w_{ik}\theta(z_{i1}, ..., x_i, ..., z_{ik}, w_{i1}, ..., y_i, ..., w_{ik}))$$

defines the action of g on N.

By essentially the same argument one may show that if M is a finitely presented R-module which is finitely generated over its endomorphism ring then the ring of definable scalars for M again coincides with its biendomorphism ring. All that is needed is that every pp-type realised in M is finitely generated modulo the theory of M together with the injectivity property that if $pp^M(\bar{c}) \subseteq pp^M(\bar{a})$ then there is an endomorphism of M taking $\bar{c}$ to $\bar{a}$ (for the fact that finitely presented modules have these properties see [Pre88b]).

We close this section by noting that the rings of definable scalars of two pure-injectives may be equal, while their rings of type-definable scalars can be at the opposite extremes allowed by Proposition 3.3. [Her93, Corollary 6.3] tells us that for any ring R and complete theory of left R-modules T we have that $R_T = R_{DT}^{op}$ where DT is the dual theory of T (the definition of the theory of right R-modules DT can be found in [Her93]). Let $R = \mathbb{Z}$ and again consider the Ziegler-rank 1 modules $\mathbb{Z}_{p^\infty}$ and $\overline{\mathbb{Z}_{(p)}}$. We know, from Theorem 2.5, that $R_{\overline{\mathbb{Z}_{(p)}}} = \mathbb{Z}_{(p)}$ and since $D(Th(\overline{\mathbb{Z}_{(p)}})) = Th(\mathbb{Z}_{p^\infty})$ we have $R_{\mathbb{Z}_{p^\infty}} = \mathbb{Z}_{(p)}$ also. By Proposition 3.6 since $\mathbb{Z}_{p^\infty}$ is tt we have $R_{\mathbb{Z}_{p^\infty}}^{\infty} = \mathbb{Z}_{(p)}$. However, for $\overline{\mathbb{Z}_{(p)}}$ there are type-definable scalars that are not definable by a pp-formula and in fact the type-definable scalars make up the whole biendomorphism ring $\overline{\mathbb{Z}_{(p)}}$. To see this we represent an element a of the ring $\overline{\mathbb{Z}_{(p)}}$ as a sum of the form $\Sigma_{n \in \mathbb{N}} a_n p^n$ where the a_n are representatives of equivalence classes of $\mathbb{Z}$ modulo the ideal (p). Every element of $Biend(\overline{\mathbb{Z}_{(p)}})$ is given

by multiplication by an element of $\overline{\mathbb{Z}_{(p)}}$ and the relation $v = wa$ where $a \in \overline{\mathbb{Z}_{(p)}}$ is defined by the $\mathbb{Z}$-pp-type

$$q_a = \{p|v - wa_0,\ \ p^2|v - w(a_0 + a_1p),\ \ p^3|v - w(a_0 + a_1p + a_2p^2), ...\}$$

and so $R^{\infty}_{\overline{\mathbb{Z}_{(p)}}} = Biend(\overline{\mathbb{Z}_{(p)}}) = \overline{\mathbb{Z}_{(p)}}$.

In Section 2 it was shown that the ring of definable scalars could be defined on closed sets of the Ziegler Spectrum. In a similar way we can define the ring of type-definable scalars for any closed set of the "Types over Formulas" topology defined in [Bur94]. This topology has the same points as the Ziegler Spectrum but in this case a basis of open sets is defined in terms of types modulo sorts (defined at the end of Section 1) rather than by pp-pairs. The closed sets of this topology turn out to be precisely those subsets of Zg_R which are closed under indecomposable direct summands of direct products.

4 Scalars in other sorts and Endomorphism Rings of Localised Functors

At least from the model-theoretic point of view it would be rather limiting to consider definable scalars only in the home sort. So here we generalise the previous discussion to arbitrary pp-sorts and we also show that the rings so obtained are exactly the endomorphism rings of localisations of finitely presented functors.

First we say what we mean by pp-sorts. Let M be any module and let $\psi \leq \phi$ be pp-formulas in n (say) free variables. Then $M^{\phi/\psi} = \phi(M)/\psi(M)$ is a group - a slice of M^n. By pp-elimination of quantifiers, all the structure on M is definable in terms of pp-formulas and every such pp-formula $\theta(\bar{v})$ induces a relation on $M^{\phi/\psi}$: given θ introduce a relation symbol R_θ of sort ϕ/ψ such that $M^{\phi/\psi} \models R_\theta(\bar{b})$ iff there is a tuple $\bar{a}$ in M with $\bar{b} = \bar{a}_\psi$ and $M \models \theta(\bar{a})$ (the notation $\bar{a}_\psi$ is shorthand for the coset $\bar{a} + \psi(M)$ of $\phi(M)/\psi(M)$). We will generally understand $M^{\phi/\psi}$ to mean the group $\phi(M)/\psi(M)$ with this **full induced structure** (see [KuPr92] for more detail). Within this structure there will be some binary relations which are total and functional - these form a ring under addition and composition and we refer to this as the **ring of definable scalars of M in sort ϕ/ψ** writing $R_M^{\phi/\psi}$. This ring depends only on the support C of M and so we may also write $R_C^{\phi/\psi}$.

There are two important points to be made about this in comparison with the ring R_M of definable scalars in the home sort. The first is that, since pp-formulas define subgroups and not submodules, in general the ring R does not act on $M^{\phi/\psi}$ (although the centre of R does - therefore there is a ring morphism from it to $R_M^{\phi/\psi}$). The second is that the $R_M^{\phi/\psi}$-module structure on $\phi(M)/\psi(M)$ may not be rich enough to define the full induced structure (see the example below): so there may well be endomorphisms of the $R_M^{\phi/\psi}$-module $\phi(M)/\psi(M)$ which do not preserve the full induced structure. In particular, **when we write $End(M^{\phi/\psi})$ we will mean** the abelian group endomorphisms of $M^{\phi/\psi}$ which preserve the full induced structure (as opposed to the possibly larger ring of $R_M^{\phi/\psi}$-module endomorphisms).

Example 4.1 *Let R be the path algebra over a field K of the quiver A_2 with the arrow pointing from vertex 1 to vertex 2. Let M be the direct sum of the indecomposable module of dimension type $(1,1)$ and the module of dimension type $(0,1)$. Let e_2 be the idempotent corresponding to the second vertex, let $\phi(v)$ be the formula $ve_2 = v$ and let $\psi(v)$ be the formula $v = 0$. Then $M^{\phi/\psi}$ is isomorphic to $K \oplus K$ as a K-vector space. The full induced structure on $\phi(M)$ includes a predicate which defines $\phi((1,1))$ - since this is the image of M under multiplication by an element of R. But it is easy to check that the ring $R_M^{\phi/\psi}$ is just the field K - so the full induced structure is strictly richer than the structure as a module over the ring of definable scalars in that sort.*

Similarly we have the concept of a ring of type-definable scalars in an arbitrary sort. We say that an element $c_\psi \in N_\psi^I$ is a **generic** for $N^{\phi/\psi}$ iff $c \in \phi(N)$ and $pp^{N^I}(c_\psi) \subseteq pp^N(d_\psi)$ for every $d_\psi \in N^{\phi/\psi}$ (equivalently $pp^{N^I}(c_\psi) \leftrightarrow \phi_\psi$ modulo $Th(N)$). We can think of $pp^N(d_\psi)$ as the set of those θ such that $N \models (\theta+\psi)(d)$ or equivalently as the set of formulas $\theta'(x_\psi)$ in the free variable x_ψ in the sort ϕ/ψ such that $N \models \theta'(d_\psi)$. We take as the elements of $(R_N^\infty)^{\phi/\psi}$ those pp-types $p(x_\psi, y_\psi)$ such that for each $a_\psi \in N^{\phi/\psi}$ we have $N \models p(a_\psi, b_\psi)$ for some unique $b_\psi \in \phi(N_\psi)$, factored by the equivalence relation $\sim_N$ where $p_1 \sim_N p_2$ iff $N \models p_1(a_\psi, b_\psi)$ holds precisely when $N \models p_2(a_\psi, b_\psi)$ does for all $a_\psi, b_\psi \in \phi(N_\psi)$. Again we will confuse pp-types with the equivalence classes that they lie in. Addition and multiplication in this ring are defined using generics (only this time using generics in the specified sort) in precisely the same way as in the definition of R_N^∞.

The following result, which is folklore, relates the endomorphism ring of a pure-injective module N to that of $N^{\phi/\psi}$.

Proposition 4.2 *Suppose that N is a pure-injective module and let ϕ/ψ be a sort. Then the natural morphism $EndN \to End(N^{\phi/\psi})$ given by restriction is a surjection.*

Proof Since every endomorphism preserves pp-formulas, endomorphisms of N restrict to endomorphisms of $N^{\phi/\psi}$. To show that this restriction is surjective, enumerate $N^{\phi/\psi}$ as $\bar{a}_\psi$ for some (perhaps infinite) tuple $\bar{a}$ in N. Let $g \in End(N^{\phi/\psi})$ and choose an inverse image, $\bar{b}$, of $g\bar{a}_\psi$ with maximal pp-type (since N is pure-injective, this pp-type is realised in N). We claim that $pp(\bar{a}) \subseteq pp(\bar{b})$.

Suppose then, that we have $\theta(\bar{a})$, hence $R_\theta(\bar{a}_\psi)$ and so $R_\theta(g\bar{a}_\psi)$. Then there is $\bar{c}$ with $\bar{c}_\psi = \bar{b}_\psi$ and $\theta(\bar{c})$. From $\psi(c_i - b_i)$ for each i we have, in the notation of [KuPr92, p.708] $\theta^\psi(\bar{b})$ and hence, by maximality of $pp(\bar{b})$ and [KuPr92, p.716] we deduce $\theta(\bar{b})$, as claimed. Therefore there is an endomorphism of N taking $\bar{a}$ to $\bar{b}$: the restriction of this endomorphism to $N^{\phi/\psi}$ is g. □

One may, following Herzog [Her93], define the category $Mod - R^{eq}$ of definable scalars. The objects are the pp-sorts, ϕ/ψ, and the morphisms are the pp-definable morphisms between sorts. So if M is such that $supp(M) = Zg_R$ then the ring of definable scalars of M in sort ϕ/ψ is just the endomorphism ring of the sort ϕ/ψ in this category. In fact, the same comment applies to any module if we replace

the category of definable scalars by a suitable localisation of it. To explain this we discuss very briefly (for a fuller account, see [Bur94] or [Pre93]) the realisation of $Mod - R^{eq}$ as a category of functors.

Denote by $(mod - R, Ab)$ the category of functors (always we mean additive functors) from the category $mod - R$ of finitely presented right R-modules to the category Ab of abelian groups. Then the category of definable scalars is equivalent [Bur94] to the category $(mod - R, Ab)^{fp}$ of finitely presented such functors, with the sort ϕ/ψ corresponding to the functor $F_{\phi/\psi}$ (which takes a module M to $\phi(M)/\psi(M)$ and has the obvious action on morphisms). In particular, the morphisms between finitely presented functors are all given by pp-formulas. So we may regard the global ring $R^{\phi/\psi}$ of definable scalars in sort ϕ/ψ as the opposite of the endomorphism ring of the functor $F_{\phi/\psi} \in (mod - R, Ab)^{fp}$. There is a duality $(mod - R, Ab)^{fp} \cong ((R - mod, Ab)^{fp})^{op}$ between the categories of finitely presented functors on right and left finitely presented modules, which takes the typical finitely presented functor $F_{\phi/\psi}$ to its dual $F_{D\psi/D\phi}$. Here $D\phi$ is the dual of the pp-formula ϕ ([Pre88a], [Her93]). For various reasons, such as the fact that $M \mapsto M \otimes -$ gives a nice embedding of $Mod - R$ into $(R - mod, Ab)$, it is often convenient to take $(R - mod, Ab)^{fp}$ as the functorial version of "eq+". Then the global ring $R^{\phi/\psi}$ of definable scalars in sort ϕ/ψ will be the endomorphism ring of the functor $F_{D\psi/D\phi}$.

Thus, for any module M, M^{eq+} becomes a right "module" over $(R-mod, Ab)^{fp}$: namely the functor from $((R - mod, Ab)^{fp})^{op}$ to Ab which takes $F_{D\psi/D\phi}$ to

$$(F_{D\psi/D\phi}, M \otimes -) \cong \phi(M)/\psi(M)$$

(see, for example, [Pre93]). It is easily checked that the natural structure that $\phi(M)/\psi(M)$ carries in this isomorphism, as a right $R^{\phi/\psi}$-module, coincides with the $R^{\phi/\psi}$-module structure induced by the isomorphism (from the fact that $R^{\phi/\psi} = End(F_{D\psi/D\phi})$).

We will see below that the "local" rings $R_M^{\phi/\psi}$ of definable scalars may be obtained as endomorphism rings of certain functors. We will also realise them as biendomorphism rings.

Now we give a very brief account of the role of localisation (again see [Her94], [Kra94], [Pre95] for more detail).

There is a bijective correspondence between the closed subsets of the Ziegler spectrum and the hereditary torsion theories on $(R - mod, Ab)$ which are of finite type (= the torsion functor commutes with directed limits). If $supp(M) = C$ then the localisation of $(R - mod, Ab)^{fp}$ at the torsion theory τ corresponding to C is equivalent to the category of definable scalars for the theory of M.

Similarly there is a bijection between the closed subsets in the "Types over formulas" topology (see [Bur94]) - a refinement of the Ziegler topology - and those hereditary torsion theories on $(R - mod, Ab)$ which are cogenerated by sets of indecomposable injectives.

The fact that any torsion theory on $(R - mod, Ab)$ of finite type is cogenerated by a set of indecomposable injective functors is just [Zie84, 6.9].

An R-module N is pure-injective iff the functor $N \otimes -$ in $(R - mod, Ab)$ is injective. The functor $N \otimes -$ may not cogenerate a torsion theory of finite type

(see example below): if it does then this will be the torsion theory corresponding to $supp(N)$ and then we say that N is an **elementary cogenerator**. One has the following ([Pre88b], [Pre95]).

The pure-injective module N is an elementary cogenerator iff every model of $Th(N)$ purely embeds in some power of N iff N realises every 1- type which is neg-isolated in $Th(N)$.

Let N be a pure-injective R-module which is weakly saturated, $|R|^+$-saturated or totally transcendental (=Σ-pure-injective). Then N is an elementary cogenerator. In particular, every module is elementarily equivalent to an elementary cogenerator.

We therefore obtain the following corollary of Corollary 3.5.

Theorem 4.3 *Suppose that N is an elementary cogenerator. Then there is a power N^I of N such that the ring of definable scalars of N coincides with the biendomorphism ring of N^I.*

Example 4.4 *The torsion theory cogenerated by an arbitrary tensor functor may be strictly smaller than the corresponding torsion theory of finite type. Let $R = \mathbb{Z}$ and consider the torsion theory cogenerated by $\overline{\mathbb{Z}_{(p)}} \otimes -$. The injective $\mathbb{Q} \otimes -$ is not torsionfree for this torsion theory - otherwise it would be a direct summand of a product of copies of $\overline{\mathbb{Z}_{(p)}} \otimes -$ and hence $\mathbb{Q}$ would be a direct summand of a product of copies of $\overline{\mathbb{Z}_{(p)}}$ which is not the case (use that $\mathbb{Q}$ is divisible). But $\mathbb{Q} \otimes -$ is torsionfree for the smallest finite type torsion theory for which $\overline{\mathbb{Z}_{(p)}} \otimes -$ is torsion-free since $\mathbb{Q}$ is in the Ziegler closure of $\overline{\mathbb{Z}_{(p)}}$.*

The ring $(R_N^\infty)^{\phi/\psi}$ exists in the functor category $(R-mod, Ab)$ as the endomorphism ring of an object that we will specify using torsion theories. Before showing this we look at another way of viewing generic elements c for a pure-injective N. Let $\mathcal{T}$ be the torsion class of the torsion theory cogenerated by the functor $N \otimes_R -$ so that $\mathcal{T}$ is the localising subcategory of $(R-mod, Ab)$ generated by $\{X : (X, N \otimes_R -) = 0\}$. There is a corresponding torsion functor τ which maps each object F to its largest subobject in $\mathcal{T}$, τF. We let Q_N be the quotient functor $Q_N : (R-mod, Ab) \to (R-mod, Ab)/\mathcal{T}$ and denote (given ϕ) by $c_\psi \otimes -$ the natural transformation $F_{D\psi/D\phi} \to N \otimes_R -$ which takes the element $a_{D\phi}$ where $a \in D\psi(M)$ to $c \otimes a \in N \otimes_R M$.

Lemma 4.5 *Let $c \in N^I$ for some set I. Then c_ψ is a generic for $N^{\phi/\psi}$ iff $Q_N(c_\psi \otimes -)$ is a monomorphism in $(R-mod, Ab)/\mathcal{T}$. In this case the kernel of $(c_\psi \otimes -)$ is $\tau F_{D\psi/D\phi}$.*

Proof ($\Rightarrow$) Let K be the kernel of the map $c_\psi \otimes - : F_{D\psi/D\phi} \to N^I \otimes_R -$ which may be written [Pre88b, 12.1] as $\Sigma F_{D\theta/D\phi}$ for some pp-formulas $D\theta$. Now $F_{D\theta/D\phi} \leq K$ iff $c_\psi \otimes a_{D\phi} = 0$ in $N^I \otimes_R M$ for every $a \in D\theta(M)$ with $M \in R-mod$. But since c_ψ is a generic element for $N^{\phi/\psi}$ we have $d_\psi \otimes a_{D\phi} = 0$ in $N \otimes_R M$ so that $d \otimes a = 0$ for all $d \in \phi(M)$ and so $(F_{D\theta/D\phi}, N \otimes_R -) = 0$, $F_{D\theta/D\phi} \in \mathcal{T}$ and so $K \in \mathcal{T}$. Now $K = \tau F_{D\psi/D\phi}$ since $F_{D\psi/D\phi}/K$ is trivially $\mathcal{T}$-torsion-free and

in particular Q_N preserves the embedding of $F_{D\psi/D\phi}/K$ into $N^I \otimes_R -$ [Pop73, Chapter 4 Section 5].

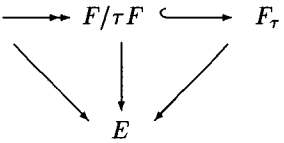
($\Leftarrow$) If we suppose that $Q_N(c_\psi \otimes -)$ is a monomorphism then the kernel K of $c_\psi \otimes -$ lies in $\mathcal{T}$. Hence if $N^I \models \theta(c_\psi)$ then $F_{D\theta/D\phi} \leq K$ and $(F_{D\theta/D\phi}, N \otimes_R -) = 0$ and so $N \models \theta(d_\psi)$ for all $d \in N$. Thus c_ψ is a generic element for $N^{\phi/\psi}$. □

Proposition 4.6 *Let $F = F_{D\psi/D\phi}$. Then $(R^\infty_N)^{\phi/\psi} = End(F_\tau)$ where F_τ, the localisation of F at τ is the largest subobject of the injective hull of $F/\tau F$ such that $Q_N(F_\tau) = Q_N(F/\tau F)$.*

Proof Let N^I_ψ contain a generic for $N^{\phi/\psi}$. Since $N^I \otimes_R -$ cogenerates the same torsion theory as $N \otimes_R -$ and since $(R^\infty_N)^{\phi/\psi} = (R^\infty_{N^I})^{\phi/\psi}$ (by an easy modification of the proof of Proposition 3.3) we may assume that N_ψ contains a generic c_ψ for itself. Now the functor $E = N \otimes_R -$ is injective and torsion-free so we have

$$(F, E) = (F/\tau F, E) = (F_\tau, E) \tag{1}$$

where the identifications are the obvious ones given by the diagram



So we can think of $\phi/\psi(N)$ as having a left-S, right-$End(F_\tau)$ bimodule structure where $S = End(N_R)$. We also notice that by Lemma 4.5, the map $c_\psi \otimes - : F \to E$ induces an embedding of F_τ in E.

We define a map $\alpha : End(F_\tau) \to (R^\infty_N)^{\phi/\psi}$ by taking $h \in End(F_\tau)$ to the type q where $q = pp^N(c_\psi, c_\psi h)$ and $c_\psi h$ is the element of $N^{\phi/\psi}$ such that $c_\psi h \otimes - : F \to E$ defines the composition

$$F \longrightarrow\!\!\!\!\!\rightarrow F/\tau F \hookrightarrow F_\tau \xrightarrow{\ h\ } F_\tau \hookrightarrow E\ .$$

To see that $q \in (R^\infty_N)^{\phi/\psi}$, suppose that $N_\psi \models q(c_\psi, b_\psi)$ so that $N_\psi \models q(0_\psi, c_\psi h - b_\psi)$ and let $g \in S$ with $g : c_\psi, c_\psi h \mapsto 0_\psi, c_\psi h - b_\psi$. Then $g.c_\psi h = gc_{\ psi}.h = 0_\psi$ and so $c_\psi h = b_\psi$.

Finally we need to show that α defines an isomorphism of rings. To show that it defines a homomorphism let $\alpha h = p$, $\alpha h' = p'$ so that $c_\psi p = c_\psi h$ and $c_\psi p' = c_\psi h'$. Suppose that $g \in S$ is such that $g : c_\psi \mapsto c_\psi h$. Then $c_\psi pp' = c_\psi h.p' = gc_\psi.p'$ $= g.c_\psi p' = g.c_\psi h' = gc_\psi.h' = c_\psi hh'$. The other properties are easy to prove. That α is a monomorphism is also clear so we need only prove that it is onto.

Now since E is an injective cogenerator of τ we have $F_\tau = \cap\{Ker(f) : f : E(F/\tau F) \to E$ with $f(F/\tau F) = 0\}$. We let $p \in (R^\infty_N)^{\phi/\psi}$ and look for a map h in $End(F_\tau)$ with the same action as p. Suppose that $g \in End(E) = End(N_R)$ maps the image of $c_\psi \otimes - : F/\tau F \hookrightarrow E$ to 0. This implies that the map $g(c_\psi \otimes -) : F \to E$ is zero by identification (1) and so $gc_\psi = 0_\psi$ (thinking of $g \in S$). So we have $g.c_\psi p = gc_\psi.p = 0_\psi$ and thus the image of $F/\tau F$ under the transformation $c_\psi p \otimes -$ is a subfunctor of F_τ by the description of F_τ as an intersection of kernels given above.

Also since E is torsion-free, the image of an extension h of $c_\psi p \otimes -$ to F_τ in the identification (1) is a subfunctor of F_τ. So we have that $c_\psi p \otimes -$ is the composition

$$F/\tau F \hookrightarrow F_\tau \xrightarrow{h} F_\tau \hookrightarrow E .$$

Therefore $c_\psi h = c_\psi p$ and $\alpha(h) = p$. This completes the proof. □

We recall (see [Ste75]) that the localisation of a ring at the torsion theory cogenerated by an injective E can be obtained as the biendomorphism ring of some power of E (which also, of course, cogenerates the same torsion theory). Bearing in mind that a module is pure-injective iff the corresponding tensor functor in $(R - mod, Ab)$ is injective we arrive at a corresponding description (noticed independently by Henning Krause) of the ring of (type-) definable scalars.

Theorem 4.7 *Let N be pure-injective and let ϕ/ψ be any sort. Then there is a power I such that the ring of type-definable scalars of N in sort ϕ/ψ is equal to the endomorphism ring of $(N^I)^{\phi/\psi}$ regarded as a left module over the endomorphism ring of N.*

To outline the proof of this, let $\mathcal{T}$ be the torsion theory cogenerated by $N \otimes -$ in $(R - mod, Ab)$. Replacing N by a power N^I if necessary, we can and will assume that N contains a generic element of sort ϕ/ψ so that $\mathcal{T}$ remains unchanged. We have that the natural morphism $EndN \to End(N^{\phi/\psi})$ given by restriction is a surjection by Proposition 4.2. One may check that the proofs of Corollary 3.5 and Theorem 4.3 work just as well in this more general context.

Corollary 4.8 *Suppose that N is an elementary cogenerator and ϕ/ψ is a sort. Then there is a power I such that the ring of definable scalars of N in sort ϕ/ψ is equal to the endomorphism ring of $(N^I)^{\phi/\psi}$ regarded as a module over $EndN$.*

Corollary 4.9 *Suppose that N is an elementary cogenerator and ϕ/ψ is a sort such that the $R^{\phi/\psi}$-structure on $N^{\phi/\psi}$ induces the full structure on $N^{\phi/\psi}$. Then there is a power I such that the ring of definable scalars in the sort ϕ/ψ for N is the biendomorphism ring of $(N^I)^{\phi/\psi}$ where this is regarded as a module over $R^{\phi/\psi}$.*

Wilfrid Hodges has pointed out that Corollary 4.9 may also be obtained using Svenonius' Theorem.

References

[Bur94] K. Burke. *Some Model-Theoretic Properties of Functor Categories for Modules.* Doctoral Thesis, University of Manchester, 1994.

[ChKe73] C.C. Chang, H.J. Keisler. *Model Theory*, North Holland, Amsterdam, 1973.

[Her93] I. Herzog. *Elementary duality of modules.* Trans. Amer. Math. Soc, **340**:37-69, 1993.

[Her94] I. Herzog. *The Ziegler Spectrum of a locally coherent Grothendieck category.* Preprint, Albert-Ludwigs-Universität, Freiburg, 1994.

[Kra94] H. Krause. *The spectrum of a locally coherent category.* Preprint, Universität Bielefeld, 1994.

[KuPr92] T.G. Kucera, M. Prest. *Imaginary modules.* J. of Symbolic Logic, **57**(2):698-723, 1992.

[Pop73] N. Popescu. *Abelian Categories with Applications to Rings and Modules.* Academic Press, London and New York, 1973.

[Pre88a] M. Prest. *Duality and pure-semisimple rings.* J. of the London Math. Soc. **38**, 1988.

[Pre88b] M. Prest. *Model theory and modules.* London Mathematical Society Lecture Note Series no.**130**, Cambridge University Press, Cambridge, 1988.

[Pre93] M. Prest. *Remarks on elementary duality.* Ann. of Pure and Applied Logic, **62**:183-205, 1993.

[Pre95] M. Prest. *The (pre)sheaf of the ring of definable scalars.* Preprint, University of Manchester, 1995.

[Pre96] M. Prest. *Epimorphisms of rings, Interpretations of modules and strictly wild algebras* Comm. Algebra, **24(2)**:517-531, 1996.

[Sch85] A. Schofield. *Representations of Rings over Skew Fields.* LMS Lecture Notes in Mathematics Vol **92**, 1985.

[Ste75] B. Stenström. *Rings of Quotients.* Springer-Verlag, New York and Heidelberg, 1975.

[Zie84] M. Ziegler. *Model theory of modules.* Ann. Pure and Applied Logic, **26**:149-213, 1984.

Authors' address:
Department of Mathematics,
University of Manchester,
Manchester M13 9PL,
England.
e-mail: burke@ma.man.ac.uk and mprest@ma.man.ac.uk

Recent Results on Simple First Order Theories

Byunghan Kim

This paper summarizes the main results of [8], [9] and some facts from [12] with condensed proofs.

The study of simple theories began with Shelah's paper "Simple unstable theories" [12] where he introduced a class of first order theories, he called simple, having $D(p, \Delta, k)$ rank. The class includes all stable theories and some unstable theories. His intention was to ask whether we can build a theory of simple theories analogous to stability theory.

Remarkable progress in the study of simple theories has been made very recently after Hrushovski and others developed notions of independence in specific unstable structures such as pseudo-finite fields ([5], [6]), fields with an automorphism ([3]) and smoothly approximable structures ([4], [7]). The independence notion in each of these unstable structures behaves similarly to nonforking in stable structures. Hence we may ask naturally the following questions in connection with simple unstable theories.

(1) Are all unstable structures mentioned above simple ?

(2) If so, then "what is the relation between the independence notion and nonforking?"

(3) Does any simple theory have a similar notion of independence?

It turns out that the answer to (1) is positive and the independence notion in each unstable structure above is exactly nonforking. Furthermore nonforking supplies a notion of independence to all simple unstable theories as well as to stable theories, primarily by the following theorem.

THEOREM (B. KIM) [8] *If T is simple, then the following hold.*

(i) *Let p be a type, and let A be a set. Then p divides over A if and only if p forks over A.*

(ii) *Nonforking satisfies symmetry and transitivity.*

In this paper we shall describe the proofs of the preceding theorem and additional results with almost no prerequisites except compactness and Ramsey's Theorem. The approach here is slightly different from [8]. However, the aim of this paper is not to suggest new results or proofs but to guide the readers who want to understand simple theories quickly. Hence we state weaker (but simpler) versions of lemmas from [8], [9] and [12] if they suffice for yielding the main results in each section of this paper.

We use standard notation. T is a complete theory in a first order language L. p denotes a consistent type, perhaps partial, unless stated otherwise. We work

in a huge $\bar{\kappa}$-saturated model $\mathcal{C}$ as usual. Sets $A, B, C, \ldots$ are subsets of $\mathcal{C}$ and models $M, N, \ldots$ are elementary submodels of $\mathcal{C}$ whose cardinalities are strictly less than $\bar{\kappa}$. An *A-automorphism* is an automorphism of $\mathcal{C}$ fixing A pointwise. If $p(\bar{x}, \bar{a}_0)$ is a type over $A\bar{a}_0$ and $tp(\bar{a}_0/A) = tp(\bar{a}_1/A)$ then $p(\bar{x}, \bar{a}_1)$ is the image of $p(\bar{x}, \bar{a}_0)$ under an A-automorphism taking $\bar{a}_0$ to $\bar{a}_1$. Recall that a sequence $I = \langle \bar{c}_i | i \in \omega \rangle$ is an *A-indiscernible sequence* (*of* $p \in S(A)$) if for each $n \in \omega$, $tp(\bar{c}_{i_0}, \ldots, \bar{c}_{i_n}/A) = tp(\bar{c}_0, \ldots, \bar{c}_n/A)$ for each $i_0 < \ldots < i_n < \omega$ (and $\bar{c}_0 \models p$). For a type p over a set B, $p|A$ denotes a restriction of p to $A(\subseteq B)$. There will be no unstated assumptions for T. For example, Fact 1.3. is true for any theory.

We wish to thank Anand Pillay for permitting us to present his work from [9].

§1. Stability, dividing and forking

A first order theory T is *unstable* if there are a formula $\varphi(\bar{x}, \bar{y})$ of L and a sequence $\langle \bar{a}_i | i < \omega \rangle$ of tuples such that for all $i, j < \omega$, $\models \varphi(\bar{a}_i, \bar{a}_j)$ if and only if $i < j$. T is *stable* if it is not unstable.

Many well-known algebraic structures are stable, such as algebraically closed fields, differentially closed fields, vector spaces and modules and so on. On the other hand, real closed fields and pseudo-finite fields (see [6]) are unstable.

One of the reasons why model theorists have been interested in stability theory is that there is a notion of nonforking which yields a uniform concept of independence in any stable structure. (See Fact 1.4.) For example, nonforking characterizes linear independence in vector spaces and algebraic independence in fields.

DEFINITION 1.1. A type p *divides over a set* A *with respect to* $k \in \omega$, if there are a formula $\varphi(\bar{x}, \bar{c})$ and a sequence $\langle \bar{c}_i | i \in \omega \rangle$ such that

(i) $p \vdash \varphi(\bar{x}, \bar{c})$,

(ii) $tp(\bar{c}/A) = tp(\bar{c}_i/A)$ for all i,

(iii) $\{\varphi(\bar{x}, \bar{c}_i) | i \in \omega\}$ is k-inconsistent, i.e. any finite subset of size k is inconsistent.

p *divides over* A if p divides over A with respect to some k.

p *forks over* A if there are formulas $\varphi_0(\bar{x}, \bar{a}_0), \ldots, \varphi_n(\bar{x}, \bar{a}_n)$ such that

(i) $p \vdash \bigvee_{0 \leq i \leq n} \varphi_i(\bar{x}, \bar{a}_i)$,

(ii) $\varphi_i(\bar{x}, \bar{a}_i)$ divides over A for each i.

REMARK 1.2. From the definitions of dividing and forking, we easily obtain the following.

(i) $\varphi(\bar{x}, \bar{c})$ divides over A if and only if there is an A-indiscernible sequence I of $tp(\bar{c}/A)$ such that $\{\varphi(\bar{x}, \bar{c}') | \bar{c}' \in I\}$ is inconsistent. (Use Ramsey's Theorem.)

(ii) $\varphi(\bar{x}, \bar{c})$ divides over A w.r.t. k if and only if for all finite $\bar{a} \subseteq A$, $\varphi(\bar{x}, \bar{c})$ divides over $\bar{a}$ w.r.t. k. (Use compactness.)

(iii) If p divides over A, then p forks over A.

(iv) If p divides (forks) over A and $p \subseteq q$, then q divides (forks) over any subset of A.

Intuitively dividing is the right notion of dependence. Suppose that $tp(\bar{b}/A\bar{c})$ divides over A. Then we may interpret this as: the set X of realizations of $tp(\bar{b}/A\bar{c})$ breaks up that of $tp(\bar{b}/A)$ into infinitely many pieces X_i, each of X_i is an A-automorphic image of X. This means in some sense that $\bar{b}$ satisfies more relations with $A\bar{c}$ than it does with A.

Why, then is forking introduced as well as dividing? With forking, we have the following Extension axiom which is quite useful in developing arguments. Moreover forking turns out to be dividing for stable T.

FACT 1.3. *Let $A \subseteq B \subseteq C$. If p in $S(B)$ does not fork over A, then there is an extension q of p in $S(C)$ such that q does not fork over A.*

FACT 1.4. *If T is a stable theory, then nonforking has the following properties.*
(1) *p does not fork over A if and only if p does not divide over A.*
(2) (i) (*Symmetry*) *$tp(\bar{c}/A\bar{b})$ does not fork over A if and only if $tp(\bar{b}/A\bar{c})$ does not fork over A.*

(ii) (*Transitivity*) *Let $A \subseteq B \subseteq C$. Then $tp(\bar{a}/C)$ does not fork over A if and only if $tp(\bar{a}/B)$ does not fork over A and $tp(\bar{a}/C)$ does not fork over B.*

(iii) (*Extension*) *Fact* 1.3.

(iv) (*Local Character*) *For any complete type p over B, there is a subset A of B such that $|A| \leq |T|$ and p does not fork over A.*

(v) (*Finite Character*) *Let $A \subseteq B$. Then $tp(\bar{a}/B)$ does not fork over A if and only if for each finite tuple $\bar{b} \in B$, $tp(\bar{a}/A\bar{b})$ does not fork over A.*

(vi) (*Boundedness*) *Let $A \subseteq B$. For any complete type p over A, there are at most $2^{|T|}$ many extensions of p in $S(B)$ which do not fork over A. If A is a model, then there is a unique nonforking extension.*

Moreover if, in a theory T, an automorphism-invariant relation between complete types and sets satisfies all the preceding axioms, then the theory is stable and the relation has to be nonforking.

The properties listed in Fact 1.4. do not all hold in general, if T is not stable.

EXAMPLE 1.5. (1) Let T be the theory of the ternary relation $R(x, y, z)$ defined on the circle C as follows. $R(x, y, z)$ holds if and only if x, y, z are points on C (x, z are not diametrically opposed points), and y lies on the shorter arc from x to z. Now

$$\models x = x \leftrightarrow R(a_0, x, a_1) \vee R(a_1, x, a_2) \vee R(a_2, x, a_0)$$

for some a_i's in C such that $R(a_i, x, a_j)$ divides over $\emptyset$. Hence $x = x$ forks over $\emptyset$ but does not divide over $\emptyset$.

(2) Assume that $(M, <)$ is a dense linear ordering without endpoints. It can easily be checked that forking is dividing in this structure. However symmetry fails. Let $a < b < c$. Then $tp(b/ac)$ divides over $\emptyset$ witnessed by $a < x < c$ whereas $tp(ac/b)$ does not divide over $\emptyset$.

§2. Simple theories

DEFINITION 2.1. T is *simple* if nonforking satisfies the Local Character axiom, i.e. for any complete type p over B, there is a subset A of B such that $|A| \leq |T|$ and p does not fork over A.

The typical example of a simple unstable structure is the random graph. A variation of the random graph, the so called bipartite random graph, is also worth mentioning. A bipartite random graph consists of two disjoint infinite subsets, say U, V with a binary relation R between U and V. For any finite disjoint subsets X and Y of U, there is $z \in V$ such that xRz for $x \in X$ and $\neg yRz$ for $y \in Y$, and vice versa. We shall see more algebraic simple unstable structures at the end of this paper.

For any stable theory, nonforking has Local Character. Hence any stable theory is simple. Also the definition yields the next lemma.

LEMMA 2.2. *If T is simple and $p \in S(A)$, then p does not fork over A.*

DEFINITION 2.3. Let $A \subseteq B$ and $p \in S(B)$. By a *Morley sequence of p over A*, we mean a B-indiscernible sequence $I = \langle \bar{a}_i : i \in \omega \rangle$ of p such that for every $i \in \omega$, $tp(\bar{a}_i / B \cup \bigcup\{\bar{a}_j | j < i\})$ does not fork over A .

By a *Morley sequence of $p \in S(B)$*, we mean a Morley sequence of p over B.

PROPOSITION 2.4. *If $p \in S(B)$ does not fork over $A(\subseteq B)$, then there exists a Morley sequence of p over A. Thus for simple T, a Morley sequence of a complete type exists.*

Proof. For any cardinal λ, Extension (Fact 1.3.) guarantees the existence of a sequence $\langle \bar{a}_\alpha | \alpha < \lambda \rangle$ such that for every $\alpha < \lambda$, $tp(\bar{a}_\alpha / B \cup \bigcup\{\bar{a}_\beta | \beta < \alpha\})$ extends p and does not fork over A. We can extend the sequence as long as we want, but there are only boundedly many types over B. Hence using the Erdös-Rado Theorem repeatedly (plus some techniques (cf. [2, Theorem 7.2.2.])) we obtain the desired indiscernible sequence from a huge λ-sequence. □

Main theorems

The following lemma plays a crucial role in proving the main proposition 2.7.

LEMMA 2.5. *Let $I = \langle \bar{c}_i : i \in \omega \rangle$ be a Morley sequence of $tp(\bar{c}_0 / A)$ and $J = \langle \bar{c}^j : j \in \omega \rangle$ an A-indiscernible sequence of $tp(\bar{c}_0 / A)$. Then there is a sequence I', an A-automorphic image of I, such that $\langle \bar{c}^j \rangle^\frown I'$ is an A-indiscernible sequence for each $j < \omega$.*

Proof. Suppose that $\bar{b}_1, ..., \bar{b}_n$ have been chosen so that
(i) $tp(\bar{c}^j \bar{b}_1 ... \bar{b}_n / A) = tp(\bar{c}_0 \bar{c}_1 ... \bar{c}_n / A)$ for every $j < \omega$,
(ii) J is $A\bar{b}_1 ... \bar{b}_n$-indiscernible.

Let us find $\bar{b}_{n+1}$ so that (i), (ii) hold for $n+1$. First we shall show

$$\cup\{p(\bar{c}^j, \bar{b}_1, ..., \bar{b}_n, \bar{x}_{n+1}) | j < \omega\} = q$$

is consistent where $p(\bar{x}_0, \bar{x}_1, ..., \bar{x}_{n+1}) = tp(\bar{c}_0\bar{c}_1...\bar{c}_{n+1}/A)$. Let

$$\Gamma = \{\varphi(\bar{c}^j, \bar{b}_1, ..., \bar{b}_n, \bar{x}_{n+1}, \bar{a}) | j < \omega\}$$

for given $\varphi(\bar{x}_0, \bar{x}_1, ..., \bar{x}_{n+1}, \bar{a}) \in p(\bar{x}_0, ..., \bar{x}_{n+1})$ with $\bar{a} \subseteq A$. Since $tp(\bar{c}_{n+1}/A\bar{c}_0...\bar{c}_n)$ does not divide over A and $\langle \bar{c}^j\bar{b}_1...\bar{b}_n | j < \omega\rangle$ is A-indiscernible, Γ is consistent. Hence q is consistent. Using Ramsey's Theorem we can select a tuple $\bar{b}_{n+1}$ satisfying q such that J is $A\bar{b}_1...\bar{b}_n\bar{b}_{n+1}$-indiscernible. Now $\langle \bar{b}_n | 1 \leq n < \omega\rangle$ is a desired sequence I'. □

That I is a Morley sequence is essential in Lemma 2.5. For consider the following example.

EXAMPLE 2.6. Let T be the theory of an equivalence relation E with infinitely many infinite equivalence classes. Let p be a unique complete one type over $\emptyset$ in T and I an indiscernible sequence of p in the same equivalence class. If $J = \langle c_i | i \in \omega\rangle$ is an indiscernible sequence of p such that $\models \neg c_iEc_j$ for each $i \neq j \in \omega$, then there is no copy of I which is a common extension of all c_i's.

PROPOSITION 2.7. *If T is simple, then the following are equivalent.*

(i) *$\varphi(\bar{x}, \bar{c})$ divides over A.*

(ii) *$\varphi(\bar{x}, \bar{c})$ forks over A.*

(iii) *For any Morley sequence I of $tp(\bar{c}/A)$, $\{\varphi(\bar{x}, \bar{c}') | \bar{c}' \in I\}$ is inconsistent.*

(iv) *There exists a Morley sequence I of $tp(\bar{c}/A)$ such that $\{\varphi(\bar{x}, \bar{c}') | \bar{c}' \in I\}$ is inconsistent.*

Proof. (iii) → (iv) → (i) →(ii) Obvious.

(i)→(iii) Assume that a Morley sequence $I = \langle \bar{c}_n : n \in \omega\rangle$ of $tp(\bar{c}/A)$ is given. We claim that $\varphi(\bar{x}, \bar{c}_0)$ divides over $A \cup \bigcup\{\bar{c}_m | 0 < m\}$ with respect to some k. Now since $\varphi(\bar{x}, \bar{c}_0)$ divides over A w.r.t. some integer k, there is an A-indiscernible sequence $\langle \bar{c}^j : j < \omega\rangle$ with $\bar{c}^0 = \bar{c}_0$ such that $\{\varphi(\bar{x}, \bar{c}^j) | j < \omega\}$ is k-inconsistent. Now we need Lemma 2.5. by which there is an A-automorphic image I' of I such that $tp(\bar{c}^jI'/A) = tp(I/A)$ for all $j < \omega$. This shows, simply by the definition of dividing, that $\varphi(\bar{x}, \bar{c}_0)$ divides over AI' w.r.t. k, hence so over $A \cup \bigcup\{\bar{c}_m | 0 < m\}$ w.r.t. k.

Let us continue to prove (i)→(iii). Suppose that $p = \{\varphi(\bar{x}, \bar{c}_n) | n < \omega\}$ is consistent. Then the preceding claim and Remark 1.2.(ii), together with compactness, yield a sequence $\langle \bar{c}'_\alpha | \alpha < |T|^+\rangle$ (reversing the order of I and extending it.) such that $p' = \{\varphi(\bar{x}, \bar{c}'_\alpha) | \alpha < |T|^+\}$ is consistent and $\varphi(\bar{x}, \bar{c}'_\alpha)$ divides over $A \cup \bigcup\{\bar{c}'_\beta | \beta < \alpha\}$ for each $\alpha < |T|^+$. This contradicts the simplicity of T by an easy argument.

(ii)→(i) Since $\varphi(\bar{x}, \bar{c})$ forks over A, there are $\psi_i(\bar{x}, \bar{b}_i)$ $(0 \leq i \leq r)$ such that $\models \varphi(\bar{x}, \bar{c}) \to \bigvee_{i \leq r} \psi_i(\bar{x}, \bar{b}_i)$ and for each i, $\psi_i(\bar{x}, \bar{b}_i)$ divides over A w.r.t. some k_i. Let $\bar{d} = \bar{b}_0...\bar{b}_r$ and $I = \langle \bar{c}_n\bar{d}_n | n \in \omega\rangle$ be a Morley sequence of $tp(\bar{c}\bar{d}/A)$ with $\bar{c}_0\bar{d}_0 = \bar{c}\bar{d}$. We first claim that $\varphi(\bar{x}, \bar{c}_0)$ forks over $A \cup \bigcup\{\bar{c}_m | 0 < m\}$. For a given i

$(0 \leq i \leq r)$, by Ramsey's Theorem, we may assume that there is an A-indiscernible sequence $\langle \bar{c}^j \bar{d}^j | j < \omega \rangle$ where $\bar{c}^0 \bar{d}^0 = \bar{c}\bar{d}$, $\bar{d}^j = \bar{b}_0^j \ldots \bar{b}_r^j$ such that $\{\psi_i(\bar{x}, \bar{b}_i^j) | j < \omega\}$ is k_i-inconsistent. Now, as in the previous proof of (i)→(iii), Lemma 2.5. enables us to show that $\psi_i(\bar{x}, \bar{b}_i)$ divides over $A \cup \bigcup\{\bar{c}_m | 0 < m\}$ w.r.t. k_i.

Finally, if $\{\varphi(\bar{x}, \bar{c}_n) | n < \omega\}$ is consistent then again similar argument to the preceding proof leads to a contradiction. Hence $\varphi(\bar{x}, \bar{c})$ divides over A. □

Proposition 2.7. says forking is equivalent to dividing for simple T. Moreover it says for simple T, in order to make sure that $\varphi(\bar{x}, \bar{c})$ does not divide over A, it suffices to check that $\{\varphi(\bar{x}, \bar{c}') | \bar{c}' \in I\}$ is consistent for an arbitrary Morley sequence I of $tp(\bar{c}/A)$. This property essentially produces symmetry and transitivity of forking as the reader will see.

EXAMPLE 2.8. Simplicity of T is essential in 2.7. For consider a dense linear ordering $(M, <)$ without end points. It is not simple. Let $a_0 < b_0$ in M. Choose sequences $\ldots a_2 < a_1 < a_0$ and $b_0 < b_1 < b_2 < \ldots$, then $\langle (a_n, b_n) | n \in \omega \rangle$ is a Morley sequence of $tp((a_0, b_0)/\emptyset)$. Now $\{a_n < x < b_n | n \in \omega\}$ is consistent whereas $a_0 < x < b_0$ divides over $\emptyset$.

THEOREM 2.9. *If T is simple, then nonforking satisfies symmetry and transitivity.*

Proof. (i) (Symmetry) Suppose that $tp(\bar{b}/A\bar{c})$ does not fork over A. We want to prove that $tp(\bar{c}/A\bar{b})$ does not fork over A. Now there is a Morley sequence $I = \langle \bar{b}_i | i \in \omega \rangle$ of $tp(\bar{b}/A\bar{c})$ over A where $\bar{b}_0 = \bar{b}$ (by 2.4.) It is easy to check that I is a Morley sequence of $tp(\bar{b}/A)$. Hence by 2.7., it suffices to show that for given $\varphi(\bar{x}, \bar{a}, \bar{b})$ in $tp(\bar{c}/A\bar{b})$ with $\bar{a} \in A$, $\{\varphi(\bar{x}, \bar{a}, \bar{b}_i) | i \in \omega\} = p$ is consistent. But by $A\bar{c}$-indiscernibility of I, $\bar{c}$ realizes p.

(ii) (Transitivity (1.4.(ii))) (→) Holds for any T.

(←) By symmetry it suffices to show that for given $\bar{c} \in C$, any $\psi(\bar{x}, \bar{d}, \bar{a})$ (with $\bar{d} \in A$) in $tp(\bar{c}/A\bar{a})$ does not fork over A. Hence again by 2.7. it is enough to obtain a Morley sequence $\langle \bar{a}_i | i \in \omega \rangle$ of $tp(\bar{a}/A)$ with $\bar{a}_0 = \bar{a}$ such that $\{\psi(\bar{x}, \bar{d}, \bar{a}_i) | i \in \omega\}$ is consistent. We note that a Morley sequence $I = \langle \bar{a}_i | i \in \omega \rangle$ of $tp(\bar{a}/B)$ over A with $\bar{a}_0 = \bar{a}$ is a desired sequence, by 2.7. and the fact that $tp(\bar{c}/B\bar{a})$ does not fork over B. □

§3. Rank and forking

We introduce the $D(p, \Delta, k)$ rank which expresses simplicity and forking. (3.3., 3.5.)

DEFINITION 3.1. Let Δ be a finite set of formulas in L, and let k be a positive integer. For any type p, $D(p, \Delta, k)$ (either a natural number or ∞) is defined by induction as follows;

(i) $D(p, \Delta, k) \geq 0$ for any consistent type p.

(ii) $D(p, \Delta, k) \geq n + 1$ if there are $\varphi(\bar{x}, \bar{y})$ in Δ and tuples $\{\bar{a}_i | i < \omega\}$ such that $D(p \cup \{\varphi(\bar{x}, \bar{a}_i)\}, \Delta, k) \geq n$ for each $i < \omega$, and $\{\varphi(\bar{x}, \bar{a}_i) | i < \omega\}$ is k-inconsistent.

(iii) $D(p, \Delta, k) = n$ if $D(p, \Delta, k) \geq n$ and $D(p, \Delta, k) \not\geq n + 1$.
(iv) $D(p, \Delta, k) = \infty$ if $D(p, \Delta, k) \geq n$ for all $n \in \omega$.

The basic properties of the rank $D(p, \Delta, k)$ are as follows.

LEMMA 3.2. (i) *Let a type p over a set A, finite Δ, k be given. Then $D(p, \Delta, k) \geq n + 1$ if and only if there are an A-indiscernible sequence $\langle \bar{a}_i | i < \omega \rangle$ and $\varphi(\bar{x}, \bar{y}) \in \Delta$ such that $D(p \cup \{\varphi(\bar{x}, \bar{a}_0)\}, \Delta, k) \geq n$ and $\{\varphi(\bar{x}, \bar{a}_i) | i < \omega\}$ is k-inconsistent.*

(ii) $D(p_1, \Delta_1, k_1) \leq D(p_2, \Delta_2, k_2)$ *if* $p_1 \vdash p_2$, $\Delta_1 \subseteq \Delta_2$, *and* $k_1 \leq k_2$.

(iii) *For every p, Δ, k, there is a finite subset q of p such that $D(p, \Delta, k) = D(q, \Delta, k)$.*

(iv) *For each finite Δ, we can find a formula $\psi(\bar{x}, \bar{y}) \in L$ such that for every type p, and every k, $D(p, \Delta, k) = D(p, \psi(\bar{x}, \bar{y}), k)$.*

(v) $D(p \cup \{\bigvee_{i<n} \varphi_i(\bar{x}, \bar{a}_i)\}, \Delta, k) = Max_{i<n} D(p \cup \{\varphi_i(\bar{x}, \bar{a}_i)\}, \Delta, k)$.

(vi) *Let Δ, k be given. For each type p over A, there is an extension $p' \in S(A)$ of p such that $D(p, \Delta, k) = D(p', \Delta, k)$.*

(vii) *If $p \in S(B)$ forks over $A (\subseteq B)$ witnessed by $\psi_i(\bar{x}, \bar{b}_i)$ and k_i $(i = 0, ..., r)$, then $D(p, \Delta, k) < \omega$ implies $D(p, \Delta, k) < D(p|A, \Delta, k)$ where $\Delta = \{\psi_i(\bar{x}, \bar{b}_i) | i = 0, ..., r\}$ and $k = Max\{k_i | i = 0, ..., r\}$.*

LEMMA 3.3. *The following are all equivalent.*

(i) *T is not simple.*

(ii) *There are a sequence of sets $\langle A_i | i \in |T|^+ \rangle$ with $A_i \subseteq A_j$ for $i \leq j$, and a complete type p over $\cup\{A_i | i \in |T|^+\}$ such that $p|A_{i+1}$ forks over A_i for all $i \in |T|^+$.*

(iii) *There is a complete type p over a set B such that for any subset A of B with $|A| \leq |T|$, p divides over A.*

(iv) *There are a sequence of sets $\langle A_i | i \in |T|^+ \rangle$ with $A_i \subseteq A_j$ for $i \leq j$, and a complete type p over $\cup\{A_i | i \in |T|^+\}$ such that $p|A_{i+1}$ divides over A_i for all $i \in |T|^+$.*

(v) *There are p, $\varphi(\bar{x}, \bar{y})$, k such that $D(p, \varphi(\bar{x}, \bar{y}), k) = \infty$.*

(vi) *T has the tree property, i.e. there exist a formula $\varphi(\bar{x}, \bar{y})$, an integer k and tuples $\bar{c}_\alpha$ with $\alpha \in \omega^{<\omega}$ such that for any $\alpha \in \omega^{<\omega}$, $\{\varphi(\bar{x}, \bar{c}_{\alpha^\frown n}) | n \in \omega\}$ is k-inconsistent, whereas $\{\varphi(\bar{x}, \bar{c}_{\beta|n}) | n \in \omega\}$ is consistent for all $\beta \in \omega^\omega$.*

Proof. (i)↔(ii), (iii)↔(iv), (v)→(vi)→(iv)→(ii) Left to the reader.
(ii)→(v) By Lemma 3.2.(vii),(iv) and the fact $|L| = |T|$. □

Lemma 3.3. says that T is simple if and only if $D(p, \Delta, k)$ is defined for any p, finite Δ, k. Furthermore it says T is simple if and only if T satisfies the Local Character axiom for nondividing. This fact does not automatically follow from the fact that forking is dividing for simple T.

LEMMA 3.4. *The following are equivalent.*

(i) *$tp(\bar{c}/A\bar{b})$ does not divide over A.*

(ii) *For any A-indiscernible sequence I with $\bar{b} \in I$, there is a tuple $\bar{c}'$ realizing $tp(\bar{c}/A\bar{b})$ such that I is an $A\bar{c}'$-indiscernible sequence.*

Proof. (ii)→(i) By the definition of dividing.

(i)→(ii) Let $p(\bar{x}, \bar{b}) = tp(\bar{c}/A\bar{b})$. Then since $p(\bar{x}, \bar{b})$ does not divide over A, it can easily be shown that $q = \cup\{p(\bar{x}, \bar{b}')|\bar{b}' \in I\}$ is consistent. Moreover using Ramsey's Theorem, we can select a tuple $\bar{c}'$ realizing q such that I is $A\bar{c}'$-indiscernible. □

PROPOSITION 3.5. *Let T be simple. $p \in S(B)$ with $A \subseteq B$. Then p does not fork over A if and only if $D(p, \psi(\bar{x}, \bar{y}), k) = D(p|A, \psi(\bar{x}, \bar{y}), k)$ for every $\psi(\bar{x}, \bar{y}) \in L$ and $k \in \omega$.*

Proof. (←) Lemma 3.2.(vii).

(→) We may assume that $B = A\bar{b}$ with finite $\bar{b}$. Let us suppose $D(p|A, \psi(\bar{x}, \bar{y}), k) \geq n+1$. We shall show $D(p, \psi(\bar{x}, \bar{y}), k) \geq n+1$ assuming the induction hypothesis for n. Using Ramsey's Theorem with the basic properties of the rank (3.2.(i),(vi)), we can assume that the following situation holds.

(1) There is an A-indiscernible sequence $\langle \bar{a}_i\bar{c}_i|i < \omega\rangle$ with $p = tp(\bar{a}_0/A\bar{b})$.

(2) $\{\psi(\bar{x}, \bar{c}_i)|i < \omega\}$ is k-inconsistent.

(3) $p|A \cup \psi(\bar{x}, \bar{c}_i) \subseteq tp(\bar{a}_i/A\bar{c}_i)$ for all $i < \omega$.

(4) $D(tp(\bar{a}_i/A\bar{c}_i), \psi(\bar{x}, \bar{y}), k) \geq n$.

Now by the forking axioms and 3.4. we may assume that $tp(\bar{b}/A\bar{a}_0\bar{c}_0)$ does not fork over A and $\langle \bar{a}_i\bar{c}_i|i < \omega\rangle$ is $A\bar{b}$-indiscernible. Since now $tp(\bar{a}_0/A\bar{c}_0\bar{b})$ does not fork over $A\bar{c}_0$, we notice $D(tp(\bar{a}_i/A\bar{c}_i\bar{b}), \psi, k) \geq n$ by the induction hypothesis. Hence $D(p \cup \psi(\bar{x}, \bar{c}_i), \psi, k) \geq n$ for each $i < \omega$, and we conclude $D(p, \psi, k) \geq n+1$. □

§4. The Independence Theorem

The aim of this section is Proposition 4.5. which is an analogous result to Fact 1.4. In particular we suggest that 4.2. (and/or) the Independence Theorem (over a model), which are weaker forms of the Boundedness axiom, can be substitutes for it for simple theories.

LEMMA 4.1. *Let T be simple and $p(\bar{x}, \bar{a}_0)$ be a type over $A\bar{a}_0$ which does not fork over A. If $I = \langle \bar{a}_i|i < \omega\rangle$ is a Morley sequence of $tp(\bar{a}_0/A)$, then $\cup\{p(\bar{x}, \bar{a}_i)|i < \omega\}$ is consistent and does not fork over A.*

Proof. Let $\varphi(\bar{x}, \bar{a}_0, \bar{c}) \in p(\bar{x}, \bar{a}_0)$ where $\bar{c} \subseteq A$. We note that $\{\varphi(\bar{x}, \bar{a}_i, \bar{c})|i < \omega\}$ is consistent by 2.7. Now it suffices to show that for given $n < \omega$, $\varphi(\bar{x}, \bar{a}_0, \bar{c}) \wedge \ldots \wedge \varphi(\bar{x}, \bar{a}_{n-1}, \bar{c})$ does not fork over A. This is guaranteed again by 2.7. since $\langle \bar{b}_r|r \in \omega\rangle$ where $\bar{b}_r = \bar{a}_{n \cdot r}\bar{a}_{n \cdot r+1}\ldots\bar{a}_{n \cdot r+n-1}\bar{c}$ is a Morley sequence of $tp(\bar{b}_0/A)$. □

Lemma 4.1. is true even if I is just an indiscernible sequence.

COROLLARY 4.2. *Let T be simple and $p(\bar{x}, \bar{a}_0)$ be a type over $A\bar{a}_0$ which does not fork over A. If $I = \langle \bar{a}_i|i < \omega\rangle$ is an A-indiscernible sequence of $tp(\bar{a}_0/A)$, then $\cup\{p(\bar{x}, \bar{a}_i)|i < \omega\}$ is consistent and does not fork over A.*

Proof. Suppose that $I' = \langle \bar{a}_{\omega+i}|i < \omega\rangle$ is a sequence such that $I^\frown I'$ is A-indiscernible. Using 3.2.(iii) and (vii), it can easily be noticed that I' is a Morley sequence of $tp(\bar{a}_\omega/AI)$. Now let us take a complete extension $p'(\bar{x}, \bar{a}_\omega)$ over $AI\bar{a}_\omega$

of $p(\bar{x}, \bar{a}_\omega)$ which does not fork over A. A basic forking calculation and 4.1. enable us to prove that $\cup\{p'(\bar{x}, \bar{a}_{\omega+i}) | i < \omega\}$ does not fork over A. Hence neither $\cup\{p(\bar{x}, \bar{a}_{\omega+i}) | i < \omega\}$ nor $\cup\{p(\bar{x}, \bar{a}_i) | i < \omega\}$ forks over A. □

We recall the notion "coheir" which plays an important role in proving 4.3. (A detailed explanation of coheirs can be found in [**10**] or [**11**].) Let $M \subseteq A$. Then $p \in S(A)$ is a *coheir* of $p|M$ if every $L(A)$ formula $\psi(\bar{x})$ in p is satisfied by some tuple in M. $I = \langle \bar{a}_i | i < \omega \rangle$ is a *coheir sequence of* $q \in S(M)$ if $tp(\bar{a}_i/M \cup \bigcup\{\bar{a}_j | j < i\})$ is a coheir of q and a subset of $tp(\bar{a}_{i+1}/M \cup \bigcup\{\bar{a}_j | j < i+1\})$ for each $i < \omega$. For any T, a coheir sequence of given $q \in S(M)$ exists and any coheir sequence of q is an M-indiscernible sequence of q. Moreover, a coheir sequence is a Morley sequence, but the converse is not true in general, even for simple unstable T. A counterexample appears in the random graph. A rather important property of a coheir sequence is as follows. Suppose that $\{I_n | n < \omega\}$ is a family of coheir sequences of $q \in S(M)$ such that $tp(I_n/M) = tp(I_0/M)$ for each $n < \omega$. Then there is a sequence I' which is a common extension of the I_n's, i.e. $I_n^\frown I'$ is M-indiscernible for each $n < \omega$.

If T is simple, we call a set of tuples $\{\bar{c}_i | i < \kappa\}$ *A-independent* if for each $i < \kappa$, $tp(\bar{c}_i/A \cup \bigcup\{\bar{c}_j | j \neq i, j < \kappa\})$ does not fork over A.

THEOREM 4.3. (the Independence Theorem) *Let T be simple and M be a model of T. Suppose that $\{\bar{a}, \bar{b}\}$ is M-independent and $p \in S(M)$. Let $p_1 \in S(M\bar{a})$ and $p_2 \in S(M\bar{b})$ be extensions of p which do not fork over M. Then $p_1 \cup p_2$ is consistent and does not fork over M.*

Proof. Suppose not, then there are formulas $\varphi(\bar{x}, \bar{y}), \psi(\bar{x}, \bar{y}) \in L(M)$ such that $\varphi(\bar{x}, \bar{a}) \in p_1$, $\psi(\bar{x}, \bar{b}) \in p_2$ and $\varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{x}, \bar{b})$ forks over M (or is inconsistent.) Now by a basic forking calculation, we can find a tuple $\bar{b}'$ realizing $tp(\bar{b}/M)$ such that $\{\bar{a}, \bar{b}'\}$ is M-independent and $\varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{x}, \bar{b}')$ is consistent and does not fork over M. Now let $I = \langle \bar{b}_i | i < \omega \rangle$ be a coheir sequence of $tp(\bar{b}/M)$ with $\bar{b}_0 = \bar{b}$. We may assume that I is $M\bar{a}$-indiscernible, by 3.4. Moreover there is an $M\bar{a}$-indiscernible sequence $I' = \langle \bar{b}^j | j < \omega \rangle$ with $\bar{b}^0 = \bar{b}'$ such that $tp(I/M) = tp(I'/M)$. Let J be a common M-indiscernible extension of I and I'. We may further assume that there is an infinite subsequence $J' = \langle \bar{c}_i | i < \omega \rangle$ of J such that $tp(\bar{a}\bar{c}_i/M)$ are all the same for all $i < \omega$. Now in order to deduce a contradiction, let us consider the following two cases.

(Case I) $\varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{x}, \bar{c}_i)$ is consistent and does not fork over M for each $\bar{c}_i \in J'$. Now with I, we already have the following additional situation.

$$\varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{x}, \bar{b}_i) \text{ forks over } M \text{ (or is inconsistent) for each } \bar{b}_i \in I$$

By indiscernibility of IJ', we obtain a sequence $\langle \bar{a}_i | i < \omega \rangle$ with $\bar{a}_0 = \bar{a}$ such that

$$\varphi(\bar{x}, \bar{a}_i) \wedge \psi(\bar{x}, \bar{c}_j) \text{ is consistent and does not fork over } M \text{ if and only if } i \leq j.$$

Furthermore Ramsey's Theorem enables us to assume that $\langle \bar{a}_i \bar{c}_i | i < \omega \rangle$ is M-indiscernible. But it can easily be seen that this contradicts Corollary 4.2.

(Case II) $\varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{x}, \bar{c}_i)$ forks over M (or is inconsistent) for each $\bar{c}_i \in J'$. Since $I'J'$ is M-indiscernible, similarly to (Case I), we obtain a contradiction. □

COROLLARY 4.4. *Let T be simple. Let $\{\bar{a}_i | i \in I\}$ be M-independent and $p \in S(M)$. Suppose that $p_i \in S(M\bar{a}_i)$ is an extension of p which does not fork over M for each $i \in I$, then $\cup\{p_i | i \in I\}$ is consistent and does not fork over M.*

(Question) Does the Independence Theorem hold over an algebraically closed set in $\mathcal{C}^{eq}$ when T is simple?

For stable T, the Independence Theorem over a model obviously follows from the uniqueness of nonforking extensions over a model. The Independence Theorem over an algebraically closed set (in $\mathcal{C}^{eq}$) also holds since there is a unique nonforking extension over the set, too. But we can not expect the Independence Theorem to hold over an arbitrary set. An equivalence relation having finitely many infinite equivalence classes trivially supplies a counterexample.

PROPOSITION 4.5. (1) *If T is simple, then nonforking satisfies* (i) *Symmetry*, (ii) *Transitivity*, (iii) *Extension*, (iv) *Local Character*, (v) *Finite Character and the following.*

(vi) (4.2.) *Suppose that $p(\bar{x}, \bar{a}_0) \in S(A\bar{a}_0)$ does not fork over A. Let $I = \langle \bar{a}_i | i < \omega \rangle$ be an A-indiscernible sequence of $tp(\bar{a}_0/A)$. Then there is a tuple $\bar{b}$ realizing $\cup\{p(\bar{x}, \bar{a}_i) | i < \omega\}$ such that $tp(\bar{b}/AI)$ does not fork over A.*

(vi)′ *(the Independence Theorem) Suppose that $\{\bar{a}, \bar{b}\}$ is M-independent and $p \in S(M)$. Let $p_1 \in S(M\bar{a})$ and $p_2 \in S(M\bar{b})$ be extensions of p which do not fork over M. Then there is an extension $p_3 \in S(M\bar{a}\bar{b})$ of $p_1 \cup p_2$ which does not fork over M.*

(2) *Assume that, in a theory T, an automorphism-invariant relation between complete types and sets satisfies the preceding axioms* (i) *to* (v) *(in terms of nonforking). If the relation satisfies* (vi)′ *then it satisfies* (vi).

(3) *If the relation in T satisfies* (i) *to* (v) *and either* (vi) *or* (vi)′, *then T is simple and the relation is nonforking.*

Proof. (1) Done.

(2) Suppose that the relation in T satisfies (i) to (v) and (vi)′. Let us temporarily say "p is free over A" if (p, A) satisfies the relation. Now if $I = \langle \bar{a}_i | i < \omega \rangle$ is an A-indiscernible sequence, then Ramsey's Theorem and (iv) enable us to find a model M containing A such that I is M-independent (w.r.t. the relation). Now suppose that $p(\bar{x}, \bar{a}_0) \in S(A\bar{a}_0)$ is free over A. Then by a routine argument we can show that $\cup\{p(\bar{x}, \bar{a}_i) | i < \omega\}$ has a complete extension over AI which is free over A.

(3) Suppose that the relation in T satisfies (i) to (v) and (vi). It is enough to show that whenever $p \in S(A\bar{b})$ then p is free over A if and only if p does not divide over A (cf. Lemma 3.3.). Let us denote $p = p(\bar{x}, \bar{b})$. Now there is an A-indiscernible and A-independent (w.r.t. the relation) sequence $I = \langle \bar{b}_i | i < \omega \rangle$ with $\bar{b}_0 = \bar{b}$ (cf. Proposition 2.4.). If $p(\bar{x}, \bar{b})$ is not free over A, then $\{p(\bar{x}, \bar{b}_i) | i < \omega\}$ should be inconsistent in order not to contradict (iv). Hence $p(\bar{x}, \bar{b})$ divides over A. Conversely, let us assume that $p(\bar{x}, \bar{b})$ is free over A. Then by (vi), $p(\bar{x}, \bar{b})$ does not divide over A. □

4.5.(3) can be applied to fields with an automorphism, pseudo-finite fields and

smoothly approximable structures where there are notions of independence fulfilling basic axioms plus the Independence Theorem over a model. Furthermore these structures are known to be unstable. Hence we conclude that each structure is simple unstable, and the independence notion in each structure is nonforking.

References

[1] J. T. Baldwin, *Fundamentals of stability theory* (Springer-Verlag, New York, 1987).

[2] C. C. Chang and H. J. Keisler, *Model theory*, 3rd edn (North-Holland, Amsterdam, 1990).

[3] Z. Chatzidakis and E. Hrushovski, 'Algebraically closed fields with an automorphism', preprint.

[4] G. Cherlin and E. Hrushovski, 'Lie coordinatised structures', in preparation.

[5] E. Hrushovski, 'PAC and related structures', manuscript (1991).

[6] E. Hrushovski and A. Pillay, 'Groups definable in local fields and pseudo-finite fields', *Israel J. Math.* 85 (1994) 203-262.

[7] W. M. Kantor, M. W. Liebeck, and H. D. Macpherson, '$\aleph_0$-categorical structures smoothly approximated by finite structures', *Proc. London Math. Soc.* (3) 59 (1989) 439-463.

[8] B. Kim, 'Forking in simple unstable theories', *J. of London Math. Soc.*, to appear.

[9] B. Kim and A. Pillay, 'Simple theories', *Ann. Pure and Applied Logic*, to appear.

[10] D. Lascar and B. Poizat, 'An introduction to forking', *J. of Symbolic Logic* 44 (1979) 330-350.

[11] A. Pillay, *An introduction to stability theory* (Clarendon Press, Oxford, 1983).

[12] S. Shelah, 'Simple unstable theories', *Ann. Math. Logic* 19 (1980) 177-203.

[13] S. Shelah, *Classification theory*, revised (North-Holland, Amsterdam, 1990).

Author's address:
The Fields Institute,
222 College Street,
Toronto, Ontario,
Canada M5T 3J1.
e-mail: bkim@fields.utoronto.ca