London Mathematical Society
Lecture Note Series 422

Groups St Andrews 2013

Edited by
C. M. Campbell, M. R. Quick, E. F. Robertson
and C. M. Roney-Dougal

CAMBRIDGE

LONDON MATHEMATICAL SOCIETY LECTURE NOTE SERIES

Managing Editor: Professor M. Reid, Mathematics Institute,
University of Warwick, Coventry CV4 7AL, United Kingdom

The titles below are available from booksellers, or from Cambridge University Press at
http://www.cambridge.org/mathematics

301 Stable modules and the D(2)-problem, F.E.A. JOHNSON
302 Discrete and continuous nonlinear Schrödinger systems, M.J. ABLOWITZ, B. PRINARI & A.D. TRUBATCH
303 Number theory and algebraic geometry, M. REID & A. SKOROBOGATOV (eds)
304 Groups St Andrews 2001 in Oxford I, C.M. CAMPBELL, E.F. ROBERTSON & G.C. SMITH (eds)
305 Groups St Andrews 2001 in Oxford II, C.M. CAMPBELL, E.F. ROBERTSON & G.C. SMITH (eds)
306 Geometric mechanics and symmetry, J. MONTALDI & T. RATIU (eds)
307 Surveys in combinatorics 2003, C.D. WENSLEY (ed)
308 Topology, geometry and quantum field theory, U.L. TILLMANN (ed)
309 Corings and comodules, T. BRZEZINSKI & R. WISBAUER
310 Topics in dynamics and ergodic theory, S. BEZUGLYI & S. KOLYADA (eds)
311 Groups: topological, combinatorial and arithmetic aspects, T.W. MÜLLER (ed)
312 Foundations of computational mathematics, Minneapolis 2002, F. CUCKER *et al* (eds)
313 Transcendental aspects of algebraic cycles, S. MÜLLER-STACH & C. PETERS (eds)
314 Spectral generalizations of line graphs, D. CVETKOVIĆ, P. ROWLINSON & S. SIMIĆ
315 Structured ring spectra, A. BAKER & B. RICHTER (eds)
316 Linear logic in computer science, T. EHRHARD, P. RUET, J.-Y. GIRARD & P. SCOTT (eds)
317 Advances in elliptic curve cryptography, I.F. BLAKE, G. SEROUSSI & N.P. SMART (eds)
318 Perturbation of the boundary in boundary-value problems of partial differential equations, D. HENRY
319 Double affine Hecke algebras, I. CHEREDNIK
320 L-functions and Galois representations, D. BURNS, K. BUZZARD & J. NEKOVÁŘ (eds)
321 Surveys in modern mathematics, V. PRASOLOV & Y. ILYASHENKO (eds)
322 Recent perspectives in random matrix theory and number theory, F. MEZZADRI & N.C. SNAITH (eds)
323 Poisson geometry, deformation quantisation and group representations, S. GUTT *et al* (eds)
324 Singularities and computer algebra, C. LOSSEN & G. PFISTER (eds)
325 Lectures on the Ricci flow, P. TOPPING
326 Modular representations of finite groups of Lie type, J.E. HUMPHREYS
327 Surveys in combinatorics 2005, B.S. WEBB (ed)
328 Fundamentals of hyperbolic manifolds, R. CANARY, D. EPSTEIN & A. MARDEN (eds)
329 Spaces of Kleinian groups, Y. MINSKY, M. SAKUMA & C. SERIES (eds)
330 Noncommutative localization in algebra and topology, A. RANICKI (ed)
331 Foundations of computational mathematics, Santander 2005, L.M PARDO, A. PINKUS, E. SÜLI & M.J. TODD (eds)
332 Handbook of tilting theory, L. ANGELERI HÜGEL, D. HAPPEL & H. KRAUSE (eds)
333 Synthetic differential geometry (2nd Edition), A. KOCK
334 The Navier–Stokes equations, N. RILEY & P. DRAZIN
335 Lectures on the combinatorics of free probability, A. NICA & R. SPEICHER
336 Integral closure of ideals, rings, and modules, I. SWANSON & C. HUNEKE
337 Methods in Banach space theory, J.M.F. CASTILLO & W.B. JOHNSON (eds)
338 Surveys in geometry and number theory, N. YOUNG (ed)
339 Groups St Andrews 2005 I, C.M. CAMPBELL, M.R. QUICK, E.F. ROBERTSON & G.C. SMITH (eds)
340 Groups St Andrews 2005 II, C.M. CAMPBELL, M.R. QUICK, E.F. ROBERTSON & G.C. SMITH (eds)
341 Ranks of elliptic curves and random matrix theory, J.B. CONREY, D.W. FARMER, F. MEZZADRI & N.C. SNAITH (eds)
342 Elliptic cohomology, H.R. MILLER & D.C. RAVENEL (eds)
343 Algebraic cycles and motives I, J. NAGEL & C. PETERS (eds)
344 Algebraic cycles and motives II, J. NAGEL & C. PETERS (eds)
345 Algebraic and analytic geometry, A. NEEMAN
346 Surveys in combinatorics 2007, A. HILTON & J. TALBOT (eds)
347 Surveys in contemporary mathematics, N. YOUNG & Y. CHOI (eds)
348 Transcendental dynamics and complex analysis, P.J. RIPPON & G.M. STALLARD (eds)
349 Model theory with applications to algebra and analysis I, Z. CHATZIDAKIS, D. MACPHERSON, A. PILLAY & A. WILKIE (eds)
350 Model theory with applications to algebra and analysis II, Z. CHATZIDAKIS, D. MACPHERSON, A. PILLAY & A. WILKIE (eds)
351 Finite von Neumann algebras and masas, A.M. SINCLAIR & R.R. SMITH
352 Number theory and polynomials, J. MCKEE & C. SMYTH (eds)
353 Trends in stochastic analysis, J. BLATH, P. MÖRTERS & M. SCHEUTZOW (eds)
354 Groups and analysis, K. TENT (ed)
355 Non-equilibrium statistical mechanics and turbulence, J. CARDY, G. FALKOVICH & K. GAWEDZKI
356 Elliptic curves and big Galois representations, D. DELBOURGO
357 Algebraic theory of differential equations, M.A.H. MACCALLUM & A.V. MIKHAILOV (eds)
358 Geometric and cohomological methods in group theory, M.R. BRIDSON, P.H. KROPHOLLER & I.J. LEARY (eds)
359 Moduli spaces and vector bundles, L. BRAMBILA-PAZ, S.B. BRADLOW, O. GARCÍA-PRADA & S. RAMANAN (eds)

360 Zariski geometries, B. ZILBER
361 Words: Notes on verbal width in groups, D. SEGAL
362 Differential tensor algebras and their module categories, R. BAUTISTA, L. SALMERÓN & R. ZUAZUA
363 Foundations of computational mathematics, Hong Kong 2008, F. CUCKER, A. PINKUS & M.J. TODD (eds)
364 Partial differential equations and fluid mechanics, J.C. ROBINSON & J.L. RODRIGO (eds)
365 Surveys in combinatorics 2009, S. HUCZYNSKA, J.D. MITCHELL & C.M. RONEY-DOUGAL (eds)
366 Highly oscillatory problems, B. ENGQUIST, A. FOKAS, E. HAIRER & A. ISERLES (eds)
367 Random matrices: High dimensional phenomena, G. BLOWER
368 Geometry of Riemann surfaces, F.P. GARDINER, G. GONZÁLEZ-DIEZ & C. KOUROUNIOTIS (eds)
369 Epidemics and rumours in complex networks, M. DRAIEF & L. MASSOULIÉ
370 Theory of p-adic distributions, S. ALBEVERIO, A.YU. KHRENNIKOV & V.M. SHELKOVICH
371 Conformal fractals, F. PRZYTYCKI & M. URBAŃSKI
372 Moonshine: The first quarter century and beyond, J. LEPOWSKY, J. MCKAY & M.P. TUITE (eds)
373 Smoothness, regularity and complete intersection, J. MAJADAS & A. G. RODICIO
374 Geometric analysis of hyperbolic differential equations: An introduction, S. ALINHAC
375 Triangulated categories, T. HOLM, P. JØRGENSEN & R. ROUQUIER (eds)
376 Permutation patterns, S. LINTON, N. RUŠKUC & V. VATTER (eds)
377 An introduction to Galois cohomology and its applications, G. BERHUY
378 Probability and mathematical genetics, N. H. BINGHAM & C. M. GOLDIE (eds)
379 Finite and algorithmic model theory, J. ESPARZA, C. MICHAUX & C. STEINHORN (eds)
380 Real and complex singularities, M. MANOEL, M.C. ROMERO FUSTER & C.T.C WALL (eds)
381 Symmetries and integrability of difference equations, D. LEVI, P. OLVER, Z. THOMOVA & P. WINTERNITZ (eds)
382 Forcing with random variables and proof complexity, J. KRAJÍČEK
383 Motivic integration and its interactions with model theory and non-Archimedean geometry I, R. CLUCKERS, J. NICAISE & J. SEBAG (eds)
384 Motivic integration and its interactions with model theory and non-Archimedean geometry II, R. CLUCKERS, J. NICAISE & J. SEBAG (eds)
385 Entropy of hidden Markov processes and connections to dynamical systems, B. MARCUS, K. PETERSEN & T. WEISSMAN (eds)
386 Independence-friendly logic, A.L. MANN, G. SANDU & M. SEVENSTER
387 Groups St Andrews 2009 in Bath I, C.M. CAMPBELL *et al* (eds)
388 Groups St Andrews 2009 in Bath II, C.M. CAMPBELL *et al* (eds)
389 Random fields on the sphere, D. MARINUCCI & G. PECCATI
390 Localization in periodic potentials, D.E. PELINOVSKY
391 Fusion systems in algebra and topology, M. ASCHBACHER, R. KESSAR & B. OLIVER
392 Surveys in combinatorics 2011, R. CHAPMAN (ed)
393 Non-abelian fundamental groups and Iwasawa theory, J. COATES *et al* (eds)
394 Variational problems in differential geometry, R. BIELAWSKI, K. HOUSTON & M. SPEIGHT (eds)
395 How groups grow, A. MANN
396 Arithmetic differential operators over the p-adic integers, C.C. RALPH & S.R. SIMANCA
397 Hyperbolic geometry and applications in quantum chaos and cosmology, J. BOLTE & F. STEINER (eds)
398 Mathematical models in contact mechanics, M. SOFONEA & A. MATEI
399 Circuit double cover of graphs, C.-Q. ZHANG
400 Dense sphere packings: a blueprint for formal proofs, T. HALES
401 A double Hall algebra approach to affine quantum Schur–Weyl theory, B. DENG, J. DU & Q. FU
402 Mathematical aspects of fluid mechanics, J.C. ROBINSON, J.L. RODRIGO & W. SADOWSKI (eds)
403 Foundations of computational mathematics, Budapest 2011, F. CUCKER, T. KRICK, A. PINKUS & A. SZANTO (eds)
404 Operator methods for boundary value problems, S. HASSI, H.S.V. DE SNOO & F.H. SZAFRANIEC (eds)
405 Torsors, étale homotopy and applications to rational points, A.N. SKOROBOGATOV (ed)
406 Appalachian set theory, J. CUMMINGS & E. SCHIMMERLING (eds)
407 The maximal subgroups of the low-dimensional finite classical groups, J.N. BRAY, D.F. HOLT & C.M. RONEY-DOUGAL
408 Complexity science: the Warwick master's course, R. BALL, V. KOLOKOLTSOV & R.S. MACKAY (eds)
409 Surveys in combinatorics 2013, S.R. BLACKBURN, S. GERKE & M. WILDON (eds)
410 Representation theory and harmonic analysis of wreath products of finite groups, T. CECCHERINI-SILBERSTEIN, F. SCARABOTTI & F. TOLLI
411 Moduli spaces, L. BRAMBILA-PAZ, O. GARCÍA-PRADA, P. NEWSTEAD & R.P. THOMAS (eds)
412 Automorphisms and equivalence relations in topological dynamics, D.B. ELLIS & R. ELLIS
413 Optimal transportation, Y. OLLIVIER, H. PAJOT & C. VILLANI (eds)
414 Automorphic forms and Galois representations I, F. DIAMOND, P.L. KASSAEI & M. KIM (eds)
415 Automorphic forms and Galois representations II, F. DIAMOND, P.L. KASSAEI & M. KIM (eds)
416 Reversibility in dynamics and group theory, A.G. O'FARRELL & I. SHORT
417 Recent advances in algebraic geometry, C.D. HACON, M. MUSTAŢĂ & M. POPA (eds)
418 The Bloch–Kato conjecture for the Riemann zeta function, J. COATES, A. RAGHURAM, A. SAIKIA & R. SUJATHA (eds)
419 The Cauchy problem for non-Lipschitz semi-linear parabolic partial differential equations, J.C. MEYER & D.J. NEEDHAM
420 Arithmetic and geometry, L. DIEULEFAIT *et al* (eds)
421 O-minimality and Diophantine geometry, G.O. JONES & A.J. WILKIE (eds)
422 Groups St Andrews 2013, C.M. CAMPBELL *et al* (eds)
423 Inequalities for graph eigenvalues, Z. STANIĆ
424 Surveys in combinatorics 2015, A. CZUMAJ *et al* (eds)

London Mathematical Society Lecture Note Series: 422

Groups St Andrews 2013

Edited by

C. M. CAMPBELL
University of St Andrews

M. R. QUICK
University of St Andrews

E. F. ROBERTSON
University of St Andrews

C. M. RONEY-DOUGAL
University of St Andrews

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107514546

First published 2015

Printed in the United Kingdom by Clays, St Ives plc

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

ISBN 978-1-107-51454-6 Paperback

CONTENTS

Approximate subgroups and super-strong approximation
Emmanuel Breuillard 1

Width questions for finite simple groups
Martin W. Liebeck 51

Profinite properties of discrete groups
Alan W. Reid 73

$GL(n,\mathbb{Z})$, $Out(F_n)$ and everything in between: automorphism groups of RAAGs
Karen Vogtmann 105

Permutation groups and transformation semigroups: results and problems
João Araújo & Peter J. Cameron 128

New progress on factorized groups and subgroup permutability
Milagros Arroyo-Jordá, Paz Arroyo-Jordá, Ana Martínez-Pastor & M. Dolores Pérez-Ramos 142

A survey on the normalizer problem for integral group rings
Andreas Bächle 152

A survey on Clifford-Fischer theory
Ayoub B.M. Basheer & Jamshid Moori 160

A generalisation on the solvability of finite groups with three class sizes for normal subgroups
Antonio Beltrán & María José Felipe 173

Automorphism groups of non-orientable Riemann surfaces
E. Bujalance, F.J. Cirre, J.J. Etayo, G. Gromadzki & E. Martínez 183

What are the $\mathcal{C}_2$-groups?
Inna Capdeboscq & Christopher Parker 194

Resurrecting Wells' exact sequence and Buckley's group action
Jill Dietz 209

Recent work on Beauville surfaces, structures and groups
Ben Fairbairn 225

Something for nothing: some consequences of the solution of the Tarski problems
Benjamin Fine, Anthony Gaglione, Gerhard Rosenberger & Dennis Spellman ... 242

The groups of projectivities in finite planes
Theo Grundhöfer ... 271

On the relation gap and relation lifting problem
Jens Harlander ... 278

Some results on products of finite subsets in groups
Marcel Herzog, Patrizia Longobardi & Mercede Maj ... 286

Formal languages and group theory
Sam A.M. Jones & Richard M. Thomas ... 306

On the Castelnuovo-Mumford regularity of the cohomology of fusion systems and of the Hochschild cohomology of block algebras
Radha Kessar & Markus Linckelmann ... 324

Recent advances on torsion subgroups of integral group rings
Wolfgang Kimmerle & Alexander Konovalov ... 331

On finite groups with small prime spectrum
Anatoly S. Kondratiev & Igor V. Khramtsov ... 348

Solvability criteria for finite loops and groups
Emma Leppälä ... 360

The rational subset membership problem for groups: a survey
Markus Lohrey ... 368

A survey of Milnor laws
Olga Macedońska ... 390

Capable p-groups
Arturo Magidin & Robert Fitzgerald Morse ... 399

On the normal structure of a finite group with restrictions on the maximal subgroups
N.V. Maslova & D.O. Revin ... 428

Certain monomial characters and their normal constituents
Gabriel Navarro & Carolina Vallejo ... 436

Recognition of finite quasi-simple groups by the degrees of their irreducible representations
Hung Ngoc Nguyen & Hung P. Tong-Viet 439

Generalized Baumslag-Solitar groups: a survey of recent progress
Derek J.S. Robinson 457

Zeta functions of groups and rings – recent developments
Christopher Voll 469

INTRODUCTION

Groups St Andrews 2013 was held at the University of St Andrews from 3rd August to 11th August 2013. This was the ninth in the series of Groups St Andrews group theory conferences organised by Colin Campbell and Edmund Robertson of the University of St Andrews. There were just under 200 mathematicians from over 20 countries involved in the meeting as well as some family members and partners. The Scientific Organising Committee of Groups St Andrews 2013 (all from St Andrews) was Colin Campbell, Max Neunhöffer, Martyn Quick, Edmund Robertson and Colva Roney-Dougal.

This time the academic business of the conference ran for seven days from Sunday 4th August to Saturday 10th August. Four main speakers delivered four talks each, surveying areas of contemporary development in group theory and related areas; Emmanuel Breuillard (Université Paris Sud 11), Martin Liebeck (Imperial College), Alan Reid (University of Texas), and Karen Vogtmann (Cornell University). There were five invited speakers delivering one-hour plenary talks: Peter Cameron (St Andrews), Radha Kessar (City University, London), Markus Lohrey (Universität Leipzig), Derek Robinson (University of Illinois at Urbana-Champaign) and Christopher Voll (University of Bielefeld). In addition there were nearly 100 contributed short talks from the delegates.

In the evenings throughout the conference there was an extensive social programme. The main conference outing was to the Royal Burgh of Falkland either to visit Falkland Palace or, as it turned out, to go on an adventure walk in the Lomond Hills. Other highlights of the social programme were a whisky tasting evening, a musical evening and the conference banquet. Once again The Daily Group Theorist was a nice feature of the conference. We thank the various editors of this, by now traditional, publication.

The support of the two main United Kingdom mathematics societies, the Edinburgh Mathematical Society and the London Mathematical Society has, once again, been an important factor in the success of these conferences. As well as supporting some of the expenses of the main speakers, the grants from these societies were used to support postgraduate students and also participants from Scheme 5 and fSU countries.

Once again all the main speakers have written substantial articles for these Proceedings. The majority of the other papers are of a survey nature. Regretably we have been limited to one volume so that, even more than has been the case in the past, we have been forced to exclude many worthwhile papers.

We would like to thank Martyn Quick and Colva Roney-Dougal not only for their editorial assistance with these Proceedings but also, along with Max Neunhöffer, for all their hard work in organising the conference.

CMC, EFR

APPROXIMATE SUBGROUPS AND SUPER-STRONG APPROXIMATION

EMMANUEL BREUILLARD

Laboratoire de Mathématiques, Bâtiment 425, Université Paris Sud 11, 91405 Orsay, France
Email: emmanuel.breuillard@math.u-psud.fr

Abstract

Surveying some of the recent developments on approximate subgroups and super-strong approximation for thin groups, we describe the Bourgain-Gamburd method for establishing spectral gaps for finite groups and the proof of the classification of approximate subgroups of semisimple algebraic groups over finite fields. We then give a proof of the super-strong approximation for mod p quotients via random matrix products and a quantitative version of strong approximation. Some applications to the group sieve are also presented. These notes are based on a series of lectures given at the 2013 Groups St Andrews meeting.

1 Introduction

In the early 1980's Matthews-Vaserstein-Weisfeiler [69], and then Nori [72] and Weisfeiler [101] (independently) proved the following theorem:

Theorem 1.1 (Strong-approximation theorem) *Suppose $\mathbb{G}$ is a connected, simply connected, semisimple algebraic group defined over $\mathbb{Q}$, and let $\Gamma \leqslant \mathbb{G}(\mathbb{Q})$ be a finitely generated Zariski-dense subgroup. Then for all sufficiently large prime numbers p, the reduction Γ_p of Γ is equal to $\mathbb{G}_p(\mathbb{F}_p)$.*

For example, if $\Gamma \leqslant \mathrm{SL}_n(\mathbb{Z})$ is a finitely generated Zariski dense subgroup, then $\Gamma_p = \mathrm{SL}_n(\mathbb{Z}/p\mathbb{Z})$ for all large enough prime numbers p. When p is large enough, the algebraic group $\mathbb{G}$ (viewed as a closed subgroup of some GL_n) admits a smooth reduction defined over $\mathbb{F}_p$, which we denote by $\mathbb{G}_p$. Since Γ is finitely generated, there are finitely many primes $p_1, \ldots, p_k$ (appearing in the denominators of the matrix entries of S) such that Γ belongs to $\mathbb{G}(\mathbb{Z}[1/p_1, \ldots, 1/p_k]) := \mathbb{G} \cap \mathrm{GL}_n(\mathbb{Z}[1/p_1, \ldots, 1/p_k])$, and the reduction modulo p map is well-defined on this subgroup if p is large enough.

The result fails if $\mathbb{G}$ is not simply connected (e.g., the image of $\mathrm{SL}_2(\mathbb{Z})$ in $\mathrm{PGL}_2(\mathbb{F}_p)$ has index 2 when $p > 2$). However every connected absolutely almost simple algebraic group admits a simply connected finite cover to which we can lift Γ and apply the theorem. This yields that $[\mathbb{G}_p(\mathbb{F}_p) : \Gamma_p]$ is nevertheless always bounded (for p large) by a constant depending only on $\mathbb{G}$ (one can take $1 + \mathrm{rank}(\mathbb{G})$, see [72, Remark 3.6]).

A similar result holds for groups defined over number fields instead of $\mathbb{Q}$. Its proof reduces to the case of $\mathbb{Q}$ by suitable restriction of scalars. See Remark 6.4 below (see also [101]).

That the result holds when Γ is an S-arithmetic group $\Gamma = \mathbb{G}(\mathbb{Z}[1/p_1, \ldots, 1/p_m])$ was known much earlier by work of Kneser [49] and Platonov [74] in particular. See [75, Chapter 7]) and [82].

Theorem 1.1 is then of particular interest when the group Γ is not a full S-arithmetic subgroup of $\mathbb{G}$ but has infinite index in one of them, while still remaining Zariski dense in $\mathbb{G}$ (S-arithmetic subgroups are Zariski dense by the Borel density theorem). Such a group is called a *thin subgroup* of $\mathbb{G}$ in recent terminology due to Peter Sarnak [91].

What we call *super-strong approximation* is the fact stated in Theorem 1.2 below that Γ not only surjects onto $\mathbb{G}_p(\mathbb{F}_p)$ for p large but that the associated Cayley graphs of $\mathbb{G}_p(\mathbb{F}_p)$ form a *family of expanders*. The goal of these notes is to give a proof of this fact, give some applications, and introduce the reader to the various techniques used in the proof.

It is of course not the purpose of this survey to give a complete introduction to expander graphs and for that matter we refer the reader to the many sources on the subject starting with Lubotzky's monograph [61] and survey [63] (see also [38] and [51, 96, 10]). Let us simply recall that to every finite k-regular graph $\mathcal{G}$ is associated a combinatorial Laplace operator acting on the (finite dimensional) space of functions on the vertices of the graph. It is defined by the formula

$$\Delta f(x) = f(x) - \frac{1}{k}\sum_{y\sim x} f(y),$$

where $y \sim x$ is a vertex connected to x by an edge. This operator is symmetric and non-negative. Its eigenvalues are real and non-negative. The eigenvalue 0 comes with multiplicity one if the graph is connected and the first nonzero eigenvalue is denoted by $\lambda_1(\mathcal{G})$ and satisfies:

$$\lambda_1(\mathcal{G}) = \inf\{\langle \Delta f, f\rangle, \|f\|_2 = 1, \sum_x f(x) = 0\}. \tag{1.1}$$

An infinite family of k-regular graphs $(\mathcal{G}_n)_{n\geqslant 1}$ is said to be a *family of expanders* if there is $\varepsilon > 0$ such that for all $n \geqslant 1$,

$$\lambda_1(\mathcal{G}_n) > \varepsilon.$$

We are now in a position to state the following strengthening of Theorem 1.1.

Theorem 1.2 (Super-strong approximation) *Suppose $\mathbb{G}$ is a connected, simply connected, semi-simple algebraic group defined over $\mathbb{Q}$, and let $\Gamma \leqslant \mathbb{G}(\mathbb{Q})$ be a Zariski-dense subgroup generated by a finite set S. Then there is $\varepsilon = \varepsilon(S) > 0$ such that for all large enough prime numbers p, the reduction Γ_p of Γ is equal to $\mathbb{G}_p(\mathbb{F}_p)$ and the associated Cayley graph $\mathrm{Cay}(\mathbb{G}_p(\mathbb{F}_p), S_p)$ is an ε-expander.*

Here S_p is the image of S by reduction modulo p. As before, the result also holds if $\mathbb{G}$ is not assumed to be simply connected, but Γ_p may then only be a subgroup of $\mathbb{G}_p(\mathbb{F}_p)$ whose index is nevertheless bounded independently of p, while $\mathrm{Cay}(\Gamma_p, S_p)$ remains an ε-expander.

This theorem is a special case of a result due to Salehi-Golsefidy and Varjú [87], which asserts that the conclusion also holds for quotient modulo a square free integer and even when the connected algebraic group $\mathbb{G}$ is only assumed to be perfect. Their proof follows the so-called Bourgain-Gamburd expansion machine, which can

be implemented in this context in part thanks to the recent results on approximate subgroups of linear groups due to Pyber-Szabó [80] and Breuillard-Green-Tao [19].

In these notes we describe the Bourgain-Gamburd method as well as the above mentioned results on approximate subgroups and finally give a complete proof of Theorem 1.2 (i.e., of super-strong approximation for mod p quotients) following a somewhat alternate route than in [87] by use of random matrix products [15].

1.1 The Lubotzky alternative and its expander version

One can formulate a version of the strong approximation theorem, which is valid for every finitely generated subgroup of $\mathrm{GL}_d(k)$, where k is an arbitrary field of characteristic zero (one can also deal with the positive characteristic case thanks to the work of Pink [73], however no super-strong version is known in positive characteristic thus far). When the group $\Gamma = \langle S \rangle$ we start with is non virtually solvable, one can show that there is a non trivial connected and simply connected semisimple algebraic group G defined over $\mathbb{Q}$ and a group homomorphism from a finite index subgroup of Γ into $\mathbb{G}(\mathbb{Q})$ with a Zariski-dense image (see [68, Prop. 16.4.13] and the discussion that follows). This allows to then apply the strong-approximation theorem 1.1 and deduce that Γ_0 admits $\mathbb{G}_p(\mathbb{F}_p)$ as a quotient for almost all p.

This information was used in a key way by Lubotzky and Mann in their work on subgroup growth [64]. For this version of strong approximation, called *the Lubotzky alternative*, we refer the reader to the notes devoted to it and its various refinements in the book by Lubotzky and Segal on subgroup growth ([68, 16.4.12], see also [48]). Strengthened by the super-strong approximation theorem, this gives the following statement:

Theorem 1.3 (Lubotzky super-alternative) *Let S be a finite symmetric subset of $\mathrm{GL}_d(k)$, where k is a field of characteristic zero. Then the subgroup $\Gamma = \langle S \rangle$ generated by S contains a subgroup Γ_0 whose index m in Γ is finite and bounded in terms of d only, such that*

- *either the subgroup Γ_0 is solvable,*
- *or there is a connected, simply connected, semisimple algebraic group $\mathbb{G}$ defined over $\mathbb{Q}$, such that for all large enough primes $p \in \mathbb{N}$, there is a surjective group homomorphism ρ_p from Γ_0 to $\mathbb{G}_p(\mathbb{F}_p)$ such that the Cayley graph $\mathrm{Cay}(\mathbb{G}_p(\mathbb{F}_p), \rho_p(S_0))$ is an ε-expander, for some $\varepsilon > 0$ independent of p, where S_0 is a subset of S^{2m} generating Γ_0.*

Note that given a group Γ generated by a symmetric set S, then every subgroup of finite index Γ_0 is finitely generated by a symmetric subset contained in S^{2m-1}, if m is the index of Γ_0 in Γ (e.g., see [19, Lemma C.1]).

A version of Theorem 1.3 for a bounded number of primes is also true: given large enough distinct primes $p_1, \ldots, p_k$, the Cayley graphs $\mathrm{Cay}(G(\mathbb{F}_{p_1}) \times \ldots \times G(\mathbb{F}_{p_k}), (\rho_{p_1} \times \ldots \times \rho_{p_k})(S))$ are ε-expanders for a uniform $\varepsilon > 0$ independent of the number of primes k. We will prove this stronger version only with an ε depending on k (but not on the choice of k primes). See Theorem 6.3 below. One needs the works of Varjú [100] and Salehi-Golsefidy-Varjú [87] to get this uniformity in the number of primes, but the proof is rather more involved. Note that at any case ε depends on S and it is an

open question whether this dependence can be removed (see [16] for partial results in this direction).

1.2 The group sieve method

Knowing that the finite quotients Cayley graphs are expanders is a very useful information for a number of applications to group theory and number theory, in particular it is the basis of the so-called Group Sieve, pioneered by Kowalski [52, 53], Rivin [83], and Lubotzky-Meiri [65, 66] and of the Affine Sieve of Bourgain-Gamburd-Sarnak [7]. See [50] and [55] for two nice expositions.

Roughly speaking, the expander property allows one to give very good bounds on the various error terms that appear when sieving modulo primes. In these notes, we will give a general statement, *the group sieve lemma* (Lemma 7.3 below), due to Lubotzky and Meiri, which allows to show that a subset Z of a given finitely generated linear group is exponentially small, provided its reduction modulo p does not occupy too large a subset of the quotient group for many primes p. For this version of the group sieve, expansion for pairs of primes is sufficient (i.e., we need that $G(\mathbb{F}_{p_1})\times G(\mathbb{F}_{p_2})$ expands for $p_1 \neq p_2$), so our version of the Lubotzky super-alternative above will be enough. Expansion for all square free moduli is necessary however, and sometimes crucial, in other situations, such as in the Affine Sieve pioneered by Bourgain-Gamburd-Sarnak [7] and further developed by Salehi-Golsefidy-Sarnak [86], Bourgain and Kontorovich [9] and others.

The conclusion of the super-strong approximation theorem (Theorem 1.2) can be reformulated in the following way: there is $\varepsilon > 0$ depending only on the generating set S such that for every real valued function f on the group $\mathbb{G}_p(\mathbb{F}_p)$, such that $\sum_{x\in\mathbb{G}_p(\mathbb{F}_p)} f(x) = 0$ and $\|f\|^2_{\ell^2} = \sum_{x\in\mathbb{G}_p(\mathbb{F}_p)} |f(x)|^2 = 1$,

$$\langle \Delta f, f\rangle > \varepsilon,$$

where

$$\langle \Delta f, f\rangle = \frac{1}{2k}\sum_{s\in S} \|s\cdot f - f\|^2_{\ell^2} = \frac{1}{2k}\sum_{s\in S}\sum_{x\in\mathbb{G}_p(\mathbb{F}_p)} |f(s^{-1}x) - f(x)|^2.$$

Let $S_p = \{s_1, \dots, s_k\}$ be the image of S under the reduction modulo p map and μ_{S_p} be the uniform probability measure on S_p, assigning equal mass $1/k$ ($= 1/|S|$ for p large enough) to each element of S_p.

$$\mu_{S_p} := \frac{1}{k}(\delta_{s_1} + \cdots + \delta_{s_k})$$

Note that $\mu_{S_p} = Id - \Delta$ as operators on $\ell^2(\mathbb{G}_p(\mathbb{F}_p))$, and hence its operator norm on $\ell^2_0(\mathbb{G}_p(\mathbb{F}_p))$, the orthogonal of constants, satisfies:

$$\|\mu_{S_p}|_{\ell^2_0}\| < 1 - \varepsilon$$

It is in this form that the theorem is used in its applications to the group sieve method. For example it allows Lubotzky and Meiri [65] to establish the following result about the scarcity of proper powers in non virtually solvable linear groups. A group element is called a proper power if it is of the form g^n for some integer $n \geqslant 2$ and some other group element g (from the same group).

Theorem 1.4 (Lubotzky-Meiri [65]) *Let $\Gamma \leqslant \mathrm{GL}_d(\mathbb{C})$ be a finitely generated subgroup and let μ_S be the uniform probability measure on a finite symmetric generating S. Assume that Γ is not virtually solvable. Then the set $\mathcal{P}_\Gamma$ of proper powers in Γ is exponentially small in the sense that there is $c = c(S) > 0$ such that for every $n \in \mathbb{N}$,*

$$\mu_S^n(\mathcal{P}_\Gamma) \leqslant e^{-cn}.$$

Here μ_S^n is the n-th convolution power of the probability measure μ_S on Γ. Equivalently, it is the distribution at time n of the simple random walk starting at the identity on the associated Cayley graph $\mathrm{Cay}(\Gamma, S)$. Or more explicitly:

$$\mu_S^n(\mathcal{P}_\Gamma) = \mathbb{P}_{w \in W_{n,k}}(\mathcal{P}_\Gamma) := \frac{|\{w, |w| = n, \overline{w} \in \mathcal{P}_\Gamma\}|}{|\{w, |w| = n\}|},$$

where $W_{n,k}$ is the set of (non reduced!) words w of length $|w| = n$ in the formal alphabet made of letters from the set S, and $\overline{w}$ its value as a group element when computed inside Γ. One can analogously count reduced words of length n in the free group and get the same result, but we note in passing that obtaining a result of this kind for the average with respect to the word metric on Γ induced by S seems out of reach at the moment, because little is known about the balls for the word metric on a group of exponential growth.

1.3 On the proof of the super-strong approximation theorem

Theorem 1.2 was first proved in the special case of subgroups of $\mathrm{SL}_2(\mathbb{Z})$ in a remarkable breakthrough by Bourgain and Gamburd [5]. They deduced the expansion by showing that the simple random walk on the finite quotient $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ must equidistribute very fast, indeed after only $O(\log p)$ steps. In doing so they reversed the traditional way of looking at things: traditionally spectral gaps estimates were proven by other methods (e.g., representation theory, property (T), etc.) and were then used to prove fast equidistribution of random walks. Bourgain and Gamburd reversed this order, first proving equidistribution and then deducing the gap (see Proposition 3.3 below for the equivalence between spectral gap and fast equidistribution).

This idea can be traced back to the seminal work of Sarnak and Xue [92], which gave a new, softer, approach toward Selberg's 3/16 theorem (i.e., the first eigenvalue of the Laplace operator on quotients of the hyperbolic plane by congruence subgroups of $\mathrm{SL}(2, \mathbb{Z})$ is at least 3/16, see [93]). They exploited, via the trace formula, the high multiplicity of the spectrum coming from the $(p-1)/2$ lower bound on the dimension of the smallest non trivial complex representation of $\mathrm{SL}_2(\mathbb{F}_p)$ (this bound goes back to Frobenius) and a soft combinatorial upper bound on the number of lattice points in a ball of radius roughly $\log p$. We refer the reader to the expository papers of P. Sarnak [90, 89], where this method and its history (in particular the role of Bernstein and Kazhdan) is described.

In his thesis [29] Gamburd pursued this method and established the first spectral gap result valid for thin groups: he showed that if a finitely generated subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is large enough in the sense that the Hausdorff dimension of its limit set on $\mathbb{P}^1(\mathbb{R})$ is at least $\frac{5}{6}$, then the spectrum of the associated (infinite volume) quotients of the hyperbolic plane modulo the congruence subgroups $\Gamma_p := \Gamma \cap \ker(\mathrm{SL}_2(\mathbb{Z}) \to$

$\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$) admits a uniform lower bound independent of p. In turn the resulting Cayley graphs of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ are expander graphs.

Bourgain and Gamburd [5] pushed the method even further to implement it for all Zariski-dense subgroups of $\mathrm{SL}_2(\mathbb{Z})$ with no restriction on the limit set. The structure of their proof retained the same patterns, playing the high multiplicity lower bound against a combinatorial upper bound via the trace formula applied to convolution powers of a fixed probability measure on the generating set. Achieving this combinatorial upper bound is the gist of their work: they brought in an important graph theoretic result (the Balog-Szemerédi-Gowers lemma, a parent of the celebrated Szemerédi regularity lemma) revisited in this context by Tao [97] to show that convolution powers of probability measures decay in ℓ^2 norm (the so-called ℓ^2-flattening) unless the measure charges significantly a certain approximate subgroup. That there exists no interesting approximate subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$ was established for this purpose by Helfgott [36]. The combinatorial upper bound (on the probability of return to the identity of the simple random walk at time roughly $\log p$), and hence the spectral gap, then reduces to establishing a certain non concentration estimate on subgroups for random walks on $\mathrm{SL}_2(\mathbb{Z})$ (see Theorem 5.1), which in this case can easily be deduced from Kesten's theorem [47].

This new method became known as the *Bourgain-Gamburd expansion machine* (see, e.g., the papers [20, 22] as well as the forthcoming book [96]). Its scope goes beyond $\mathrm{SL}_2(\mathbb{F}_p)$ and, quite remarkably, it can potentially be applied to any finite group (see Proposition 3.1 for a precise formulation of the method and its ingredients). It was understood early on that the scheme of the proof in [5] was general enough that it could be made to work in the general setting of Theorem 1.2, provided one could establish each step in the right generality. The bounds on the dimension of complex representations are well-known thanks to classical work of Landazuri-Seitz [57]. The graph theoretic lemma needs no modification in the general setting. The remaining two items however require deeper consideration. The classification of approximate groups, first established by Helfgott for $\mathrm{SL}_2(\mathbb{F}_p)$ and $\mathrm{SL}_3(\mathbb{F}_p)$, was finally completed in the general case by Pyber and Szabó [80] and independently by Breuillard-Green-Tao [19]. Regarding the upper bounds on the probability of hitting a subgroup, there are two known ways to achieve them. The first is to use the theory of random matrix products, and this was done in subsequent work of Bourgain-Gamburd [6], but only in the special case of subgroups of $\mathrm{SL}_n(\mathbb{Z})$, because the estimates from the theory of random matrix products required to deal with the general case were lacking. The second consists in applying a ping-pong argument akin to the proof of the Tits alternative [99], and this was performed by Varjú in his thesis [100] and subsequently by Salehi-Golsefidy and Varjú in their joint work [87], in which they establish Theorem 1.2 in full generality.

In the remainder of these notes we will prove Theorem 1.2 following each of these steps very closely. The only novelty in our proof lies in the last step: thanks to [15], we now understand how to use random matrix products to prove in the desired generality the required upper bounds for the probability of hitting a subgroup (the non-concentration estimates). This approach is somewhat more direct than the one taken by Salehi-Golsefidy and Varjú in [87], and it is very close to what Green, Tao and I had in mind, when we announced a proof of Theorem 1.2 in [18, Theorem 7.3]

in the special case of absolutely simple groups over $\mathbb{Z}$, but never came to the point of writing it up in full.

As already mentioned Salehi-Golsefidy and Varjú [87] actually proved a strong version of Theorem 1.2 showing the expansion property also for the quotients modulo a square free integer, and assuming only that $\mathbb{G}$ is perfect (which is also a necessary condition for expansion). See Theorem 6.5 below. That strong version is crucial for certain applications to sieving in orbits (à la Bourgain-Gamburd-Sarnak [7]), but its proof is much more involved. Often it is enough to have Theorem 1.2, or its extension to two or a bounded number of primes, which is not more costly. That will be the case for the applications presented in this paper. This, I thought, was enough justification for writing a complete proof of super-strong approximation for prime moduli in one place.

1.4 Outline of the article

In Section 2 we present a proof of the strong approximation theorem of Matthews, Vassertein and Weisfeiler following Nori's proof. Our treatment yields a quantitative version in the sense that it gives a upper bound on the first p for which the surjectivity of the reduction mod p holds in terms of the height of the generating set. Section 3 is devoted to the Bourgain-Gamburd machine: we state very general conditions on the Cayley graph of an arbitrary finite group that are sufficient to establish a spectral gap. Section 4 is devoted to approximate subgroups of linear groups over finite fields. We prove there the theorem of Pyber-Szabó and Breuillard-Green-Tao. In Section 5 we discuss random matrix products and a general non-concentration on subgroups result for random walks on linear groups. Finally in Section 6 we combine the results of the preceding three sections to complete the proof of the super-strong approximation theorem in the case of mod p quotients (Theorems 1.2 and 6.3). The final section is devoted to applications to the group sieve method and results of Aoun, Jouve-Kowalski-Zywina, Lubotzky-Meiri, Lubotzky-Rosenzweig and Prasad-Rapinchuk on generic properties elements in non virtually solvable linear groups.

2 Nori's theorem and a quantitative version of strong approximation

It was Matthews, Vaserstein and Weisfeiler [69] who first proved the strong approximation theorem for Zariski-dense subgroups, i.e., Theorem 1.1, in the case when G is absolutely simple. Their proof made use of the (brand new at the time) classification of finite simple groups. Another, classification-free proof was found roughly at the same time and independently by M. Nori, yielding also the case G semisimple, as a consequence of the following general result proved in [72].

Theorem 2.1 (Nori [72]) *Let H be a subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$, and H^+ the subgroup generated by its elements of order p. If p is larger than some constant $c(n)$ depending only on n, then there is a connected algebraic subgroup $\widetilde{H}$ of GL_n defined over $\mathbb{F}_p$ such that H^+ coincides with $\widetilde{H}(\mathbb{F}_p)^+$. Moreover there is a normal abelian subgroup $A \leqslant H$ such that $[H : AH^+]$ is bounded in terms of n only.*

Observe that if $p \geqslant n$, then elements of order p in $\mathrm{GL}_n(\mathbb{F}_p)$ are precisely the unipotent matrices: indeed $x^p = 1$ is equivalent to $(x-1)^p = 0$ for $x \in \mathrm{GL}_n(\mathbb{F}_p)$ and

hence to $x = 1 + n$, where n is a nilpotent matrix. As Nori explains in [72, Remark 3.6.], the index of $\widetilde{H}(\mathbb{F}_p)^+$ in $\widetilde{H}(\mathbb{F}_p)$ is bounded by a function of n only. So the meaning of Nori's theorem is that finite subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$ generated by elements of order p are essentially algebraic subgroups, if $p > c(n)$.

The key feature of Nori's theorem is that no assumption whatsoever is made on the subgroup H. Hence Nori's theorem can be seen as a description of *arbitrary* subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$. It can be viewed as complementing the celebrated theorem of Camille Jordan [44] on finite subgroups of $\mathrm{GL}_n(K)$ whose order is prime to the characteristic of the field K: such a group admits an abelian subgroup whose index is bounded by some function of n only. Nori's theorem explains what happens when the characteristic divides the order of the finite group: recall that a finite group has an element of prime order p if and only if its order is a multiple of p (Cauchy's theorem).

Jordan's theorem is usually quoted for subgroups of $\mathrm{GL}_n(\mathbb{C})$, but this stronger version can be derived easily by lifting the group to $\mathbb{C}$ (see [72, Theorem C]). In fact Jordan had already proved this stronger version in his original paper: his proof is purely algebraic and applies to any finite subgroup of $\mathrm{GL}_n(K)$ all of whose elements are semisimple (or equivalently to finite subgroups without a non trivial unipotent element), where K is any algebraically closed field (see [11] for a discussion).

Textbooks presenting Jordan's theorem usually give a different, more geometric treatment, due to Frobenius, Bieberbach and Blichfeldt. Jordan's own argument seems to have been forgotten for more than a hundred years until Larsen and Pink [59] rediscovered it and generalized it considerably to obtain a classification of all finite subgroups of GL_d in every characteristic. The Larsen-Pink theorem is more general than Nori's result stated above in that it applies to finite subgroups of GL_d regardless of the field and the size of the characteristic. We will comment on the Larsen-Pink theorem further below, when we discuss approximate subgroups of linear groups. The proof of the Larsen-Pink theorem, which by the way is also independent of the classification of finite simple groups, plays a key role in the structure theorem for approximate subgroups of linear groups (see Theorem 4.5 below).

For the applications to strong and super-strong approximation, we will not need the full force of Theorem 2.1 above. Rather the following important special case will be sufficient.

Theorem 2.2 (Sufficiently Zariski-dense subgroups) *There is $M = M(d)$ such that the following holds. Let $p > M$ be a prime number and $\mathbb{G}_p \leqslant \mathrm{GL}_d$ be a semisimple simply connected algebraic group defined over $\mathbb{F}_p$. If a subgroup $H \leqslant \mathbb{G}_p(\mathbb{F}_p)$ is not contained in a proper algebraic subgroup of $\mathbb{G}_p$ of complexity at most M, then it must be equal to $\mathbb{G}_p(\mathbb{F}_p)$.*

We say informally that a closed algebraic subvariety of GL_d has complexity at most M if it can be defined as the vanishing locus of a finite set of polynomials such that the sum of their degrees in each variable is at most M. See [19] for background on this notion. It is particularly useful in positive characteristic: saying that a finite subgroup of $\mathrm{GL}_d(\overline{\mathbb{F}_p})$ is algebraic is meaningless, because every finite subgroup is an algebraic subset with several (possibly many) irreducible components. However putting a bound on the complexity forces a bound on the number of irreducible components [19, Lemma A.4] and hence restricts the class of finite subgroups drastically

and leads to interesting statements, such as the above.

We now sketch Nori's proof of Theorem 2.2. A similar argument is due to Gabber, see [46, Thm 12.4.1]. Pushing this idea a bit further allows Nori to also prove Theorem 2.1.

Proof (sketch) If H had no non trivial unipotent element, it would have an abelian subgroup of bounded index by Jordan's theorem. But this would violate the assumption that H is sufficiently Zariski-dense. So H contains a unipotent element, which we may write in the form $h = \exp \xi$, for some nilpotent matrix ξ. The $\mathbb{F}_p$-span V_H of all H-conjugates of ξ is invariant under the adjoint action of H. The assumption that H is sufficiently Zariski-dense implies that V_H must be the full $\mathbb{F}_p$-Lie algebra of $\mathbb{G}_p$ in $gl_d(\mathbb{F}_p)$. Pick unipotent elements $h_1, \dots, h_d \in H$ such that the corresponding ξ_i's form a basis of $\mathrm{Lie}(\mathbb{G}_p)$.

Now consider the map $\Phi : \mathbb{F}_p^{\dim \mathbb{G}} \to \mathbb{G}_p(\mathbb{F}_p)$, $(t_1, \dots, t_d) \mapsto h_1^{t_1} \cdot \ldots \cdot h_d^{t_d}$. Note that Φ is a polynomial map whose degree is bounded in terms of d only. Its image lies in H. We claim that there is a constant $c = c(d) > 0$ such that $|\operatorname{Im} \Phi| \geqslant cp^d$. Indeed, the Jacobian of Φ is not identically zero, so outside its vanishing locus (a proper subvariety, hence a subset of size $O(p^{d-1})$) the fibers of Φ are of bounded cardinality. This implies the desired bound.

Now since there are positive constants c_1, c_2 such that $c_1 p^d \leqslant |\mathbb{G}_p(\mathbb{F}_p)| \leqslant c_2 p^d$ (e.g., see [72, Lemma 3.5.]), we get that the index $[\mathbb{G}_p(\mathbb{F}_p) : H]$ is bounded. However since $\mathbb{G}$ is simply connected, $\mathbb{G}_p(\mathbb{F}_p)$ is an almost direct product of quasi-simple groups and thus has no subgroups of bounded index when p is large (Kneser-Tits for $\mathbb{F}_p$, see [75], see also Remark 3.4). Hence $H = \mathbb{G}_p(\mathbb{F}_p)$. □

Nori's proof of strong approximation (i.e., of Theorem 1.1) is based on Theorem 2.2 alone. We will explain this argument below. It turns out that this argument even yields a quantitative lower bound on the first prime number for which we can claim that $\Gamma_p = \mathbb{G}_p(\mathbb{F}_p)$ in terms of the height of the generating set of Γ. Namely:

Theorem 2.3 (Strong approximation, quantitative version)
Suppose $\mathbb{G} \leqslant \mathrm{GL}_d$ is a connected, simply connected, semisimple algebraic group defined over $\mathbb{Q}$. Then there are constants $p_0, C_0 \geqslant 1$ such that if $S \subset \mathbb{G}(\mathbb{Q})$ is a finite symmetric set generating a Zariski-dense subgroup $\Gamma = \langle S \rangle$ of $\mathbb{G}$, and M_S denotes the maximal height of an element of S, then for every prime number $p > \max\{p_0, M_S^{C_0}\}$, the reduction Γ_p of Γ is equal to $\mathbb{G}_p(\mathbb{F}_p)$.

Here the height $H(s)$ of an element $s \in \mathrm{GL}_d(\mathbb{Q})$ is defined naively as the maximum of the numerators and denominators appearing in the expressions of the matrix coefficients of s as irreducible fractions. The bound p_0 is related to the bound $c(n)$ from Nori's theorem and to p_M from Lemma 2.7 below. There is very little control on this bound in general (see [87, Appendix] for a discussion of this issue).

Several other proofs and extensions of Theorem 1.1 (to groups defined over number fields, to positive characteristic, etc.) have since been found. For those we refer the reader to the original articles, in particular [101], [72], [41], [73], and to the chapter on strong approximation in the recent book by Lubotzky and Segal [68] or in Nikolov's

lecture notes in [48, chapter II]. We also recommend reading Rapinchuk's recent survey [82], which gives a thorough overview of strong approximation.

We now pass to the derivation of Theorem 2.3 from Nori's theorem. First, we replace the naive height with another height, which is better suited for our purposes since it is sub-additive. Given $a \in \mathrm{GL}_d(\mathbb{Q})$, set

$$h(a) := \sum_{p,\infty} \log^+ \|a\|_p,$$

where the sum is over all prime numbers p as well as the infinite place ∞. Here $\log^+ := \max\{\log, 0\}$, and $\|a\|_p$ denotes $\max_{ij} |a_{ij}|_p$, the maximum p-adic absolute value of a matrix entry a_{ij} of a, while $\|a\|_\infty$ is the operator norm of a for the standard Euclidean norm on $\mathbb{R}^d$. The following is straightforward:

Lemma 2.4 (a) *The height $h(a)$ is sub-additive, i.e., for all $a, b \in \mathrm{GL}_d(\mathbb{Q})$,*

$$h(ab) \leqslant h(a) + h(b),$$

and (b) *it is comparable to the naive height $H(a)$, namely, for all a,*

$$H(a) \leqslant e^{h(a)} \leqslant d(H(a))^{d^2}.$$

We conclude that for all $a_1, \ldots, a_n \in \mathrm{GL}_d(\mathbb{Q})$,

$$H(a_1 \cdot \ldots \cdot a_n) \leqslant d^n (H(a_1) \cdot \ldots \cdot H(a_n))^{d^2} \tag{2.1}$$

Combined with the next lemma, this inequality allows us to assume, in the proof of Theorem 2.3 that Γ is generated by two elements, i.e., that $S := \{1, a^{\pm 1}, b^{\pm 1}\}$.

Lemma 2.5 (Reduction to 2 generators) *Let $\mathbb{G}$ be a semisimple algebraic group over $\mathbb{C}$. Then there is $c > 0$ such that given any finite symmetric subset $S \subset \mathbb{G}(\mathbb{C})$, with $1 \in S$, generating a Zariski dense subgroup of $\mathbb{G}$, the bounded power S^c contains two elements a, b which alone already generate a Zariski-dense subgroup.*

Proof This is Proposition 1.8. from [13]. The proof is fairly classical, and relies on Jordan's theorem and the Eskin-Mozes-Oh escape from subvarieties lemma (see, e.g., [19, Lemma 3.11]). □

Lemma 2.6 (Generating is an algebraic condition) *Let $\mathbb{G} \leqslant \mathrm{GL}_d$ be a semisimple algebraic group defined over $\mathbb{Q}$. There is a proper closed algebraic subvariety $\mathbf{X} \leqslant \mathbb{G} \times \mathbb{G}$ defined over $\mathbb{Q}$, whose points are precisely the pairs of elements in $\mathbb{G}$ which are contained in a proper algebraic subgroup of $\mathbb{G}$.*

Proof This is well-known (see, e.g., [35, Theorem 11.6]). We work over an algebraic closure of $\mathbb{Q}$ and show that $\mathbf{X}$ is a closed algebraic subset. Since $\mathbf{X}$ is invariant under Galois automorphisms, it will automatically be defined over $\mathbb{Q}$. We claim that there are finitely many absolutely irreducible finite dimensional non trivial modules of $\mathbb{G}$, say $\rho_1, \ldots, \rho_k$ such that a subgroup $\Gamma \leqslant \mathbb{G}$ is not Zariski-dense if and only if $\rho_i(\Gamma)$ fixes a line in the representation space V_i of ρ_i for some $i = 1, \ldots, k$. And this happens

if and only if $\rho_i(\Gamma)$ fixes a non trivial subspace of V_i for some $i = 1, \dots, k$. This last condition clearly forms an algebraic condition, because it is equivalent to saying that $\rho_i(\Gamma)$ does not span the ring of endomorphisms of V_i. Moreover the span of $\rho_i(\Gamma)$ is spanned by the $\rho_i(w(a,b))$'s for a bounded set of words w. So we indeed have an algebraic condition on the pair a, b. Finally $\mathbf{X}$ is proper, because every semisimple algebraic group can be generated by two elements (see, e.g., [56]).

To prove the claim, note that if $\mathbb{H}$ is a proper closed algebraic subgroup of $\mathbb{G}$, then either it is finite in projection to one of the simple factors of $\mathbb{G}$, or its Lie algebra is not preserved under the adjoint action of $\mathbb{G}$ on $\mathrm{Lie}(\mathbb{G})$. Let $j(d)$ the bound from Jordan's theorem, so that every finite subgroup of GL_d has a normal abelian subgroup of index at most $j(d)$. For each simple factor $\mathbb{G}_i$ pick an irreducible module whose dimension is larger than $j(d)$, so that no finite subgroup of $\mathbb{G}_i$ can act irreducibly on it. We thus have found finitely many irreducible modules, say $\pi_1, \dots, \pi_m$ of $\mathbb{G}$ with the property that if a subgroup acts irreducibly on each of them, it must be Zariski-dense. Adding to this list all the non trivial irreducible submodules of the wedge powers $\Lambda^* \pi_i$, we obtain the desired list of modules $\rho_1, \dots, \rho_k$. □

Now, reducing modulo a large prime p, we obtain:

Lemma 2.7 (Generating mod p) *With the assumptions of the previous lemma, there is $M_0 \geqslant 1$ such that for all $M \geqslant M_0$, there is $p_M > 0$ such that if $p > p_M$ is a prime number, the reduction of $\mathbf{X}$ mod p is a proper algebraic subvariety of $\mathbf{X}_p \leqslant \mathbb{G}_p \times \mathbb{G}_p$ defined over $\mathbb{F}_p$ whose points are precisely the pairs of elements in $\mathbb{G}_p$ which are contained in a proper algebraic subgroup of $\mathbb{G}_p$ of complexity at most M.*

Proof First observe that there is a bound M_0 such that every proper algebraic subgroup of $\mathbb{G}$ is contained in a proper algebraic subgroup of complexity at most M_0. This follows from the discussion in the proof of Lemma 2.6, since a proper algebraic subgroup will either stabilize a subalgebra of $\mathrm{Lie}(\mathbb{G})$ which is not an ideal, or will stabilize a proper subspace of some V_i. Each of these stabilizers have bounded complexity. Now to prove the lemma we argue by contradiction. If no such p_M can be found, there must be an infinite sequence of primes $p_i < p_{i+1}$ and pairs $(a_i, b_i) \in \mathbb{G}_{p_i}(\overline{\mathbb{F}_{p_i}})$ such that either for all i, $(a_i, b_i) \in \mathbf{X}_{p_i}$ and are not contained in a proper algebraic subgroup of $\mathbb{G}_{p_i}$ of complexity at most M, or for all i, $(a_i, b_i) \notin \mathbf{X}_{p_i}$ and are contained in a proper algebraic subgroup of $\mathbb{G}_{p_i}$ of complexity at most M. The ultraproduct of the $\mathbf{X}_{p_i}$ coincides with $\mathbf{X} \otimes_{\mathbb{Q}} K$, where K is the ultraproduct of the finite fields $\mathbb{F}_{p_i}$. This gives rise to a pair (a, b) in the associated ultraproduct, which, in the first case, belongs to $\mathbf{X}(K)$ and generates a Zariski-dense subgroup, and in the second case does not belong to $\mathbf{X}(K)$ and yet generates a subgroup contained in a proper algebraic subgroup of complexity at most M. In both cases we have a contradiction with the definition of $\mathbf{X}$ in Lemma 2.6. For more details on similar ultraproduct arguments, we refer the reader to the appendix of [19]. □

Now comes the point where Nori's theorem is used in the form of Corollary 2.2: when $\mathbb{G}_p$ is simply connected every subgroup of $\mathbb{G}_p(\mathbb{F}_p)$ which is not contained in an algebraic subgroup of bounded complexity must be all of $\mathbb{G}_p(\mathbb{F}_p)$.

We may then complete the proof of Theorem 2.3. Pick polynomial functions $(P_k)_{k=1,\dots,k_0}$, $P_k = P_k((a_{ij}, b_{ij}))$, in pairs of matrices (a, b) in GL_d, which generate the

radical ideal of polynomial functions vanishing on $\mathbf{X}$ in $\mathbb{G} \times \mathbb{G}$. We may assume that the P_k's have integer coefficients. If $S = \{1, a^{\pm 1}, b^{\pm 1}\} \subset \mathbb{G}(\mathbb{Q})$ generates a Zariski-dense subgroup of $\mathbb{G}$, then $(a, b) \notin \mathbf{X}$ and there must exist k such that $P_k(a, b) \neq 0$. We may bound the height of $P_k(a, b)$ in terms of the heights of a and b and the heights of the coefficients of P_k. Hence

$$H(P_k(a, b)) \leqslant O(H(a)H(b))^{O(1)} \leqslant (2M_S)^C,$$

for some constant C depending only on $\mathbb{G}$ and not on k, a, b, and where $M_S = \max\{H(a), H(b)\}$. This means that if $p > (2M_S)^C$, then $P_k(a, b)$ does not vanish modulo p. Now Lemma 2.6, combined with Nori's theorem (in the form of Corollary 2.2), tells us that if additionnally p is larger than a constant depending on $\mathbb{G}$ only, then the reduction mod p of the pair (a, b) generates all of $\mathbb{G}_p(\mathbb{F}_p)$ and we are done. This ends the proof of Theorem 2.3.

3 The Bourgain-Gamburd expansion machine

Bourgain and Gamburd, in their groundbreaking paper [5], came up with a new method to establish the expander property for Cayley graphs of finite groups. They applied it to prove Theorem 1.2 in the special case of subgroups of $\mathrm{SL}_2(\mathbb{Z})$, but their method is very general. We call it the *Bourgain-Gamburd expansion machine*. In this section we give an overview of this machine, suitable for the proof of Theorem 1.2 in full generality.

Let G_0 be a finite group, and $S_0 = \{s_1, \ldots, s_k\}$ be a symmetric generating set for G_0. As before we write:

$$\mu = \mu_{S_0} := \frac{1}{k}(\delta_{s_1} + \cdots + \delta_{s_k})$$

for the uniform probability measure on the set S, where δ_x is the Dirac mass at x. For us a probability measure on G_0 is the same thing as a function on G_0 taking non-negative values at each element of G_0 and summing to 1.

We write

$$\mu^n := \mu * \cdots * \mu$$

for the n-fold convolution power of μ with itself, where the convolution $\mu_1 * \mu_2$ of two functions $\mu_1, \mu_2 \colon G_0 \to \mathbb{R}^+$ is given by the formula

$$\mu_1 * \mu_2(g) := \sum_{x \in G_0} \mu_1(gx^{-1})\mu_2(x). \tag{3.1}$$

The function $x \mapsto \mu^n(x)$ is a probability measure describing the distribution of a random walk of length n starting at the identity in G_0 and with generators from S. In particular, if A is a subset of G_0,

$$\mu^n(A) = \mathbb{P}_{w \in W_{n,k}}(w(a_1, \ldots, a_k) \in A), \tag{3.2}$$

where $W_{n,k}$ is the space of all formal words (not necessarily reduced) on k generators of length exactly n. We can now state a version of the Bourgain-Gamburd machine, adapted from [22] and [100].

Proposition 3.1 (Bourgain-Gamburd machine) *Suppose that G_0 is a finite group, that $S_0 \subseteq G_0$ is a symmetric generating subset, and that there are constants $0 < \kappa, \beta < 1$ such that the following properties hold for every quotient G of G_0.*

(i) (High multiplicity). *For every faithful representation $\rho\colon G \to \mathrm{GL}_d(\mathbb{C})$ of G, $\dim \rho \geqslant |G|^{\beta}$;*

(ii) (Classification of Approximate Subgroups). *For every $\varepsilon > 0$, there is $\delta = \delta(\varepsilon)$, $0 < \delta < \varepsilon$, with the property that every $|G|^{\delta}$-approximate subgroup A of G, is either of size $|A| \geqslant |G|^{1-\varepsilon}$ or is contained in at most $[G : H]^{\varepsilon}/|G|^{\delta}$ left cosets of a subgroup $H \leqslant G$;*

(iii) (Non-concentration estimate). *Let S be the image of S_0 in G. There is some even number $n \leqslant \log |G|$ such that for all subgroups $H \leqslant G$,*

$$\mu_S^n(H) \leqslant [G : H]^{-\kappa}.$$

Then the first non zero eigenvalue of the Cayley graph $\mathrm{Cay}(G_0, S_0)$ satisfies

$$\lambda_1 \geqslant \beta \cdot e^{-C/\delta},$$

where $\delta := \delta(\varepsilon) > 0$ with $\varepsilon := \min\{\beta, \kappa\}/4$ and C is an absolute constant.

We will discuss approximate subgroups in the next section. It suffices for now to say that by definition, given a parameter $K \geqslant 1$, a *K-approximate subgroup* of G_0 is a finite symmetric set A containing 1 such that $AA \subset XA$ for some subset $X \subset G_0$ of size at most K.

Remark 3.2 We already observed that if the Cayley graph $\mathcal{G}(G_0, S_0)$ is an ε-expander, then so are all induced quotient Cayley graphs corresponding to a quotient group $G := G_0/H$, for any normal subgroup $H \leqslant G_0$. It is therefore very natural that the Assumptions (i) to (iii) are made on all quotients of G_0.

As mentioned earlier, Assumption (ii), the classification of approximate subgroups of G_0, and (iii), the nonconcentration estimate, really constitute the beefy parts of the proof of the expander property. They will be dealt with in the next sections. We also remark that (iii) is the only condition of the three that actually involves the set S. Finally we stress that the lower bound on λ_1 obtained here is independent of the size k of S.

An interesting feature of (iii) is that, unlike (i) and (ii), it is *necessary* in order to verify the expander property, because the simple random walk on an expander graph will equidistribute in logarithmic time. Indeed we have the following basic lemma (recall the definition of ε-expanders in (1.1) above).

Lemma 3.3 (Random walk characterization of expanders) *Let G_0 be a finite group and S_0 a symmetric generating subset not contained in a coset of a subgroup of index 2 of G_0.*

- *if the Cayley graph $\mathcal{G}(G_0, S_0)$ is an ε-expander, then there is $C = C_{\varepsilon} > 0$ such that for every $n \geqslant C \log |G_0|$,*

$$\max_{x \in G_0} \left| \mu_{S_0}^n(x) - \frac{1}{|G_0|} \right| \leqslant \frac{e^{-n/C}}{|G_0|^{10}}, \tag{3.3}$$

- *if* (3.3) *holds for some* $n \leqslant C \log |G_0|$, *and* $C > 20$, *then* $\mathcal{G}(G_0, S_0)$ *is an* ε*-expander, with* $\varepsilon = 10/C$.

Proof Let $T_\mu = 1 - \Delta$ be the operator $f \mapsto \mu * f$ on $\ell^2(G_0)$. To prove the second item, pick an eigenfunction f of the Laplacian with eigenvalue λ_1 and note that $\|(T_\mu^n - (1/|G_0|)Id)f\|_2 \leqslant \|f\|_2/|G|^{10}$ forcing $(1 - \lambda_1) \leqslant |G_0|^{-10/n}$. As for the first item, note that the left hand side of (3.3) is bounded by $\|T_\mu\|^n \leqslant \|T_\mu\|^{C_\varepsilon \log |G_0|} = 1/|G_0|^{-C_\varepsilon \log(1/\|T_\mu\|)}$. The assumption on S_0 and the fact that $\mathcal{G}$ is the Cayley graph of a group ensure that it is not bi-partite and that $\|T_\mu\| \leqslant e^{-c_\varepsilon}$, for some $c_\varepsilon > 0$ depending only on ε and $|S_0|$ (see [22, Prop. E.1]). The result follows with $C_\varepsilon = 10/c_\varepsilon$. □

To see that (iii) is necessary, simply note that $\mu^{nm}(H) \geqslant (\mu^n(H))^m$ and apply the first item in the above lemma to evaluate $\mu^{nm}(H)$ using some m between C_ε and $2C_\varepsilon$ say.

Remark 3.4 According to result of Landazuri-Seitz [57], Assumption (i) is always verified when G_0 is a simple or quasi-simple group of Lie type of bounded rank, with the parameter $\beta > 0$ depending only on the rank. See Prop. 6.1 below. Looking at the action by translation on $\ell^2(G_0/H)$, where H is an arbitrary subgroup of G_0, this implies that every proper subgroup of G_0 has index at least $|G_0|^c$ for some $c > 0$ depending only on the rank of G_0.

We now pass to the proof of Proposition 3.1. The following basic observation relates the eigenvalues of the Laplace operator Δ on the Cayley graph, with the probability of return to the identity of the simple random walk. Let $1 = \alpha_0 > \alpha_1 \geqslant \ldots \geqslant \alpha_{|G_0|-1}$ be the eigenvalues of the convolution operator

$$T_\mu : f \mapsto \mu * f$$

on $\ell^2(G_0)$. Since $T_\mu = T_{\mu_{S_0}} = Id - \Delta$, the first non trivial eigenvalue of Δ, is just $\lambda_1 = 1 - \alpha_1$.

Now observe that the eigenspace of T_μ corresponding to the eigenvalue α_1 is invariant under G_0 and thus forms a linear representation of G_0. Up to replacing G_0 with its image modulo of the kernel of this representation, and μ with the corresponding push-forward measure, we may assume that G_0 acts faithfully on this eigenspace. And hence, applying Assumption (i), that the dimension of this eigenspace is at least $|G_0|^\beta$.

Thus we seek a lower bound on $1 - \alpha_1$. For this, we write the following naive *trace formula*, which consists in expressing the trace of $T_{\mu^n} = T_\mu^n$ in two ways (this key idea is analogous to what is done in the context of discrete groups in Sarnak-Xue [92] and Gamburd [29]). Firstly:

$$tr(T_{\mu^n}) = \sum_{x \in G_0} \langle (T_\mu)^n \delta_x, \delta_x \rangle = |G_0| \langle (T_\mu)^n \delta_1, \delta_1 \rangle = |G_0| \mu^n(1),$$

where $\mu^n(1)$ is the value at the identity of the probability measure μ^n. Here δ_x denotes the Dirac mass at x and $\langle \cdot, \cdot \rangle$ the ℓ^2 scalar product on G_0. And secondly:

$$tr(T_{\mu^n}) = \alpha_0^n + \alpha_1^n + \ldots + \alpha_{|G_0|-1}^n.$$

We will now play the multiplicity lower bound on α_1 against the combinatorial upper bound on $\mu^n(1)$. Since α_1^n appears at least $|G_0|^\beta$ times in the above sum, discarding all other eigenvalues (note that n is even and hence $\alpha_i^n \geqslant 0$), we get the following:

Observation 1 If $\mu^n(1) \leqslant 1/|G_0|^{1-\beta/2}$ for some even integer $n \leqslant C_1 \log|G_0|$, then the first non trivial eigenvalue α_1 of T_μ satisfies

$$\alpha_1 \leqslant e^{-\beta/2C_1}.$$

Assumption (iii) only guarantees the existence of an even integer $n_0 \leqslant \log|G_0|$ such that $\mu^{n_0}(1) \leqslant 1/|G_0|^\kappa$ for some positive κ which may be smaller than $1-\beta/2$. So in order to conclude, we need to show that $\mu^n(1)$ will decay from $1/|G_0|^\kappa$ at time $n = n_0 \leqslant \log|G_0|$ to $1/|G_0|^{1-\beta/2}$ at a not much larger time $n = n_1 \leqslant C_1 \log|G_0|$ for some constant C_1 depending only on the constants at hand and not on the size of G_0.

Before going further, let us record the following simple remarks:

Remark 3.5 When n tends to infinity $\mu^n(1)$ converges to $1/|G_0|$, the uniform distribution on G_0.

Remark 3.6 Since μ is assumed symmetric,

$$\mu^{2n}(1) = \sum_{x\in G_0} \mu^n(x)\mu^n(x^{-1}) = \|\mu^n\|_2^2 \tag{3.4}$$

Remark 3.7 For every subgroup $H \leqslant G_0$, the sequence $\mu^{2n}(H)$ is non-increasing: indeed $\mu^{2n}(H) = \|f_{n,H}\|_2^2$, where $f_{n,H} : G_0/H \to \mathbb{R}$, $gH \mapsto \mu^n(gH)$, and $f_{n+1,H} = T_\mu f_{n,H}$, while T_μ is a contraction in ℓ^2.

The key ingredient in establishing this final decay of $\mu^n(1)$ from $1/|G_0|^\kappa$ to $1/|G_0|^{1-\beta/2}$ is the following ℓ^2-flattening lemma, due to Bourgain-Gamburd. It says in substance that the only reason why the convolution of a probability measure with itself would not decay in ℓ^2-norm is because it gave a lot of mass to (a coset of) an approximate subgroup.

Lemma 3.8 (ℓ^2-flattening lemma) *There is absolute constant $R > 0$ such that the following holds. Let $K \geqslant 2$ and $\nu : G_0 \to \mathbb{R}^+$ be a probability measure on a finite group G_0 which satisfies*

$$\|\nu * \nu\|_2 \geqslant \frac{1}{K}\|\nu\|_2,$$

where convolution is defined in (3.1)*. Then there is a K^R-approximate subgroup A of G_0 with*

$$K^{-R}\frac{1}{\|\nu\|_2^2} \leqslant |A| \leqslant K^R\frac{1}{\|\nu\|_2^2}$$

and such that for each $x \in A$,

$$\nu * \nu^{-1}(x) \geqslant \frac{1}{K^R|A|}.$$

Here $\|\nu\|_2$ denotes the ℓ^2 norm on G_0, i.e., $\|\nu\|_2^2 := \sum_{x\in G_0} \nu(x)^2$, and ν^{-1} denotes the symmetric of ν, namely the probability measure $\nu^{-1}(x) := \nu(x^{-1})$. Observe that the last condition implies immediately that there is $g \in G_0$ such that $\nu(Ag) \geqslant 1/K^R$.

Proof The proof of the ℓ^2-flattening lemma is really the core of the Bourgain-Gamburd machine. It is derived from a powerful combinatorial tool, the Balog-Szemerédi-Gowers lemma (see Lemma 4.4 below), due in this context to Tao ([97], [98, §2.5, 2.7]), but which originates from the work of Balog-Szemerédi [3] and from Szemerédi's celebrated regularity lemma for large graphs. A simple derivation of the above ℓ^2-flattening lemma, based on Tao's version of the Balog-Szemerédi-Gowers lemma, namely Lemma 4.4 below, is given by Varjú in [100, Lemma 15] and we refer the reader to it for the details. He can also consult [22, Appendix A]. The basic idea is to decompose ν into approximate level sets $\nu = \sum_i 1_{A_i}\nu$, where $A_i = \{x \in G_0; 2^{i-1}\|\nu\|_2^2 < \nu(x) \leqslant 2^i\|\nu\|_2^2\}$ and show that for some suitable pair A_{i_1}, A_{i_2} the number of collisions $\|1_{A_{i_1}} * 1_{A_{i_2}}\|_2^2$ is large enough to be able to apply Lemma 4.4. □

Applying this lemma to a symmetric measure ν with $K = |G_0|^{\delta/R}$, we obtain the following direct consequence:

Corollary 3.9 *Let $0 < \delta, \varepsilon \leqslant 1/4$ and let ν be a symmetric probability measure on a finite group G_0 such that $|G_0|^{2\varepsilon} \leqslant 1/\|\nu\|_2^2 \leqslant |G_0|^{1-2\varepsilon}$. Then*

$$\|\nu * \nu\|_2 \leqslant \frac{1}{|G_0|^{\delta/R}}\|\nu\|_2,$$

unless there is a $|G_0|^\delta$-approximate subgroup A of G_0 with $|G_0|^\varepsilon \leqslant |A| \leqslant |G_0|^{1-\varepsilon}$ such that $\nu(gA) \geqslant 1/|G_0|^\delta$ for some $g \in G_0$.

Here R is the absolute constant from Lemma 3.8. We are going to apply this corollary several times to the convolution powers μ^n with even n between $\log|G_0|$ and $C_1 \log|G_0|$. After only a bounded number of applications of the corollary, $\mu^n(1)$ will be at least as small as $1/|G_0|^{1-\beta/2}$ and we will be done by Observation 1 above.

So we set $\varepsilon = \min\{\beta, \kappa\}/4$, where $0 < \beta \leqslant 1$ is the exponent of quasirandomness given by Assumption (i) from Proposition 3.1 and $\kappa > 0$ is given by Assumption (iii). Let $\delta = \delta(\varepsilon)$ be given by Assumption (ii) of Proposition 3.1 (the Classification of Approximate Subgroups).

We will now apply the above corollary to any ν of the form $\nu = \mu^n$ for some even $n \geqslant \log|G_0|$. Assume that $\|\nu\|_2^2 \geqslant 1/|G_0|^{1-\beta/2}$. Then $1/\|\nu\|_2^2 \leqslant |G_0|^{1-2\varepsilon}$, and if $\|\nu\|_2^2 \leqslant 1/|G_0|^{2\varepsilon}$, we may apply Corollary 3.9, which gives

$$\|\nu * \nu\|_2 \leqslant \frac{\|\nu\|_2}{|G_0|^{\delta/R}}, \tag{3.5}$$

unless there is a $|G_0|^\delta$-approximate group A in G_0 with $|A| \leqslant |G_0|^{1-\varepsilon}$ such that $\nu(gA) \geqslant 1/|G_0|^\delta$ for some $g \in G_0$. By Assumption (ii) of Proposition 3.1, A must be contained in at most $[G : H]^\varepsilon/|G|^\delta$ left cosets of a proper subgroup H. Hence at least one coset xH of H charges ν a lot, i.e., $\nu(xH) \geqslant 1/[G_0 : H]^\varepsilon$. However $\nu^2(H) \geqslant \nu(xH)^2$ since ν is symmetric, and hence

$$\nu^2(H) \geqslant 1/[G_0 : H]^{2\varepsilon}. \tag{3.6}$$

Since $n \mapsto \mu^{2n}(H)$ is non-increasing (see Remark 3.7 above), Assumption (iii) of Proposition 3.1 implies that $\nu^2(H) \leqslant 1/[G_0 : H]^\kappa$. However $\kappa > 2\varepsilon$, so this clearly contradicts (3.6).

Therefore (3.5) always holds as long as $1/|G_0|^{2\varepsilon} \leqslant \|\nu\|_2^2 \leqslant 1/|G_0|^{1-\beta/2}$. As a consequence, we need to apply (3.5) at most a bounded number of times starting from $\nu = \mu^{2n_0}$ with $n_0 = [\log |G_0|]$ say to reach the desired upper bound. Note that the bound $1/|G_0|^{2\varepsilon} \geqslant \|\mu^{2n_0}\|_2^2$ holds thanks to Remark 3.7, (3.4) and Assumption (iii) applied to $H = \{1\}$, because $\kappa > 2\varepsilon$. Now apply successively T times Corollary 3.9 to get:

$$\|(\mu^{2n_0})^{2^T}\|_2 \leqslant \frac{\|\mu^{2n_0}\|_2}{|G_0|^{T\delta/R}} \leqslant \frac{1}{|G_0|^{T\delta/R}} \leqslant \frac{1}{|G_0|^{1-\beta/2}},$$

provided $T\delta/R \geqslant 1 - \beta/2$.

This yields a constant C_1 such that $\mu^{2m}(1) \leqslant 1/|G_0|^{1-\beta/2}$ for some $m \geqslant C_1 \log |G_0|$, where an upper bound for C_1 is

$$C_1 \leqslant 2^{R(1-\beta/2)/\delta}.$$

Together with Observation 1, this finishes the proof of Proposition 3.1 with a rather explicit spectral gap, $\alpha_1 \leqslant e^{-\beta/2C_1}$. Working out the above expression yields the following dependence of the gap in terms of the parameters involved:

$$\lambda_1 \geqslant \beta \cdot e^{-C/\delta},$$

for some absolute constant $C > 0$. Recall that $\delta := \delta(\varepsilon)$ is the function given in Assumption (ii) with $\varepsilon := \min\{\beta, \kappa\}/4$.

4 Approximate subgroups of linear groups

In this section, we give a very brief introduction to approximate subgroups. The first paragraph gives a definition and some general facts, including the relation with small tripling and the Balog-Szemerédi-Gowers lemma. Those are needed only to understand the proof of the ℓ^2-flattening lemma, Lemma 3.8, stated in the last section.

Next we describe the classification of approximate subgroups of simple algebraic groups required to deal with Assumption (ii) of the Bourgain-Gamburd machine (Prop. 3.1 above) and prove Theorem 4.5 below, a structure theorem [19, 80] for approximate subgroups of linear groups. Its proof is purely algebro-geometric and requires nothing on approximate subgroups besides the definition. For further introductory material on approximate groups see [97, 17, 14].

4.1 General facts about approximate groups

The notion of an approximate subgroup of an ambient group G was introduced by Terry Tao in [97] in connection with the work of Bourgain-Gamburd [5] and the Balog-Szemerédi-Gowers theorem alluded to above in the proof of the ℓ^2-flattening lemma (Lemma 3.8). Here is a definition:

Definition 4.1 (Approximate subgroup) A (finite) subset A of a group G is said to be a K-approximate subgroup of G (here $K \geqslant 1$ is a parameter) if A is symmetric

(i.e., $a \in A \Rightarrow a^{-1} \in A$), contains the identity, and if there is a symmetric subset $X \subset G$ of size $|X| \leqslant K$ such that

$$AA \subset XA.$$

Although the definition makes sense without the assumption that A is finite, we will always put this assumption throughout these notes whenever we speak of an approximate subgroup.

Note that $AA = (AA)^{-1} \subset AX$, so we always have $AA \subset XA \cap AX$. Clearly if $K = 1$ this notion coincides with the requirement that A be a finite subgroup of G.

Although Tao was the first to define approximate subgroups in a non-commutative context, their study in $(\mathbb{Z}, +)$, or $(\mathbb{R}, +)$, is an old subject, part of *additive combinatorics* (see [70], [98] for modern expositions), culminating with the so-called Freiman-Ruzsa theorem [27, 85], which gives a structure theorem for approximate subgroups of $\mathbb{Z}$, or more generally (Green-Ruzsa [33]) abelian groups:

Theorem 4.2 (Freiman-Ruzsa, Green-Ruzsa) *Let G be an abelian group and $A \subset G$ be a K-approximate subgroup of G. Then there is a finite subgroup $H \leqslant G$ and a centered multidimensional progression $P \subset G$ of dimension at most $d(K)$ such that A is contained in at most $C(K)$ translates of the subset HP and $|HP| \leqslant C(K)|A|$. The constants $d(K)$ and $C(K)$ depend only on K and not on G nor A.*

By definition a centered *multidimensional progression* of dimension at most d is a subset $P \leqslant G$ of the form $\pi(B)$, where $\pi : \mathbb{Z}^d \to G$ is a group homomorphism and B is a box in $\mathbb{Z}^d$, namely a subset of the form $\prod_{i=1}^{d}[-N_i, N_i]$, where the N_i's are non-negative integers. It is easy to see that B is a 2^d-approximate subgroup, indeed BB is the box with sides $[-2N_i, 2N_i]$ and thus can be covered by the translates of B centered at each of the 2^d corners of the box B. Passing to the quotient via π, we get that P too is a 2^d-approximate subgroup, and finally that for every finite subgroup $H \leqslant G$, the so-called *coset-progression* HP is also a 2^d-approximate subgroup.

For the proof of this theorem, we refer the reader to the book by Tao and Vu [98] as well as the article [33] and the original references therein.

Two remarks are in order:

- The bounds $d(K)$ and $C(K)$ can be made quantitative, and good estimates on them are useful for applications as we will see below. Conjecturally (Freiman-Ruzsa conjecture), $d(K) = O(\log K)$ while $C(K) = O(K^{O(1)})$. See Sanders [88] for the best currently available bounds.
- The conclusion is quite special to abelian groups. A very general structure theorem was recently obtained in [21] valid for approximate subgroups of arbitrary groups, but it yields no explicit bounds on $C(K)$. As we will see below, when G is a finite simple group of bounded rank, then a polynomial bound can be given on $C(K)$ provided A generates G. Obtaining here a polynomial bound is crucial for the applications to the Bourgain-Gamburd expansion machine, i.e. to Assumption (ii) of Prop. 3.1.

As follows immediately from their definition, approximate subgroups do not grow much under self multiplication, namely the product set $A^k := A \cdot \ldots \cdot A$ of A with

itself k times has size at most $|X|^{k-1}|A|$. An important observation (due to Tao using related ideas of Ruzsa) is that we have the following converse:

Proposition 4.3 (Small tripling) *Let A be a finite subset of a group G such that $|AAA| \leqslant K|A|$ for some parameter $K \geqslant 1$. Then $B := (A \cup A^{-1} \cup \{1\})^2$ is a $c(K)$-approximate subgroup of size $|B| \leqslant c(K)|A|$, where $c(K) = O(K^{O(1)})$ and the implied constants are absolute. In particular $|A^n| \leqslant O(K^{O(n)})|A|$.*

Proof The proof is elementary. It is a simple application of the Ruzsa inequality and Ruzsa covering lemma. See [97, Theorem 3.9] or [12, Prop 2.2]. □

We remark that it is necessary to take the 3-fold power of A in the assumption of this proposition. It is not true if we only assume that $|AA| \leqslant K|A|$ (take $A = \{x\} \cup H$, where $x \in G$ and H is a large subgroup such that $xHx^{-1} \cap H = \{1\}$). Nevertheless one can still show in this case that A is covered by $O(K^{O(1)})$ left translates of an $O(K^{O(1)})$-approximate subgroup of G of size at most $O(K^{O(1)})|A|$ (see [97, Theorem 4.6]).

A deeper fact, recorded in the lemma below, is that one can still identify an approximate subgroup "near" the finite set A assuming only that A does not grow under self multiplication in the following statistical sense:

$$\|1_A * 1_A\|_2^2 = |\{(a,b,c,d) \in A \times A \times A \times A; ab = cd\}| \geqslant |A|^3/K.$$

The left hand side is called *the multiplicative energy* of the set A with itself and is sometimes denoted by $E(A,A)$. It is the ℓ^2-norm squared of the convolution product of the indicator function of A in G with itself and is easily seen to be equal to the expression in the middle (number of "collisions" $ab = cd$). In other words: this condition means that the probability that $ab = cd$, when a, b, c and d are chosen at random in A is at least $1/K|A|$. Clearly if A is a subgroup, this probability if exactly $1/|A|$. Also easy to see is the remark that if $|AA| \leqslant K|A|$, then $\|1_A * 1_A\|_2^2 \geqslant |A|^3/K$, indeed setting $r(x) := |\{(a,b) \in A \times A; ab = x\}|$ we have $\sum r(x)^2 = \|1_A * 1_A\|_2^2$, $\sum r(x) = |A|^2$ and $|\{x, r(x) > 0\}| = |AA|$, hence applying Cauchy-Schwarz:

$$|A|^4 = \Big(\sum r(x)\Big)^2 \leqslant |AA| \Big(\sum r(x)^2\Big) \leqslant K|A| \cdot \|1_A * 1_A\|_2^2.$$

Lemma 4.4 (Balog-Szemerédi-Gowers-Tao lemma) *Suppose A_1, A_2 are finite subsets of a group G such that $|A_1| \leqslant K|A_2|$ and $|A_2| \leqslant K|A_1|$ and assume that*

$$\|1_{A_1} * 1_{A_2}\|_2^2 \geqslant (|A_1||A_2|)^{3/2}/K,$$

then there is a $O(K^{O(1)})$-approximate subgroup $A \subset G$ of size $O(K^{O(1)})|A_1|$ such that a subset of A_1 of size at least $|A_1|/O(K^{O(1)})$ is contained in some left translate of A and similarly a subset of A_2 of size at least $|A_2|/O(K^{O(1)})$ is contained in some right translate of A.

Proof We will not give the proof of this important combinatorial result here. Rather we refer the reader to the book by Tao and Vu [98, §2.5, 2.7] and Tao's paper [97, Theorem 5.4]. See also [12, Corollaries 4.5, 4.6] for a somewhat different argument. □

Note that we cannot claim that A_1 itself is contained in few translates of A, because if the condition $\|1_{A_1'} * 1_{A_2'}\|_2^2 > (|A_1'||A_2'|)^{3/2}/O(K^{O(1)})$ holds for some subsets A_1', A_2' each making a proportion $\geqslant 1/O(K^{O(1)})$ of A_1 and A_2 respectively, then $\|1_{A_1} * 1_{A_2}\|_2^2 \geqslant \|1_{A_1'} * 1_{A_2'}\|_2^2 \geqslant (|A_1||A_2|)^{3/2}/O(K^{O(1)})$. For example if $A_1 = A_2 = \{1, \dots, N\} \cup \{2, 2^2, \dots, 2^N\}$, then $\|1_{A_1} * 1_{A_1}\|_2^2 \geqslant \|1_{\{1,\dots,N\}} * 1_{\{1,\dots,N\}}\|_2^2 \geqslant N^3$, while A_1 is not contained in a bounded number of translates of multidimensional arithmetic progression in $\mathbb{Z}$, hence not contained in a bounded number of translates of an approximate subgroup of $\mathbb{Z}$ (using Theorem 4.2).

4.2 Classification of approximate subgroups of $\mathbb{G}(\mathbb{F}_q)$

The main result here is the following:

Theorem 4.5 (Classification theorem) *Let $K, M \geqslant 2$. Assume that $\mathbb{G}$ is an absolutely simple algebraic group of complexity at most M defined over an algebraically closed field. If A is a finite K-approximate subgroup of $\mathbb{G}$ which is C-sufficiently Zariski-dense in $\mathbb{G}$, then either $|A| \leqslant K^C$, or $\langle A \rangle$ is finite and of cardinality at most $K^C|A|$. Here $C = C(M) > 0$ is a constant depending only on M and* $\dim \mathbb{G}$.

The rest of this subsection is devoted to the proof of this theorem and some of its corollaries.

Remark Although this will not be used later on, we may replace K^C in the above theorem by $CK^{3\dim \mathbb{G}+3}$, where C depends again on M and $\dim \mathbb{G}$.

Recall that an affine algebraic variety is said to have complexity at most M if it is the vanishing locus of a finite set of polynomials whose sum of their total degree is at most M. This notion can be extended to all algebraic varieties (see [19, Appendix A] for background). Recall further that a subset of $\mathbb{G}$ is called M-sufficiently Zariski-dense if it is not contained in a proper algebraic subvariety of complexity at most M.

This result was obtained by Green, Tao and the author in [19, Theorem 5.5]. The proof of a closely related statement (in fact Corollary 4.7 below) was derived independently at the same time by Pyber and Szabó, see [80] and [81] for their point of view.

Simple and quasi-simple groups of Lie type are of the form $G = \mathbb{G}(\mathbb{F}_q)^\sigma/Z$, where $\mathbb{G}$ is a simply connected absolutely simple algebraic group defined and split over the prime field $\mathbb{F}_p$, σ is a Frobenius map, i.e., the composition of a field automorphism and a graph automorphism, and Z is a central subgroup (whose cardinal is bounded in terms of $\dim \mathbb{G}$ only). It is not difficult (for example using the Lang-Weil bounds or the related and easier Schwarz-Zippel estimates) to check that the subgroups $\mathbb{G}(\mathbb{F}_q)^\sigma$ of fixed points of σ are C-sufficiently Zariski-dense in $\mathbb{G}$ whenever q is larger than a constant depending only on C and $\dim \mathbb{G}$ (see [22, Proposition 5.4] for details). Thus a consequence of Theorem 4.5 is the following:

Corollary 4.6 *Let G be a (non-abelian) finite simple (or quasisimple) group of Lie type and suppose that A is a K-approximate subgroup of G. Then either $|A| \leqslant K^C$, or*

$|A| \geqslant |G|/K^C$, or A is contained in a proper subgroup of G. Here $C > 0$ is a constant depending only on the rank of G, not on the size of the associated finite field.

Proof By the discussion above, we may assume that G is a sufficiently Zariski-dense subgroup of a simple algebraic group $\mathbb{G}$ of bounded complexity. It only remains to check that if A generates G, then there is a bounded m such that A^m is sufficiently Zariski-dense and then apply Theorem 4.5 to A^m. This fact goes back to Eskin-Mozes-Oh [26, Prop. 3.2]. It is a basic tool called since *escape from subvarieties*, which can be proved with explicit bounds using Bezout's theorem. It can also easily be proved (without an explicit bound on m) using ultraproducts: if no such m existed we could form the ultraproduct of possible counter-examples, yielding a subset of $\mathbb{G}(K)$, where K is the corresponding ultraproduct of fields, which generates a subgroup which is not Zariski-dense, hence is contained in a proper algebraic subgroup of $\mathbb{G}(K)$. But this means that most (for the ultrafilter) counter-examples are contained in that algebraic subgroup, contradicting the assumption. See [19, Lemma 3.11] for more details regarding this argument. □

Another related statement is the following, sometimes called the *product theorem*, because it guarantees that any generating subset of G grows under products:

Corollary 4.7 (Product theorem) *Let G be a (non-abelian) finite simple (or quasi-simple) group of Lie type and $A \subset G$ an arbitrary generating finite subset, then*

$$|AAA| \geqslant \min\{|A|^{1+\varepsilon}, |G|\},$$

where $\varepsilon > 0$ is a constant depending only on the rank of G, not on the size of the associated finite field.

This result was obtained by Pyber and Szabó [80, Theorem 4]. We show now how to derive it from the classification of approximate subgroups, i.e., Theorem 4.5.

Proof Let $K = |A|^{\varepsilon}$ and apply Proposition 4.3 to get a $(2|A|)^{C\varepsilon}$-approximate subgroup B containing A, where $C > 0$ is an absolute constant. By Corollary 4.6, either $|A| \leqslant K^C$, or $|A| \geqslant |G|/K^C$. The first case is ruled out if $\varepsilon < 1/2C^2$, because that would force $|A| = 1$. In the second case $|A| \geqslant |G|^{1-\delta}$ for $\delta > 0$ which can be taken arbitrarily small provided ε is small enough. Then a general result of Nikolov-Pyber [71], based on an observation of Gowers [32] using the quasirandomness of G (i.e., Proposition 6.1 below), implies that $AAA = G$. See [14, Corollary 2.3.] for a detailed proof of this last step using basic representation theory of finite groups. □

Corollary 4.7 was first proved by Helfgott [36] in the special cases of $\mathrm{SL}_2(\mathbb{F}_p)$, for the prime field $\mathbb{F}_p$ only, using some *ad hoc* matrix computations based on the sum-product phenomenon from additive combinatorics (i.e., the Bourgain-Katz-Tao theorem [8]). Helfgott later settled the case of $\mathrm{SL}_3(\mathbb{F}_p)$ in [37]. Earlier work of Elekes and Király [25] had dealt with the analogous result for $\mathrm{SL}_2(\mathbb{R})$. Although these elementary methods fail to extend to the general case, they have the merit of being somewhat more explicit on the ε (see, e.g., [54, 23]).

Remark Our lower bound on ε is not explicit. However, if one assumes further that the subset A is C-sufficiently Zariski-dense in the ambient simple algebraic group $\mathbb{G}$ (i.e., is not contained in any proper algebraic subvariety of degree, or complexity, at most C for some non explicit C depending only on $\mathbb{G}$), then ε can be taken to be $1/(3\dim\mathbb{G}+4)$. See Remark 4.11 below. The constant C itself (and hence the ε of Corollary 4.7) can be made effective (although not really explicit) using effective algebraic geometry bounds as done by Pyber-Szabó in [80]. The treatment in [21] was not effective, because we used ultrafilters to achieve these uniform bounds.

We will sketch below the proof of Theorem 4.7. The proof is germane to the proof of the Larsen-Pink theorem [59] on the classification of finite subgroups of $\mathbb{G}$. Let us first state a version of the Larsen-Pink theorem appropriate to our discussion (see [59, Theorem 0.5] and [42]).

Theorem 4.8 (Larsen-Pink theorem) *Let F be an algebraically closed field and $\mathbb{G}$ be an absolutely simple simply connected algebraic group of complexity at most M defined and split over the prime field of F. If Γ is a finite subgroup of $\mathbb{G}$ which is C-sufficiently Zariski-dense in $\mathbb{G}$, then the field F has positive characteristic p and Γ is a conjugate of the subgroup $\mathbb{G}(\mathbb{F}_q)$ for a finite field $\mathbb{F}_q \leqslant F$, q a power of p. Here $C = C(M) > 0$ is a constant depending only on M and $\dim\mathbb{G}$.*

This theorem is a strict generalization of Nori's Theorem 2.2 discussed earlier in the case of simple algebraic groups. However the proof by Larsen and Pink is very different from Nori's counting argument sketched in Theorem 2.2 above. While Nori was building the algebraic subgroup from below taking products of unipotent elements and using crucially that p is large, Larsen and Pink argue differently and cut the group from above so to speak by computing the approximate size of the centralizers in Γ of any subset of elements. This allows them to eventually find many unipotent elements (using an argument similar to the original argument of Jordan [44, 11]) including a minimal one which will generate the additive subgroup of the finite field $\mathbb{F}_q$ that we are required to build from Γ alone.

In order to compute the correct size of centralizers, Larsen and Pink establish first a very general inequality, the Larsen-Pink non-concentration estimate, which gives an a priori upper bound on the intersection of Γ with any algebraic subvariety of bounded complexity. Namely:

Proposition 4.9 (Larsen-Pink non-concentration estimate [59, Thm 4.2]) *Under the assumptions of Theorem 4.8, consider a closed algebraic subvariety $\mathcal{V}$ of $\mathbb{G}$ of complexity at most M. Then if Γ is a finite subgroup of $\mathbb{G}$ which is C-sufficiently Zariski-dense in $\mathbb{G}$,*

$$|\Gamma \cap \mathcal{V}| \leqslant C|\Gamma|^{\dim\mathcal{V}/\dim\mathbb{G}}, \tag{4.1}$$

where $C = C(M) > 0$ is a constant depending only on M and $\dim\mathbb{G}$.

Before we say more about the proof of this proposition and its relation to approximate subgroups, let us explain what it entails for centralizers. Define q_Γ as the positive real number $|\Gamma|^{1/\dim\mathbb{G}}$. Let Z_a be the centralizer in $\mathbb{G}$ of an element $a \in \Gamma$. The orbit-stabilizer formula tells us that

$$|Z_a \cap \Gamma| \cdot |\{\gamma a \gamma^{-1}; \gamma \in \Gamma\}| = |\Gamma|,$$

so

$$q_\Gamma^{\dim \mathbb{G}} = |\Gamma| \leqslant |Z_a \cap \Gamma|\,|\mathcal{V}_a \cap \Gamma| \leqslant |Z_a \cap \Gamma| \cdot C q_\Gamma^{\dim \mathcal{V}_a},$$

where $\mathcal{V}_a$ is the conjugacy class of a in $\mathbb{G}$, which is a constructible set in G, being the image of $\mathbb{G}$ under the map $g \mapsto gag^{-1}$. We applied the Larsen-Pink inequality (4.1) to the Zariski closure of $\mathcal{V}_a$, which also has dimension $\dim \mathcal{V}_a = \dim \mathbb{G} - \dim Z_a$. Now applying (4.1) once again but this time to Z_a we obtain:

$$\frac{1}{C} q_\Gamma^{\dim Z_a} \leqslant |Z_a \cap \Gamma| \leqslant C q_\Gamma^{\dim Z_a}. \tag{4.2}$$

The constant C depends only on the complexity of Z_a and the closure of $\mathcal{V}_a$, which are both bounded in terms of $\dim \mathbb{G}$ and the complexity of $\mathbb{G}$ only and are in particular independent of a (see, e.g., [19, Appendix A] for general facts on the complexity of algebraic varieties). So we see that the Larsen-Pink inequality (4.1) not only gives an upper bound, but also a lower bound of the same order of magnitude on the size of centralizers.

The proof of Theorem 4.5 rests on the same key idea. The main step consists in extending the Larsen-Pink inequality (4.1) to the setting of approximate subgroups:

Proposition 4.10 (Larsen-Pink for approximate subgroups) *Let $K, M \geqslant 2$. Assume that $\mathbb{G}$ is an absolutely simple algebraic group of complexity at most M defined over an algebraically closed field. If A is a finite K-approximate subgroup of $\mathbb{G}$ which is C-sufficiently Zariski-dense in $\mathbb{G}$, then for every closed algebraic subvariety $\mathcal{V}$ of $\mathbb{G}$ of complexity at most M,*

$$|A \cap \mathcal{V}| \leqslant C K^C |A|^{\dim \mathcal{V}/\dim \mathbb{G}}, \tag{4.3}$$

where $C = C(M) > 0$ is a constant depending only on M and $\dim \mathbb{G}$.

This is a strict generalization of (4.1), indeed we recover Proposition 4.9 in the special case when $K = 1$ (i.e., A is a subgroup). The possibility of an extension to approximate groups of the Larsen-Pink estimate is an idea of Hrushovski, who proved a qualitative version of (4.3) in his ground-breaking paper on approximate groups [39]. The polynomial dependence of the constant (in CK^C) is proved in [19, Thm 4.1] using a variation of the argument we are about to present. Helfgott in [37] proved a special case of this inequality when $\mathcal{V} = T$ is a maximal torus.

Proof We follow the Larsen-Pink strategy for proving Proposition 4.9, see [59, Thm 4.2]. Since a bound on complexity implies a bound on the number of irreducible components (see [19, Appendix A]), it is enough to prove (4.3) for irreducible varieties. Clearly the estimate (4.3) holds when $\mathcal{V}$ has dimension 0 or dimension $\dim \mathbb{G}$, so we may pick a possible counter-example to (4.3) of minimal positive dimension, say $\mathcal{V}^-$ and another one of minimal co-dimension, say $\mathcal{V}^+$. The basic idea of the proof, which relies crucially on the hypothesis that $\mathbb{G}$ is simple, is that we should be able to find $a \in A$ such that the product $\mathcal{W} := \mathcal{V}^- a \mathcal{V}^+ a^{-1}$ is a constructible set of dimension $> \dim \mathcal{V}^+$ and thus hopefully will contain too many elements of $AaAa^{-1}$. Hence, since A is an approximate subgroup, some translate of $\mathcal{W}$ will contain too many elements of A, contradicting the choice of $\mathcal{V}^+$.

To effect this strategy rigorously, one cannot just proceed as outlined above, because $A\times A$ could concentrate of a singular subvariety of $\mathcal{V}^-\times\mathcal{V}^+$ made of non-generic fibers of the product map

$$\begin{aligned}\Phi\colon \mathcal{V}^-\times\mathcal{V}^+ &\to \mathcal{W},\\ (x,y) &\mapsto xaya^{-1}.\end{aligned} \tag{4.4}$$

So instead we will prove first a weaker version of (4.3) in which the exponent $1/\dim\mathbb{G}$ is replaced by some $\alpha\in[1/\dim\mathbb{G},1]$. And then improve that estimate by showing that, given any fixed $\beta\geqslant 1/\dim\mathbb{G}$, if the bound (4.5) below holds for all subvarieties and for some $\alpha\leqslant\beta+\varepsilon$, where $\varepsilon=1/(\dim\mathbb{G})^2$, then it also holds for $\alpha=\beta$ and all subvarieties:

$$|A\cap\mathcal{V}|\leqslant O(K^{O(1)})|A|^{\alpha\dim\mathcal{V}}. \tag{4.5}$$

Since (4.5) holds obviously when $\alpha=1$ and all subvarieties and since if (4.5) holds for $\alpha=\alpha_0$, then it holds for all $\alpha\geqslant\alpha_0$, this will eventually prove that (4.5) holds for $\alpha=1/\dim\mathbb{G}$, so that (4.3) holds as desired.

Let us proceed as announced. We fix $\beta\geqslant 1/\dim\mathbb{G}$ and assume that (4.5) holds for all $\alpha\geqslant\beta+\varepsilon$. Pick irreducible subvarieties $\mathcal{V}^-$ and $\mathcal{V}^+$ as above of minimal and maximal dimension providing counter-examples to (4.5) for $\alpha=\beta$. This means that $|A\cap\mathcal{V}^+|$ is much bigger than $CK^C|A|^{\beta\dim\mathcal{V}^+}$ and similarly for $\mathcal{V}^-$. By Lemma 4.12 below, we may find $a\in A$ such that $\mathcal{W}:=\mathcal{V}^-a\mathcal{V}^+a^{-1}$ is a constructible set of dimension $>\dim\mathcal{V}^+$. Consider the product map Φ defined in (4.4) above. Let $\mathcal{S}\leqslant\mathcal{V}^-\times\mathcal{V}^+$ be a singular subvariety of strictly smaller dimension outside of which each point lies on a fiber of the right dimension namely $d:=\dim\mathcal{V}^-+\dim\mathcal{V}^+-\dim\mathcal{W}$. By assumption $d<\dim\mathcal{V}^-$. Basic algebraic geometry (cf. [95, I.6.3]) tells us that $\mathcal{S}$ and the fibers are closed algebraic subvarieties, and it is possible to prove by abstract nonsense (see [19, Appendix A]) that their complexity is bounded in terms of those of $\mathcal{V}^\pm$ alone.

Then we see that $A\times A$ must concentrate on $\mathcal{S}$, i.e., $|(A\times A)\cap\mathcal{S}|>\frac{1}{2}|(A\times A)\cap(\mathcal{V}^-\times\mathcal{V}^+)|$, since otherwise, decomposing $(\mathcal{V}^-\times\mathcal{V}^+)\setminus\mathcal{S}$ into fibers of Φ we would get:

$$\begin{aligned}CK^C|A|^{\beta(\dim\mathcal{V}^++\dim\mathcal{V}^-)} &\ll \frac{1}{2}|A\cap\mathcal{V}^-|\cdot|A\cap\mathcal{V}^+|\\ &\leqslant \sum_{z\in\mathcal{W}\cap\Phi(A\times A)}|\Phi^{-1}(z)\cap(A\times A)|\\ &\ll |\mathcal{W}\cap A^4|\,|A|^{\beta d}\end{aligned}$$

implying that some translate of $\mathcal{W}$ intersects A in a subset of size much larger than $O(K^{O(1)})|A|^{\beta\dim\mathcal{W}}$ and thus contradicting the choice (maximality) of $\mathcal{V}^+$. It was licit to bound $|\Phi^{-1}(z)\cap(A\times A)|$ as we did above because of the minimality of $\dim\mathcal{V}^-$ and the fact that $d=\dim\Phi^{-1}(z)<\dim\mathcal{V}^-$.

So we are reduced to the case when $A\times A$ concentrates on the singular subvariety $\mathcal{S}$, which is of dimension at most $\dim\mathcal{V}^-+\dim\mathcal{V}^+-1$. Passing to a proper subvariety of smaller dimension if necessary, we may assume that $\mathcal{S}$ is a subvariety of smallest possible dimension on which $A\times A$ concentrates (i.e., $|A\cap\mathcal{V}^-|\cdot|A\cap\mathcal{V}^+|\ll O(1)|(A\times$

$A) \cap \mathcal{S}|$). If its projection to the second factor $\mathcal{V}^+$ is contained in a proper closed subvariety of $\mathcal{V}^+$, then we use (4.5) for $\beta + \varepsilon$, to write

$$|(A \times A) \cap \mathcal{S}| \leqslant O(K^{O(1)})|A \cap \mathcal{V}^-| \cdot |A|^{(\beta+\varepsilon)(\dim \mathcal{V}^+ - 1)},$$

which is a contradiction since $(\beta+\varepsilon)(\dim \mathcal{V}^+ - 1) < \beta \dim \mathcal{V}^+$. So we may assume that the projection of $\mathcal{S} \leqslant \mathcal{V}^- \times \mathcal{V}^+$ to the second factor $\mathcal{V}^+$ contains an open dense set of $\mathcal{V}^+$, i.e., the projection is dominant, and hence away from a proper closed subvariety $\mathcal{S}_0$ of $\mathcal{S}$ (on which $A \times A$ cannot concentrate by minimality of $\mathcal{S}$) the fibers of this projection have dimension at most $\dim \mathcal{V}^- - 1$. Hence:

$$\begin{aligned} |(A \times A) \cap \mathcal{S}| &\leqslant O(1)|(A \times A) \cap \mathcal{S} \setminus \mathcal{S}_0| \\ &\leqslant O(1) \sum_{a \in A \cap \mathcal{V}^+} |(A \times \{a\}) \cap \mathcal{S} \setminus \mathcal{S}_0| \\ &\leqslant O(K^{O(1)})|A \cap \mathcal{V}^+| \cdot |A|^{\beta(\dim \mathcal{V}^- - 1)}, \end{aligned}$$

which is again contradictory. This establishes that (4.5) holds for $\alpha = \beta$ and thus by induction that (4.3) holds unconditionally. □

Remark 4.11 A careful analysis of the above argument shows that the exponent of K in (4.3) can be taken to be $3 \dim \mathbb{G}$, while the multiplicative constant C, depends on the complexity of $\mathcal{V}$, and is less explicit owing to the less explicit nature of our notion of complexity and the way it bounds the number of irreducible components (as proved in [19, Appendix A] using ultraproducts). Similarly the threshold of "sufficient Zariski-density" of A is non explicit.

Let $M \geqslant 2$ and $\mathbb{G}$ as above an absolutely simple connected algebraic group $\mathbb{G}$ of complexity at most M. In the above proof, we made use of the following lemma.

Lemma 4.12 (Finding a transverse conjugate) *There is $C = C(M) > 0$ such that the following holds. If A is a C-sufficiently Zariski-dense finite subset of $\mathbb{G}$ of complexity at most M, then for any two closed algebraic subvarieties $\mathcal{V}_1, \mathcal{V}_2$ in $\mathbb{G}$ of complexity at most M and positive dimension and co-dimension, there is $a \in A$ such that the constructible set $\mathcal{V}_1 a \mathcal{V}_2 a^{-1}$ has dimension strictly bigger than $\dim \mathcal{V}_2$ and complexity $O_M(1)$ (i.e., a constant depending on M only).*

Proof We may assume both varieties to be irreducible. If no such a can be found, then for every $x_1 \in \mathcal{V}_1$ the closed irreducible subvarieties $x_1 a \mathcal{V}_2$ and the closure of $\mathcal{V}_1 a \mathcal{V}_2$ have same dimension, hence are equal. This means that $x_1 a \mathcal{V}_2 a^{-1} = x_1' a \mathcal{V}_2 a^{-1}$ for all $x_1, x_1' \in \mathcal{V}_1$. Hence that $x_1^{-1} x_1'$ lies in the stabilizer in $\mathbb{G}$ of the subvariety $a \mathcal{V}_2 a^{-1}$, namely $\mathcal{V}_1^{-1} \mathcal{V}_1$ lies in $a \mathbb{H} a^{-1}$, where $\mathbb{H}$ is the closed algebraic subgroup $\{g \in \mathbb{G}; g\mathcal{V}_2 = \mathcal{V}_2\}$. Since $\mathcal{V}_2$ is a proper subvariety, $\mathbb{H}$ is a proper subgroup, and since $\mathbb{G}$ is simple $\bigcap_{a \in \mathbb{G}} a \mathbb{H} a^{-1}$ is finite. We claim that, because A is assumed sufficiently Zariski-dense, $\bigcap_{a \in A} a \mathbb{H} a^{-1}$ is finite too; this will contradict the assumption that $\mathcal{V}_1$ has positive dimension and prove the lemma.

To see the claim, observe that if $\mathbf{Y} \leqslant \mathbb{H}$ is an algebraic subvariety of complexity at most M', then $\{g \in \mathbb{G}, \mathbf{Y} \leqslant g \mathbb{H} g^{-1}\}$ is a subvariety of complexity at most $O_{M'}(1)$. So if $M' = O_M(1)$, then there will be $a \in A$ outside it. Applying this remark several times

to each of the irreducible components $\mathbf{Y}$ of the intersections $\mathbb{H} \cap a_1 \mathbb{H} a_1^{-1} \cap \ldots \cap a_i \mathbb{H} a_i^{-1}$, $i \leqslant k$, we can build $k = O_M(1)$ elements $a_i \in A$ such that $\bigcap_{1 \leqslant i \leqslant k} a_i \mathbb{H} a_i^{-1}$ has dimension 0. □

Having the Larsen-Pink estimate for approximate groups (Proposition 4.10) at our disposal, we are ready to prove our main theorem. So we now pass to the proof of Theorem 4.5. For this it will be convenient to make the following definition:

Definition A maximal torus T of $\mathbb{G}$ will be called an *involved torus* if $A^2 \cap T$ contains at least one regular element.

Recall that a maximal torus is a connected closed algebraic subgroup of $\mathbb{G}$ containing only semisimple elements (i.e., elements that are diagonalizable in some hence any embedding of $\mathbb{G}$ in GL_d) and maximal for this property. Maximal tori are all conjugate. We refer to Borel's book [4] or Humphreys [43] for background on algebraic groups. A semisimple element is called regular if its centralizer has a maximal torus of finite index. Regular semisimple elements form a Zariski open subset of $\mathbb{G}$. In particular, since the approximate group A is assumed to be sufficiently Zariski-dense in Theorem 4.5, we see that A contains a regular semisimple element. We also recall that every maximal torus T is of bounded index in its normalizer $N(T)$.

We observe at the outset that the number of involved tori is finite, indeed of size at most $|A^2|$, because a regular semisimple element can be contained in at most one maximal torus (the connected component of its centralizer). As in the Larsen-Pink theorem, we set $q_A := |A|^{1/\dim \mathbb{G}}$. We need to prove that either q_A is $O(K^{O(1)})$, or $\langle A \rangle$ is finite and $q_A / q_{\langle A \rangle}$ is $O(K^{O(1)})$.

Claim 1 If T is an involved maximal torus, then

$$1/O(K^{O(1)}) q_A^{\dim T} \leqslant |T \cap A^2| \leqslant O(K^{O(1)}) q_A^{\dim T}. \tag{4.6}$$

Proof The argument is the same as the one used to prove (4.2) above applying the Larsen-Pink inequality to both the centralizer and the conjugacy class, and yields the desired estimate for the centralizer $Z(a_0)$ of a regular semisimple $a_0 \in A^2 \cap T$ instead of T. Namely looking at the fibers of the map $A \to A^3 \cap \mathcal{V}_{a_0}$, $a \mapsto a a_0 a^{-1}$, where $\mathcal{V}_{a_0}$ is the conjugacy class of a_0 in $\mathbb{G}$, we see that

$$|A| \leqslant |A^2 \cap Z(a_0)| \cdot |A^3 \cap \mathcal{V}_{a_0}|,$$

but each factor in the right handside is, respectively, at most $O(K^{O(1)}) q_A^{\dim Z(a_0)}$ and $O(K^{O(1)}) q_A^{\dim \mathbb{G} - \dim Z(a_0)}$, so the product is $O(K^{O(1)})|A|$. We thus obtain (4.6) with $Z(a_0)$ in place of T. But $Z(a_0)$ is an algebraic subgroup with bounded complexity and T is its connected component, hence T has bounded index in $Z(a)$. This easily implies that $|A^2 \cap T| \geqslant |A^2 \cap Z(a_0)| / O(K^{O(1)})$ (indeed A will intersect some translate of $Z(a_0)$ in a set of size $\geqslant q_A^{\dim T} / O(K^{O(1)})$, hence also some translate of T in a comparable size). This establishes (4.6). □

Claim 1 above is really the beef of the proof: assuming only that $T \cap A^2$ has one regular element, we get that it has at least $q_A^{\dim T}$ regular elements up to a $O(K^{O(1)})$

factor. Indeed, the non-regular elements in T are concentrated on a bounded union of proper algebraic subgroups of bounded complexity (the subtori corresponding to the vanishing of some root: in SL_n this corresponds to the subgroups of diagonal matrices having at least one double eigenvalue). So applying the Larsen-Pink inequality (4.3) to this bounded union T_{sing} of subtori, we see that $|A^2 \cap T_{sing}| \leqslant O(K^{O(1)}) q_A^{\dim T - 1}$. This means that there are at least $q_A^{\dim T}/O(K^{O(1)})$ elements in $A^2 \cap T$ lying outside of T_{sing}.

Claim 2 Unless q_A is $O(K^{O(1)})$, for every maximal torus T of $\mathbb{G}$, if T is involved in A, so is aTa^{-1} for every $a \in A$.

Proof This follows easily from Claim 1 and the above remark. Note that $|A^2 \cap aTa^{-1}| = |a^{-1}A^2a \cap T|$. However aA^2a^{-1} being contained in A^4 must lie in at most K^3 left translates of A. Hence A^2 is contained in at most K^3 left translates of $a^{-1}Aa$. This means that one of these translates must intersect T in a set of size at least $q_A^{\dim T}/O(K^{O(1)})$. Hence $|a^{-1}A^2a \cap T| \geqslant q_A^{\dim T}/O(K^{O(1)})$, which implies by the remark above, that $a^{-1}A^2a$ contains a regular semisimple element of T, unless $q_A \leqslant O(K^{O(1)})$. This proves the claim. □

Obviously this lemma implies that all conjugates gTg^{-1}, $g \in \langle A \rangle$, are involved.

Claim 3 Unless q_A is $O(K^{O(1)})$, $\langle A \rangle$ is finite.

Proof As remarked earlier, since A is sufficiently Zariski-dense, it must contain a regular semisimple element, so there is at least one involved torus. Since every regular semisimple element is contained in at most one torus, there are only finitely many involved tori. By Claim 2, unless $q_A = O(K^{O(1)})$, they are permuted under conjugation by $\langle A \rangle$. In particular the Zariski-closure $\mathbb{H}$ of $\langle A \rangle$ intersects the normalizer $N(T)$, and hence T itself in a subgroup of finite index. We claim that if $\langle A \rangle$ is infinite, and hence the connected component of the identity $\mathbb{H}^0$ has positive dimension, then there is a closed connected algebraic subgroup $S \leqslant T$ of bounded complexity and containing $\mathbb{H}^0$ such that $\mathbb{H} \leqslant N(S)$. This will yield the desired contradiction, because $N(S)$ has then bounded complexity. Starting with $S = T$ observe that if $\mathbb{H}$ does not normalize S, then there is $h \in \mathbb{H}$ such that $S \cap hSh^{-1}$ has dimension $< \dim S$ and bounded complexity. Hence so does the connected component S_1 of $S \cap hSh^{-1}$. Since $\mathbb{H}^0$ is normalized by h, this S_1 also contains $\mathbb{H}^0$. Reiterate with $S := S_1$. This process ends after at most $\dim T$ steps and the claim follows. □

The proof of Theorem 4.5 now follows in a few lines from Claims 1 and 3 and the Larsen-Pink inequality by counting the number $\mathcal{T}$ of involved tori. Since every regular semisimple element is contained in at most one maximal torus, Claim 1 implies that

$$\mathcal{T} \leqslant O(K^{O(1)})|A^2|/q_A^{\dim T} \leqslant O(K^{O(1)}) q_A^{\dim \mathbb{G} - \dim T}.$$

On the other hand, the subgroup $\langle A \rangle$ acts by conjugation on the (finite) set of involved tori by Claim 2. So

$$\mathcal{T} \geqslant |\langle A \rangle|/|\langle A \rangle \cap N(T)| \geqslant |\langle A \rangle|/O(1) q_{\langle A \rangle}^{\dim T} = q_{\langle A \rangle}^{\dim \mathbb{G} - \dim T}/O(1),$$

where the second inequality follows from the original Larsen-Pink inequality (Proposition 4.9) applied to the sufficiently Zariski-dense subgroup $\langle A\rangle$. So $q_A/q_{\langle A\rangle} = O(K^{O(1)})$ as desired. Finally note that the Larsen-Pink estimate was used only for subvarieties (tori, conjugacy classes, etc.) whose complexity is bounded in terms of the complexity of $\mathbb{G}$ only. Hence the threshold of sufficient Zariski density required in these applications of (4.3) is uniform. This ends the proof of Theorem 4.5.

4.3 Verifying Assumption (ii) of the Bourgain-Gamburd machine

Suppose $G_0 = \mathbb{G}(\mathbb{F}_q)$, where $\mathbb{G}$ is an absolutely simple algebraic group defined over the finite field $\mathbb{F}_q$. Then Corollary 4.6 proved in the previous subsection implies that Assumption (ii) of the Bourgain-Gamburd machine (i.e., Proposition 3.1) holds for G_0 with a function $\delta(\varepsilon)$ given by $\delta(\varepsilon) = \varepsilon \min\{\beta, 1/(C+1)\}$, where C is the constant from Corollary 4.6 (distinguishing the cases $H = 1$ and $H \neq 1$ and using Remark 3.4).

In order to deal with products of a bounded number of quasi-simple groups of Lie type of bounded rank, one needs the following rather straightforward extension of Theorem 4.7, based on Goursat's lemma about subgroups of direct products of groups.

Theorem 4.13 (Approximate subgroups of semisimple groups) *Let G be an (almost direct) product of finite simple (or quasisimple) groups of Lie type and suppose that A a K-approximate subgroup of G. Then either $|A| \geqslant |G|/K^C$, or A is contained in at most K^C left cosets of a proper subgroup H of G, where $C > 0$ is a constant depending only on the rank of G.*

We refer the reader to [22, Theorem 8.1] for a detailed proof.

If $\mathbb{G}$ is a semisimple algebraic group defined over $\mathbb{Q}$ its reduction $\mathbb{G}_p$ modulo p is well-defined for all but finitely many primes p. When $\mathbb{G}$ is simply connected, then $\mathbb{G}_p(\mathbb{F}_p)$ is an almost direct product of quasi-simple groups of Lie type over $\mathbb{F}_q$, where q is a bounded power of p. It then follows from Remark 3.4 that every proper subgroup H of a quotient G of $\mathbb{G}_p(\mathbb{F}_p)$ has index at least $|G|^\eta$ in G for some $\eta = \eta(\mathbb{G}) > 0$ independent of p and of the quotient G. We may then take as above $\delta(\varepsilon) = \min\{\eta, 1/(C+1)\}\varepsilon$, where $C \geqslant 1$ is the constant in the above proposition and apply this proposition to $K = |G|^\delta$ to obtain Assumption (ii) of the Bourgain-Gamburd machine (Proposition 3.1).

More generally we can handle a bounded number of quasi-simple factors. Namely if G is a (almost direct) product of at most r quasi-simple groups of Lie type of dimension at most d (so for instance if G is the reduction modulo $q := p_1 \cdot \ldots \cdot p_r$, for some distinct large primes $p_1, \ldots, p_r$ of some Zariski-dense subgroup of $\mathbb{G}(\mathbb{Q})$), then Assumption (ii) of the Bourgain-Gamburd machine is still satisfied with say $\delta := \min\{\eta(d)/2r, 1/(2C)\}\varepsilon$. Here $\eta(d) > 0$ denotes the constant of quasi-randomness (see Remark 3.4) such that every proper subgroup of a quasi-simple group S of Lie type of dimension at most d has index at least $|S|^\eta$, r is the number of quasi-simple factors of G and C is the constant from Theorem 4.13.

To verify that these constants indeed work, split G as a product G_1G_2, where G_1 is the product of the quasi-simple factors of size at most $|G|^{\varepsilon/2r}$. Given a

$|G|^\delta$-approximate subgroup A of G, apply Theorem 4.13 to $\pi_2(A)$, the projection of A to G_2. Then either $|\pi_2(A)| \geqslant |G_2|/|G|^{C\delta}$, in which case $|A| \geqslant |\pi_2(A)| \geqslant |G|/(|G_1||G|^{C\delta}) \geqslant |G|^{1-\varepsilon}$, because $|G_1| \leqslant |G|^{\varepsilon/2}$ and $C\delta \leqslant \varepsilon/2$; or $\pi_2(A)$ is covered by at most $|G|^{C\delta}$ translates of a proper subgroup of G_2. However proper subgroups of G_2 have index at least $|S|^\eta$, where S is a quasi-simple factor of G_2, hence have index at least $|G|^{\eta\varepsilon/2r}$. It follows that A itself is covered by at most $|G|^{C\delta} \leqslant [G:H]^\varepsilon/|G|^\delta$ translates of a proper subgroup H of G. We are done.

To summarize the above discussion, we have proved in particular:

Corollary 4.14 *If $\mathbb{G}$ is a semisimple simply connected algebraic group defined over $\mathbb{Q}$ and $p_1, \dots, p_r$ distinct large enough primes, then Assumption (ii) of Proposition 3.1 holds for $G_0 := \prod_{i=1}^r \mathbb{G}_{p_i}(\mathbb{F}_{p_i})$ with $\delta = \varepsilon/Dr$, for some constant $D > 0$ depending only on the dimension of the algebraic group $\mathbb{G}$ and not on the p_i's.*

5 Random matrix products

The theory of random matrix products is a well developed part of probability theory on groups. It aims at understanding the statistical behavior of products of n matrices chosen at random when n tends to infinity. It is customary to restrict attention to the case when the matrices are independent and chosen according to the same probability distribution.

In order to establish the non-concentration estimate in the Bourgain-Gamburd machine (i.e., Assumption (iii) in Proposition 3.1) we will need the following result:

Theorem 5.1 (Probability of return to a subgroup [15]) *Let $\mathbb{G}$ be a connected semisimple algebraic group over a field K of characteristic zero and $\Gamma \leqslant \mathbb{G}(K)$ a Zariski-dense subgroup generated by a finite set S. Let μ be a probability measure on S with $\mu(s) > 0$ for each $s \in S$. Then there is a positive constant $c > 0$ such that for every integer $n \geqslant 1$,*

$$\mu^n(\mathbb{H}) < e^{-cn},$$

uniformly for every proper closed algebraic subgroup $\mathbb{H}$ of $\mathbb{G}$.

We will not go here into all the details of the proof of Theorem 5.1 and instead refer the reader to [15]. However we will indicate how the theory of random matrix products is used to derive it. Theorem 5.1 is deduced from the following fact proved in [15].

Proposition 5.2 (Probability of fixing a line) *Let K be a local field of characteristic zero and μ a probability measure on $\mathrm{GL}_d(K)$ such that $\max\{\|g\|, \|g^{-1}\|\}^\varepsilon$ is μ-integrable for some $\varepsilon > 0$. Assume that the support of μ generates a subgroup Γ_μ which is not relatively compact in projection to $PGL_d(K)$ and does not preserve any finite union of proper vector subspaces of K^d. Then there is $c > 0$ such that for every $n \geqslant 1$ and every line $x \in \mathbb{P}(K^d)$,*

$$\mu^n(\{g \in \mathrm{GL}_d(K); g(x) = x\}) < e^{-cn}.$$

The condition that the support of μ does not preserve any finite union of proper subspaces is usually called *strong irreducibility*. It is equivalent to asking that every subgroup of finite index in Γ_μ acts irreducibly, or that the connected component of the Zariski-closure of Γ_μ acts irreducibly. This condition was introduced by Furstenberg in the 1960's in his study of random matrix products [28]: he showed that under the conditions of the proposition, if μ is supported on $\mathrm{SL}_d(k)$, then the first Lyapunov exponent of μ is positive, namely:

$$\lim \frac{1}{n} \int \log \|g\| \, d\mu^n(g) > 0.$$

Another key theorem in the theory of random matrix products is the simplicity of the Lyapunov spectrum, due to Guivarc'h and Raugi [34]. It states that under the assumptions of proposition, if the subgroup Γ_μ is proximal (by definition this means that the semigroup $K\Gamma_\mu$ contains a rank one matrix in its closure in the algebra of $d \times d$ matrices $M_d(K)$), then the second Lyapunov exponent is strictly smaller than the first. In other words the random matrix product will almost surely contract almost all of the projective space $\mathbb{P}(k^d)$ into an exponentially small neighborhood of a point. From this the conclusion of Proposition 5.2 can be easily obtained. However this requires the proximality assumption and this assumption does not always hold. It holds for measures μ supported on Zariski-dense subgroups of $\mathrm{SL}_d(\mathbb{R})$ due to work of Goldsheid-Margulis [30] and this was used by Bourgain and Gamburd in their work [6]. But it does not hold in general in particular if we replace $\mathbb{R}$ with a p-adic field. So one needs to avoid this assumption if one wishes to establish Proposition 5.2 in full generality (and this generality is require to get Theorem 5.1). This is what is done in [15].

Let us now explain how to derive Theorem 5.1 from Proposition 5.2. First we claim that we may assume that K is a local field and that Γ is not relatively compact in $\mathbb{G}(K)$. To see it, first note that we may assume K to be finitely generated over $\mathbb{Q}$, since K can be taken to be generated by the matrix entries of the elements of the finite generating set S. Now pick a semisimple element of infinite order in Γ (it always exists, because Γ is Zariski-dense in $\mathbb{G}$) and let λ be one of its eigenvalues of infinite order. Find an absolute value on an algebraic closure of K, which is not equal to one on λ and consider the associated completion to obtain the desired local field. This argument is standard, details can be found in [99, Lemma 4.1].

Note that passing to a finite extension of K if necessary, we may assume that $\mathbb{G}$ is K-split, so that each absolutely irreducible module of $\mathbb{G}$ can be defined over K. Next, we claim that there are a finite number of absolutely irreducible finite dimensional representations of $\mathbb{G}$, say $\rho_1, \ldots, \rho_k$, each of dimension at least 2, such that every proper closed algebraic subgroup $\mathbb{H}$ of $\mathbb{G}$ must stabilize a line in one of these representations. This claim was already verified in the proof of Lemma 2.6 above.

Finally, note that we may apply Proposition 5.2 to each $\rho_i(\Gamma)$, because $\rho_i(\Gamma)$ acts strongly irreducibly on the representation space of ρ_i and is not relatively compact modulo scalars, because it is non relatively compact and of determinant 1 since $\mathbb{G}$ is semisimple. Since there are only finitely many ρ_i's to consider, we get the desired uniformity in $\mathbb{H}$ and Theorem 5.1 is proved.

6 Proof of the super-strong approximation theorem

In this section we verify that the ingredients of the expansion machine (i.e., Proposition 3.1) are all met under the assumptions of Theorem 1.2 and complete the proof of this theorem.

In view of Proposition 3.1, we see that Theorem 1.2 will follow from Proposition 3.1 applied to the groups $G_0 := \mathbb{G}_p(\mathbb{F}_p)$ with generating sets S_p, where S_p is the reduction modulo p of the generating set S of the Zariski-dense subgroup $\Gamma \leqslant \mathbb{G}(\mathbb{Q})$, provided the three assumptions of Proposition 3.1 are fulfilled. We saw in subsection 4.3 that Assumption (ii), the classification of approximate subgroups, is satisfied. Let us now consider Assumption (i).

Proposition 6.1 (High multiplicity/Quasirandomness) *Let $\mathbb{G}$ be a semisimple and simply connected algeraic group defined over $\mathbb{Q}$ and p a large enough prime. Then every non-trivial irreducible representation $\rho : \mathbb{G}_p(\mathbb{F}_p) \to \mathrm{GL}_d(\mathbb{C})$ of $G = \mathbb{G}_p(\mathbb{F}_p)$ has dimension at least $|G|^\beta$, where $\beta > 0$ depends only on the dimension of $\mathbb{G}$.*

Proof As observed by Sarnak-Xue [92] and Gamburd [29], this goes back to Frobenius in the case of SL_2. In [57] Landazuri and Seitz proved that all non-trivial irreducible projective representations of a finite simple group of Lie type have dimension at least $|G|^\beta$ for some $\beta > 0$ depending only on the rank, which implies the analogous claim for irreducible linear representations of any quasi-simple group. Actually, since we do not need the best possible β, we can arrive to this conclusion rather quickly if we observe that (see, e.g., [62, Theorem 4.1]) with the exception of the Suzuki groups, every quasi-simple finite group of Lie type contains a copy of either $\mathrm{SL}_2(\mathbb{F}_q)$ or $\mathrm{PSL}_2(\mathbb{F}_q)$. But both the Suzuki case and $\mathrm{PSL}_2(\mathbb{F}_q)$, can be handled easily (see [57, Lemma 4.1]).

Now $\mathbb{G}_p(\mathbb{F}_p)$ is an almost direct product of quasi-simple groups over $\mathbb{F}_q$, with $q = p^f$ and f is bounded in terms of the dimension of $\mathbb{G}$ only. So any non trivial linear representation of $\mathbb{G}_p(\mathbb{F}_p)$ gives rise to a representation of a quasi-simple group of Lie type over $\mathbb{F}_{p^f}$ with f and rank bounded in terms of $\dim \mathbb{G}$ only. The proposition follows. □

Remark Tim Gowers called a group *quasi-random* if it has the property sought for in this proposition. In such groups large subsets behave in a quasi-random way in the sense that the (non-abelian) non trivial characters of the indicator function of a subset are always very small [32]. This was used by Gowers to show that product-free sets (i.e., subsets $A \subset G$ not containing any x, y, z with $xy = z$) in such groups are small.

It now remains to verify Assumption (iii) of the Bourgain-Gamburd machine. This is usually the most difficult step. Here it will follow easily from the combination of the quantitative version of the strong approximation theorem proved in Section 2 and the large deviation estimates from the theory of random matrix products recalled in the previous section.

We may assume that $\mathbb{G} \leqslant \mathrm{GL}_d$ and this allows us to define the height $H(\gamma)$ of an element of $\Gamma \leqslant \mathbb{G}(\mathbb{Q})$ as in Theorem 2.3. In follows from (2.1) that for every $n \geqslant 1$

and every $\gamma \in S^n$,

$$H(\gamma) \leqslant (dM_S)^{nd^2}, \tag{6.1}$$

where we recall that M_S is defined as

$$M_S = \max\{H(s), s \in S\} \tag{6.2}$$

and the height $H(s)$ is the naive height (maximum of the numerator and denominator of each matrix entries written as an irreducible fraction).

Fix $\tau > 0$ to be determined below. Let p_0 be defined as in Theorem 2.3 and $p > p_0$ be any prime number. Choose an even integer n between $\tau \log p$ and $2\tau \log p$. Now let H be a proper subgroup of $\mathbb{G}_p(\mathbb{F}_p)$, and $S_{H,n}$ be the subset of all elements in S^n whose reduction modulo p lies in H. From (6.1) we see that if $\tau < 1/(2C_0 d^2 \log(dM_S))$, then

$$p > (M_{S_{H,n}})^{C_0},$$

where C_0 is the constant arising in Theorem 2.3. Hence Theorem 2.3 applies to the symmetric set $S_{H,n}$ and we conclude that the subgroup generated by $S_{H,n}$ is not Zariski-dense in $\mathbb{G}$. Let $\mathbb{H}$ be its Zariski-closure.

From Theorem 5.1 we know that in Γ and for all $n \geqslant 1$,

$$\mu_S^n(\mathbb{H}) \leqslant e^{-cn},$$

where $c > 0$ is a positive constant independent of the choice of $\mathbb{H}$. However, the reduction mod p map from Γ to $\mathbb{G}_p(\mathbb{F}_p)$ is injective on all elements of height at most p, and hence on S^n, thanks to our choice of n (of size roughly $\tau \log p$). Therefore

$$\mu_{S_p}^n(H) = \mu_S^n(\mathbb{H}) \leqslant e^{-cn} \leqslant 1/p^{\tau c} \leqslant 1/|\mathbb{G}_p(\mathbb{F}_p)|^{\kappa},$$

where we have set $\kappa = c\tau/2d^2$, because $|\mathbb{G}_p(\mathbb{F}_p)| \leqslant p^{d^2}$. In particular we see that the exponent κ can be taken of the form $c_1 c/\log M(S)$, where $c_1 > 0$ depends only on $\mathbb{G}$ and c is the constant from Theorem deviation. This establishes the non-concentration estimate needed in the Bourgain-Gamburd machine (Assumption (iii)) and ends the proof of the super-strong approximation theorem (Theorem 1.2).

Remark 6.2 (Explicit estimate on the gap) The proposed proof of Theorem 5.1 is non effective (it uses the ergodic theorem in Proposition 5.2) and hence gives no explicit lower bound on c. However it is likely that c is in fact independent of the choice of S provided $|S|$ is bounded. In that case the estimate given by Proposition 3.1 would give the following lower bound for the spectral gap:

$$\lambda_1 \geqslant 1/M_S^{O(1)},$$

where M_S (see (6.2)) is the maximal height of an element of S and the implied constant depends only on $\mathbb{G}$ and the cardinal of S. See [54] for an explicit upper bound on the implied constant in the special case when S belongs to $\mathrm{SL}_2(\mathbb{Z})$.

6.1 Several prime factors

The case of several (but boundedly many primes) can be handled at little additional cost. Assumptions (i) and (ii) of Prop. 3.1 have already been verified in this more general setting (see §4.3). Assumption (iii) follows in the same way as before by projecting the proper subgroup H to the largest simple factor where it remains proper. The corresponding bound on κ and thus on the λ_1 will depend on the number of prime factors involved.

Hence we get the following improved version of Theorem 1.2.

Theorem 6.3 *Suppose $\mathbb{G}$ is a connected and simply connected semisimple algebraic group defined over $\mathbb{Q}$ and let $\Gamma \leqslant \mathbb{G}(\mathbb{Q})$ be a Zariski-dense subgroup generated by a finite set S. Let also $r \in \mathbb{N}$. Then there is $\varepsilon = \varepsilon(S, r) > 0$ such that for all large enough distinct prime numbers $p_1, \ldots, p_r$, the projection of Γ in the finite group $G_0 := \prod_{i=1}^{r} \mathbb{G}_{p_i}(\mathbb{F}_{p_i})$ is surjective and the induced Cayley graph of G_0 is an ε-expander.*

Note that the spectral gap in this result depends on r but not on the choice of the r primes $p_1, \ldots, p_r$. Here again, if $\mathbb{G}$ is not assumed simply connected, then the projection of Γ to G_0 may not be surjective, but it has bounded index (depending only on $\mathbb{G}$ and r) in G_0 and the induced Cayley graph of the image remains an ε-expander. One reduces easily to the simply connected case by lifting to Γ to the simply connected cover of $\mathbb{G}$ (see, e.g., [68, p.399–418]).

Remark 6.4 (Groups defined over a number field) If $\mathbb{Q}$ is replaced by a number field K, then a similar result holds, which can be reduced to the case of $\mathbb{Q}$. If one wants to take quotients modulo prime ideals $\mathcal{P}$ of the ring of integers $\mathcal{O}_K$ of K, then one needs to be careful that the corresponding reduction may not be surjective on $\mathbb{G}(\mathcal{O}_K/\mathcal{P})$ (e.g., $\mathrm{SL}_2(\mathbb{Z})$ is Zariski-dense in SL_2, but maps onto $\mathrm{SL}_2(\mathbb{F}_p)$ and not onto $\mathrm{SL}_2(\mathcal{O}_K/\mathcal{P}) \simeq \mathrm{SL}_2(\mathbb{F}_{p^f})$ for any prime $\mathcal{P}$ with residual degree $f > 1$.)

To palliate this problem, one needs either to pass to a smaller number field (e.g., the one generated by the traces of the elements of Γ) or to consider the Zariski-closure of the embedding of Γ under the restriction of scalars of $\mathbb{G}$ from K to $\mathbb{Q}$. This Zariski closure will be semisimple and Theorem 6.3 will apply. In case $\mathbb{G}$ is not simply connected, one can lift to the simply connected cover. At any case it will always be the case that if Γ is a Zariski-dense subgroup of $\mathbb{G}(K)$ for some number field K and semisimple algebraic K-group $\mathbb{G}$, then the quotients of Γ modulo prime ideals of $\mathcal{O}_K$ will be expanders. This follows readily, by restriction of scalars, from Theoremn 6.3.

Bourgain-Gamburd-Sarnak [7] for SL_2, Varjú [100] for SL_d and Salehi-Golsefidy-Varjú [87] in general for $\mathbb{G}$ perfect, went much further by establishing that the spectral gap can be made independent of r (for a given S). This however requires to prove Assumption (ii) of the Bourgain-Gamburd machine in this setting, hence to understand approximate subgroups of large products of quasi-simple finite groups of bounded rank. This lies much deeper and requires a delicate multi-scale analysis. They prove:

Theorem 6.5 (Salehi-Golsefidy-Varjú [87]) *Let $q_0 \in \mathbb{N}$ and $\Gamma = \langle S \rangle$ be a finitely generated subgroup of $\mathrm{GL}_d(\mathbb{Z}[1/q_0])$. Assume that the connected component of the Zariski closure of Γ is perfect. Then there is $\varepsilon = \varepsilon(d, S) > 0$ such that the Cayley*

graphs of the quotients $\pi_q(\Gamma)$ induced by the generating set S are ε-expanders uniformly over all square-free integers q co-prime to q_0. Here π_q is the reduction modulo q defined on rational numbers with denominator co-prime to q.

To finish, let us quote the following related by-product of the Bourgain-Gamburd method.

Proposition 6.6 ([22, Prop. 8.4]) *Let $r \in \mathbb{N}$ and $\varepsilon > 0$. Suppose $G = G_1 G_2$, where G_1 and G_2 are products of at most r finite simple (or quasisimple) groups of Lie type of rank at most r. Suppose that no simple factor of G_1 is isomorphic to a simple factor of G_2. If $x_1 = x_1^{(1)} x_1^{(2)}, \dots, x_k = x_k^{(1)} x_k^{(2)}$ are chosen so that $\{x_1^{(1)}, \dots, x_k^{(1)}\}$ and $\{x_1^{(2)}, \dots, x_k^{(2)}\}$ are both ε-expanding generating subsets in G_1 and G_2 respectively, then $\{x_1, \dots, x_k\}$ is δ-expanding in G for some $\delta = \delta(\varepsilon, r) > 0$.*

The assumption that no simple factor of G_1 be isomorphic to a simple factor of G_2 is necessary here, because otherwise $\{x_1, \dots, x_k\}$ may not generate. However what if we suppose it generates, is the conclusion still true without the assumption that G_1 and G_2 have no isomorphic factors (e.g., if $G_1 = G_2 = \mathrm{SL}_2(\mathbb{F}_p)$)? This is an open question.

7 The group sieve method

One of the leitmotives of the subject matter in this paper is the ability to study finite simple groups of Lie type as quotients of certain infinite linear groups and thereby to do geometry and analysis on infinite groups in order to derive properties of finite groups, such as the expander property of their Cayley graph. The purpose of the sieve method is to achieve the converse: to study infinite linear groups from the properties of their finite quotients.

In this concluding section, we describe this method, first by showing a very simple application of Theorem 1.2 to random matrix product theory, where only one prime is required, and then by describing the *group sieve lemma* of Lubotzky and Meiri and two of its applications to the study of *generic properties* in infinite linear groups.

7.1 Large deviations for subvarieties

One of the simplest example showing the power of Theorem 1.2 is the following theorem. It says that random walks on linear groups do not concentrate much on any algebraic subvariety.

Theorem 7.1 (Subvarieties are exponentially small) *Let K be a field of characteristic zero, $\Gamma \leqslant \mathrm{GL}_d(K)$ a non virtually solvable finitely generated subgroup and μ a probability measure whose support S is a finite symmetric generating subset of Γ. Let $\mathbb{G}$ be the Zariski closure of Γ, and R its solvable radical. Suppose $\mathcal{V}$ is an algebraic subvariety in GL_d such that $\dim(R(\mathcal{V} \cap \mathbb{G})) < \dim \mathbb{G}$. Then we have for all $n \geqslant 1$:*

$$\mu^n(\Gamma \cap \mathcal{V}) \leqslant c_0(\mathcal{V}) \cdot e^{-cn},$$

where $c_0(\mathcal{V}) > 0$ is a constant depending only on the complexity (i.e., degree) of $\mathcal{V}$, and $c > 0$ depends only on μ.

Note that we have already shown a special case of this theorem in Theorem 5.1 above. Theorem 5.1 claimed essentially the same result when the subvariety $\mathcal{V}$ is assumed to be an algebraic subgroup. Although a direct approach similar to the proof of the Larsen-Pink inequality (Prop. 4.10) might be successful in deriving Theorem 7.1 from Theorem 5.1, the sieve method here can be implemented without any effort (modulo standard reductions) and yields Theorem 7.1 as a direct consequence of the super-strong approximation theorem (Theorem 1.2) as we now show. This was already observed (and proved in a special case) in the original work of Bourgain-Gamburd [6, Corollary 1.1].

Proof We first reduce to the case when the Zariski-closure $\mathbb{G}$ of Γ is semisimple and defined over $\mathbb{Q}$. Taking the quotient modulo the solvable radical R, we may assume that $\mathbb{G}$ is semisimple (with connected component of the identity $\mathbb{G}^0$). Now since Γ is finitely generated, we may assume that the field K is finitely generated over $\mathbb{Q}$, hence is a finite algebraic extension of a purely transcendental extension of $\mathbb{Q}$ with a finite transcendence basis. One may then specialize and pick algebraic values for this transcendence basis in such way that the connected component of the Zariski closure of the resulting image group Γ', now a subgroup of $\mathrm{GL}_d(\overline{\mathbb{Q}})$, is still $\mathbb{G}^0$ (this follows from Lemma 2.6, see also [58] for a related statement). Now taking the restriction of scalars to $\mathbb{Q}$ we have reduced to the case when $K = \mathbb{Q}$ and $\mathbb{G}^0$ is semisimple.

It is enough to prove the result for n even, and hence replacing S with S^2 we may assume that 1 belongs to S (note that the subgroup generated by S^2 has finite index in Γ). Let then $\Gamma_0 := \Gamma \cap \mathbb{G}^0$. It is a subgroup of finite index in Γ which is Zariski dense in $\mathbb{G}^0$. Now pick a large prime p. For p large enough, we know by the super-strong approximation theorem (Theorem 1.2) that $(\Gamma_0)_p$, the reduction mod p of Γ_0, has bounded index in $\mathbb{G}^0_p(\mathbb{F}_p)$ and that its induced Cayley graph is an ε-expander for some $\varepsilon > 0$ independent of p. It follows that the reduction mod p of Γ, itself is a finite group G_p containing $(\Gamma_0)_p$ as a subgroup of bounded index and hence is also an ε'-expander for some $\varepsilon' > 0$ independent of p and depending only on ε, $\mathbb{G}$, and the index of Γ_0 in Γ. Moreover S_p is not contained in a coset of a proper subgroup of G_p, because $1 \in S_p$. By the random walk characterization of expanders (see Lemma 3.3 above), this means that random walks at any time larger than $C_\varepsilon \log p$ are very well equidistributed in the sense that if $n = [C_{\varepsilon'} \log |G_p|]$ say

$$\left| \mu_p^n(x) - \frac{1}{|G_p|} \right| \leqslant 1/|G_p|^{10}$$

for every $x \in G_p$. In particular

$$\mu^n(\mathcal{V}) \leqslant \mu_p^n(\mathcal{V} \bmod p) \leqslant \frac{|\mathcal{V}_p|}{|G_p|} + 1/|G_p|^9,$$

However the assumption on $\mathcal{V}$ implies that $|\mathcal{V}_p| \leqslant c_0(\mathcal{V}) p^{\dim \mathbb{G}-1}$ while $|G_p| = \Omega(p^{\dim \mathbb{G}})$ (see the Schwarz-Zippel lemma in [22]). If follows that

$$\mu^n(\mathcal{V}) \leqslant c_0(\mathcal{V}) \cdot O(1/p),$$

with the implied constant depending only on $\mathbb{G}^0$. Now given any large n, one needs only pick a prime p such that n is roughly of size $C_\varepsilon \log |G_p|$ and the result follows. □

For another method towards Theorem 7.1 and related partial results see the work of Aoun [1].

We now pass to a corollary of Theorem 7.1. In [2], R. Aoun showed a probabilistic version of the Tits alternative: he proved that two independent random walks on a non virtually solvable linear group eventually generate a free subgroup. In other words a generic pair of elements always generates a free subgroup. Combining Theorem 7.1 with Lemma 2.6 we can now assert that a generic pair of elements generates a Zariski-dense free subgroup, namely:

Corollary 7.2 (A generic pair generates a Zariski-dense free subgroup) *Under the assumptions of Theorem 7.1 assume further that the Zariski closure of* $\Gamma = \langle S \rangle$ *is connected semisimple. Let* $\mathcal{E}$ *be the set of pairs* (a, b) *in* $\Gamma \times \Gamma$ *such that the subgroup* $\langle a, b \rangle$ *is either not free, or not Zariski dense in* Γ*. Then there is* $c = c(\mu) > 0$

$$\mu^n \times \mu^n(\mathcal{E}) \leqslant e^{-cn}.$$

Proof Aoun's theorem [2] tells us that for some $c > 0$, $\mu^n \times \mu^n(\mathcal{NF}) \leqslant e^{-cn}$, where $\mathcal{NF}$ is the set of non-free pairs. Now applying Theorem 7.1 to the group $\Gamma \times \Gamma$ in $\mathbb{G} \times \mathbb{G}$ the measure $\mu \times \mu$ and subvariety $\mathcal{V} = \mathbf{X}$ from Lemma 2.6, we get the desired result. □

For related results, see Aoun's work [1] and Rivin's [84].

7.2 The group sieve lemma

The spectral gap for mod p quotients has been exploited by Rivin [83] and Kowalski [52] to perform sieving on arithmetic lattice subgroups. Prior to the new results on thin groups such as the super-strong approximation theorem, the spectral gap was known in a variety of cases for mod p or mod n quotients of arithmetic subgroups. Thanks to super-strong approximation (i.e., Theorem 1.2 or [87]), we can now perform this sieving on arbitrary Zariski-dense subgroups (i.e., thin subgroups).

In Theorem 7.1 we used only one prime number to show our non concentration estimate. The power of the sieve consists in taking advantage of several primes and using as a guiding principle that primes are essentially independent.

Lubotzky and Meiri [65] formulated the following elegant lemma, which gives a simple set of conditions to be fulfilled in order to get further genericity results (akin to Theorem 7.1 above) that may require more than one prime.

Lemma 7.3 (Group sieve lemma) *Let* $\Gamma = \langle S \rangle$ *be a group generated by a finite symmetric set* S *and* $\{N_i\}_{1 \leqslant i \leqslant N}$ *be a finite sequence of finite index normal subgroups. Set* $\pi_i : \Gamma \to \Gamma/N_i$ *the projection maps. Let* $\mathcal{Z} \subset \Gamma$ *be a subset of* Γ *and assume that there are positive constants* D, ε, α*, with* $\alpha \in (0, 1)$*, such that*

- $\mathrm{Cay}(\Gamma/(N_i \cap N_j), S \bmod N_i \cap N_j)$ *for* $i \neq j$ *are* ε*-expanders;*
- $\Gamma/(N_i \cap N_j) \simeq \Gamma/N_i \times \Gamma/N_j$ *for* $i \neq j$*;*
- $|\Gamma/N_i| \leqslant N^D$ *for all* $i = 1, \ldots, N$*;*
- $|\pi(\mathcal{Z})| \leqslant (1 - \alpha)|\pi_i(\Gamma)|$ *for all* $i = 1, \ldots, N$*.*

Then there is a constant $B = B(\varepsilon, D, \alpha) > 0$ such that for all $n \geqslant B \log N$,

$$\mu_S^n(\mathcal{Z}) \leqslant \frac{1}{N}.$$

As before we have denoted by μ_S the uniform probability measure on the finite symmetric generating set S. Note that only the last assumption involves the set $\mathcal{Z}$. In applying this lemma, typically the π_i will be the reduction maps modulo a prime p_i. It is crucial that the constant $B(\varepsilon, D, \alpha)$ depends only on these three parameters and not on Γ, nor the choice of the sequence $\{N_i\}_i$.

The proof of this lemma is quite short, but before we give it in full, let us comment on it a little. Let $S_n := Y_1 \cdot \ldots \cdot Y_n$ be the product of n independent random variables $Y_1, \ldots, Y_n$ on Γ all distributed according to the same probability distribution μ_S (the uniform distribution on the generating set S). The key feature of an expander graph is that random walks on them become equidistributed very fast. By the first item in the above lemma, the Cayley graph of $\Gamma/(N_i \cap N_j)$ is an ε-expander. Clearly this also implies that the quotients Γ/N_i and Γ/N_j are ε-expanders. Hence the distributions of $\pi_i(S_n)$ and $\pi_j(S_n)$ are very close to the uniform distribution on Γ/N_i and Γ/N_j respectively as long as $n \geqslant C_\varepsilon \log |\pi_i(\Gamma)|$, so in particular if $n \geqslant C_\varepsilon D \log N$ (thanks to the third item). By the second item the natural injection from $\Gamma/(N_i \cap N_j)$ to $\Gamma/N_i \times \Gamma/N_j$ is surjective. This implies that the joint distribution $(\pi_i(S_n), \pi_j(S_n))$ is also close to the uniform distribution, and hence that $\pi_i(S_n)$ and $\pi_j(S_n)$ are almost independent as random variables.

Suppose for a second that they were actually independent. Then quite obviously, using the fourth item in the last inequality:

$$\mathbb{P}(S_n \in \mathcal{Z}) \leqslant \mathbb{P}\left(\pi_i(S_n) \in \pi_i(\mathbb{Z}) \ \forall i \leqslant e^{n/C_\varepsilon D}\right) \leqslant (1-\alpha)^{e^{n/C_\varepsilon D}},$$

where $\mathbb{P}(\Omega)$ denotes the probability of the event Ω. We would thus get a super-exponential decay of the probability of belonging to $\mathcal{Z}$.

Of course joint independence is too much to hope for, but the expander property on $\Gamma/N_i \times \Gamma/N_j$ implies that the $\pi_i(S_n)$ are pairwise almost independent. Now the following classical result from basic probability theory (the second moment method) allows us to take advantage of this pairwise almost independence in order to derive a meaningful upper bound on $\mathbb{P}(S_n \in \mathcal{Z})$.

Lemma 7.4 *Let $X \geqslant 0$ be a real random variable with $\mathbb{E}(X^2) < \infty$ and $T \geqslant 1$ a parameter.*

(i) (1st moment method) $\mathbb{P}(X \leqslant T \cdot \mathbb{E}(X)) \geqslant 1 - 1/T$;

(ii) (2nd moment method) $\mathbb{P}(X \geqslant (1/T) \cdot \mathbb{E}(X)) \geqslant (1-1/T)^2 \mathbb{E}(X)^2 / \mathbb{E}(X^2)$.

Proof The first item is an instance of Chebychev's inequality:

$$\mathbb{P}(X \geqslant T \cdot \mathbb{E}(X)) \cdot T \cdot \mathbb{E}(X) \leqslant \mathbb{E}(X 1_{X \geqslant T \cdot \mathbb{E}(X)}),$$

while the second follows from Cauchy-Schwarz:

$$\Big(1 - \frac{1}{T}\Big)\mathbb{E}(X) \leqslant \mathbb{E}\Big(X 1_{X \geqslant \frac{1}{T} \cdot \mathbb{E}(X)}\Big) \leqslant \mathbb{E}(X^2)^{1/2} \mathbb{P}\Big(X \geqslant \frac{1}{T}\mathbb{E}(X)\Big)^{1/2}$$

□

Applying this lemma to the variable $X := \sum_{i=1}^N 1_{A_i^c}$, ($A_i^c$ being the complement of the event A_i), we obtain:

Fact (Exploiting pairwise almost independence): If $\{A_i\}_{1\leqslant i\leqslant N}$ are N events on a probability space, such that for some $\alpha, \delta > 0$,

- $\mathbb{P}(A_i) \leqslant 1-\omega$ for each $i = 1, \ldots, N$, and
- $\mathbb{P}(A_i \cap A_j) \leqslant \mathbb{P}(A_i)\mathbb{P}(A_j) + \delta$ for all $i \neq j$,

then

$$\mathbb{P}\Big(\bigcap_{1\leqslant i\leqslant N} A_i\Big) \leqslant \frac{1}{\omega^2}\left(\delta + \frac{3}{N}\right).$$

Proof Indeed, $\mathbb{P}(A_i^c) = 1 - \mathbb{P}(A_i) \geqslant \omega$, so $\mathbb{E}(X) \geqslant \omega N$ and by Lemma 7.4

$$1 - \mathbb{P}\Big(\bigcap_1^N A_i\Big) = \mathbb{P}(X \geqslant 1) \geqslant \mathbb{P}\left(X \geqslant \frac{1}{\omega N}\cdot \mathbb{E}(X)\right) \geqslant \left(1 - \frac{1}{\omega N}\right)^2 \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)},$$

while

$$\mathbb{E}(X^2) = \sum_i \mathbb{P}(A_i^c) + \sum_{i\neq j} \mathbb{P}(A_i^c \cap A_j^c) \text{ and } \mathbb{E}(X)^2 = \sum_i \mathbb{P}(A_i^c)^2 + \sum_{i\neq j} \mathbb{P}(A_i^c)\mathbb{P}(A_j^c).$$

Hence using that $\mathbb{P}(A_i^c \cap A_j^c) \leqslant \mathbb{P}(A_i^c)\mathbb{P}(A_j^c) + \delta$,

$$\mathbb{E}(X^2) - \mathbb{E}(X)^2 \leqslant \sum_i \mathbb{P}(A_i^c)\mathbb{P}(A_i) + \delta N(N-1) \leqslant N(1-\omega) + \delta N^2,$$

from which we deduce (using that $\mathbb{E}(X) \geqslant N\omega$) that

$$1 - \mathbb{P}\Big(\bigcap_1^N A_i\Big) \geqslant \left(1 - \frac{1}{\omega N}\right)^2 \left(1 - \frac{N(1-\omega) + \delta N^2}{(\omega N)^2}\right) \geqslant 1 - \frac{1}{\omega^2}\left(\delta + \frac{3}{N}\right)$$

as desired. □

We can now complete the proof of the group sieve lemma (i.e., Lemma 7.3):

Proof Note that we may assume that n is even, and thus replacing S by S^2 if necessary we may assume that S contains 1. Then by the random walk characterization of ε-expanders (Lemma 3.3) we know that the random walk $S_n = Y_1 \cdot \ldots \cdot Y_n$ is almost equidistributed in projection to each $\pi_i(\Gamma)$ as long as $n \geqslant C_\varepsilon \log|\Gamma/N_i|$, hence as soon as $n \geqslant C_\varepsilon D \log N$. In particular for all $x \in \pi_i(\Gamma)$:

$$\left|\mathbb{P}(\pi_i(S_n) = x) - \frac{1}{|\pi_i(\Gamma)|}\right| \leqslant \frac{e^{-n/C_\varepsilon}}{|\pi_i(\Gamma)|^{10}} \tag{7.1}$$

and for $i \neq j$, $x \in \pi_i(\Gamma)$ and $y \in \pi_j(\Gamma)$

$$\left|\mathbb{P}((\pi_i(S_n), \pi_j(S_n)) = (x,y)) - \frac{1}{|\pi_i(\Gamma)|\cdot|\pi_j(\Gamma)|}\right| \leqslant \frac{e^{-n/C_\varepsilon}}{|\pi_i(\Gamma)|^{10}|\pi_j(\Gamma)|^{10}} \tag{7.2}$$

Let A_i be the event "$\pi_i(S_n) \in \pi_i(\mathcal{Z})$". From (7.1) and (7.2) we get for $i \neq j$

$$\left|\mathbb{P}(A_i) - \frac{|\pi_i(\mathcal{Z})|}{|\pi_i(\Gamma)|}\right| \leqslant \frac{e^{-n/C_\varepsilon}}{|\pi_i(\Gamma)|^9},$$
$$\left|\mathbb{P}(A_j) - \frac{|\pi_j(\mathcal{Z})|}{|\pi_j(\Gamma)|}\right| \leqslant \frac{e^{-n/C_\varepsilon}}{|\pi_j(\Gamma)|^9},$$
$$\left|\mathbb{P}(A_i \cap A_j) - \frac{|\pi_i(\mathcal{Z})|}{|\pi_i(\Gamma)|}\frac{|\pi_j(\mathcal{Z})|}{|\pi_j(\Gamma)|}\right| \leqslant \frac{e^{-n/C_\varepsilon}}{|\pi_i(\Gamma)|^9|\pi_j(\Gamma)|^9}.$$

Hence

$$|\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j)| \leqslant 3e^{-n/C_\varepsilon}.$$

Recall further that by assumption $|\pi_i(\mathcal{Z})|/|\pi_i(\Gamma)| \leqslant 1-\alpha$ hence

$$\mathbb{P}(A_i) \leqslant 1 - \alpha + e^{-n/C_\varepsilon} \leqslant 1 - \alpha/2,$$

for n large enough. Setting $B(\varepsilon, D, \alpha) = 10C_\varepsilon D/\alpha^2$ (say), the lemma now follows by applying the Fact above with $\omega := \alpha/2$, $\delta = 3e^{-n/C_\varepsilon}$. □

In the next subsection, we give an application of this group sieve lemma to a counting problem in infinite linear groups.

To conclude we note that the pairwise almost independence given by the assumption that the Cayley graphs of $\Gamma/(N_i \cap N_j) \simeq \Gamma/N_i \times \Gamma/N_j$ are expanders corresponds to the super-strong approximation theorem for products of two prime factors (i.e. when $r = 2$ in Theorem 6.3). The result of Salehi-Golsefidy and Varjú [87] shows uniform expansion for an arbitrary (growing) number of prime factors. This corresponds to joint almost independence of the sequence $\pi_i(S_n)$ instead of pairwise. Clearly this is a much stronger property to have at one's disposal and it is crucial in the Affine Sieve of Bourgain-Gamburd-Sarnak [7] and Salehi-Golsefidy-Sarnak [86].

7.3 Proper powers in linear groups are scarce

In [65] Lubotzky and Meiri use the group sieve lemma (Lemma 7.3) above to establish the following result:

Theorem 7.5 (Proper powers are exponentially small, [65]) *Under the assumptions of Theorem 7.1, let $\mathcal{P}$ be the proper powers in Γ, i.e., the set of elements in $\gamma \in \Gamma$ such that there is $\gamma_0 \in \Gamma$ and $k \geqslant 2$ such that $\gamma = \gamma_0^k$. Then $\mathcal{P}$ is exponentially small, namely there is $c > 0$ such that for all $n \geqslant 1$,*

$$\mu^n(\mathcal{P}) \leqslant e^{-cn}.$$

An old result of Malcev (see [60] and references therein) says that for each $n \geqslant 1$, the set of n-th powers in any finitely generated nilpotent group contains a finite index subgroup, and thus cannot be exponentially small. So Theorem 7.5 can be seen as a strong quantitative converse to Malcev's theorem. Prior attempts to prove this result, see Hrushovski-Kropholler-Lubotzky-Shalev [40], could only go as far as proving that for each k, the set of k-powers in Γ does not contain a finite index subgroup of Γ.

We sketch the proof in the special case when Γ is a Zariski-dense subgroup of $\mathrm{SL}_d(\mathbb{Z})$.

Proof We want to apply the group sieve lemma to the subset $\mathcal{Z} := \mathcal{P}$ of proper powers. The projection maps π_i will be the reduction maps modulo large primes p_i to be chosen carefully. By the strong approximation theorem (Theorem 1.1 above) Γ maps onto $\mathrm{SL}_d(\mathbb{F}_p)$ for all large enough prime p.

In a finite group every element of order at least 3 is a proper power, so we have to restrict attention to m-powers (i.e., elements in the image of the map $g \mapsto g^m$) for each given m. Luckily we do not need to consider all m's, but only those with $m \leqslant Cn$ for some $C = C(S) > 0$. The reason is that if $\gamma \in \mathrm{SL}_d(\mathbb{Z})$ has an eigenvalue λ of modulus > 1, then it is of modulus $> 1 + \delta$ for some δ depending only on the dimension d (indeed eigenvalues are roots of the characteristic polynomial, which has degree d and integer coefficients: if all eigenvalues were say $\leqslant 2$ in modulus, then the coefficients would be bounded, leaving only finitely many possibilities for λ). So for every $m \geqslant 2$,

$$\|\gamma^m\| \geqslant |\lambda|^m \geqslant (1+\delta)^m,$$

while every element in the support of the measure μ^n has size at most M_S^n, where $M_S = \max\{\|s\|, s \in S\}$. So if an element g in the support of μ^n is a proper power γ_0^m, then either $m = O(n)$ or g has all its eigenvalues of modulus 1. Kronecker's lemma tells us that if the roots of a monic polynomial of degree d in $\mathbb{Z}[X]$ have all modulus 1, they must be roots of unity of degree at most d. Hence $g^{d!}$ must be a unipotent element, i.e., $(g^{d!} - 1)^d = 0$. However $\mathcal{V} := \{g \in \mathrm{SL}_d; g^{d!} \text{ is unipotent}\}$ is a proper algebraic subvariety of SL_d, and hence Theorem 7.1 tells us that this set is exponentially small and can be ignored. It follows that

$$\mu^n(\mathcal{P}) \leqslant \sum_{m \leqslant C(S)n} \mu^n\{\mathcal{P}_m\} + O(e^{-cn})$$

where $\mathcal{P}_m$ is the set of m-powers. We will then apply the group sieve lemma to each $\mathcal{P}_m$ separately.

Now given $m \geqslant 2$, how many m-powers are there in $SL_d(\mathbb{F}_p)$? If m is co-prime to the order of $\mathrm{SL}_d(\mathbb{F}_p)$, then every element is an m-power. So we wish to choose p in such a way that there are not too many m-powers. For example, assume that $p \equiv 1 \mod m$, so that m divides the order of the multiplicative group of $\mathbb{F}_p$, which is a cyclic group of order $p-1$. In $\mathbb{Z}/(p-1)\mathbb{Z}$ there are precisely $(p-1)/m$ multiples of m. So there are exactly $((p-1)/m)^{d-1}$ m-powers in the subgroup of $\mathrm{SL}_d(\mathbb{F}_p)$ made of diagonal matrices, which is a subgroup isomorphic to $(\mathbb{Z}/(p-1)\mathbb{Z})^{d-1}$. In particular at least $(p-1)^{d-1}/2$ of the diagonal matrices are not m-powers. Among them at most $(p-1)^{d-2}$ have two identical diagonal entries, i.e., at least $(p-1)^{d-1}/3$ of them (for p large) have distinct eigenvalues and thus a centralizer which is as small as possible, that is equal to the full diagonal group. In each conjugacy class of such a diagonal matrix, there are no more than $d!$ other such matrices. Taking the union of the conjugacy classes of these elements thus yields at least $|\mathrm{SL}_d(\mathbb{F}_p)|/3d!$ different elements that are not m-powers. Thus we have shown that for large p and any $m \geqslant 2$ with $p \equiv 1 \mod m$

$$|\{m\text{-powers in } \mathrm{SL}_d(\mathbb{F}_p)| \leqslant \left(1 - \frac{1}{3d!}\right) |\mathrm{SL}_d(\mathbb{F}_p)|$$

To apply Lemma 7.3 need now choose a sequence of distinct primes $\{p_i\}_{i=1,\dots,N}$ with N of exponential size in n. We choose one sequence of primes for each $m \leqslant Cn$.

Dirichlet's theorem ensures that there are infinitely many primes congruent to 1 mod m. More follows from the proof: there is in fact a positive density of such primes among the primes. However we need a uniform estimate as m is allowed to vary from 1 to n, while the primes we sieve with will be of exponential size in n. We need that there are exponentially many primes of exponential size congruent to 1 mod m uniformly in $m \leqslant Cn$. So one needs a fairly precise quantitative version of Dirichlet's theorem: we need to know that the number $\pi(x; m, 1)$ of primes congruent to 1 mod m and less than x is at least say $\sqrt{x}$ *uniformly* over all moduli $m \leqslant \log x$. The Siegel-Walfisz theorem says that the prime number theorem in arithmetic progressions is accurate uniformly for values m going up to $(\log x)^A$ for any given $A \geqslant 1$. But it is non-effective in the sense that the first x for which the estimate begins to be meaningful is not explicitly computable in terms of A due to the possible presence of Siegel zeros. In our case, we need only a much weaker lower bound on the number of such primes and the estimate

$$\pi(x; m, 1) = \frac{x}{\phi(m)} \left(1 + O(e^{-O((\log x)^{1/5})})\right)$$

holds uniformly for all $m \leqslant (\log x)^{3/2}$ with effective implied constants in the big O's, where $\phi(m)$ denotes the Euler function (see (7) on page 123 of Davenport's book [24]). In particular $\pi(x; m, 1) \geqslant \sqrt{x}$ for all $m \leqslant (\log x)^{3/2}$ and x large enough.

We can now finish the proof of Theorem 7.5 (in our special case of Zariski-dense subgroups of $\mathrm{SL}_d(\mathbb{Z})$). Let $B = B(\varepsilon, D, \alpha) > 0$ be the constant from the group sieve lemma (Lemma 7.3). Set $\alpha = 1/3d!$, $D = 2d^2$, and $\varepsilon = \varepsilon(S) > 0$ is given by the super-strong approximation theorem (Theorem 6.3 for $r = 2$ primes). Given a large n, and some $m \leqslant C(S)n$, by the above there are at least $\sqrt{x}$ distinct primes congruent to 1 mod m and smaller than $x := e^{2n/B}$. Pick a subset of roughly $N = e^{n/B}$ of them, and apply Lemma 7.3 to conclude that

$$\mu^n(\mathcal{P}_m) \leqslant e^{-n/B}$$

for each $m \leqslant C(S)n$. The result follows. □

Remark In the proof we used an effective version of the prime number theorem in progressions as opposed to the Siegel-Walfisz theorem, which is non-effective. This has only some sense if all other constants involved are indeed effective. The expander constant $\varepsilon > 0$ depends on the approximate subgroup constant δ from Proposition 3.1. It is effective since all the algebraic geometry bounds used in Section 4 are effective, although not really explicit (see in particular [80] where an attempt has been made to make some of these constants more explicit). Finally the first prime starting from which the super-strong approximation theorem holds is also effective as it relies on Nori's theorem (see the appendix of [87]) although far from explicit. So it is fair to say that the rate of exponential decay in Theorem 7.5, though effective, is far from explicit.

7.4 The generic Galois group is the Weyl group

Given a matrix in $\mathrm{SL}_d(\mathbb{Z})$, one may look at its characteristic polynomial and ask if it is irreducible over $\mathbb{Q}$. This amounts to say that the Galois group of the polynomial

acts transitively on the roots. More generally when is the Galois group equal to the full group of all permutations of the roots? When is it only a proper subgroup?

Prasad and Rapinchuk [77] have shown that given a Zariski-dense subgroup Γ of $\mathrm{SL}_d(\mathbb{Z})$, the subset of elements in Γ whose characteristic polynomial is irreducible, or even has full Galois group, is itself Zariski-dense in Γ, and even contains an entire coset of a certain finite index subgroup (see [77, Remark 6]). They proved their result in a much greater generality (for an arbitrary semisimple group) and we refer the reader to [78] and to the excellent surveys [79] and [76, §9] for a description of their work and several further interesting results on how to find many elements in Γ with various constraints on their characteristic polynomial.

Their method is also based on the study of the mod p quotients of Γ. By Jordan's lemma (see below Lemma 7.7), the Galois group is maximal if and only if it has elements from every conjugacy class of the symmetric group. It is thus enough to find one prime number per conjugacy class for which the associated Frobenius element modulo p is in that conjugacy class.

The same idea, this time combined with the group sieve lemma (Lemma 7.3) and the super-strong approximation theorem (Theorem 1.2), can be applied to show the following somewhat stronger result, due to Jouve, Kowalski and Zywina [45], which asserts that, the set of elements in Γ whose characteristic polynomial is not all of $\mathfrak{S}_d$ is exponentially small in the above sense of random walks: the probability that a random walk at time n hits this subset decays exponentially with n. Note that combined with Theorem 7.1, this also implies that the subset of elements in Γ with full Galois group is Zariski-dense.

Theorem 7.6 *Let $d \geqslant 2$ and $\Gamma = \langle S \rangle \leqslant \mathrm{SL}_d(\mathbb{Z})$ be a Zariski-dense subgroup of SL_d. Let as above μ_S denote the uniform probability measure on the symmetric set S. Then there is $c = c(S) > 0$ such that for all $n \geqslant 1$,*

$$\mu_S^n(\{\gamma \in \Gamma, \mathrm{Gal}(\gamma) \neq \mathfrak{S}_d\}) \leqslant e^{-cn}.$$

Here $\mathrm{Gal}(\gamma)$ denotes the Galois group of the extension $K_\gamma|\mathbb{Q}$, where K_γ is the splitting field of the characteristic polynomial of γ and $\mathfrak{S}_d$ denotes the symmetric group of all permutations of d elements. In particular

$$\mu_S^n(\{\gamma \in \Gamma, \pi_\gamma \text{ not } \mathbb{Q}\text{-irreducible}\}) \leqslant e^{-cn}.$$

We also refer the reader to the earlier work of Rivin [83, 84] for related statements and generalizations to other geometric contexts. And to the subsequent work of Gorodnik and Nevo [31], which proves a similar result (for arithmetic groups only) when counting with respect to a height function of $M_d(\mathbb{Z})$ instead of the random walk average.

Theorem 7.6 was proved by Jouve, Kowalski and Zywina [45] in the special case when Γ has finite index in $\mathrm{SL}_d(\mathbb{Z})$. When [45] was written the super-strong approximation theorem was still in limbo. Now that we have Theorems 1.2 and 6.3 at our disposal, we can use them in the argument from [45] and the whole proof goes through verbatim yielding Theorem 7.6 above. We give below the complete proof (see also [67]).

Jouve, Kowalski and Zywina proved their result in the wider generality of arithmetic subgroups of arbitrary connected semisimple groups (see below). Likewise, combined with the super-strong approximation, their argument extends to all Zariski-dense subgroups. It remains an open problem however to extend the Gorodnik-Nevo result to Zariski-dense subgroups.

In [67] Lubotzky and Rosenzweig extended these results to cover also non-connected semisimple algebraic groups and showed the interesting phenomenon that each coset of the connected component has its own generic Galois group, which may be different from the Weyl group of the connected component.

We now pass to the proof of Theorem 7.6.

Proof The method is based on the following classical lemma of Jordan:

Lemma 7.7 (Jordan) *Let G be a finite group and H a subgroup. If H is a proper subgroup of G, then some conjugacy class of G is disjoint from H.*

In other words, the only subgroup of G intersecting every conjugacy class is G itself. Looking at the action by left translations on the set of left cosets G/H, we see that the lemma is equivalent to the following assertion: every transitive subgroup of $\mathfrak{S}_d$ ($d \geqslant 2$) must contain a permutation with no fix points. For the proof of this simple lemma and a number of pretty applications to number theory, we refer the reader to Serre's short note [94].

We will apply this lemma with $G = \mathfrak{S}_d$ and $H = \mathrm{Gal}(\gamma)$. Set $\mathcal{Z} := \{\gamma \in \Gamma; \mathrm{Gal}(\gamma) \neq \mathfrak{S}_d\}$ and $\mathcal{Z}_C := \{\gamma \in \Gamma; \mathrm{Gal}(\gamma) \cap C = \varnothing\}$, where C denotes a conjugacy class in the symmetric group $\mathfrak{S}_d$. A conjugacy class C of $\mathfrak{S}_d$ is given by a partition of d as $d = d_1 + \ldots + d_k$ for integers $d_i \geqslant 1$. Jordan's lemma then tell us that

$$\mathcal{Z} = \bigcup_C \mathcal{Z}_C,$$

where the union ranges over all conjugacy classes of $\mathfrak{S}_d$. Thus for proving Theorem 7.6 it will suffice to show that each $\mathcal{Z}_C$ is exponentially small. We will apply the group sieve lemma (Lemma 7.3 above) to show precisely this.

As is well-known, to every prime p not dividing the discriminant of π_γ, one can associate a particular conjugacy class of $\mathrm{Gal}(\gamma)$, the Frobenius conjugacy class $\mathrm{Frob}_p(\pi_\gamma)$. The prime ideals above p in the splitting field K_γ are permuted transitively by $\mathrm{Gal}(\gamma)$. Each stabilizer subgroup is in bijection with the Galois group of the reduced polynomial $\pi_\gamma \bmod p$ in $\mathbb{F}_p[X]$, which is a cyclic group generated by the Frobenius element mapping x to x^p in the corresponding residue field extension $\mathbb{F}_p[X]/(\pi_\gamma \bmod p)$. The corresponding elements in each stabilizer (decomposition) subgroup form the conjugacy class $\mathrm{Frob}_p(\pi_\gamma)$ in $\mathrm{Gal}(\gamma)$. The Frobenius element permutes the roots of $\pi_\gamma \bmod p$ and its decomposition into a product of disjoint cycles corresponds to the factorization

$$\pi_{\gamma \bmod p} = \pi_\gamma \bmod p = P_1 \cdot \ldots \cdot P_k$$

into irreducible polynomials in $\mathbb{F}_p[X]$ with one cycle of length $\deg(P_i)$ for each $i = 1, \ldots, k$. It determines a conjugacy class of $\mathfrak{S}_d$ identified by the partition of d given by $d = \deg(P_1) + \ldots + \deg(P_k)$.

Let C be a conjugacy class of $\mathfrak{S}_d$ determined by a partition $d = d_1 + \ldots + d_k$ of d. From the above discussion, we see that if $\gamma \in \mathcal{Z}_C$ and p is a prime, then either the discriminant of π_γ is divisible by p and γ mod p has a multiple eigenvalue, or γ mod p is contained in the set of elements $g \in \mathrm{SL}_d(\mathbb{F}_p)$ whose characteristic polynomial is without multiple roots (i.e., g is regular semisimple) and whose factorization into irreducible polynomials in $\mathbb{F}_p[X]$ determines a partition of d different from the partition associated to C.

The set of elements with a multiple eigenvalue (i.e., non regular semisimple elements) forms a proper subvariety of SL_d of bounded degree (it is defined by the vanishing of the gcd of the characteristic polynomial and its derivative). The Lang-Weil bound, or the easier Schwarz-Zippel estimate (see [22]), allows us to assert that this set has size $O(p^{d^2-2})$, while $\mathrm{SL}_d(\mathbb{F}_p)$ has size at least $\Omega(p^{d^2-1})$, and is thus negligible. Consider now the second set.

To apply the group sieve lemma (Lemma 7.3) to the set $\mathcal{Z}_C$, it thus remains to show a uniform upper bound on the proportion of $\mathrm{SL}_d(\mathbb{F}_p)$ the set of such elements can occupy. Or, equivalently, to prove a uniform lower bound on the size of the set $\Omega_{p,C}$ of regular semisimple elements in $\mathrm{SL}_d(\mathbb{F}_p)$ whose characteristic polynomial admits a factorization of the form dictated by the partition of d associated to C.

It is easy to obtain such a lower bound. Every monic polynomial with constant term $(-1)^d$ is the characteristic polynomial of some matrix in $\mathrm{SL}_d(\mathbb{F}_p)$, e.g., the companion matrix of the polynomial. So, given C, just pick a polynomial whose irreducible factors are distinct and whose degrees d_i's are such that $d = d_1 + \ldots + d_k$ is the partition associated to C. Let g be the associated companion matrix. It belongs to $\Omega_{p,C}$ and so do all its conjugates. It is a regular semisimple element of $\mathrm{SL}_d(\mathbb{F}_p)$ and thus it belongs to a unique maximal torus T. All other regular semisimple elements in T have the same associated partition of d, because they generate the same commutative subalgebra of matrices over $\mathbb{F}_p$. It follows that $\Omega_{p,C}$ contains $\bigcup_{g\in \mathrm{SL}_d(\mathbb{F}_p)} gT^{reg}g^{-1}$, where T^{reg} denotes the subset of regular elements in T (i.e., with distinct eigenvalues). Hence

$$|\Omega_{p,C}| \geqslant \frac{|\,\mathrm{SL}_d(\mathbb{F}_p)|}{|N(T)/T|} - |\{g \in SL_d(\mathbb{F}_p); g \text{ not regular semisimple}\}|,$$

where $N(T)$ is the normalizer of T. Now $N(T)/T$ is the Weyl group of SL_d, thus isomorphic to $\mathfrak{S}_d$. As already mentioned the set of non regular semisimple elements in $\mathrm{SL}_d(\mathbb{F}_p)$ is negligible (being of size $O(|SL_d(\mathbb{F}_p)|/p)$). Hence $|\Omega_{p,C}| \geqslant |SL_d(\mathbb{F}_p)|/2d!$ say when p is large enough.

To conclude the proof of Theorem 7.6, it remains to apply the group sieve lemma (Lemma 7.3) to the sets $\mathcal{Z}_C$ for each conjugacy class C of $\mathfrak{S}_d$ and to the group Γ with projection homomorphisms π_i given by the reduction modulo N primes p_i of size at most N^2 say, where N is chosen of size $e^{n/B}$, with $B = B(\varepsilon, D, \alpha) > 0$ is the constant given by Lemma 7.3 with $D := 3d^2$, $\alpha := 1/2d!$ say, and $\varepsilon = \varepsilon(S) > 0$ is given by the super-strong approximation theorem for two primes (Theorem 6.3). This ends the proof. □

In their paper Jouve, Kowalski and Zywina prove (the correct modified version of) Theorem 7.6 in the more general setting where the ambient group is a connected

semisimple algebraic group defined (and not necessarily split) over a number field. Again while they treated only arithmetic subgroups, because the super-strong approximation theorem was not available to them, their method extends and applies to all Zariski dense subgroups. This was worked out by Lubotzky and Rosenzweig [67], who also described in full the most general situation, when the field of definition is only assumed to be finitely generated over $\mathbb{Q}$ and, most interestingly, the algebraic group may not be connected nor semisimple. Without reaching out for the greatest generality, we will only state their theorem for split connected semisimple groups defined over a field of characteristic zero. In order to do so we first give some background on the Galois action on tori (see also [76], [45]).

Let the ambient group $\mathbb{G}$ be a connected semisimple algebraic group defined and split over some finitely generated field K of characteristic zero. This means that $\mathbb{G}$ admits a maximal torus T_0 which is defined and diagonalizable in any linear representation of $\mathbb{G}$ over K. To every regular semisimple element g in $\mathbb{G}(K)$ corresponds the unique maximal K-torus T_g it contains. A priori T_g is not diagonalizable over K, but there is a smallest finite extension of K, the splitting field K_{T_g} of T_g such that T_g is conjugate over K_{T_g} to the K-split (i.e., diagonalizable) torus T_0. The Galois group $\mathrm{Gal}(g)$ of the extension $K_{T_g}|K$ acts on the group $X(T_g)$ of characters of T_g. The group $X(T_g)$ is the free abelian group of rank $r = \mathrm{rank}(\mathbb{G})$ made of algebraic homomorphisms from T_g to the multiplicative group $\mathbb{G}_m$. The Galois action of $\mathrm{Gal}(g)$ on X_{T_g} is via the formula

$$\sigma(\chi(t)) = {}^{\sigma}\chi(\sigma(t)).$$

This action is faithful and thus $\mathrm{Gal}(g)$ can be viewed as a finite subgroup of $\mathrm{Aut}(X(T_g)) \simeq \mathrm{GL}_r(\mathbb{Z})$.

The Weyl group $W(T_g) := N(T_g)/Z(T_g)$ of T_g, where $N(T_g)$ is the normalizer and $Z(T_g) = T_g$ the centralizer of T_g, can also be viewed as a subgroup of $\mathrm{Aut}(X(T_g))$ using the action by conjugation of the normalizer $N(T_g)$, namely

$$\chi \mapsto (t \mapsto \chi(n^{-1}tn)),$$

for $n \in N(T_g)$ and $t \in T_g$.

Under the identification, it turns out that $\mathrm{Gal}(g)$ becomes a subgroup of the Weyl group $W(T_g)$: indeed fixing a K-split maximal torus T_0, there is an element $x \in \mathbb{G}(\overline{K})$ such that $T_g = xT_0x^{-1}$, because all maximal tori are conjugate over the algebraic closure $\overline{K}$ of K. Now from the fact that T_g is defined over K, we see that $n_\sigma := \sigma(x)x^{-1}$ belongs to $N(T_g)$, and that ${}^{\sigma}\chi(t) = \chi(n_\sigma^{-1}tn_\sigma)$ for all $t \in T_g$. Recall that the isomorphism class of $W(T)$ is independent of T, it is the Weyl group $W_{\mathbb{G}}$ of $\mathbb{G}$. When $\mathbb{G} = \mathrm{SL}_d$, then $W_{\mathbb{G}} \simeq \mathfrak{S}_d$.

We can now state the theorem of Jouve, Kowalski and Zywina [45] in the version proved by Lubotzky and Rosenzweig [67] (i.e., for Zariski-dense subgroups over fields of characteristic zero and not merely arithmetic groups over number fields).

Theorem 7.8 *Let $\mathbb{G}$ be a connected semisimple algebraic group defined and split over K, a finitely generated field extension of $\mathbb{Q}$. Suppose $\Gamma \leqslant \mathbb{G}(K)$ is a Zariski-dense subgroup and μ a symmetric probability measure whose support is a finite generating subset of Γ. Then there is $c > 0$ such that*

$$\mu^n(\gamma \in \Gamma; \mathrm{Gal}(\gamma) \lneq W(T_\gamma)) \leqslant e^{-cn}.$$

Here again, this implies (via Theorem 7.1) that the set of elements γ with $\mathrm{Gal}(\gamma) = W(T_\gamma)$ is Zariski-dense in Γ, a fact first established by Prasad and Rapinchuk in [77].

The proof of Theorem 7.8 follows the same sieving argument as in the special case of subgroups of $\mathrm{SL}_d(\mathbb{Z})$ presented above. Using a specialization argument Lubotzky and Rosenzweig reduce to the case when K is a number field. Then the group sieve lemma together with the super-strong approximation theorem (applied to the reduction of scalars of $\mathbb{G}$ from K to $\mathbb{Q}$, see Remark 6.4) apply in a similar way.

If $\mathbb{G}$ is not split over the base field K, or if it is not connected, then the theorem still holds, but the generic Galois group of an element γ may no longer be the Weyl group (in the connected non split case, the Weyl group appears only as a subgroup) and it will depend (only) on the coset of the connected component of $\mathbb{G}$ it lives in. See [76], [45] and [67] for this and further information about the generic $\mathrm{Gal}(\gamma)$.

Acknowledgements It is a pleasure to thank Florent Jouve, Alex Lubotzky, Laci Pyber, Andrei Rapinchuk, Lior Rosenzweig, Peter Sarnak and Terry Tao for their comments on an earlier version of this article. I am also grateful to Edmund Robertson and Colin Campbell for their patience and the gentle reminders that helped me finish this article.

References

[1] R. Aoun, Transience of algebraic varieties in linear groups and application to generic Zariski density, *Annales de l'Institut Henri Poincaré* (to appear), arXiv:1103.0944.

[2] R. Aoun, Random subgroups of linear groups are free, *Duke Math. J.* **160** (2011), 117–173.

[3] A. Balog & E. Szemerédi, A statistical theorem of set addition, *Combinatorica* **14** (1994), 263–268.

[4] A. Borel, *Linear Algebraic Groups (2nd ed.)*, Grad. Texts Math. **126**, Springer-Verlag, New York, 1991.

[5] J. Bourgain & A. Gamburd, Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$, *Ann. of Math. (2)* **167** (2008), 625–642.

[6] J. Bourgain & A. Gamburd, Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$, II, *J. Eur. Math. Soc.* **11** (2009), 1057–1103, with an appendix by Bourgain.

[7] J. Bourgain, A. Gamburd, & P. Sarnak, Affine linear sieve, expanders, and sum-product, *Invent. Math.* **179** (2010), 559–644.

[8] J. Bourgain, N. Katz & T. Tao, A sum-product estimate in finite fields, and applications, *Geom. Funct. Anal.* **14** (2004), 27–57.

[9] J. Bourgain & A. Kontorovich, On Zaremba's conjecture, *C. R. Math. Acad. Sci. Paris* **349** (2011), 493–495.

[10] E. Breuillard, Expander graphs, property τ and approximate groups, in *2012 PCMI Summer School Lecture Notes, Park City-IAS*, Amer. Math. Soc.

[11] E. Breuillard, An exposition of Camille Jordan's original proof of his theorem on finite subgroups of invertible matrices, available from author's website `http://www.math.u-psud.fr/~breuilla/Jordan.pdf`.

[12] E. Breuillard, Lecture notes on approximate groups, from a course given at IHP, Paris 2011, available from author's website.

[13] E. Breuillard, A height gap theorem for finite subsets of $\mathrm{GL}_d(\overline{\mathbb{Q}})$ and nonamenable subgroups, *Ann. of Math. (2)* **174** (2011), 1057–1110.

[14] E. Breuillard, A brief introduction to approximate groups, in *Thin Groups and Super Stong Appoximation*, eds. E. Breuillard & H. Oh, MSRI Publications 61, 2014.

[15] E. Breuillard, A non concentration estimate for random matrix products, 2014, in preparation.

[16] E. Breuillard & A. Gamburd, Strong uniform expansion in $\mathrm{SL}(2,p)$, *Geom. Funct. Anal.* **20** (2010), 1201–1209.
[17] E. Breuillard, B. Green & T. Tao, Small doubling in groups, to appear in Erdős centennial volume.
[18] E. Breuillard, B. Green & T. Tao, Linear approximate groups, *Electron. Res. Announc. Math. Sci.* **17** (2010), 57–67.
[19] E. Breuillard, B. Green & T. Tao, Approximate subgroups of linear groups, *Geom. Funct. Anal.* **21** (2011), 774–819.
[20] E. Breuillard, B. Green & T. Tao, Suzuki groups as expanders, *Groups Geom. Dyn.* **5** (2011), 281–299.
[21] E. Breuillard, B. Green & T. Tao, The structure of approximate groups, *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), 115–221.
[22] E. Breuillard, R. Guralnick, B. Green & T. Tao, Expansion in simple groups of Lie type, *J. Europ. Math. Soc.* (to appear).
[23] J. Button & C. Roney-Dougal, An explicit upper bound for the Helfgott δ in $\mathrm{SL}(2,p)$, arXiv:1401.2863, 2013 (preprint).
[24] H. Davenport, *Multiplicative Number Theory (3rd Edition)*, Grad. Texts Math. **74**, Springer-Verlag, New York, 2000, revised with preface by H.L. Montgomery.
[25] G. Elekes & Z. Király, On the combinatorics of projective mappings, *J. Algebraic Combin.* **14** (2001), 183–197.
[26] A. Eskin, S. Mozes & H. Oh, On uniform exponential growth for linear groups, *Invent. Math.* **160** (2005), 1–30.
[27] G.A. Freĭman, *Foundations of a Structural Theory of Set Addition*, Translations of Math. Monographs **37**, Amer. Math. Soc., Providence, RI, 1973 (Translated from the Russian).
[28] H. Furstenberg, Noncommuting random products, *Trans. Amer. Math. Soc.* **108** (1963), 377–428.
[29] A. Gamburd, On the spectral gap for infinite index "congruence" subgroups of $\mathrm{SL}_2(\mathbf{Z})$, *Israel J. Math.* **127** (2002), 157–200.
[30] I.Y. Gol′dsheĭd & G.A. Margulis, Lyapunov exponents of a product of random matrices, *Uspekhi Mat. Nauk* **44** (1989), 13–60.
[31] A. Gorodnik & A. Nevo, Splitting fields of elements in arithmetic groups, *Math. Res. Lett.* **18** (2011), 1281–1288.
[32] W.T. Gowers, Quasirandom groups, *Combin. Probab. Comput.* **17** (2008), 363–387.
[33] B. Green & I.Z. Ruzsa, Freiman's theorem in an arbitrary abelian group, *J. Lond. Math. Soc. (2)* **75** (2007), 163–175.
[34] Y. Guivarc'h & A. Raugi, Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence, *Z. Wahrsch. Verw. Gebiete* **69** (1985), 187–242.
[35] R.M. Guralnick & P.H. Tiep, Decompositions of small tensor powers and Larsen's conjecture, *Represent. Theory* **9** (2005), 138–208 (electronic).
[36] H.A. Helfgott, Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math. (2)* **167** (2008), 601–623.
[37] H.A. Helfgott, Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, *J. Eur. Math. Soc.* **13** (2011), 761–851.
[38] S. Hoory, N. Linial & A. Wigderson, Expander graphs and their applications, *Bull. Amer. Math. Soc. (N.S.)* **43** (2006), 439–561 (electronic).
[39] E. Hrushovski, Stable group theory and approximate subgroups, *J. Amer. Math. Soc.* **25** (2012), 189–243.
[40] E. Hrushovski, P.H. Kropholler, A. Lubotzky & A. Shalev, Powers in finitely generated groups, *Trans. Amer. Math. Soc.* **348** (1996), 291–304.
[41] E. Hrushovski & A. Pillay, Definable subgroups of algebraic groups over finite fields, *J. Reine Angew. Math.* **462** (1995), 69–91.
[42] E. Hrushovski & F. Wagner, Counting and dimensions, in *Model Theory with Appli-*

cations to Algebra and Analysis, Vol. 2, London Math. Soc. Lecture Note Ser. **350**, 161–176. Cambridge Univ. Press, Cambridge, 2008.

[43] J.E. Humphreys, *Linear Algebraic Groups*, Graduate Texts Math. **21**, Springer-Verlag, New York-Heidelberg, 1975.

[44] C. Jordan, Mémoire sur les équations différentielles linéaires à intégrale algébrique, *J. Reine Angew. Math.* **84** (1878), 89–215.

[45] F. Jouve, E. Kowalski & D. Zywina, Splitting fields of characteristic polynomials of random elements in arithmetic groups, *Israel J. Math.* **193** (2013), 263–307.

[46] N.M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Annals of Math. Studies **116**, Princeton University Press, Princeton, NJ, 1988.

[47] H. Kesten, Symmetric random walks on groups, *Trans. Amer. Math. Soc.* **92** (1959), 336–354.

[48] B. Klopsch, N. Nikolov & C. Voll, *Lectures on Profinite Topics in Group Theory*, London Math. Soc. Student Texts **77**, Cambridge Univ. Press, Cambridge, 2011.

[49] M. Kneser, Starke Approximation in algebraischen Gruppen, I, *J. Reine Angew. Math.* **218** (1965), 190–203.

[50] A. Kontorovich, Levels of distribution and the affine sieve, *Annales de la Faculte des Sci. Toulouse*, 2014, to appear.

[51] E. Kowalski, Lecture notes on expander graphs, http://www.math.ethz.ch/~kowalski/expanders.html.

[52] E. Kowalski, *The Large Sieve and its Applications: Arithmetic geometry, random walks and discrete groups*, Cambridge Tracts Math. **175**, CUP, Cambridge, 2008.

[53] E. Kowalski, Crible en expansion, Séminaire Bourbaki: Vol. 2010/2011. Exposés 1027–1042. *Astérisque* **348**, Exp. No. 1028, vii, 17–64, 2012.

[54] E. Kowalski, Explicit growth and expansion for SL_2, *Int. Math. Res. Not.* **24** (2013), 5645–5708.

[55] E. Kowalski, Sieve in discrete groups, especially sparse, in *Thin Groups and Super Stong Appoximation*, eds. E. Breuillard & H. Oh, MSRI Publications **61**, 2014.

[56] M. Kuranishi, Two elements generations on semi-simple Lie groups, *Kōdai Math. Sem. Rep.* **1** (1949), 9–10.

[57] V. Landazuri & G.M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418–443.

[58] M. Larsen & A. Lubotzky, Normal subgroup growth of linear groups: the (G_2, F_4, E_8)-theorem, in *Algebraic Groups and Arithmetic*, 441–468, Tata Inst. Fund. Res., Mumbai, 2004.

[59] M. Larsen & R. Pink, Finite subgroups of algebraic groups, *J. Amer. Math. Soc.* **24** (2011), 1105–1158.

[60] J.C. Lennox & J. Wiegold, Converse of a theorem of Mal′cev on nilpotent groups, *Math. Z.* **139** (1974), 85–86.

[61] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, Birkhäuser Verlag, Basel, 2010, with appendix by J.D D. Rogawski, reprint of 1994 edition.

[62] A. Lubotzky, Finite simple groups of Lie type as expanders, *J. Eur. Math. Soc.* **13** (2011), 1331–1341.

[63] A. Lubotzky, Expander graphs in pure and applied mathematics, *Bull. Amer. Math. Soc. (N.S.)* **49** (2012), 113–162.

[64] A. Lubotzky & A. Mann, On groups of polynomial subgroup growth, *Invent. Math.* **104** (1991), 521–533.

[65] A. Lubotzky & C. Meiri, Sieve methods in group theory I: Powers in linear groups, *J. Amer. Math. Soc.* **25** (2012), 1119–1148.

[66] A. Lubotzky & C. Meiri, Sieve methods in group theory II: the mapping class group, *Geom. Dedicata* **159** (2012), 327–336.

[67] A. Lubotzky & L. Rosenzweig, The Galois group of random elements of linear groups,

Amer. J. Math., to appear. arXiv:1205.5290
[68] A. Lubotzky & D. Segal, *Subgroup Growth*, Progress Math. **212**, Birkhäuser Verlag, Basel, 2003.
[69] C.R. Matthews, L.N. Vaserstein & B. Weisfeiler, Congruence properties of Zariski-dense subgroups, I, *Proc. London Math. Soc. (3)* **48** (1984), 514–532.
[70] M.B. Nathanson, *Additive Number Theory: Inverse problems and the geometry of sumsets*, Graduate Texts Math. **165**, Springer-Verlag, New York, 1996.
[71] N. Nikolov & L. Pyber, Product decompositions of quasirandom groups and a Jordan type theorem, *J. Eur. Math. Soc.* **13** (2011), 1063–1077.
[72] M.V. Nori, On subgroups of $\mathrm{GL}_n(\mathbf{F}_p)$, *Invent. Math.* **88** (1987), 257–275.
[73] R. Pink, Strong approximation for Zariski dense subgroups over arbitrary global fields, *Comment. Math. Helv.* **75** (2000), 608–643.
[74] V. Platonov, The problem of strong approximation and the Kneser-Tits hypothesis for algebraic groups, *Izv. Akad. Nauk SSSR Ser. Mat.* **33** (1969), 1211–1219.
[75] V. Platonov & A. Rapinchuk, *Algebraic Groups and Number Theory*, Pure & Applied Math. **139**, Academic Press Inc., Boston, MA, 1994 (transl. from 1991 Russian original by R. Rowen).
[76] G. Prasad & A. Rapinchuk, Generic elements in Zariski dense subgroups and isospectral locally symmetric spaces, in *Thin Groups and Super Stong Appoximation*, ed. E. Breuillard & H. Oh, MSRI Publications **61**, 2014.
[77] G. Prasad & A. Rapinchuk, Existence of irreducible $\mathbb{R}$-regular elements in Zariski-dense subgroups, *Math. Res. Lett.* **10** (2003), 21–32.
[78] G. Prasad & A. Rapinchuk, Weakly commensurable arithmetic groups and isospectral locally symmetric spaces, *Publ. Math. Inst. Hautes Études Sci.* **109** (2009), 113–184.
[79] G. Prasad & A. Rapinchuk, Number-theoretic techniques in the theory of Lie groups and differential geometry, in *Fourth International Congress of Chinese Mathematicians*, AMS/IP Stud. Adv. Math. **48**, 231–250, Amer. Math. Soc., Providence, RI, 2010.
[80] L. Pyber & E. Szabó, Growth in finite simple groups of lie type of bounded rank (preprint), 2010, arXiv:1005.1858.
[81] L. Pyber & E. Szabó, Growth in linear groups, in *Thin Groups and Super Stong Appoximation*, eds. E. Breuillard & H. Oh, MSRI Publications **61**, 2014.
[82] A. Rapinchuk, On strong approximation for algebraic groups, in *Thin Groups and Super Stong Appoximation*, eds. E. Breuillard & H. Oh, MSRI Publications **61**, 2014.
[83] I. Rivin, Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms, *Duke Math. J.* **142** (2008), 353–379.
[84] I. Rivin, Zariski density and genericity, *Int. Math. Res. Not.* **19** (2010), 3649–3657.
[85] I.Z. Ruzsa, Generalized arithmetical progressions and sumsets, *Acta Math. Hungar.* **65** (1994), 379–388.
[86] A. Salehi Golsefidy & P. Sarnak, The affine sieve, *J. Amer. Math. Soc.* **26** (2013), 1085–1105.
[87] A. Salehi Golsefidy & P.P. Varjú, Expansion in perfect groups, *Geom. Funct. Anal.* **22** (2012), 1832–1891.
[88] T. Sanders, The structure theory of set addition revisited, *Bull. Amer. Math. Soc. (N.S.)* **50**, (2013), 93–127.
[89] P. Sarnak, Diophantine problems and linear groups, in *Proc. Internat. Congress of Mathematicians, Vol. I, II (Kyoto, 1990)*, 459–471, Math. Soc. Japan, Tokyo, 1991.
[90] P. Sarnak, Selberg's eigenvalue conjecture, *Notices Amer. Math. Soc.* **42** (1995), 1272–1277.
[91] P. Sarnak, Notes on thin groups, in *Thin Groups and Super Stong Appoximation*, eds. E. Breuillard & H. Oh, MSRI Publications**61**, 2014.
[92] P. Sarnak & X.X. Xue, Bounds for multiplicities of automorphic representations, *Duke Math. J.* **64** (1991), 207–227.

[93] A. Selberg, On the estimation of Fourier coefficients of modular forms, in *Proc. Sympos. Pure Math., Vol. VIII*, 1–15, Amer. Math. Soc., Providence, RI, 1965.
[94] J.-P. Serre, On a theorem of Jordan, *Bull. Amer. Math. Soc.* **40** (2003), 429–440 (electronic).
[95] I.R. Shafarevich, *Basic Algebraic Geometry 1: Varieties in projective space (3rd Ed.)*, Springer, Heidelberg, third edition, 2013.
[96] T. Tao, *Expansion in groups of Lie type*, http://terrytao.wordpress.com/books/expansion-in-finite-simple-groups-of-lie-type.
[97] T. Tao, Product set estimates for non-commutative groups, *Combinatorica* **28** (2008), 547–594.
[98] T. Tao & V.H. Vu, *Additive Combinatorics*, Cambridge Studies Adv. Math. **105**, Cambridge University Press, Cambridge, 2010.
[99] J. Tits, Free subgroups in linear groups, *J. Algebra* **20** (1972), 250–270.
[100] P.P. Varjú, Expansion in $\mathrm{SL}_d(\mathcal{O}_K/I)$, I square-free, *J. Eur. Math. Soc.* **14** (2012), 273–305.
[101] B. Weisfeiler, Strong approximation for Zariski-dense subgroups of semisimple algebraic groups, *Ann. of Math. (2)* **120** (1984), 271–315.

WIDTH QUESTIONS FOR FINITE SIMPLE GROUPS

MARTIN W. LIEBECK

Department of Mathematics, Imperial College, London SW7 2BZ, UK
Email: m.liebeck@imperial.ac.uk

Abstract

Let G be a finite group generated by a collection $\mathcal{S}$ of subsets of G. Define the width of G with respect to $\mathcal{S}$ to be the minimal integer n such that G is equal to the union of a product of n subsets in $\mathcal{S}$, together with all subproducts. For example, when $\mathcal{S}$ consists of a single subset, the width is just the diameter of the Cayley graph of G with respect to this subset. This article contains a discussion of a variety of problems concerning the width of simple groups, mainly in the following cases: (1) the case where $\mathcal{S}$ consists of a single subset; (2) the case where $\mathcal{S}$ is closed under conjugation. There are many examples of special interest. Particular emphasis is given to recent results and problems concerning the "word width" of simple groups – namely, the width in the case where $\mathcal{S}$ consists of all values in G of a fixed word map. Also discussed are combinatorial interpretations of some width problems, such as the estimation of diameters of orbital graphs.

1 Introduction

Let G be a finite group, and suppose $\mathcal{S}$ is a collection of subsets of G such that G is generated by their union. Every element $g \in G$ has an expression $g = t_1 \dots t_k$ where $t_i \in T_i \in \mathcal{S}$. Hence it is possible to write G as the union of a product $T_1 \cdots T_d := \{t_1 \dots t_d : t_i \in T_i\}$, together with all subproducts $T_{i_1} \cdots T_{i_k}$ $(i_1 < \cdots < i_k)$, where each $T_i \in \mathcal{S}$ and repeats are allowed among the T_i. We define the *width* of G with respect to $\mathcal{S}$ to be the minimal such positive integer d, and denote this by $\mathrm{width}(G, \mathcal{S})$.

In this article, we consider the problem of finding, or bounding, the width of finite groups in various cases of interest, mainly when G is a finite non-abelian simple or almost simple group. We remind the reader that the finite non-abelian simple groups are the alternating groups of degree at least 5, the simple groups of Lie type over finite fields, and the 26 sporadic groups; and an *almost simple* group is a group G such that $S \lhd G \leq \mathrm{Aut}(S)$ for some non-abelian simple group S. For brevity in the text below, whenever we say a group G is simple, we mean that G is a finite non-abelian simple group.

Examples Here are two contrasting examples of such width problems.

1. Let $G = S_n$, the symmetric group of degree n, and let $\mathcal{S} = \{T\}$, where T is the set of all transpositions. Then $\mathrm{width}(G, \mathcal{S})$ is the minimal value of d such that $S_n = T^d \cup T^{d-1} \cup \cdots \cup \{1\}$ (where $T^k := \{t_1 \dots t_k : t_i \in T\}$). Since every permutation can be expressed as a product of at most $n-1$ transpositions, and such an expression for an n-cycle requires precisely this number, the width in

this example is $n-1$.

2. Again let $G = S_n$, but this time let $\mathcal{S} = \{\langle t_1\rangle, \ldots, \langle t_k\rangle\}$, where $t_1, \ldots, t_k$ are all the transpositions in G (and $k = \binom{n}{2}$). Here the width problem is more subtle than in the previous example: $\mathrm{width}(G, \mathcal{S})$ is the minimal value of d for which we can write $S_n = \langle t_{i_1}\rangle \cdots \langle t_{i_d}\rangle$ (repeats allowed). Notice that the right hand side has at most 2^d elements while the left has $n!$, so the width d must be at least the order of $n \log n$. The question of whether the width in this example does have this order of magnitude is not so easy; we shall give the answer in Section 3.2 (see the proof of Theorem 3.9).

All the width questions we shall discuss in these lectures are of one of the two types in the above examples:

(a) the case where $\mathcal{S}$ consists of a single generating subset S of G

(b) the case where $\mathcal{S}$ consists of a conjugacy class of subsets of G: that is,

$$\mathcal{S} = \{A^g : g \in G\}$$

for some subset A of G.

In case (a), the width is just the diameter of the Cayley graph of G with respect to S. We shall discuss recent developments on this topic for simple groups in the next section. There are many interesting questions arising from case (b), and these will be the focus of the remaining sections.

2 Width, Cayley graphs and orbital graphs

Let G be a finite group with a generating set S which is symmetric – that is, closed under taking inverses – and does not contain the identity. The *Cayley graph* $\Gamma(G, S)$ is defined to be the graph with vertex set G and edges $\{g, gs\}$ for all $g \in G, s \in S$. It is connected and regular of valency $|S|$, and G acts regularly on $\Gamma(G, S)$ by left multiplication. Because of the transitive action of G, the diameter of $\Gamma(G, S)$, denoted by $\mathrm{diam}(G, S)$, is equal to the maximum distance between the identity element and any $g \in G$, and so $\mathrm{diam}(G, S) = \max\{l(g) : g \in G\}$, where $l(g)$ is the length of the shortest expression for g as a product of elements of S. If $d = \mathrm{diam}(G, S)$, then d is minimal such that $G = S^d \cup S^{d-1} \cup \cdots \cup \{1\}$, and hence

$$\mathrm{diam}(G, S) = \mathrm{width}(G, \{S\}).$$

Also $|G| \leq \sum_{r=0}^{d} |S|^r < |S|^{d+1}$. Hence

$$\mathrm{diam}(G, S) > \frac{\log |G|}{\log |S|} - 1. \tag{1}$$

Examples

1. Let $G = C_n = \langle x \rangle$, a cyclic group of order n, and let $S = \{x, x^{-1}\}$. Then $\Gamma(G,S)$ is an n-gon. So $\mathrm{diam}(G,S)$ is $[\frac{n}{2}]$, whereas $\frac{\log|G|}{\log|S|}$ is $\frac{\log n}{\log 2}$.
2. Let $G = S_n$ and S be the set of all transpositions. Here $\mathrm{diam}(G,S)$ is $n-1$, while $\frac{\log|G|}{\log|S|}$ is roughly $\frac{n}{2}$.
3. Let $G = S_n$ and $S = \{(1\,2), (1\,2\cdots n)^{\pm 1}\}$. In this case $\mathrm{diam}(G,S)$ is roughly n^2, while $\frac{\log|G|}{\log|S|}$ is of the order of $n \log n$. The same orders of magnitude apply to a similar generating set for A_n consisting of a 3-cycle and an n- or $(n-1)$-cycle and their inverses.
4. Let $G = SL_n(q)$ and S be the set of transvections. Then $\mathrm{diam}(G,S) \approx n$ and $\frac{\log|G|}{\log|S|} \approx \frac{n}{2}$.
5. Let $G = SL_n(p)$ (p prime) and $S = \{x^{\pm 1}, y^{\pm 1}\}$ where

$$x = \begin{pmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & . & & \\ & & & . & \\ & & & & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & & . & \\ & & & & . \\ \pm 1 & & & & \end{pmatrix}$$

   Then $\frac{\log|G|}{\log|S|} \sim n^2 \log p$, and also $\mathrm{diam}(G,S) \sim n^2 \log p$.

All these examples are elementary except the last, where the fact that $\mathrm{diam}(G,S) \leq Cn^2 \log p$ for some constant C is a result of Kassabov and Riley [32].

2.1 Babai's Conjecture

Define $\mathrm{diam}(G)$ to be the maximum of $\mathrm{diam}(G,S)$ over all generating sets S. The main conjecture in the field is due to Babai, and appears as Conjecture 1.7 in [6]:

Babai's Conjecture *There is a constant c such that* $\mathrm{diam}(G) < (\log|G|)^c$ *for any non-abelian finite simple group G.*

It can be seen from Example 3 above that c must be at least 2 for the conjecture to hold.

There have been spectacular recent developments on Babai's conjecture, both for groups of Lie type and for alternating groups. We shall discuss these separately.

2.1.1 Groups of Lie type

For a long time, even $SL_2(p)$ (p prime) was a mystery as far as proving Babai's conjecture was concerned. Probably the first small (symmetric) generating set one thinks of for this group is

$$S = \{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\pm 1}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\pm 1}\}.$$

Babai's conjecture asserts that $\mathrm{diam}(G,S) < (\log p)^c$ for these generators. Surely this must be easy?

In fact it is not at all easy, and was proved by the following beautiful but indirect method (see [51]). First observe that the matrices in S, when regarded as integer matrices, generate $SL_2(\mathbb{Z})$. Now let $\Gamma(p)$ denote the congruence subgroup which is the kernel of the natural map from $SL_2(\mathbb{Z}) \to SL_2(p)$. If $\mathbb{H}$ is the upper half plane and $X(p)$ denotes the Riemann surface $\Gamma(p)\backslash\mathbb{H}$, denote by $\lambda_1(X(p))$ the smallest eigenvalue for the Laplacian on $X(p)$. A theorem of Selberg [61] gives $\lambda_1(X(p)) \geq \frac{3}{16}$ for all p, and this can be used to show that the Cayley graphs $\{\Gamma_p = \Gamma(SL_2(p), S) : p \text{ prime}\}$ have their second largest eigenvalues bounded away from the valency, and hence that they form a family of *expander graphs*. This means that there is an *expansion* constant $c > 0$, independent of p, such that for every set A consisting of fewer that half the total number of vertices in Γ_p, we have $|\delta A| > c|A|$, where δA is the boundary of A – that is, the set of vertices not in A that are joined to some vertex in A. From the expansion property it is easy to deduce that Γ_p has logarithmic diameter, so that $\mathrm{diam}(\Gamma(SL_2(p), S) < c \log p$, a strong form of Babai's conjecture.

One can adopt essentially the same method for the generators

$$\{\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}^{\pm 1}, \begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}^{\pm 1}\}$$

of $SL_2(p)$, since, while these do not generate $SL_2(\mathbb{Z})$, they do generate a subgroup of finite index therein. But what if we replace the 2's in these generators with 3's? Then the matrices generate a subgroup of infinite index in $SL_2(\mathbb{Z})$, and the above method breaks down. This question became known as Lubotzky's 1-2-3 problem, and was not solved until the breakthrough achieved by Helfgott [23]:

Theorem 2.1 *Babai's conjecture holds for $G = SL_2(p)$. That is,*

$$\mathrm{diam}(SL_2(p)) < (\log p)^c,$$

where c is an absolute constant.

Helfgott deduced this from his key proposition: for any generating set S of $G = SL_2(p)$, either $|S^3| > |S|^{1+\epsilon}$, or $S^k = G$, where $\epsilon > 0$ and k do not depend on p. (Later it was observed that one can take $k = 3$ here.) The heart of his proof is to relate the growth of powers of subsets A of G with the growth of the corresponding set of scalars $B = \mathrm{tr}(A) = \{\mathrm{tr}(x) : x \in A\}$ in $\mathbb{F}_p$ under sums and products. By doing this he could tap into the theory of additive combinatorics, using results such as the following, taken from [10]: if B is a subset of $\mathbb{F}_p$ with $p^\delta < |B| < p^{1-\delta}$ for some $\delta > 0$, then $|B \cdot B| + |B + B| > |B|^{1+\epsilon}$, where $\epsilon > 0$ depends only on δ.

Following Helfgott's result, there was a tremendous surge of progress in this area. Many new families of expanders were constructed in [9]. Helfgott himself extended his result to $SL_3(p)$ in [24], and this has now been proved for all groups of Lie type of bounded rank in [11, 58]. As a consequence, we have

Theorem 2.2 *If $G = G(q)$ is a simple group of Lie type of rank r, then* $\mathrm{diam}(G) < (\log |G|)^{c(r)}$ *where $c(r)$ depends only on r.*

Again, the theorem is proved via a growth statement: for any generating set S of $G(q)$, either $|S^3| > |S|^{1+\epsilon}$, or $S^3 = G$, where $\epsilon > 0$ depends only on r. From this one gets a strong version of the previous result which takes the size of the generating set S into account:

Theorem 2.3 *If $G = G(q)$ is a simple group of Lie type of rank r, and S is a generating set of G, then $G = S^d$ for some $d \leq (\frac{\log |G|}{\log |S|})^{c(r)}$, where $c(r)$ depends only on r.*

These results, and particularly their developments into the theory of expanders, have many wonderful and surprising applications. For a survey of these developments and some of the applications, see [53].

Finally, let us remark that Babai's conjecture remains open for groups of Lie type of unbounded rank.

2.1.2 Alternating groups

For the alternating groups A_n, Babai's conjecture is that there is a constant C such that $\mathrm{diam}(A_n) < n^C$. Until very recently, the best bound for $\mathrm{diam}(A_n)$ was that obtained by Babai and Seress in [5], where it was proved that

$$\mathrm{diam}(A_n) < \exp((1 + o(1)) \cdot (n \log n)^{1/2}) = \exp((1 + o(1)) \cdot (\log |A_n|)^{1/2}).$$

Various other partial results appeared at regular intervals, such as that in [3], where it was shown that if the generating set S contains a permutation of degree at most $0.33n$, then $\mathrm{diam}(A_n, S)$ is polynomially bounded. But no real progress was made on Babai's conjecture until a recent breakthrough of Helfgott and Seress [25]:

Theorem 2.4 *We have $\mathrm{diam}(A_n) \leq \exp(O((\log n)^4 \log \log n))$, where the implied constant is absolute.*

This does not quite prove Babai's conjecture, but it does prove that $\mathrm{diam}(A_n)$ is "quasipolynomial" (where a quasipolynomial function $f(n)$ is one for which $\log f(n)$ is polynomial in $\log n$), which represents a big step forward. The same paper also gives a bound of the same magnitude for the diameter of any transitive subgroup of S_n.

2.2 Orbital graphs

Here we discuss another class of graphs for which the diameter has an interpretation in terms of width.

Denote by (G, X) a permutation group G on a finite set X. Suppose G is transitive on X, and let $X^{\{2\}}$ denote the set of unordered pairs of elements of X. For each orbit Δ of G on $X^{\{2\}}$, there is a corresponding *orbital graph* having vertex set X and edge set Δ. These are precisely the non-empty graphs on X for which G acts transitively on edges. A well known criterion of D.G. Higman (see [26, 1.12]) states that G is primitive on X if and only if all of its orbital graphs are connected. For G primitive on X, define $\mathrm{diam}(G, X)$ to be the maximum of the diameters of all the orbital graphs.

The diameters of orbital graphs of primitive groups have an interpretation in terms of width. Indeed, let Δ be an orbit of G on $X^{\{2\}}$ as above, and let $\{x, xg\} \in \Delta$, where $g \in G$. Notice that also $\{x, xg^{-1}\} \in \Delta$. Write $H = G_x$. For each i, the set of vertices at distance i from x in the corresponding orbital graph is contained in

$$\{xg^{\pm 1}h_1g^{\pm 1}h_2 \cdots g^{\pm 1}h_i \,:\, h_i \in H\}.$$

It follows that if we define $w = \mathrm{width}(G, \mathcal{S})$ where $\mathcal{S} = H \cup \{g, g^{-1}\}$, then the diameter of the orbital graph lies between w and $[\frac{1}{2}w]$. (Both extremes are possible: for example the diameter is w when $H = 1$ and G is cyclic of prime order.)

For a positive integer d, denote by $\mathcal{C}_d$ the class of all finite primitive permutation groups (G, X) for which $\mathrm{diam}(G, X) \leq d$. In [42], the following problem is addressed.

Problem 2.5 For each d, describe the class of finite primitive groups $\mathcal{C}_d$.

The motivation in [42] is mainly model-theoretical and stems from the fact that for groups of bounded orbital diameter, primitivity is implied by a first order expressible condition in the language of permutation groups (whereas for permutation groups in general, primitivity is not a first order property). This means, for example, that the primitivity condition extends to ultraproducts.

In [42], the above problem is solved "asymptotically"; as discussed in detail in [42], this leads to the solution of a number of related model-theoretic problems, such as the description of primitive infinite ultraproducts of finite permutation groups, and of primitive ω-saturated pseudofinite permutation groups.

We present part of the main result of [42] in Theorem 2.6 below, which describes the classes of *simple* groups in $\mathcal{C}_d$. This time, unlike the previous section, there is a satisfactory result for groups of unbounded rank.

In order to state the theorem we need to define some terminology. We say that the primitive group (G, X) with G simple is a *standard t-action* if one of the following holds:

(a) $G = A_n$ and $X = I^{\{t\}}$, the set of t-subsets of $I = \{1, ..., n\}$

(b) $G = Cl_n(q)$, a classical group with natural module $V = V_n(q)$ of dimension n over $\mathbb{F}_q$, and X is an orbit of subspaces of dimension or codimension t in V; the subspaces are arbitrary if $G = PSL_n(q)$, and otherwise are totally singular, non-degenerate, or, if G is orthogonal and q is even, non-singular 1-spaces (in which case $t = 1$)

(c) $G = Sp_{2m}(q)$, q is even, and a point stabilizer in G is $O^{\pm}_{2m}(q)$ (here we take $t = 1$).

If $G(q)$ is a simple group of Lie type over $\mathbb{F}_q$, then a *subfield subgroup* is a group $G(q_0)$ embedded naturally in $G(q)$, where $\mathbb{F}_{q_0}$ is a subfield of $\mathbb{F}_q$. For convenience in the statement below we define the *rank* of an alternating group A_n to be n.

We say that a class $\mathcal{C}$ of finite primitive permutation groups is *bounded* if $\mathcal{C} \subseteq \mathcal{C}_d$ for some d. All bounds implicit in the statement below are in terms of d, where $\mathcal{C} \subseteq \mathcal{C}_d$.

Theorem 2.6 *Let $\mathcal{C}$ be an infinite class of finite simple primitive permutation groups, and suppose $\mathcal{C}$ is bounded.*

(i) *If $\mathcal{C}$ consists of simple groups of unbounded ranks, then the groups in $\mathcal{C}$ of sufficiently large rank are alternating or classical groups in standard t-actions, where t is bounded.*

(ii) *If $\mathcal{C}$ consists of simple groups G of bounded rank, then point stabilizers G_x have unbounded orders; moreover, if $G = G(q)$, of Lie type over $\mathbb{F}_q$, and G_x is a subfield subgroup $G(q_0)$, then $|\mathbb{F}_q : \mathbb{F}_{q_0}|$ is bounded.*

Conversely, any class of simple primitive groups satisfying the conclusions of (i) *or* (ii) *is bounded.*

One of the most interesting parts of this result is the converse statement for part (ii): if $\mathcal{C}$ is a class consisting of simple primitive permutation groups of Lie type of bounded Lie rank with unbounded point stabilizers (and also satisfying the given condition on subfields), then $\mathcal{C}$ is a bounded class. For example, if $\mathcal{C}$ consists of the groups $E_8(q)$ (q varying) acting on the coset space $E_8(q)/H(q)$ for some maximal subgroup $H(q)$ arising from a maximal connected subgroup $H(K)$ of the simple algebraic group $E_8(K)$, where $K = \bar{\mathbb{F}}_q$ (for example $H(K) = D_8(K)$ or $A_1(K)$ – see [47]), then the diameters of all the orbital graphs are bounded by an absolute constant. In fact this now follows fom Theorem 2.3, but a direct proof using a substantial amount of model theory can be found in [42].

It would be interesting to have a more explicit solution to Problem 2.5, for example for some small values of d. Work is under way on this.

3 Conjugacy width

We now turn to a discussion of the width of simple groups G with respect to a conjugacy class of subsets – that is, $\text{width}(G, \mathcal{S})$ where $\mathcal{S} = \{A^g : g \in G\}$ for some subset A of G which we take to be of size at least 2. The following lemma shows that in this case no subproducts are required in the definition of width.

Lemma 3.1 *If $A \subseteq G$ with $|A| \geq 2$, and $\mathcal{S} = \{A^g : g \in G\}$, then*

$$\text{width}(G, \mathcal{S}) = \min\{n : G = A^{g_1} \cdots A^{g_n}, g_i \in G\}.$$

Proof This is clear if $1 \in A$. If not, let $a \in A$, set $B = a^{-1}A$, and observe that G is a product of n conjugates of A if and only if it is a product of n conjugates of B. □

Examples When G is simple there are many interesting cases to consider. Here are some examples. In the first four, $\mathcal{S}$ consists of a single normal subset of G (i.e. a subset closed under conjugation), so we are back in the Cayley graph case of the previous section.

1. $\mathcal{S} = \{I(G)\}$, where $I(G)$ is the set of involutions in G: here $\text{width}(G, \mathcal{S})$ is the minimal n such that every element of G is a product of n involutions.
2. $\mathcal{S} = \{C(G)\}$, where $C(G) = \{[x, y] : x, y \in G\}$ is the set of commutators in G: here $\text{width}(G, \mathcal{S})$ is often called the *commutator width* of G.
3. $\mathcal{S} = \{P_k(G)\}$, where $k \geq 2$ and $P_k(G) = \{x^k : x \in G\}$ is the set of k^{th} powers in G.

4. (Generalizing Examples 2,3): $\mathcal{S} = \{w(G)\}$, where $w = w(x_1, \ldots, x_k)$ is a fixed word in the free group F_k of rank k and $w(G) = \{w(g_1, \ldots, g_k) : g_i \in G\}$.
5. $G = S_n$ and $\mathcal{S} = \{\langle t_1 \rangle, \ldots, \langle t_k \rangle\}$, where $t_1, \ldots t_k$ are all the transpositions in G (and $k = \binom{n}{2}$), as in Example 2 in Section 1.
6. $\mathcal{S} =$ the set of Sylow p-subgroups of G, where p is a prime dividing $|G|$.

Clearly if $\mathcal{S} = \{A^g : g \in G\}$ as above, then $\text{width}(G, \mathcal{S}) \geq \log|G|/\log|A|$. In [43] the following conjecture was posed.

Conjecture 3.2 *There is an absolute constant c such that for any finite non-abelian simple group G and any subset $A \subseteq G$ with $|A| \geq 2$, we have*

$$\text{width}(G, \mathcal{S}) \leq c\frac{\log|G|}{\log|A|},$$

where $\mathcal{S} = \{A^g : g \in G\}$.

This conjecture has been proved in a number of special cases, as we shall describe below, but it is open in general.

3.1 Normal subsets

In the case where $\mathcal{S}$ consists of a single normal subset of G, Conjecture 3.2 was proved in [50]:

Theorem 3.3 *There is an absolute constant $k > 0$ such that for any finite non-abelian simple group G, and any non-identity normal subset $S \subseteq G$, we have $G = S^n$ for all $n \geq k \log|G|/\log|S|$.*

In particular the diameter of the Cayley graph $\Gamma(G, S)$ is at most $k\frac{\log|G|}{\log|S|}$, so this proves Babai's conjecture in this case in a strong form.

The *covering number* of a finite simple group G is the minimal positive integer n such that $C^n = G$ for all conjugacy classes C of G (see [2]). Theorem 3.3 implies an upper bound for the covering number which is linear in the rank of G; further such bounds can be found in [14, 39], and the precise covering number of $PSL_n(q)$ for $n \geq 3, q \geq 4$ is shown to be n in [40]. However Theorem 3.3 carries much more information than these bounds, since it takes into account the size of the class.

Let us now examine the implications of Theorem 3.3 for Examples 1–4 above.

3.1.1 Involutions

As in Example 1 above, let $S = I(G)$, the set of involutions in G. To get a feeling for how big $\frac{\log|G|}{\log|S|}$ is, consider $G = PSL_{2m}(q)$ with q odd, m even, and let $t \in G$ be the involution which is the image modulo scalars of the matrix $\text{diag}(I_m, -I_m)$. Then the size of the conjugacy class t^G is roughly $|GL_{2m}(q) : GL_m(q) \times GL_m(q)|$, which is approximately q^{4m^2}/q^{2m^2}, and so $|t^G|$ is of the order of $|G|^{1/2}$. Therefore $\log|G|/\log|S|$ is about 2 in this case. It can be shown that there is an absolute

constant $c > 0$ such that $|I(G)| > c|G|^{1/2}$ for all finite simple groups G (see [49, 4.2,4.3]). Hence Theorem 3.3 implies the following.

Corollary 3.4 *There is an absolute constant N such that every element of every finite non-abelian simple group is a product of N involutions.*

It would be quite interesting to know the minimal value of N. It is certainly more than 2: groups in which every element is a product of two involutions are known as *strongly real* groups, and the strongly real simple groups have been classified (see [64, 59]).

3.1.2 Images of word maps

As in Example 4 above, let $w = w(x_1, \ldots, x_k)$ be a fixed non-identity word in the free group F_k of rank k and for a group G define $w(G) = \{w(g_1, \ldots, g_k) : g_i \in G\}$. Let us consider the implications of Theorem 3.3 in the case where G is simple and $S = w(G)$.

We need information about the size of the set $w(G)$. This can be 1 for some simple groups G – for example if $w = x_1^k$ and the exponent of G divides k. The first question to consider is whether there could be a word w for which $w(G) = \{1\}$ for *all* (finite non-abelian) simple groups G. The answer is no: for suppose w is a non-identity word such that $w(SL_2(p)) = \{1\}$ for all primes p. Let ϕ_p be the natural map $SL_2(\mathbb{Z}) \to SL_2(p)$. Then $\bigcap_p \mathrm{Ker}(\phi_p) = 1$, hence also $w(SL_2(\mathbb{Z})) = 1$. However $SL_2(\mathbb{Z})$ contains a free subgroup of rank 2, so this is impossible. Since many simple groups of Lie type over $\mathbb{F}_p$ contain $SL_2(p)$, the assertion follows.

In fact a much stronger assertion about the nontriviality of $w(G)$ for simple groups G holds, as proved in [30]:

Theorem 3.5 *Given any nontrivial word w, there is a constant N_w depending only on w, such that $w(G) \neq \{1\}$ for all simple groups G of order greater than N_w.*

For simple groups of order greater than N_w, how large is $w(G)$? The following gives a weak lower bound. Better bounds will be discussed in Section 4.

Lemma 3.6 *For any non-identity word w, there is a constant $\delta_w > 0$ such that $|w(G)| > |G|^{\delta_w}$ for all simple groups G of order greater than N_w.*

Proof Consider first $G = A_n$. Choose $k = k(w)$ minimal such that $w(A_k) \neq 1$, and let $1 \neq a \in w(A_k)$. Take n to be large in terms of k. If $r = [\frac{n}{k}]$, then G contains a subgroup $H \cong (A_k)^r$. Let $x \in H$ be the image under this isomorphism of the element $(a, \ldots, a) \in A_k^r$. Then $x \in w(H)$ and x moves at least $3r$ points in $\{1, \ldots, n\}$. Now the conjugacy class x^G is contained in $w(G)$, and an elementary calculation shows that $|x^G|$ is at least of the order of $|G|^{1/2k}$, which gives the conclusion in this case.

The case where $G = Cl_n(q)$, a classical group of unbounded dimension n over a finite field $\mathbb{F}_q$, is similar, using a subgroup H of the form $(Cl_k(q))^r$ in the above argument. And when G is a group of Lie type of bounded rank, the fact that any nontrivial conjugacy class has size at least q gives the result. □

As before, Theorem 3.3 implies the following consequence.

Corollary 3.7 *Let w be a nontrivial word. Then there is a constant $c = c(w)$ such that for any simple group G of order greater than N_w, we have $G = w(G)^c$ (that is, every element of G is a product of c elements of $w(G)$).*

We shall discuss some recent vast improvements of this result in Section 4.

3.1.3 Remarks on the proof of Theorem 3.3

The proof in [50] is quite technical, but it may be instructive to illustrate two of the main steps with the following example. Let $G = PSL_n(q)$ with $n \geq 3$ and let $C = x^G$, where

$$x = \begin{pmatrix} J_k & \\ & I_{n-k} \end{pmatrix},$$

J_k being the $k \times k$ Jordan block matrix with 1's on and directly above the diagonal and 0's elsewhere. Assume also that n is large compared to k. The centralizer of x can be found in [48, 7.1], and it follows that $|C|$ is roughly $q^{(k-1)(2n-k)}$. Hence $\frac{\log |G|}{\log |C|}$ is of the order of $\frac{n}{2(k-1)}$.

The first step in the proof is the elementary but useful observation that

$$\begin{pmatrix} I_{k-1} & & \\ & J_k & \\ & & I_{n-2k+1} \end{pmatrix} \begin{pmatrix} J_k & \\ & I_{n-k} \end{pmatrix} = \begin{pmatrix} J_{2k-1} & \\ & I_{n-2k+1} \end{pmatrix}.$$

Applying this repeatedly, we can obtain the matrix J_n as a product of approximately $\frac{n}{k-1}$ conjugates of x; in other words, $J_n \in C^{n/(k-1)}$. Set $y := J_n$.

The second step is to apply some character theory of the group G. The following observation essentially goes back to Frobenius, and applies to conjugacy classes in arbitrary finite groups: for $g \in G$, and an integer $l \geq 2$, the number of ways of writing g as a product of l conjugates of y is

$$\frac{|y^G|^l}{|G|} \sum_{\chi \in Irr(G)} \frac{\chi(y)^l \chi(g^{-1})}{\chi(1)^{l-1}}, \tag{2}$$

where $\mathrm{Irr}(G)$ denotes the set of irreducible characters of G. At this point we apply some basic facts about the irreducible characters χ of $G = PSL_n(q)$:

(a) $|\chi(y)| \leq |C_G(y)|^{1/2} = |C_G(J_n)|^{1/2} \leq q^{n/2}$;

(b) for $\chi \neq 1_G$, the degree $\chi(1) \geq q^{n-1} - 1$;

(c) $|\mathrm{Irr}(G)| < q^{n-1} + 3q^{n-2}$.

Indeed, (a) is trivial, (b) follows from [33] and (c) from [16, 3.6]. Let Σ denote the sum in (2). The contribution to Σ of the trivial character $\chi = 1_G$ is 1. Hence using (a)–(c), we see that

$$|\Sigma| \geq 1 - \frac{(q^{n-1} + 3q^{n-2})q^{nl/2}}{(q^{n-1} - 1)^{l-2}}.$$

Assuming that $n \geq 10$, it follows that $\Sigma \neq 0$ provided $l \geq 7$. Hence $G = (y^G)^7$ under this assumption. Since $y = J_n \in C^{n/(k-1)}$, we therefore have

$$G = (y^G)^7 = C^{7n/(k-1)}.$$

The conclusion of Theorem 3.3 follows in this case.

3.1.4 Commutators

Applying Corollary 3.7 to the commutator word, it follows that every element of every finite simple group is a product of a bounded number of commutators. In fact a much stronger result is true:

Theorem 3.8 (The Ore Conjecture) *Every element of every finite simple group is a commutator.*

This conjecture emerged from a 1951 paper of Ore [56], after which many partial results were obtained, notably those of Thompson [63] for special linear groups, and of Ellers and Gordeev [13] proving the result for groups of Lie type over sufficiently large fields $\mathbb{F}_q$ ($q \geq 8$ suffices). The proof was finally completed in [44]. This was largely based on character theory, via an elementary classical result, again due to Frobenius, that for an element g of a finite group G, the number of solutions $(x, y) \in G \times G$ to the equation $g = [x, y]$ is equal to

$$|G| \sum_{\chi \in Irr(G)} \frac{\chi(g)}{\chi(1)}.$$

Thus g is a commutator if and only if this sum is nonzero. The aim is to show that for G simple, the term coming from the trivial character (namely 1) is greater than the sum of moduli the remaining terms, in other words that

$$\sum_{\chi \neq 1_G} \frac{|\chi(g)|}{\chi(1)} < 1. \tag{3}$$

Here is a sketch of the proof from [44] of Theorem 3.8 for the family of symplectic groups $G = Sp_{2n}(2)$. The argument proceeds by induction. The base cases for the induction are $Sp_{2n}(2)$ with $n \leq 6$, and these were handled computationally; of course $Sp_2(2)$ and $Sp_4(2)$ are non-perfect, so Theorem 3.8 does not apply to them.

Let $g \in G$, and write g in block-diagonal form

$$g = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & X_k \end{pmatrix} \in Sp_{2n_1}(2) \times \cdots \times Sp_{2n_k}(2) < G,$$

where $\sum n_i = n$, this decomposition being as refined as possible. If each X_i is a commutator in $Sp_{2n_i}(2)$ then g is a commutator in G. Hence induction gives the conclusion except when either

(1) $k = 1$, or

(2) one of the factors $Sp_{2n_i}(2)$ is $Sp_2(2)$ or $Sp_4(2)$.

We call g *unbreakable* if (1) or (2) holds for every such block-diagonal decomposition of g. Thus to prove the theorem for this case it suffices to show that every unbreakable element g of $G = Sp_{2n}(2)$ with $n \geq 7$ is a commutator.

The first step is to prove that the unbreakable element g has small centralizer, namely

$$|C_G(g)| < 2^{2n+15}.$$

For example, if g is unipotent its unbreakability means that it can have few Jordan blocks, and the possiblities for the centralizers of such elements are given by [48, Chapter 7].

Next, a result of Guralnick and Tiep [21] shows that there is a collection $\mathcal{W}$ of 5 irreducible characters of G such that

(i) $\chi(1) \geq \frac{1}{6}(2^n - 1)(2^n - 2)$ for $\chi \in \mathcal{W}$, and

(ii) $\chi(1) \geq 2^{4n-7}$ for $1 \neq \chi \in \mathrm{Irr}(G) \setminus \mathcal{W}$.

Set

$$\Sigma_1(g) = \sum_{\chi \in \mathcal{W}} \frac{|\chi(g)|}{\chi(1)}, \;\; \Sigma_2(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G) \setminus \mathcal{W}} \frac{|\chi(g)|}{\chi(1)}.$$

Letting $k(G)$ denote the number of conjugacy classes of G, it follows from [16, 3.13] that $k(G) \leq (15.2) \cdot 2^n$. Also $\sum_{\chi \in Irr(G)} |\chi(g)|^2 = |C_G(g)|$ by the orthogonailty relations, from which the Cauchy-Schwartz inequality implies that

$$\sum_{\chi \in Irr(G)} |\chi(g)| \leq k(G)^{1/2} |C_G(g)|^{1/2}.$$

Plugging all this into the expression defining $\Sigma_2(g)$, we obtain

$$\Sigma_2(g) < \frac{\sqrt{15.2} \cdot 2^{n/2} \cdot |C_G(g)|^{1/2}}{2^{4n-7}} < \frac{\sqrt{15.2} \cdot 2^{n/2} \cdot 2^{n+7.5}}{2^{4n-7}} < 0.5.$$

Bounding $\Sigma_1(g)$ depends on some detailed analysis of the values $\chi(g)$ for the characters $\chi \in \mathcal{W}$, from which one shows that $\Sigma_1(g) < 0.2$.

Hence $\Sigma_1(g) + \Sigma_2(g) < 0.7$, which implies that (3) holds, and hence g is a commutator, as required.

This example gives the flavour of the proof of Theorem 3.8, but it must be said that other families of classical groups over small fields do not yield so easily as this. Indeed the unitary groups presented too many technical obstacles for us to handle them in this fashion, and we used a completely different method for these.

3.2 Bounded subsets

Conjecture 3.2 has been proved for bounded subsets in [43, Theorem 3]:

Theorem 3.9 *There is an absolute constant c such that if G is a finite non-abelian simple group, and A is any subset of G of size at least 2, then G is a product of N conjugates of A for some $N \leq c \log |G|$.*

We shall sketch a proof of this result for alternating groups, and refer the reader to [43] for the rest of the proof. Suppose then that $G = A_n$.

First we claim that, in proving the conjecture for a subset A, we may assume that $1 \in A$. Indeed, let $a \in A$ and $B = a^{-1}A$. Then $1 \in B$, and if G is a product of N conjugates of B then it is also a product of N conjugates of A. Secondly, we claim we may assume there exists $x \neq 1$ such that $1, x, x^{-1} \in A$. Indeed, suppose $1 \in A$ and let $x \in A$ be a non-identity element (whose existence follows from the assumption $|A| \geq 2$). Then $1, x, x^2 \in A^2$, hence $x^{-1}, 1, x \in x^{-1}A^2$. Assuming the conjecture holds for sets containing $x^{-1}, 1, x$ we deduce that G is a product of say $N \leq c \log |G| / \log |A^2| \leq c \log |G| / \log |A|$ conjugates of $x^{-1}A^2$, hence it is a product of N conjugates of A^2, so G is a product of $2N \leq 2c \log |G| / \log |A|$ conjugates of A.

So assume that $1, x, x^{-1} \in A \subseteq G$ for some $x \neq 1$. It is easy to choose a 3-cycle $y \in A_n$ such that $[x, y] \neq 1$ has support of size at most 5. Let $C = x^{A_n}$, the conjugacy class of x. Since $[x, y] = x^{-1}x^y \in C^{-1}C$, we see that $C^{-1}C$ contains either a 3-cycle, a 5-cycle or a double transposition. In all cases we deduce that $(C^{-1}C)^2$ contains all double transpositions in A_n. Since $x, x^{-1} \in A$, some product of 4 conjugates of A contains $\{1, t\}$ for a double transposition $t \in A_n$.

At this point a straightforward argument shows that it is sufficient to establish the result for the subset $\{1, \tau\}$ of S_{n-2}, where τ is a transposition – in other words, that S_{n-2} is a product of $cn \log n$ conjugates of $T := \{1, \tau\}$ (this is Example 2 in Section 1).

This is not as obvious as it might seem. The key to it is a lemma of Abert [1, Lemma 4]: for positive integers a, b, we have $S_{ab} = ABA$, where A is a conjugate of the natural subgroup $(S_a)^b$ and B is a conjugate of $(S_b)^a$. For notational convenience, replace $n - 2$ by n, and let 2^l be the largest power of 2 that is less than or equal to n. Then $\frac{n}{2} < 2^l \leq n$. Repeated application of Abert's lemma shows that S_{2^l} is a product of $2l - 1$ conjugates of $(S_2)^{2^{l-1}}$, hence of $(2l - 1)2^{l-1}$ conjugates of T. Since it is routine to see that for $\frac{n}{2} < k \leq n$, S_n is a product of at most 8 conjugates of S_k, it follows that S_n is a product of at most $(2l - 1)2^{l+2}$ conjugates of T, and the conclusion follows.

3.3 Bounded rank

Conjecture 3.2 has also been proved for simple groups of Lie type of bounded rank, in [18, Theorem 1.3]:

Theorem 3.10 *Fix a positive integer r. There exists a constant $c = c(r)$ such that if G is a finite simple group of Lie type of rank r and A is a subset of G of size at least 2, then G is a product of N conjugates of A for some $N \leq c \log |G| / \log |A|$.*

It is possible to get some of the way towards this result quite quickly, as follows. Firstly, as observed in the sketch proof of Theorem 3.9 above, we can assume that $1 \in A$. Next, by a result in [22], for $1 \neq x \in A$, there are $m \leq 8(2r + 1)$ conjugates of x that generate G; call them $x^{g_1}, \ldots, x^{g_m}$. Write $S = A^{g_1} \cdots A^{g_m}$. Then S generates G, so by the Product Theorem 2.3, $G = S^d$ for some $d \leq (\frac{\log |G|}{\log |S|})^{c(r)}$, and hence G is a product of $(\frac{\log |G|}{\log |S|})^{c_1(r)}$ conjugates of A.

Getting rid of the exponent $c_1(r)$ takes a lot more effort, and this is the main content of [18]. Along the way, they prove an interesting growth result for conjugates ([18, 1.4]): for G and A as in the theorem above, either $A^3 = G$ or there exists $g \in G$ such that $|AA^g| > |A|^{1+\epsilon}$, where $\epsilon > 0$ depends only on the rank r.

3.4 Sylow subgroups

The width of simple groups with respect to a class of Sylow p-subgroups has only been addressed in the case of groups of Lie type, where p is the natural characteristic.

Theorem 3.11 *If G is a simple group of Lie type over a field of characteristic p, then G is a product of 5 Sylow p-subgroups.*

This was first proved in [46] with a bound of 25 instead of 5; the improvement to 5 was announced in [4]. The proof in [46] uses the BN-structure of G, and shows that if $U \in \mathrm{Syl}_p(G)$ is the unipotent radical of a Borel subgroup B, and V is the unipotent radical of the opposite Borel, then $G = UVUV \cdots VU$ (25 terms). The reduction to 5 terms was achieved by using what has become known as the "Gowers trick", a very useful tool in the theory of width:

Proposition 3.12 *Let $n > 2$ be an integer and let G be a finite group and let k be the minimal degree of a nontrivial complex character of G. Suppose that $A_i \subseteq G$, $i = 1, 2, \ldots, n$ are such that $\frac{|A_i|}{|G|} \geq k^{-(n-2)/n}$. Then $G = A_1 \cdot A_2 \cdots A_n$.*

This can often be used when G is a group of Lie type, since these have relatively large minimal nontrivial character degrees (see [33]).

This result has an application to the width of finite linear groups. The starting point is an elegant result of Hrushovski and Pillay [27], proved using model theory (and not using the classification of finite simple groups):

Theorem 3.13 *Let p be a prime, n a positive integer, and suppose G is a subgroup of $GL_n(p)$ that is generated by elements of order p. Then $G = \langle x_1 \rangle \langle x_2 \rangle \cdots \langle x_k \rangle$ for some elements x_i of order p, where $k = k(n)$ depends only on n.*

Note that the result is trivial if p is bounded in terms of n. It was generalized as follows in [46]:

Theorem 3.14 *There is a function $f : \mathbb{N} \to \mathbb{N}$ such that the following holds. Let n be a positive integer, p a prime with $p \geq f(n)$, and F a field of characteristic p. If G is a finite subgroup of $GL_n(F)$ generated by elements of order p, then G is a product of 5 of its Sylow p-subgroups.*

Again this is proved without the classification, but using the marvellous theorem of Larsen and Pink [34] as a substitute: if S is a finite simple subgroup of $GL_n(F)$, where F is a field of characteristic p, then either S is of Lie type in characteristic p, or $|S|$ is bounded in terms of n. Bounds for the function $f(n)$ in the above theorem are not addressed in [46], but using the classification Guralnick [19] showed that $f(n) = n+3$ works; this is best possible, as can be seen from the example of the alternating group

$A_p < GL_{p-2}(p)$ (via the action on the fully deleted permutation module for A_p over $\mathbb{F}_p$) – clearly A_p is not a product of a bounded number of its Sylow p-subgroups.

4 Word maps

In this section we develop further the theory of word maps on simple groups, introduced in Section 3.1.2. Let $w = w(x_1, \ldots, x_k)$ be a nontrivial word in the free group F_k of rank k, and for a group G, denote also by $w : G^k \to G$ the *word map* sending $(g_1, \ldots, g_k) \to w(g_1, \ldots, g_k)$ for $g_i \in G$. Write $w(G)$ for the image of this map.

We shall focus on word maps on finite (non-abelian) simple groups G. Recall from Theorem 3.5 that there is a constant N_w such that $w(G) \neq \{1\}$ for simple groups G with $|G| > N_w$.

Questions Here are a few natural questions one might ask about word maps:

1. How large is $w(G)$? Previously we saw in Lemma 3.6 that $|w(G)| > |G|^{\delta_w}$ for some $\delta_w > 0$ depending only on w. Can one do better than this?
2. What is the w-width of G, i.e. the width of G with respect to $w(G)$? We saw in Corollary 3.7 that it is bounded above by a constant $c(w)$. Is it possible to improve this?
3. For $g \in G$, define $P_w(g)$ to be the probability that $w(g_1, \ldots, g_k) = g$ for $g_i \in G$ chosen uniformly at random; so
$$P_w(g) = \frac{|w^{-1}(g)|}{|G|^k}.$$
What can one say about the probability distribution P_w on G? Is it always close to the uniform distribution? Or are there words w for which P_w is highly non-uniform?
4. Regarding Question 3, consider for example $G = SL_2(p)$ with p prime. The proportion of elements of order p in G is precisely $\frac{1}{p}$, so one cannot design an algorithm in computational group theory that is based on finding an element of order p in G by random search. But can one find a fiendishly clever word w for which $\sum_{g \in C} P_w(g) >> \frac{1}{p}$, where C is the set of elements of order p? Such a word would be very interesting computationally.

4.1 Size

Sometimes $w(G) = G$ for all simple groups G – for example for the commutator word $w = [x_1, x_2]$, by the Ore Conjecture (Theorem 3.8); and sometimes $w(G) \neq G$ – for example for $w = x_1^2$, or any power word $w = x_1^k$ for which $\mathrm{hcf}(k, |G|) \neq 1$. Nevertheless, the following result of Larsen and Shalev [36, 2.1 and 1.10] shows that images of word maps on simple groups are always large:

Theorem 4.1 *Let w be a nontrivial word and r a positive integer. There exist positive constants $N(w)$ and $c(r)$ depending only on w and r respectively, such that the following hold.*

(i) *If G is a simple group of Lie type of rank at most r, then $|w(G)| > c(r)|G|$ provided $|G| > N(w)$.*

(ii) *If G is an alternating group A_n, then $|w(G)| > n^{-4}|G|$ provided $n > N(w)$.*

In fact a result stronger than (i) is proved in [36, 1.12]: one can take $c(r) = cr^{-1}$ for some absolute constant c, provided G is not of type PSL or PSU.

There are some interesting tools used in the proof of the above theorem. For (i), a crucial ingredient is a result of Borel [8], which states that if $G = G(q)$ is of Lie type over $\mathbb{F}_q$, and $\bar{G} = G(\bar{\mathbb{F}}_q)$ is the corresponding simple algebraic group over the algebraic closure $\bar{\mathbb{F}}_q$, then the word map $w : \bar{G}^k \to \bar{G}$ is dominant, which is to say that it has dense image. Further arguments from algebraic geometry are used to deduce part (i).

The proof of part (ii) involves a neat application of the celebrated result of Vinogradov [65] that every sufficiently large odd integer is a sum of three primes. So let n be large, and write $n = p_1 + p_2 + p_3 + 3 + \delta$ with p_i primes and $\delta \in \{0, 1\}$. The group $L_i := PSL_2(p_i)$ has a 2-transitive action of degree $p_i + 1$, so we can embed $L_1 \times L_2 \times L_3 < A_n$ in a natural way. A by-product of the proof of part (i) is that $w(L_i)$ contains an element x_i of order $\frac{p_i-1}{2}$, and x_i acts in the degree $p_1 + 1$ representation as a product of two cycles of length $\frac{p_i-1}{2}$ and two fixed points. Hence $x := x_1x_2x_3 \in w(A_n)$ has 6 long cycles and 6 or 7 fixed points. Then $|C_{A_n}(x)|$ is of the order of n^6, which shows that $|w(A_n)|$ is at least of the order of $n^{-6}|A_n|$. Improving the exponent to -4 (in fact to $-29/9$ in [36, 1.10]) takes more work.

There are some related results that should be mentioned here, which show that if one omits the condition that G is sufficiently large in terms of w in the above theorem, then $w(G)$ can be an arbitrary subset of G subject to the obvious necessary condition that it contains the identity and is invariant under $\mathrm{Aut}(G)$. Indeed, in [52], Lubotzky proves:

Theorem 4.2 *Let G be a finite non-abelian simple group, and let A be a subset of G such that $1 \in A$ and A is invariant under $\mathrm{Aut}(G)$. Then there is a word $w = w(x_1, x_2)$ in the free group of rank 2 such that $w(G) = A$.*

Explicit constructions of such words can be found in [31], and further results of this type in [41].

4.2 Width

Recall that for a word w and a simple group G such that $w(G) \neq 1$, the *w-width* of G is the width of G with respect to $w(G)$. A rather crude bound for w-width was given in Corollary 3.7. Can this be improved?

We pointed out at the beginning of the last section that this width is greater than 1 if w is a power word x_1^k. Hence the following remarkable result, the culmination of several papers of Shalev together with Larsen and Tiep [35, 36, 38, 62], is the best possible one of its kind.

Theorem 4.3 *For any nontrivial word w there is a constant N_w such that $w(G)^2 = G$ for all finite non-abelian simple groups G of order greater than N_w.*

Thus the w-width of all sufficiently large simple groups is at most 2. The proof that it is at most 3, originally a result in [62], was simplified for groups of Lie type in [55] using the Gowers trick (Proposition 3.12). Here is their idea in the bounded rank case. Proposition 3.12 with $n = 3$ implies that if G is a finite group with minimal nontrivial character degree k, and $A \subseteq G$ with $|A| \geq k^{-1/3}|G|$, then $G = A^3$. Letting $G = G(q)$ be a simple group of Lie type of rank r over $\mathbb{F}_q$, we have $k \geq aq^r$ for some positive absolute constant a by [33]. Fixing r, we have $|w(G)| > (aq^r)^{-1/3}|G|$ for sufficiently large q by Theorem 4.1(i), and hence $G = w(G)^3$, giving the claimed result for groups of bounded rank.

The problem of determining w-width was termed the "Waring problem" for simple groups by Shalev, by analogy with the celebrated Waring problem in number theory: this concerns the determination of the function $g : \mathbb{N} \to \mathbb{N}$, where $g(k)$ is defined to be minimal such that every positive integer is the sum of $g(k)$ k^{th} powers. (So $g(k)$ could be thought of as the additive width of $\mathbb{N}$ with respect to the set of k^{th} powers.)

In direct analogy with Waring's problem, let us consider the width of the power word x_1^k for simple groups G, where $k \geq 2$. By Theorem 4.3, the width is 2 for sufficiently large G. But this is not the case for *all* G – for example the word x_1^{30} is trivial on A_5. For which values of k could the width be 2 for all simple groups G? Clearly not when k is the exponent of a simple group. An obvious family of positive integers that are not equal to the exponent of a simple group are those which are divisible by at most two primes (by Burnside's p^aq^b theorem). For such integers we have the following result from [20]:

Theorem 4.4 *Let p, q be primes and a, b positive integers, and let $N = p^aq^b$. Then the word map $(x, y) \to x^N y^N$ is surjective on all finite (non-abelian) simple groups.*

4.3 Surjective and non-surjective words

If w has width 1 on G (i.e. $w(G) = G$), we call w a surjective word on G. Some words are surjective on *all* groups: these are precisely the words w in the free group F_k such that $w \in x^{e_1} \cdots x_k^{e_k} F_k'$, where $e_1, \ldots, e_k$ are integers with highest common factor 1 (see [60, 3.1.1]).

We have already observed that there are words that are non-surjective on finite simple groups, such as power words x_1^r. On the other hand, there are various special words that have been proved to be surjective on all finite simple groups: these include the commutator word (Theorem 3.8) and the word $x_1^N x_2^N$ for $N = p^aq^b$ (Theorem 4.4).

Could it be that the only words that are non-surjective on large simple groups are power words of the form $w = v^m$ $(m \geq 2)$? An affirmative answer was stated as a conjecture in [7, 7.14]. However it is not the case:

Theorem 4.5 *Define the word*

$$w = x_1^2 [x_1^{-2}, x_2^{-1}]^2 \in F_2.$$

Then the word map $(x, y) \to w(x, y)$ is non-surjective on $PSL_2(p^{2r+1})$ for all non-negative integers r and all odd primes $p \neq 5$ such that $p^2 \not\equiv 1 \bmod 16$ and $p^2 \not\equiv$

1 mod 5.

For example, w is non-surjective on $PSL_2(3^{2r+1})$ for all r.

This result was proved in [29], as part of a non-surjectivity theorem for the family of words of the form $x_1^2\,[x_1^{-2}, x_2^{-1}]^k$ with $2k+1$ prime.

Here is a sketch of the proof of Theorem 4.5. Let $G = SL_2(K)$ with K a field. The starting point is the observation, going back to Fricke and Klein (see [15]) that for any word $w = w(x_1, x_2)$, there is a polynomial $P_w(s,t,u)$ such that for all $x, y \in G$,

$$\mathrm{Tr}(w(x,y)) = P_w(\mathrm{Tr}(x), \mathrm{Tr}(y), \mathrm{Tr}(xy)).$$

We call P_w the *trace* polynomial of w. A proof of this fact, providing a constructive method of computing P_w for a given word w, can be found in [57, 2.2]. The method is based on the following identities for traces of 2×2 matrices A, B of determinant 1:

$$\begin{array}{ll} (1) & \mathrm{Tr}(AB) = \mathrm{Tr}(BA) \\ (2) & \mathrm{Tr}(A^{-1}) = \mathrm{Tr}(A) \\ (3) & \mathrm{Tr}(A^2B) = \mathrm{Tr}(A)\mathrm{Tr}(AB) - \mathrm{Tr}(B). \end{array}$$

As an example, let us compute P_c for the commutator word $c = [x_1, x_2]$. First observe that

$$\begin{array}{rl} \mathrm{Tr}(x^2y^2) & = \mathrm{Tr}(x)\mathrm{Tr}(xy^2) - \mathrm{Tr}(y^2) \ \ ((\text{by } (3)) \\ & = \mathrm{Tr}(x)(\mathrm{Tr}(y)\mathrm{Tr}(yx) - \mathrm{Tr}(x)) - \mathrm{Tr}(y)^2 + 2 \\ & = stu - s^2 - t^2 + 2, \end{array}$$

where $s = \mathrm{Tr}(x), t = \mathrm{Tr}(y), u = \mathrm{Tr}(xy)$. Hence

$$\begin{array}{rl} \mathrm{Tr}(x^{-1}y^{-1}xy) & = \mathrm{Tr}((x^{-1}y^{-1})^2yxxy) \\ & = \mathrm{Tr}(x^{-1}y^{-1})\mathrm{Tr}(xy) - \mathrm{Tr}(yxxy) \ \ ((\text{by } (3)) \\ & = \mathrm{Tr}(yx)\mathrm{Tr}(xy) - \mathrm{Tr}(x^2y^2) \ \ ((\text{by } (1),(2)). \end{array}$$

It follows that $P_c = s^2 + t^2 + u^2 - stu - 2$.

If one plays around with the polynomials P_w for various words w, they do not appear to have any obvious (or non-obvious) nice behaviour. However, for the magic word $w = x_1^2\,[x_1^{-2}, x_2^{-1}]^2$ in Theorem 4.5, the polynomial P_w turns out to have a miraculous property. We compute that

$$P_w = s^{10} - 2s^9tu - 10s^8 + 2s^8t^2 + s^8t^2u^2 + \cdots - 6s^2u^2 - 2,$$

a polynomial with 29 terms, of degree 12. What is this miraculous property?

Claim Let p be a prime with $p \neq 2, 5$, $p^2 \not\equiv 1 \bmod 16$ and $p^2 \not\equiv 1 \bmod 5$, and let $F = \mathbb{F}_{p^{2r+1}}$. Then

$$P_w(s,t,u) \neq 0 \quad \text{for all } s,t,u \in F.$$

It follows from this that for any $x, y \in SL_2(F)$ we have $\mathrm{Tr}(w(x,y)) = P_w(s,t,u) \neq 0$. Hence the image of w contains no matrices of trace 0, and it follows that w is non-surjective on $PSL_2(F)$, proving Theorem 4.5.

Proof of Claim The claim follows from the following amazing factorization. Letting ζ be a primitive 5^{th} root of unity, P_w factorizes over $\mathbb{Z}[\zeta + \zeta^{-1}]$ as follows:

$$\begin{aligned} P_w(s,t,u) = \quad & (s^2-2) \ \times \\ & (s^4 - s^3tu + s^2t^2 - 4s^2 + 2 + \zeta + \zeta^{-1}) \ \times \\ & (s^4 - s^3tu + s^2t^2 - 4s^2 + 2 + \zeta^2 + \zeta^{-2}). \end{aligned}$$

Let $s, t, u \in F$. If the first factor $s^2 - 2$ is 0, then F has a square root of 2, which is not the case by the assumption that $p^2 \not\equiv 1 \bmod 16$. And if one of the other factors is 0, then $\zeta + \zeta^{-1} \in F$, which is also impossible since $p^2 \not\equiv 1 \bmod 5$. Hence $P_w(s,t,u) \neq 0$, proving the claim and the theorem.

One might ask how we came up with the magic word w in Theorem 4.5. The answer is that we computed (by machine) the polynomials P_v for v in a list of representatives of minimal length for certain automorphism classes of words in F_2, generated using [12]. We then tested whether these polynomials were surjective on a selection of small fields. Nothing of interest came up until the length of the representatives reached 14 (which is the length of the magic w). We noticed that P_w was nonzero on the fields $\mathbb{F}_3$ and $\mathbb{F}_{27}$. The rest is history.... It is interesting (to me) to note that although, as I have said, computation played a key role in our discovery of the family of non-surjective words, the final proofs in [29] are completely theoretical and make no use at all of machine computation.

In principle one can try to use the same method to look for non-surjective words on higher rank groups. For example, for a word map w on $G = SL_3(K)$, the trace of $w(x,y)$ for $x, y \in G$ can be expressed as a polynomial in the variables $\mathrm{Tr}(x^{\pm 1})$, $\mathrm{Tr}(y^{\pm 1})$, $\mathrm{Tr}((xy)^{\pm 1})$, $\mathrm{Tr}((x^{-1}y)^{\pm 1})$, $\mathrm{Tr}([x,y])$ (see [28, 4.6]). Again, there is an algorithm for computing these polynomials, so as above one can test for non-surjectivity on small fields in the hope of coming up with promising words. No such promising words have come up in tests so far, and indeed it may be that there are no magic words to be found for higher ranks. In this direction we propose the following conjecture:

Conjecture 4.6 *Let w be a nontrivial word, and assume that w is not a proper power (i.e. there is no word v such that $w = v^m$ with $m \geq 2$). Then there is a constant $r = r(w)$ such that w is surjective on all simple groups of Lie type of rank at least r and all alternating groups of degree at least r.*

4.4 Probability

Recall that for a nontrivial word $w \in F_k$ and a finite group G, we define the probability distribution P_w on G by

$$P_w(g) = \frac{|w^{-1}(g)|}{|G|^k} \ \ (g \in G).$$

Let U be the uniform distribution on G (so $U(g) = \frac{1}{|G|}$ for all $g \in G$). For an infinite family $\mathcal{F}$ of groups, we say that the word map w is almost uniform on $\mathcal{F}$ if for groups $G \in \mathcal{F}$ we have

$$||P_w - U||_1 := \sum_{g \in G} |P_w(g) - U(g)| \to 0 \ \text{ as } \ |G| \to \infty.$$

When $\mathcal{F}$ is the finite simple groups, various word maps have been shown to be almost uniform: the commutator word $[x_1, x_2]$ in [17]; and the words $x_1^a x_2^b$ in [37].

Does there exist a word map that is highly non-uniform on a family of simple groups? Currently there is not much evidence for or against this. However as observed by Macpherson and Tent in [54, 4.10], one can say the following. For a word w and a family $G(q)$ of groups of a fixed Lie type, as $q \to \infty$ the fibres $w^{-1}(g)$ have cardinalities of the order of cq^d with d a non-negative integer, where the number of possibilites for c, d is bounded; the same applies to the cardinality of $w^{-1}(C)$ for a conjugacy class C. It follows, for example, that for a word map $w = w(x_1, \ldots, x_k)$ on the family $PSL_2(p)$ (p prime), as $p \to \infty$ the probability that $w(g_1, \ldots, g_k)$ has order p for random g_i is of the order of $\frac{1}{p^c}$ for $c = 1, 2$ or 3. In particular, it cannot be of an order of magnitude greater than $\frac{1}{p}$, giving a disappointingly negative answer to Question 4 stated at the beginning of this section.

References

[1] M. Abert, Symmetric groups as products of abelian subgroup, *Bull. London Math. Soc.* **34** (2002), 451–456.

[2] Z. Arad, J. Stavi & M. Herzog, Powers and products of conjugacy classes in groups, in: *Products of conjugacy classes in groups*, 6–51, Lecture Notes in Math. **1112**, Springer, Berlin, 1985.

[3] L. Babai, R. Beals & A. Seress, On the diameter of the symmetric group: polynomial bounds, Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 1108–1112, ACM, New York, 2004.

[4] L. Babai, N. Nikolov & L. Pyber, Product growth and mixing in finite groups, In: Proc. 19th Ann. Symp. on Discrete Algorithms (SODA'08), ACM–SIAM 2008, 248–257.

[5] L. Babai & A. Seress, On the diameter of Cayley graphs of the symmetric group, *J. Combin. Theory Ser. A* **49** (1988), 175–179.

[6] L. Babai & A. Seress, On the diameter of permutation groups, *European J. Combin.* **13** (1992), 231–243.

[7] T. Bandman, S. Garion & B. Kunyavskii, Equations in simple matrix groups: algebra, geometry, arithmetic, dynamics, *Cent. Eur. J. Math.* **12** (2014), 175–211.

[8] A. Borel, On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164.

[9] J. Bourgain & A. Gamburd, Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$, *Ann. of Math.* **167** (2008), 625–642.

[10] J. Bourgain, N. Katz & T. Tao, A sum-product estimate in finite fields, and applications, *Geom. Funct. Anal.* **14** (2004), 27–57.

[11] E. Breuillard, B. Green & T. Tao, Approximate subgroups of linear groups, *Geom. Funct. Anal.* **21** (2011), 774–819.

[12] B. Cooper & E. Rowland, Growing words in the free group on two generators, *Illinois J. Math.* **55** (2011), 417–426.

[13] E.W. Ellers & N. Gordeev, On the conjectures of J. Thompson and O. Ore, *Trans. Amer. Math. Soc.* **350** (1998), 3657–3671.

[14] E.E. Ellers, N. Gordeev & M. Herzog, Covering numbers for Chevalley groups, *Israel J. Math.* **111** (1999), 339–372.

[15] R. Fricke & F. Klein, Vorlesungen über die Theorie der Automorphen Funktionen, 1 and 2, Teubner, Leipzig, 1897 and 1912.

[16] J. Fulman & R. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups with applications to derangements, *Trans. Amer. Math. Soc.* **364** (2012), 3023–3070.

[17] S. Garion & A. Shalev, Commutator maps, measure preservation, and T-systems, *Trans.*

Amer. Math. Soc. **361** (2009), 4631–4651.
[18] N. Gill, I. Short, L. Pyber & E. Szabó, On the product decomposition conjecture for finite simple groups, *Groups Geom. Dyn.* **7** (2013), 867–882.
[19] R.M. Guralnick, Small representations are completely reducible, *J. Algebra* **220** (1999), 531–541.
[20] R.M. Guralnick, M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, Surjective word maps and Burnside's p^aq^b theorem, preprint.
[21] R. Guralnick & P.H. Tiep, Cross characteristic representations of even characteristic symplectic groups, *Trans. Amer. Math. Soc.* **356** (2004), 4969–5023.
[22] J.I. Hall, M.W. Liebeck & G.M. Seitz, Generators for finite simple groups, with applications to linear groups, *Quart. J. Math. Oxford* **43** (1992), 441–458.
[23] H.A. Helfgott, Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math.* **167** (2008), 601–623.
[24] H.A. Helfgott, Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, *J. Eur. Math. Soc.* **13** (2011), 761–851.
[25] H.A. Helfgott & A. Seress, On the diameter of permutation groups, *Ann. of Math.* **179** (2014), 611–658.
[26] D.G. Higman, Intersection matrices for finite permutation groups, *J. Algebra* **6** (1967), 22–42.
[27] E. Hrushovski & A. Pillay, Definable subgroups of algebraic groups over finite fields, *J. Reine Angew. Math.* **462** (1995), 69–91.
[28] S. Jambor, An $L3 - U3$-quotient algorithm for finitely presented groups, PhD Thesis, RWTH Aachen University (2012).
[29] S. Jambor, M.W. Liebeck & E.A. O'Brien, Some word maps that are non-surjective on infinitely many finite simple groups, *Bull. Lond. Math. Soc.* (2013), **45**, 907–910.
[30] G.A. Jones, Varieties and simple groups, *J. Austral. Math. Soc.* **17** (1974), 163–173.
[31] M. Kassabov & N. Nikolov, Words with few values in finite simple groups, *Q. J. Math.* **64** (2013), 1161–1166.
[32] M. Kassabov & T.R. Riley, Diameters of Cayley graphs of Chevalley groups, *European J. Combin.* **28** (2007), 791–800.
[33] V. Landazuri & G.M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418–443.
[34] M.J. Larsen & R. Pink, Finite subgroups of algebraic groups, *J. Amer. Math. Soc.* **24** (2011), 1105–1158.
[35] M. Larsen & A. Shalev, Characters of symmetric groups: sharp bounds and applications, *Invent. Math.* **174** (2008), 645–687.
[36] M. Larsen & A. Shalev, Word maps and Waring type problems, *J. Amer. Math. Soc.* **22** (2009), 437–466.
[37] M. Larsen & A. Shalev, On the distribution of values of certain word maps, *Trans. Amer. Math. Soc.*, to appear.
[38] M. Larsen, A. Shalev & P.H. Tiep, Waring problem for finite simple groups, *Ann. of Math.* **174** (2011), 1885–1950.
[39] R. Lawther & M.W. Liebeck, On the diameter of a Cayley graph of a simple group of Lie type based on a conjugacy class, *J. Comb. Theory Ser. A* **83** (1998), 118–137.
[40] A. Lev, The covering number of the group $PSL_n(F)$, *J. Algebra* **182** (1996), 60–84.
[41] M. Levy, Images of word maps in almost simple groups and quasisimple groups, *Internat.J. Algebra Comput.* **24** (2014), 47–58.
[42] M.W. Liebeck, H.D. Macpherson & K. Tent, Primitive permutation groups of bounded orbital diameter, *Proc. Lond. Math. Soc.* **100** (2010), 216–248.
[43] M.W. Liebeck, N. Nikolov & A. Shalev Product decompositions in finite simple groups, *Bull. London Math. Soc.* **44** (2012), 469–472.
[44] M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, The Ore conjecture, *J. Eur. Math. Soc.* **12** (2010), 939–1008.
[45] M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, Products of squares in finite simple

groups, *Proc. Amer. Math. Soc.* **140** (2012), 21–33.
[46] M.W. Liebeck & L. Pyber, Finite linear groups and bounded generation, *Duke Math. J.* **107** (2001), 159–171.
[47] M.W. Liebeck & G.M. Seitz, The maximal subgroups of positive dimension in exceptional algebraic groups, *Mem. Amer. Math. Soc.* **169** (2004), No. 802, 1–227.
[48] M.W. Liebeck & G.M. Seitz, *Unipotent and nilpotent classes in simple algebraic groups and Lie algebras*, Amer. Math. Soc. Surveys and Monographs **180** (2012).
[49] M.W. Liebeck & A. Shalev, Classical groups, probabilistic methods, and the $(2,3)$-generation problem, *Ann. of Math.* **144** (1996), 77–125.
[50] M.W. Liebeck & A. Shalev, Diameters of simple groups: sharp bounds and applications, *Ann. of Math.* **154** (2001), 383–406.
[51] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, Progress in Mathematics, vol. 125, Birkhäuser Verlag, Basel, 1994.
[52] A. Lubotzky, Images of word maps in finite simple groups, *Glasg. Math. J.* **56** (2014), 465–469.
[53] A. Lubotzky, Expander Graphs in Pure and Applied Mathematics, *Bull. Amer. Math. Soc.* **49** (2012), 113–162.
[54] H.D. Macpherson & K. Tent, Pseudofinite groups with NIP theory and definability in finite simple groups, Groups and model theory, 255–267, *Contemp. Math.* **576**, Amer. Math. Soc., Providence, RI, 2012.
[55] N. Nikolov & L. Pyber, Product decompositions of quasirandom groups and a Jordan type theorem, *J. Eur. Math. Soc.* **13** (2011), 1063–1077.
[56] O. Ore, Some remarks on commutators, *Proc. Amer. Math. Soc.* **2** (1951), 307–314.
[57] W. Plesken & A. Fabiańska, An L_2-quotient algorithm for finitely presented groups, *J. Algebra* **322** (2009), 914–935.
[58] L. Pyber & E. Szabó, Growth in finite simple groups of Lie type of bounded rank, preprint, arXiv:1005.1858.
[59] J. Ramo, Strongly real elements of orthogonal groups in even characteristic, *J. Group Theory* **14** (2011), 9–30.
[60] D. Segal, *Words: notes on verbal width in groups*, London Math. Soc. Lecture Note Series **361**, Cambridge University Press, Cambridge, 2009.
[61] A. Selberg, On the estimation of Fourier coefficients of modular forms, *Proc. Symp. Pure Math.* **8** (1965), 1–15.
[62] A. Shalev, Word maps, conjugacy classes, and a noncommutative Waring-type theorem, *Ann. of Math.* **170** (2009), 1383–1416.
[63] R.C. Thompson, Commutators in the special and general linear groups, *Trans. Amer. Math. Soc.* **101** (1961), 16–33.
[64] P.H. Tiep & A. Zalesskii, Real conjugacy classes in algebraic groups and finite groups of Lie type, *J. Group Theory* **8** (2005), 291–315.
[65] I.M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, (translated, revised and annotated by K. F. Roth and Anne Davenport), Interscience Publishers, London and New York (1954).

PROFINITE PROPERTIES OF DISCRETE GROUPS

ALAN W. REID

Department of Mathematics, University of Texas, Austin, TX 78712, USA
Email: areid@math.utexas.edu

1 Introduction

This paper is based on a series of 4 lectures delivered at Groups St Andrews 2013. The main theme of the lectures was *distinguishing finitely generated residually finite groups by their finite quotients*. The purpose of this paper is to expand and develop the lectures.

The paper is organized as follows. In §2 we collect some questions that motivated the lectures and this article, and in §3 discuss some examples related to these questions. In §4 we recall profinite groups, profinite completions and the formulation of the questions in the language of the profinite completion. In §5, we recall a particular case of the question of when groups have the same profinite completion, namely Grothendieck's question. In §6 we discuss how the methods of L^2-cohomology can be brought to bear on the questions in §2, and in §7, we give a similar discussion using the methods of the cohomology of profinite groups. In §8 we discuss the questions in §2 in the context of groups arising naturally in low-dimensional topology and geometry, and in §9 discuss parafree groups. Finally in §10 we collect a list of open problems that may be of interest.

Acknowledgement: The material in this paper is based largely on joint work with M. R. Bridson, and with M. R. Bridson and M. Conder and I would like to thank them for their collaborations. I would also like to thank the organizers of Groups St Andrews 2013 for their invitation to deliver the lectures, for their hopsitality at the conference, and for their patience whilst this article was completed. This work was supported in part by NSF grants.

2 The motivating questions

We begin by recalling some terminology. A group Γ is said to be *residually finite* (resp., *residually nilpotent*, *residually-p*, *residually torsion-free-nilpotent*) if for each non-trivial $\gamma \in \Gamma$ there exists a finite group (resp., nilpotent group, p-group, torsion-free-nilpotent group) Q and a homomorphism $\phi : \Gamma \to Q$ with $\phi(\gamma) \neq 1$.

2.1. If a finitely-generated group Γ is residually finite, then one can recover any finite portion of its Cayley graph by examining the finite quotients of the group. It is therefore natural to wonder whether, under reasonable hypotheses, the set

$$\mathcal{C}(\Gamma) = \{G : G \text{ is a finite quotient of } \Gamma\}$$

might determine Γ up to isomorphism.

Assuming that the groups considered are residually finite is a natural condition to impose, since, first, this guarantees a rich supply of finite quotients, and secondly, one can always form the free product $\Gamma * S$ where S is a finitely generated infinite simple group, and then, clearly $\mathcal{C}(\Gamma) = \mathcal{C}(\Gamma * S)$. Henceforth, unless otherwise stated, all groups considered will be residually finite.

The basic motivating question of this work is the following due to Remesselenikov:

Question 1: *If F_n is the free group of rank n, and Γ is a finitely-generated, residually finite group, then does $\mathcal{C}(\Gamma) = \mathcal{C}(F_n)$ imply that $\Gamma \cong F_n$?*

This remains open at present, although in this paper we describe progress on this question, as well as providing structural results about such a group Γ (should it exist) as in Question 1.

Following [31], we define *the genus* of a finitely generated residually finite group Γ to be:

$$\mathcal{G}(\Gamma) = \{\Delta : \mathcal{C}(\Delta) = \mathcal{C}(\Gamma)\}.$$

This definition is taken, by analogy with the theory of quadratic forms over $\mathbf{Z}$ where two integral quadratic forms can be locally equivalent (i.e., at all places of $\mathbf{Q}$), but not globally equivalent over $\mathbf{Z}$.

Question 2: *Which finitely generated (respectively, finitely presented) groups Γ have $\mathcal{G}(\Gamma) = \{\Gamma\}$?*

Question 3: *Which finitely generated (respectively, finitely presented) groups Γ have $|\mathcal{G}(\Gamma)| > 1$?*

Question 4: *How large can $|\mathcal{G}(\Gamma)|$ be for finitely generated (resp., finitely presented) groups?*

Question 5: *What group theoretic properties are shared by (resp., are different for) groups in the same genus?*

In addition, if $\mathcal{P}$ is a class of groups, then we define

$$\mathcal{G}(\Gamma, \mathcal{P}) = \{\Delta \in \mathcal{P} : \mathcal{C}(\Delta) = \mathcal{C}(\Gamma)\},$$

and can ask the same questions upon restricting to groups in $\mathcal{P}$.

2.2. Rather than restricting the class of groups in a genus, we can ask to distinguish finitely generated groups by restricting the quotient groups considered. A particularly interesting case of this is the following. Note first that, a group Γ is residually nilpotent if and only if $\bigcap \Gamma_n = 1$, where Γ_n, the n-th term of the *lower central series* of Γ, defined inductively by setting $\Gamma_1 = \Gamma$ and defining $\Gamma_{n+1} = \langle\, [x, y] : x \in \Gamma_n, y \in \Gamma \,\rangle$.

Two residually nilpotent groups Γ and Λ are said to have the same *nilpotent genus* if they have the same lower central series quotients; i.e., $\Gamma/\Gamma_c \cong \Lambda/\Lambda_c$ for all $c \geq 1$. Residually nilpotent groups with the same nilpotent genus as a free group are termed *parafree*. In [10] Gilbert Baumslag surveyed the state of the art concerning groups of

the same nilpotent genus with particular emphasis on the nature of parafree groups. We will discuss this in more detail in §9 below.

3 Some examples

We begin with a series of examples where one can say something about Questions 1–4.

3.1. We first prove the following elementary result.

Proposition 3.1 *Let Γ be a finitely generated abelian group, then $\mathcal{G}(\Gamma) = \{\Gamma\}$.*

Proof Suppose first that $\Delta \in \mathcal{G}(\Gamma)$ and Δ is non-abelian. We may therefore find a commutator $c = [a, b]$ that is non-trivial. Since Δ is residually finite there is a homomorphism $\phi : \Delta \to Q$, with Q finite and $\phi(c) \neq 1$. However, $\Delta \in \mathcal{G}(\Gamma)$ and so Q is abelian. Hence $\phi(c) = 1$, a contradiction.

Thus Δ is abelian. We can assume that $\Gamma \cong \mathbf{Z}^r \oplus T_1$ and $\Delta \cong \mathbf{Z}^s \oplus T_2$, where T_i $(i = 1, 2)$ are finite abelian groups. It is easy to see that $r = s$, for if $r > s$ say, we can choose a large prime p such that p does not divide $|T_1||T_2|$, and construct a finite quotient $(\mathbf{Z}/p\mathbf{Z})^r$ that cannot be a quotient of Δ.

In addition if T_1 is not isomorphic to T_2, then some invariant factor appears in T_1 say, but not in T_2. One can then construct a finite abelian group that is a quotient of T_1 (and hence Γ_1) but not of Γ_2. □

Note that the proof of Proposition 3.1 also proves the following.

Proposition 3.2 *Let Γ be a finitely generated group, and suppose that $\Delta \in \mathcal{G}(\Gamma)$. Then $\Gamma^{\mathrm{ab}} \cong \Delta^{\mathrm{ab}}$. In particular $b_1(\Gamma) = b_1(\Delta)$.*

3.2. Remarkably, moving only slightly beyond $\mathbf{Z}$ to groups that are virtually $\mathbf{Z}$, the situation is dramatically different. The following result is due to Baumslag [9]. We include a sketch of the proof.

Theorem 3.3 *There exists non-isomorphic meta-cyclic groups Γ_1 and Γ_2 for which $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$. Indeed, both of these groups are virtually $\mathbf{Z}$ and defined as extensions of a fixed finite cyclic group F by $\mathbf{Z}$.*

Sketch Proof What Baumslag actually proves in [9] is the following, and this is what we sketch a proof of:

> $(*)$ *Let F be a finite cyclic group with an automorphism of order n, where n is different from 1, 2, 3, 4 and 6. Then there are at least two non-isomorphic cyclic extensions of F, say Γ_1 and Γ_2 with $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$.*

Recall that the automorphism group of a finite cyclic group of order m is an abelian group of order $\phi(m)$. So in $(*)$ we could take F to be a cyclic group of order 11, which has an automorphism of order 5.

Now let $F = \langle a \rangle$ be a cyclic group of order m, and assume that it admits an automorphism α of order n as in $(*)$. Assume that $\alpha(a) = a^r$. Now some elementary number theory (using that $\phi(m) > 2$ by assumption) shows that we can find an integer ℓ such that $(\ell, n) = 1$, and

$$\text{(i) } \alpha^\ell \neq \alpha, \text{ and (ii) } \alpha^\ell \neq \alpha^{-1}.$$

Now define $\Gamma_1 = \langle\, a, b \mid a^m = 1, b^{-1}ab = a^r \,\rangle$ to be the split extension of F induced by α and $\Gamma_2 = \langle\, a, c \mid a^m = 1, c^{-1}ac = a^{r^\ell} \,\rangle$ be the split extension of F induced by α^ℓ. The key claims to be established are that Γ_1 and Γ_2 are non-isomorphic, and that they have the same genus.

That the groups are non-isomorphic can be checked directly as follows. If $\theta : \Gamma_1 \to \Gamma_2$ is an isomorphism, then θ must map the set of elements of finite order in Γ_1 to those in Γ_2; that is to say θ preserves F, and so induces an automorphism of F. Thus $\theta(a) = a^s$ where $(s, m) = 1$. Moreover since the quotients $\Gamma_i/F \cong \mathbf{Z}$ for $i = 1, 2$, it follows that $\theta(b) = c^\epsilon a^t$ where $\epsilon = \pm 1$ and t is an integer. Now consider $\theta(a^r)$. When $\theta(b) = ca^t$ we get:

$$\alpha(a^s) = a^{rs} = \theta(a^r) = \theta(bab^{-1}) = (ca^t)a^s(ca^t)^{-1} = \alpha^\ell(a^s),$$

and it follows that $\alpha = \alpha^\ell$. A similar argument holds when $\theta(b) = c^{-1}a^t$ to show $\alpha^{-1} = \alpha^\ell$, both of which are contradictions to (ii) above.

We now discuss proving that the groups are in the same genus. Setting $P = \Gamma_1 \times \mathbf{Z}$, Baumslag [9] shows that P is isomorphic to $\Gamma_2 \times \mathbf{Z}$. That Γ_1 and Γ_2 have the same genus now follows from a result of Hirshon [34] (see also [9]) where it is shown that (see Theorem 9 of [34]):

Proposition 3.4 *Suppose that A and B are groups with $A \times \mathbf{Z} \cong B \times \mathbf{Z}$, then $\mathcal{C}(A) = \mathcal{C}(B)$.*

3.3. The case of nilpotent groups more generally is well understood due to work of Pickel [52]. We will not discuss this in any detail, other than to say that, in [52] it is shown that for a finitely generated nilpotent group Γ, $\mathcal{G}(\Gamma)$ consists of a finite number of isomorphism classes of nilpotent groups, and moreover, examples where the genus can be made arbitrarily large are known (see for example [58] Chapter 11). Similar results are also known for polycyclic groups (see [29] and [58]).

3.4. From the perspective of this article, more interesting examples where the genus has cardinality greater than 1 (although still finite) are given by examples of lattices in semi-simple Lie groups. We refer the reader to [4] and [5] for details but we will provide a sketch of some salient points.

Let Γ be a lattice in a semi-simple Lie group, for example, in what follows we shall take $\Gamma = \mathrm{SL}(n, R_k)$ where R_k denotes the ring of integers in a number field k. A natural, obvious class of finite quotients of Γ, are those of the form $\mathrm{SL}(n, R_k/I)$ where $I \subset R_k$ is an ideal. Let π_I denote the reduction homomorphism $\Gamma \to \mathrm{SL}(n, R_k/I)$, and $\Gamma(I)$ the kernel. Note that by Strong Approximation for SL_n (see [53] Chapter 7.4 for example) π_I is surjective for all I. A *congruence subgroup* of Γ is any subgroup

$\Delta < \Gamma$ such that $\Gamma(I) < \Delta$ for some I. A group Γ is said to have the *Congruence Subgroup Property* (henceforth abbreviated to CSP) if every subgroup of finite index is a congruence subgroup.

Thus, if Γ has CSP, then $\mathcal{C}(\Gamma)$ is known precisely, and in effect, to determine $\mathcal{C}(\Gamma)$ is reduced to number theory. Expanding on this, since R_k is a Dedekind domain, any ideal I factorizes into powers of prime ideals. If $I = \prod \mathcal{P}_i^{a_i}$, then it is known that $\mathrm{SL}(n, R_k/I) = \prod \mathrm{SL}(n, R_k/\mathcal{P}_i^{a_i})$. Thus the finite groups that arise as quotients of $\mathrm{SL}(n, R_k)$ are determined by those of the form $\mathrm{SL}(n, R_k/\mathcal{P}_i^{a_i})$. Hence we are reduced to understanding how a rational prime p behaves in the extension $k/\mathbf{Q}$. This idea, coupled with the work of Serre [59] which has shed considerable light on when Γ has CSP, allows construction of non-isomorphic lattices in the same genus.

Example: Let $k_1 = \mathbf{Q}(\sqrt[8]{3^7})$ and $k_2 = \mathbf{Q}(\sqrt[8]{48})$. Let $\Gamma_1 = \mathrm{SL}(n, R_{k_1})$ and $\Gamma_2 = \mathrm{SL}(n, R_{k_2})$ $(n \geq 3)$. Then Γ_1 and Γ_2 have CSP (by [59]), are non-isomorphic (by rigidity) and $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$. The reason for the last statement is that the fields k_1 and k_2 are known to be *adelically equivalent* (see [36]); i.e. their Adele rings are isomorphic. This can be reformulated as saying that if V_i $(i = 1, 2)$ are the sets of valuations associated to the prime ideals in k_1 and k_2, then there is a bijection $\phi : V_1 \to V_2$ such that for all $\nu \in V_1$ we have isomorphisms $(k_1)_\nu \cong (k_2)_{\phi(\nu)}$. This has, as a consequence, the desired identical splitting behavior of rational primes in k_1 and k_2.

3.5. Unlike in the previous subsections, there are recent examples of Bridson [14] of finitely presented groups Γ for which $\mathcal{G}(\Gamma)$ is infinite. This will be discussed further in §5.1.

4 Profinite methods

An important reformulation of the discussion in §2 uses the language of profinite groups. In particular, the language of profinite completions is a particularly convenient formalism for organizing finite quotients of a discrete group. For completeness we provide some discussion of profinite groups and profinite completions of discrete groups. We refer the reader to [56] for a more detailed account of the topics covered here.

4.1. A *directed set* is a partially ordered set I such that for every $i, j \in I$ there exists $k \in I$ such that $k \geq i$ and $k \geq j$. An *inverse system* is a family of sets $\{X_i\}_{\{i \in I\}}$, where I is a directed set, and a family of maps $\phi_{ij} : X_i \to X_j$ whenever $i \geq j$, such that:

- $\phi_{ii} = id_{X_i}$;
- $\phi_{ij}\phi_{jk} = \phi_{ik}$, whenever $i \geq j \geq k$.

Denoting this system by (X_i, ϕ_{ij}, I), the *inverse limit* of the inverse system (X_i, ϕ_{ij}, I) is the set

$$\varprojlim X_i = \left\{ (x_i) \in \prod_{i \in I} X_i \;\middle|\; \phi_{ij}(x_i) = x_j, \text{ whenever } i \geq j \right\}.$$

We record the following standard facts about the inverse limit (see [56] Chapter 1 for further details):

(i) Let (X_i, ϕ_{ij}, I) be an inverse system of non-empty compact, Hausdorff, totally disconnected topological spaces (resp. topological groups) over the directed set I, then $\varprojlim X_i$ is a non-empty, compact, Hausdorff, totally disconnected topological space (resp. topological group).

(ii) Let (X_i, ϕ_{ij}, I) be an inverse system. A subset $J \subset I$ is defined to be *cofinal*, if for each $i \in I$, there exists $j \in J$ with $j \geq i$. If J is cofinal we may form an inverse system (X_j, ϕ_{ij}, J) obtained by omitting those $i \in I \setminus J$. The inverse limit $\varprojlim X_j$ can be identified with the image of $\varprojlim X_i$ under the projection map $\prod_{i\in I} X_i$ onto $\prod_{j\in J} X_j$.

4.2. Returning to the world of group theory, let Γ be a finitely generated group (not necessarily residually finite for this discussion), and let $\mathcal{N}$ denote the collection of all finite index normal subgroups of Γ. Note that $\mathcal{N}$ is non-empty as $\Gamma \in \mathcal{N}$, and we can make $\mathcal{N}$ into directed set by declaring that

$$\text{For } M, N \in \mathcal{N},\ M \leq N \text{ whenever } M \textit{ contains } N.$$

In this case, there are natural epimorphisms $\phi_{NM} : \Gamma/N \to \Gamma/M$, and the inverse limit of the inverse system $(\Gamma/N, \phi_{NM}, \mathcal{N})$ is denoted $\widehat{\Gamma}$ and defined to be to the *profinite completion* of Γ.

Note that there is a natural map $\iota : \Gamma \to \widehat{\Gamma}$ defined by

$$g \mapsto (gN) \in \varprojlim \Gamma/N,$$

and it is easy to see that ι is injective if and only if Γ is residually finite.

An alternative, perhaps more concrete way of viewing the profinite completion is as follows. If, for each $N \in \mathcal{N}$, we equip each Γ/N with the discrete topology, then $\prod\{\Gamma/N : N \in \mathcal{N}\}$ is a compact space and $\widehat{\Gamma}$ can be identified with $\overline{j(\Gamma)}$ where $j : \Gamma \to \prod\{\Gamma/N : N \in \mathcal{N}\}$ is the map $g \mapsto (gN)$.

4.3. From §4.1, $\widehat{\Gamma}$ is a compact topological group, and so a subgroup U is open if and only if it is closed of finite index. In addition, a subgroup $H < \widehat{\Gamma}$ is closed if and only if it is the intersection of all open subgroups of $\widehat{\Gamma}$ containing it. More recently, it is a consequence of a deep theorem of Nikolov and Segal [50] that if Γ is a finitely generated group, then every finite index subgroup of $\widehat{\Gamma}$ is open. Thus a consequence of this is the following elementary lemma (in which $\mathrm{Hom}(G, Q)$ denotes the set of homomorphisms from the group G to the group Q, and $\mathrm{Epi}(G, Q)$ denotes the set of epimorphisms).

Lemma 4.1 *Let Γ be a finitely-generated group and let $\iota : \Gamma \to \widehat{\Gamma}$ be the natural map to its profinite completion. Then, for every finite group Q, the map $\mathrm{Hom}(\widehat{\Gamma}, Q) \to \mathrm{Hom}(\Gamma, Q)$ defined by $g \mapsto g \circ \iota$ is a bijection, and this restricts to a bijection $\mathrm{Epi}(\widehat{\Gamma}, Q) \to \mathrm{Epi}(\Gamma, Q)$.*

We record the following corollary for later use.

Corollary 4.2 *If* Γ_1 *is finitely-generated and* $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$*, then*

$$|\mathrm{Hom}(\Gamma_1, Q)| = |\mathrm{Hom}(\Gamma_2, Q)|$$

for every finite group Q*.*

4.4. The first Betti number of a finitely generated group is

$$b_1(\Gamma) = \dim_{\mathbf{Q}} [(\Gamma/[\Gamma,\Gamma]) \otimes_{\mathbf{Z}} \mathbf{Q}].$$

Given any prime p, one can detect $b_1(\Gamma)$ in the p-group quotients of Γ, since it is the *greatest integer* b *such that* Γ *surjects* $(\mathbf{Z}/p^k\mathbf{Z})^b$ *for every* $k \in \mathbf{N}$. We exploit this observation as follows:

Lemma 4.3 *Let* Λ *and* Γ *be finitely generated groups. If* Λ *is isomorphic to a dense subgroup of* $\widehat{\Gamma}$*, then* $b_1(\Lambda) \geq b_1(\Gamma)$*.*

Proof For every finite group A, each epimorphism $\widehat{\Gamma} \to A$ will restrict to an epimorphism on both Γ and Λ (since by density Λ cannot be contained in a proper closed subgroup). But the resulting map $\mathrm{Epi}(\widehat{\Gamma}, A) \to \mathrm{Epi}(\Lambda, A)$ need not be surjective, in contrast to Lemma 4.1. Thus if Γ surjects $(\mathbf{Z}/p^k\mathbf{Z})^b$ then so does Λ (but perhaps not vice versa). □

4.5. We now discuss the profinite topology on the discrete group Γ, its subgroups and the correspondence between the subgroup structure of Γ and $\widehat{\Gamma}$. We begin by recalling the profinite topology on Γ. This is the topology on Γ in which a base for the open sets is the set of all cosets of normal subgroups of finite index in Γ. Now given a tower $\mathcal{T}$ of finite index normal subgroups of Γ:

$$\Gamma > N_1 > N_2 > \ldots > N_k > \ldots$$

with $\bigcap N_k = 1$, this can be used to define an inverse system and thereby determines a completion of $\widehat{\Gamma}_{\mathcal{T}}$ (in which Γ will inject). Now if the inverse system determined by $\mathcal{T}$ is cofinal (recall §4.1) then the natural homomorphism $\widehat{\Gamma} \to \widehat{\Gamma}_{\mathcal{T}}$ is an isomorphism. That is to say $\mathcal{T}$ determines the full profinite topology of Γ.

The following is important in connecting the discrete and profinite worlds (see [56] 3.2.2, where here we use [50] to replace "open" by "finite index").

Notation Given a subset X of a profinite group G, we write $\overline{X}$ to denote the closure of X in G.

Proposition 4.4 *If* Γ *is a finitely generated residually finite group, then there is a one-to-one correspondence between the set* $\mathcal{X}$ *of subgroups of* Γ *that are open in the profinite topology on* Γ*, and the set* $\mathcal{Y}$ *of all finite index subgroups of* $\widehat{\Gamma}$*.*

Identifying Γ *with its image in the completion, this correspondence is given by:*

- *For* $H \in \mathcal{X}$*,* $H \mapsto \overline{H}$*.*

- *For $Y \in \mathcal{Y}$, $Y \mapsto Y \cap \Gamma$.*

If $H, K \in \mathcal{X}$ and $K < H$ then $[H : K] = [\overline{H} : \overline{K}]$. Moreover, $K \lhd H$ if and only if $\overline{K} \lhd \overline{H}$, and $\overline{H}/\overline{K} \cong H/K$.

The following corollary of this correspondence will be useful in what follows.

Corollary 4.5 *Let Γ be a finitely-generated group, and for each $d \in \mathbf{N}$, let M_d denote the intersection of all normal subgroups of index at most d in Γ. Then the closure $\overline{M}_d$ of M_d in $\widehat{\Gamma}$ is the intersection of all normal subgroups of index at most d in $\widehat{\Gamma}$, and hence $\bigcap_{d\in\mathbb{N}} \overline{M}_d = 1$.*

Proof If N_1 and N_2 are the kernels of epimorphisms from Γ to finite groups Q_1 and Q_2, then $\overline{N_1 \cap N_2}$ is the kernel of the extension of $\Gamma \to Q_1 \times Q_2$ to $\widehat{\Gamma}$, while $\overline{N}_1 \times \overline{N}_2$ is the kernel of the map $\widehat{\Gamma} \to Q_1 \times Q_2$ that one gets by extending each of $\Gamma \to Q_i$ and then taking the direct product. The uniqueness of extensions tells us that these maps coincide, and hence $\overline{N_1 \cap N_2} = \overline{N}_1 \cap \overline{N}_2$. The claims follow from repeated application of this observation. □

If now $H < \Gamma$, the profinite topology on Γ determines some pro topology on H and therefore some completion of H. To understand what happens in certain cases that will be of interest to us, we recall the following. Since we are assuming that Γ is residually finite, H injects into $\widehat{\Gamma}$ and determines a subgroup $\overline{H}$. Hence there is a natural epimorphism $\widehat{H} \to \overline{H}$. This need not be injective. For this to be injective (i.e. the full profinite topology is induced on H) we require the following to hold:

> *For every subgroup H_1 of finite index in H, there exists a finite index subgroup $\Gamma_1 < \Gamma$ such that $\Gamma_1 \cap H < H_1$.*

There are some important cases for which injectivity can be arranged. Suppose that Γ is a group and H a subgroup of Γ, then Γ is called H-separable if for every $g \in G \setminus H$, there is a subgroup K of finite index in Γ such that $H \subset K$ but $g \notin K$; equivalently, the intersection of all finite index subgroups in Γ containing H is precisely H. The group Γ is called *LERF* (or *subgroup separable*) if it is H-separable for every finitely-generated subgroup H, or equivalently, if every finitely-generated subgroup is a closed subset in the profinite topology.

It is important to note that even if the subgroup H of Γ is separable, it need not be the case that the profinite topology on Γ induces the full profinite topology on H. Stronger separability properties do suffice, however, as we now indicate.

Lemma 4.6 *Let Γ be a finitely-generated group, and H a finitely-generated subgroup of Γ. Suppose that Γ is H_1-separable for every finite index subgroup H_1 in H. Then the profinite topology on Γ induces the full profinite topology on H; that is, the natural map $\widehat{H} \to \overline{H}$ is an isomorphism.*

Proof Since Γ is H_1 separable, the intersection of all subgroups of finite index in Γ containing H_1 is H_1 itself. From this it easily follows that there exists $\Gamma_1 < \Gamma$ of finite index, so that $\Gamma_1 \cap H = H_1$. The lemma follows from the discussion above. □

Subgroups of finite index obviously satisfy the conditions of Lemma 4.6, and if Γ is LERF, the conditions of Lemma 4.6 are also satisfied. Hence we deduce the following.

Corollary 4.7 (1) *If Γ is residually finite and H is a finite-index subgroup of Γ, then the natural map from $\widehat{H}$ to $\overline{H}$ is an isomorphism.*

(2) *If Γ is LERF and H is a finitely generated subgroup of Γ, then the natural map from $\widehat{H}$ to $\overline{H}$ is an isomorphism.*

Another case of what the profinite topology does on a subgroup that will be of interest to us is the following. Let Γ be a residually finite group that is the fundamental group of a graph of groups. Let the edge groups be denoted by G_e and the vertex groups by G_v. The profinite topology on Γ is said to be *efficient* if it induces the full profinite topology on G_v and G_e for all vertex and edge groups, and G_v and G_e are closed in the profinite topology on Γ. The main example we will make use of is the following which is well-known:

Lemma 4.8 *Suppose that Γ is a free product of finitely many residually finite groups $G_1, \ldots, G_n$. Then the profinite topology on Γ is efficient.*

Proof Since Γ is residually finite, the trivial group is closed in the profinite topology. To see that each G_i is closed in the profinite topology we prove that Γ is G_i-separable. To that end let G denote one of the G_i, and let $g \in \Gamma \setminus G$. Since $g \notin G$, the normal form for g contains at least one element $a_k \in G_k \neq G$. Since G_k is residually finite there is a finite quotient A of G_k for which the image of a_k is non-trivial. Using the projection homomorphism $G_1 * \ldots * G_n \to G_k \to A$ defines a homomorphism for which the image of G is trivial but the image of g is not. This proves the vertex groups are closed.

To see that the full profinite topology is induced on each G_i, we need to show that for each G_i, $i = 1, \ldots, n$, the following condition holds (recall the condition for injectivity given above). For every subgroup H of finite index in G_i, there exists a finite index subgroup $H_i < \Gamma$ such that $H_i \cap G_i < H$. Let G denote one of the G_i's and assume that $H < G$ is a finite index subgroup. We can also assume that H is a normal subgroup. Then using the projection homomorphism $\Gamma = G_1 * \ldots * G_n \to G/H$ whose kernel K defines a finite index of subgroup of Γ with $K \cap G = H$ as required. □

Note that in the situation of Lemma 4.8, it also follows that $\widehat{\Gamma} \cong \widehat{G}_1 \amalg \widehat{G}_2 \ldots \amalg \widehat{G}_n$ where $\amalg$ indicates the profinite amalgamated product. We refer the reader to [56] Chapter 9 for more on this.

4.6. We now prove one of the key results that we make use of. This is basically proved in [25] (see also [56] pp. 88–89), the mild difference here, is that we employ [50] to replace *topological isomorphism* with *isomorphism*.

Theorem 4.9 *Suppose that Γ_1 and Γ_2 are finitely-generated abstract groups. Then $\widehat{\Gamma}_1$ and $\widehat{\Gamma}_2$ are isomorphic if and only if $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$.*

Proof If $\widehat{\Gamma}_1$ and $\widehat{\Gamma}_2$ are isomorphic then the discussion following the correspondence provided by Proposition 4.4 shows that $\mathcal{C}(\Gamma_1) = \mathcal{C}(\Gamma_2)$.

For the converse, we argue as follows. For each $n \in \mathbf{N}$ let

$$U_n = \bigcap\{U : U \text{ is a normal subgroup of } \Gamma_1 \text{ with } [\Gamma_1 : U] \leq n\}, \quad \text{and}$$
$$V_n = \bigcap\{V : V \text{ is a normal subgroup of } \Gamma_2 \text{ with } [\Gamma_2 : V] \leq n\}.$$

Then $\Gamma_1/U_n \in \mathcal{C}(\Gamma_1)$ and $\Gamma_2/V_n \in \mathcal{C}(\Gamma_2)$. Hence there exists a normal subgroup $K < \Gamma_1$ so that $\Gamma_1/K \cong \Gamma_2/V_n$. Now it follows that K is an intersection of normal subgroups of index $\leq n$, and so $U_n < K$. Hence $|\Gamma_2/V_n| = |\Gamma_1/K| \leq |\Gamma_1/U_n|$. On reversing the roles of Γ_1 and Γ_2 reverses this inequality from which it follows that $\Gamma_2/V_n \cong \Gamma_1/U_n$.

Now for each such n, let A_n denote the set of all isomorphisms Γ_1/U_n onto Γ_2/V_n. For each n this is a finite non-empty set with the property that for $m \leq n$ and $\alpha \in A_n$, then α induces a unique homomorphism $f_{nm}(\alpha) : \Gamma_1/U_m \to \Gamma_2/V_m$ such that the following diagram commutes.

$$\begin{array}{ccc} \Gamma_1/U_n & \longrightarrow & \Gamma_1/U_m \\ \alpha\big\downarrow & & \big\downarrow f_{nm}(\alpha) \\ \Gamma_2/V_n & \longrightarrow & \Gamma_2/V_m \end{array}$$

It follows that $\{A_n, f_{nm}\}$ is an inverse system of (non-empty) finite sets, and so the inverse limit $\varprojlim A_n$ exists and defines an isomorphism of the inverse systems $\varprojlim \Gamma_1/U_n$ and $\varprojlim \Gamma_2/V_n$. Also note that since U_n and V_n are co-final, the discussion in §4.5 shows that they induce the full profinite topology on Γ_1 and Γ_2 respectively and so we have:

$$\widehat{\Gamma}_1 \cong \varprojlim \Gamma_1/U_n \cong \varprojlim \Gamma_2/V_n \cong \widehat{\Gamma}_2$$

as required. □

Thus statements about $\mathcal{C}(\Gamma)$ and $\mathcal{G}(\Gamma)$ can now be rephrased in terms of the profinite completion. For example,

$$\mathcal{G}(\Gamma) = \{\Delta : \widehat{\Delta} \cong \widehat{\Gamma}\}.$$

4.7. We now give some immediate applications of Theorem 4.9 and the previous discussion in the context of the motivating questions.

Lemma 4.10 *Let $\phi : \Gamma_1 \to \Gamma_2$ be an epimorphism of finitely-generated groups. If Γ_1 is residually finite and $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$, then ϕ is an isomorphism.*

Proof Let $k \in \ker \phi$. If k were non-trivial, then since Γ_1 is residually finite, there would be a finite group Q and an epimorphism $f : \Gamma_1 \to Q$ such that $f(k) \neq 1$. This map f does not lie in the image of the injection $\mathrm{Hom}(\Gamma_2, Q) \hookrightarrow \mathrm{Hom}(\Gamma_1, Q)$ defined by $g \mapsto g \circ \phi$. Thus $|\mathrm{Hom}(\Gamma_1, Q)| > |\mathrm{Hom}(\Gamma_2, Q)|$, contradicting Corollary 4.2. □

Definition 4.11 The *rank* $d(\Gamma)$ of a finitely-generated group Γ is the least integer k such that Γ has a generating set of cardinality k. The *rank* $\widehat{d}(G)$ of a profinite group G is the least integer k for which there is a subset $S \subset G$ with $k = |S|$ and $\langle S \rangle$ is dense in G.

If Γ_1 is assumed to be a finitely generated free group of rank r and Γ_2 a finitely generated group with $d(\Gamma_2) = r$ and $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$, then it follows immediately from Lemma 4.10 that Γ_2 is isomorphic to a free group of rank r (using the natural epimorphism $\Gamma_1 \to \Gamma_2$).

Indeed, one can refine this line of argument as follows. In the following proposition, we do not assume that Γ is residually finite.

Proposition 4.12 *Let Γ be a finitely-generated group and let F_n be a free group. If Γ has a finite quotient Q such that $d(\Gamma) = d(Q)$, and $\widehat{\Gamma} \cong \widehat{F}_n$, then $\Gamma \cong F_n$.*

Proof First $\widehat{\Gamma} \cong \widehat{F}_n$, so Q is a quotient of F_n. Hence $n \geq d(Q)$. But $d(Q) = d(\Gamma)$ and for every integer $s \geq d(\Gamma)$ there exists an epimorphism $F_s \to \Gamma$. Thus we obtain an epimorphism $F_n \to \Gamma$, and application of the preceding lemma completes the proof. □

Corollary 4.13 *Let Γ be a finitely-generated group. If Γ and its abelianisation have the same rank, then $\widehat{\Gamma} \cong \widehat{F}_n$ if and only if $\Gamma \cong F_n$.*

Proof Every finitely-generated abelian group A has a finite quotient of rank $d(A)$. □

As an application of Corollary 4.13 we give a quick proof that that free groups and surface groups are distinguished by their finite quotients. For if Γ is a genus $g \geq 1$ surface group, then Γ and its abelianization have rank $2g$. Corollary 4.13 then precludes such a group having the same profinite completion as a free group.

Another application is the following. Another natural generalization of free groups are right angled Artin groups. Let K be a finite simplicial graph with vertex set $V = \{v_1, \ldots, v_n\}$ and edge set $E \subset V \times V$. Then the *right angled Artin group* (or $RAAG$) associated with K is the group $A(K)$ given by the following presentation:

$$A(K) = \langle\, v_1, \ldots, v_n \mid [v_i, v_j] = 1 \text{ for all } i, j \text{ such that } \{v_i, v_j\} \in E \,\rangle.$$

For example, if K is a graph with n vertices and no edges, then $A(K)$ is the free group of rank n, while if K is the complete graph on n vertices, then $A(K)$ is the free abelian group $\mathbf{Z}^n$ of rank n.

If the group Γ has a presentation of the form $\langle\, A \mid R \,\rangle$ where A is finite and all of the relators $r \in R$ lie in the commutator subgroup of the free group $F(A)$, then both Γ and its abelianisation (which is free abelian) have rank $|A|$. The standard presentations of RAAGs have this form.

Proposition 4.14 *If Γ is a right-angled Artin group that is not free, then there exists no free group F such that $\widehat{F} \cong \widehat{\Gamma}$.*

4.8. We shall also consider other pro-completions, and we briefly recall these. The *pro-(finite nilpotent)* completion, denoted $\widehat{\Gamma}_{\mathrm{fn}}$, is the inverse limit of the finite nilpotent quotients of Γ. Given a prime p, the the *pro-p completion* $\widehat{\Gamma}_p$ is the inverse limit of the finite p-group quotients of Γ. As above we have natural homomorphisms $\Gamma \to \widehat{\Gamma}_{\mathrm{fn}}$

and $\Gamma \to \widehat{\Gamma}_p$ and these are injections if and only if Γ is residually nilpotent in the first case and residually p in the second.

Note that in this language, two finitely generated residually nilpotent groups with the same nilpotent genus have isomorphic pro-(finite nilpotent) completions. This can proved in a similar manner as Proposition 4.4 using only the finite nilpotent quotients. Note that it is proved in [6] (before the general case of [50]) that for a finitely generated group Γ, *every* subgroup of finite index in $\widehat{\Gamma}_{\mathrm{fn}}$ is open. Moreover, finitely generated groups in the same nilpotent genus also have isomorphic pro-p completions for all primes p.

5 Grothendieck Pairs and Grothendieck Rigidity

A particular case of when discrete groups have isomorphic profinite completions is the following (which goes back to Grothendieck [28]).

5.1. Let Γ be a residually finite group and let $u : P \hookrightarrow \Gamma$ be the inclusion of a subgroup P. Then $(\Gamma, P)_u$ is called a *Grothendieck Pair* if the induced homomorphism $\widehat{u} : \widehat{P} \to \widehat{\Gamma}$ is an isomorphism but u is not. (When no confusion is likely to arise, it is usual to write (Γ, P) rather than $(\Gamma, P)_u$.) Grothendieck [28] asked about the existence of such pairs of finitely presented groups and the first such pairs were constructed by Bridson and Grunewald in [15]. The analogous problem for finitely generated groups had been settled earlier by Platonov and Tavgen [54]. Both constructions rely on versions of the following result (cf. [54], [15] Theorem 5.2 and [13]).

We remind the reader that the *fibre product* $P < \Gamma \times \Gamma$ associated to an epimorphism of groups $p : \Gamma \to Q$ is the subgroup $P = \{(x, y) : p(x) = p(y)\}$.

Proposition 5.1 *Let $1 \to N \to \Gamma \to Q \to 1$ be a short exact sequence of groups with Γ finitely generated and let P be the associated fibre product. Suppose that $Q \neq 1$ is finitely presented, has no proper subgroups of finite index, and $H_2(Q, \mathbf{Z}) = 0$. Then*

(1) *$(\Gamma \times \Gamma, P)$ is a Grothendieck Pair;*

(2) *if N is finitely generated then (Γ, N) is a Grothendieck Pair.*

More recently in [14], examples of Grothendieck Pairs were constructed so as to provide the first examples of finitely-presented, residually finite groups Γ that contain an infinite sequence of non-isomorphic finitely presented subgroups P_n so that the inclusion maps $u_n : P_n \hookrightarrow \Gamma$ induce isomorphisms of profinite completions. In particular, this provides examples of finitely presented groups Γ for which $\mathcal{G}(\Gamma)$ is infinite.

5.2. There are many classes of groups Γ that can never have a subgroup P for which (Γ, P) is a Grothendieck Pair; as in [40], we call such groups *Grothendieck Rigid*.

Before proving the next theorem, we make a trivial remark that is quite helpful. Suppose that $H < \Gamma$ and Γ is H-separable, then (Γ, H) is not a Grothendieck Pair. The reason for this is that being separable implies that H is contained in (infinitely many) proper subgroups of Γ of finite index. In particular $\overline{H} < \widehat{\Gamma}$ is contained in proper subgroups of finite index in $\widehat{\Gamma}$. On the other hand if (Γ, H) is a Grothendieck Pair, H is dense in $\widehat{\Gamma}$ and so cannot be contained in a closed subgroup (of finite index)

of $\widehat{\Gamma}$. With this remark in place, we prove our next result. Recall that a group Γ is called *residually free* if for every non-trivial element $g \in \Gamma$ there is a homomorphism ϕ_g from Γ to a free group such that $\phi_g(g) \neq 1$, and Γ is *fully residually free* if for every finite subset $X \subseteq \Gamma$ there is a homomorphism from Γ to a free group that restricts to an injection on X.

Theorem 5.2 *Let Γ be a finitely generated group isomorphic to either: a Fuchsian group, a Kleinian group, the fundamental group of a geometric 3-manifold, a fully residually free group. Then Γ is Grothendieck Rigid.*

Proof This follows immediately from the discussion above, and the fact that such groups are known to be LERF. For Fuchsian groups see [57], for Kleinian groups this follows from [2] and [62] and for fully residually free groups [60]. If M is a geometric 3-manifold, then the case when M is hyperbolic follows from the remark above, and when M is a Seifert fibered space see [57]. For those modelled on SOL geometry, separability of subgroups can be established directly and the result follows. □

Remark The case of finite co-volume Kleinian groups was proved in [40] without using the LERF assumption. Instead, character variety techniques were employed. In §8.2 we will establish Grothendieck Rigidity for prime 3-manifolds that are not geometric.

6 L^2-Betti numbers and profinite completion

Proposition 3.2 established that the first Betti number of a group is a profinite invariant. The goal of this section is to extend this to the first L^2-Betti number, and to give some applications of this.

We refer the reader to Lück's paper [47] for a comprehensive introduction to L^2-Betti numbers. For our purposes, it is best to view these invariants not in terms of their original (more analytic) definition, but instead as asymptotic invariants of towers of finite-index subgroups. This is made possible by the Lück's Approximation Theorem [46]:

Theorem 6.1 *Let Γ be a finitely presented group, and let $\Gamma = \Gamma_1 > \Gamma_2 > \ldots > \Gamma_m > \ldots$ be a sequence of finite-index subgroups that are normal in Γ and intersect in the identity. Then for all $p \geq 0$, the p-th L^2-Betti number of Γ is given by the formula*

$$b_p^{(2)}(\Gamma) = \lim_{m\to\infty} \frac{b_p(\Gamma_m)}{[\Gamma : \Gamma_m]}.$$

An important point to note is that this limit does not depend on the tower, and hence is an invariant of Γ. We will mostly be interested in $b_1^{(2)}$.

Example 6.2 Let F be a free group of rank r. Euler characteristic tells us that a subgroup of index d in F is free of rank $d(r-1)+1$, so by Lück's Theorem $b_1^{(2)}(F_r) = r-1$. A similar calculation shows that if Σ is the fundamental group of a closed surface of genus g, then $b_1^{(2)}(\Sigma) = 2g-2$.

Proposition 6.3 *Let Λ and Γ be finitely presented residually finite groups and suppose that Λ is a dense subgroup of $\widehat{\Gamma}$. Then $b_1^{(2)}(\Gamma) \leq b_1^{(2)}(\Lambda)$.*

Proof For each positive integer d let M_d be the intersection of all normal subgroups of index at most d in Γ, and let $L_d = \Lambda \cap \overline{M}_d$ in $\widehat{\Gamma}$. We saw in Corollary 4.5 that $\bigcap_d \overline{M}_d = 1$, and so $\bigcap_d L_d = 1$. Since Λ and Γ are both dense in $\widehat{\Gamma}$, the restriction of $\widehat{\Gamma} \to \widehat{\Gamma}/\overline{M}_d$ to each of these subgroups is surjective, and hence

$$[\Lambda : L_d] = [\widehat{\Gamma} : \overline{M}_d] = [\Gamma : M_d].$$

Now L_d is dense in $\overline{M}_d$, while $\widehat{M_d} = \overline{M}_d$, so Lemma 4.3 implies that $b_1(L_d) \geq b_1(M_d)$, and then we can use the towers (L_d) in Λ and (M_d) in Γ to compare L^2-Betti numbers and find

$$b_1^{(2)}(\Gamma) = \lim_{d\to\infty} \frac{b_1(M_d)}{[\Gamma : M_d]} \leq \lim_{d\to\infty} \frac{b_1(L_d)}{[\Lambda : L_d]} = b_1^{(2)}(\Lambda),$$

by Lück's approximation theorem. □

This has the following important consequence:

Corollary 6.4 *Let Γ_1 and Γ_2 be finitely-presented residually finite groups. If $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$, then $b_1^{(2)}(\Gamma_1) = b_1^{(2)}(\Gamma_2)$.*

If one assumes only that the group Γ is *finitely generated*, then one does not know if the above limit exists, and when it does exist one does not know if it is independent of the chosen tower of subgroups. However, a weaker form of Lück's approximation theorem for $b_1^{(2)}$ was established for finitely generated groups by Lück and Osin [48].

Theorem 6.5 *If Γ is a finitely generated residually finite group and (N_m) is a sequence of finite-index normal subgroups with $\bigcap_m N_m = 1$, then*

$$\limsup_{m\to\infty} \frac{b_1(N_m)}{[\Gamma : N_m]} \leq b_1^{(2)}(\Gamma).$$

6.1. We now give some applications of Proposition 6.3 in the context of Question 1 (and the analogous questions for Fuchsian groups). First we generalize the calculation in Example 6.2.

Proposition 6.6 *If Γ is a lattice in $\mathrm{PSL}(2,\mathbf{R})$ with rational Euler characteristic $\chi(\Gamma)$, then $b_1^{(2)}(\Gamma) = -\chi(\Gamma)$.*

Proof It follows from Lück's approximation theorem that if H is a subgroup of index index d in Γ (which is finitely-presented) then $b_1^{(2)}(H) = d\, b_1^{(2)}(\Gamma)$. Rational Euler characteristic is multiplicative in the same sense. Thus we may pass to a torsion-free subgroup of finite index in Γ, and assume that it is either a free group F_r of rank r, or the fundamental group Σ_g of a closed orientable surface of genus g. The free group case was dealt with above, and so we focus on the surface group case.

Thus if Γ_d is a subgroup of index d in Γ, then it is a surface group of genus $d(g-1)+1$. The first Betti number in this case is $2d(g-1)+1$ and so $b_1(\Gamma_d) = 2 - d\,\chi(\Gamma)$. Dividing by $d = |\Gamma : \Gamma_d|$ and taking the limit, we find $b_1^{(2)}(\Gamma) = -\chi(\Gamma)$. □

With this result and Proposition 6.3 we have the following. The only additional comment to make is that the assumption that the Fuchsian group Γ_1 is non-elementary implies it is not virtually abelian, and so $b_1^{(2)}(\Gamma_1) \neq 0$.

Corollary 6.7 *Let Γ_1 be a finitely generated non-elementary Fuchsian group, and Γ_2 a finitely presented residually finite group with $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$. Then $b_1^{(2)}(\Gamma_2) = b_1^{(2)}(\Gamma_1) = -\chi(\Gamma_1) \neq 0$.*

Another standard result about free groups is that if F is a finitely generated free group of rank ≥ 2, then any finitely generated non-trivial normal subgroup of F has finite index (this also holds more generally for Fuchsian groups and limit groups, see [17] for the last statement). As a further corollary of Propositon 6.4 we prove the following.

Corollary 6.8 *Let Γ be a finitely presented residually finite group in the same genus as a finitely generated free group, and let $N < \Gamma$ be a non-trivial normal subgroup. If N is finitely generated, then Γ/N is finite.*

Proof Proposition 3.1 shows that the genus of the infinite cyclic group contains only itself, and so we can assume that Γ lies in the genus of a non-abelian free group. Thus, by Corollary 6.4, $b_1^{(2)}(\Gamma) \neq 0$. The proof is completed by making use of the following theorem of Gaboriau (see [27] Theorem 6.8):

Theorem 6.9 *Suppose that*

$$1 \to N \to \Gamma \to \Lambda \to 1$$

is an exact sequence of groups where N and Λ are infinite. If $b_1^{(2)}(N) < \infty$, then $b_1^{(2)}(\Gamma) = 0$.

□

Indeed, using Theorem 6.5, Corollary 6.8 can be proved under the assumption that Γ is a finitely generated residually finite group. In this case, the argument establishes that if Γ is in the same genus as a finitely generated free group F, then $b_1^{(2)}(\Gamma) \geq b_1^{(2)}(F)$ and we can still apply [27].

As remarked upon earlier, Question 1 is still unresolved, and in the light of this, Corollary 6.8 provides some information about the structural properties of a finitely generated group in the same genus as a free group. In §8.3, we point out some other properties that occur assuming that a group Γ is in the same genus as a finitely generated free group.

Remark Unlike the case of surface groups, if M is a closed 3-manifold, then typically $b_1^{(2)}(\pi_1(M)) = 0$. More precisely, we have the following from [44]. Let $M = M_1 \# M_2 \# \ldots \# M_r$ be the connect sum of closed (connected) orientable prime 3-manifolds and that $\pi_1(M)$ is infinite. Then

$$b_1^{(2)}(\pi_1(M)) = (r-1) - \sum_{j=1}^{r} \frac{1}{|\pi_1(M_j)|},$$

where in the summation, if $\pi_1(M_j)$ is infinite, the term in the sum is understood to be zero.

6.2. Corollary 6.4 establishes that $b_1^{(2)}$ is an invariant for finitely presented groups in the same genus. A natural question arises as to whether anything can be said about the higher L^2-Betti numbers. Using the knowledge of L^2-Betti numbers of locally symmetric spaces (see [21]), it follows that the examples given §3.4 will have all $b_p^{(2)}$ equal. On the other hand, using [4] examples can be constructed which *do not* have all $b_p^{(2)}$ being equal. Further details will appear elsewhere.

7 Goodness

In this section we discuss how cohomology of profinite groups can be used to inform about Questions 1–5.

7.1. We begin by recalling the definition of the *continuous cohomology* of profinite groups (also known as *Galois cohomology*). We refer the reader to [59] and [56, Chapter 6] for details about the cohomology of profinite groups.

Let G be a profinite group, M a discrete G-module (i.e., an abelian group M equipped with the discrete topology on which G acts continuously) and let $C^n(G, M)$ be the set of all continous maps $G^n \to M$. One defines the coboundary operator $d : C^n(G, M) \to C^{n+1}(G, M)$ in the usual way thereby defining a complex $C^*(G, M)$ whose cohomology groups $H^q(G; M)$ are called the continuous cohomology groups of G with coefficients in M.

Note that $H^0(G; M) = \{x \in M : gx = x \ \forall \ g \in G\} = M^G$ is the subgroup of elements of M invariant under the action of G, $H^1(G; M)$ is the group of classes of continuous crossed homomorphisms of G into M and $H^2(G; M)$ is in one-to-one correspondence with the (equivalence classes of) extensions of M by G.

7.2. Now let Γ be a finitely generated group. Following Serre [59], we say that a group Γ is *good* if for all $q \geq 0$ and for every finite Γ-module M, the homomorphism of cohomology groups

$$H^q(\widehat{\Gamma}; M) \to H^q(\Gamma; M)$$

induced by the natural map $\Gamma \to \widehat{\Gamma}$ is an isomorphism between the cohomology of Γ and the continuous cohomology of $\widehat{\Gamma}$.

Example 7.1 Finitely generated free groups are good.

To see this we argue as follows. As is pointed out by Serre ([59] p. 15), for any (finitely generated) discrete group Γ, one always has isomorphisms $H^q(\widehat{\Gamma}; M) \to H^q(\Gamma; M)$ for $q = 0, 1$. Briefly, using the description of H^0 given above (and the discrete setting), isomorphism for H^0 follows using denseness of Γ in $\widehat{\Gamma}$ and discreteness of M. For H^1, this follows using the description of H^1 as crossed homomorphisms.

If Γ is now a finitely generated free group, since $H^2(\widehat{G}; M)$ is in one-to-one correspondence with the (equivalence classes of) extensions of M by $\widehat{\Gamma}$, it follows that $H^2(\widehat{\Gamma}; M) = 0$ (briefly, like the case of the discrete free group there are no interesting extensions).

The higher cohomology groups $H^q(\widehat{\Gamma}; M)$ $(q \geq 3)$ can also be checked to be zero. For example, since $H^q(\Gamma; M) = 0$ for all $q \geq 2$, the induced map $H^q(\widehat{\Gamma}; M) \to H^q(\Gamma; M)$ is surjective for all $q \geq 2$, and it now follows from a lemma of Serre [59] (see Ex 1 Chapter 2, and also Lemma 2.1 of [41]) that $H^q(\widehat{\Gamma}; M) \to H^q(\Gamma; M)$ is injective for all $q \geq 2$. We also refer the reader to the discussion below on cohomological dimension for another approach.

Goodness is hard to establish in general. One can, however, establish goodness for a group Γ that is LERF if one has a well-controlled splitting of the group as a graph of groups [30]. In addition, a useful criterion for goodness is provided by the next lemma due to Serre (see [59, Chapter 1, Section 2.6])

Lemma 7.2 *The group Γ is good if there is a short exact sequence*

$$1 \to N \to \Gamma \to H \to 1,$$

such that H and N are good, N is finitely-generated, and the cohomology group $H^q(N, M)$ is finite for every q and every finite Γ-module M.

We summarize what we will need from this discussion.

Theorem 7.3 *The following classes of groups are good.*

- *Finitely generated Fuchsian groups.*
- *The fundamental groups of compact 3-manifolds.*
- *Fully residually free groups.*
- *Right angled Artin groups.*

Proof The first and third parts are proved in [30] using LERF and well-controlled splittings of the group, and the fourth is proved in [41]. The second was proved by Cavendish in his PhD thesis [23]. We will sketch the proof when M is closed.

Note first that by [30] free products of residually finite good groups are good, so it suffices to establish goodness for prime 3-manifolds. As is shown in [30] goodness is preserved by commensurability, and so finite groups are clearly good. Thus it remains to establish goodness for prime 3-manifolds with infinite fundamental group. For geometric closed 3-manifolds, goodness will follow immediately from Lemma 7.2 (using the first part of the theorem) when $\Gamma = \pi_1(M)$ and M is a Seifert fibered space or has SOL geometry. For hyperbolic 3-manifolds the work of Agol [2] and Wise [62] shows that any finite volume hyperbolic 3-manifold has a finite cover that fibers over the circle, and once again by Lemma 7.2 (and the first part of the theorem) we deduce

goodness. For manifolds with a non-trivial JSJ decomposition, goodness is proved in [61]. □

7.3. Let G be a profinite group. Then the *p-cohomological dimension* of G is the least integer n such that for every finite (discrete) G-module M and for every $q > n$, the p-primary component of $H^q(G; M)$ is zero, and this is denoted by $\mathrm{cd}_p(G)$. The cohomological dimension of G is defined as the supremum of $\mathrm{cd}_p(G)$ over all primes p, and this is denoted by $\mathrm{cd}(G)$.

We also retain the standard notation $\mathrm{cd}(\Gamma)$ for the cohomological dimension (over $\mathbf{Z}$) of a *discrete* group Γ. A basic connection between the discrete and profinite versions is given by

Lemma 7.4 *Let Γ be a discrete group that is good. If $\mathrm{cd}(\Gamma) \leq n$, then $\mathrm{cd}(\widehat{\Gamma}) \leq n$.*

Proof If $\mathrm{cd}(\Gamma) \leq n$ then $H^q(\Gamma, M) = 0$ for every Γ-module M and every $q > n$. By goodness this transfers to the profinite setting in the context of finite modules. □

Discrete groups of finite cohomological dimension (over $\mathbf{Z}$) are torsion-free. In connection with goodness, we are interested in conditions that allow one to deduce that $\widehat{\Gamma}$ is also torsion-free. For this we need the following result that mirrors the behavior of cohomological dimension for discrete groups (see [59, Chapter 1 §3.3]).

Proposition 7.5 *Let p be a prime, let G be a profinite group, and H a closed subgroup of G. Then $\mathrm{cd}_p(H) \leq \mathrm{cd}_p(G)$.*

This quickly yields the following that we shall use later.

Corollary 7.6 *Suppose that Γ is a residually finite, good group of finite cohomological dimension over $\mathbf{Z}$. Then $\hat{\Gamma}$ is torsion-free.*

Proof If $\widehat{\Gamma}$ were not torsion-free, then it would have an element x of prime order, say q. Since $\langle x \rangle$ is a closed subgroup of $\widehat{\Gamma}$, Proposition 7.5 tells us that $\mathrm{cd}_p(\langle x \rangle) \leq \mathrm{cd}_p(\widehat{\Gamma})$ for all primes p. But $H^{2k}(\langle x \rangle; \mathbf{F}_q) \neq 0$ for all $k > 0$, so $\mathrm{cd}_q(\langle x \rangle)$ and $\mathrm{cd}_q(\widehat{\Gamma})$ are infinite. Since Γ is good and has finite cohomological dimension over $\mathbf{Z}$, we obtain a contradiction from Lemma 7.4. □

Note that this can be used to exhibit linear groups that are not good. For example, in [45], it is shown that there are torsion-free subgroups $\Gamma < \mathrm{SL}(n, \mathbf{Z})$ $(n \geq 3)$ of finite index, for which $\widehat{\Gamma}$ contains torsion of all possible orders. As a corollary of this we have:

Corollary 7.7 *For all $n \geq 3$, any subgroup of $\mathrm{SL}(n, \mathbf{R})$ commensurable with $\mathrm{SL}(n, \mathbf{Z})$ is not good.*

7.4. When the closed subgroup is a p-Sylow subgroup G_p (i.e., a maximal closed pro-p subgroup of G) then we have the following special case of Proposition 7.5 (see [56] §7.3). Note that cohomology theory of pro-p groups is easier to understand than general profinite groups, and so the lemma is quite helpful in connection with computing cohomology of profinite groups.

Lemma 7.8 *Let G_p be a p-Sylow subgroup of G. Then:*

- $\mathrm{cd}_p(G) = \mathrm{cd}_p(G_p) = \mathrm{cd}(G_p)$.
- $\mathrm{cd}(G) = 0$ *if and only if* $G = 1$.
- $\mathrm{cd}_p(G) = 0$ *if and only* $G_p = 1$.

Example 7.9 Let F be a finitely generated free group. Since a p-Sylow subgroup of $\widehat{F}$ is $\mathbf{Z}_p$, Lemma 7.8 gives an efficient way to establish that $\mathrm{cd}(\widehat{F}) = 1$.

7.5. In this subsection we point out how goodness (in fact a weaker property suffices) provides a remarkable condition to establish residual finiteness of extensions. First suppose that we have an extension:

$$1 \to N \to E \to \Gamma \to 1.$$

Using right exactness of the profinite completion (see [56] Proposition 3.2.5), this short exact sequence always determines a sequence:

$$\widehat{N} \to \widehat{E} \to \widehat{\Gamma} \to 1.$$

To ensure that the induced homorphism $\widehat{N} \to \widehat{E}$ is injective is simply again the statement that the full profinite topology is induced on N. As was noticed by Serre [59], this is guaranteed by goodness. Indeed the following is true, the proof of which we discuss below (the proof is sketched in [59] and see also [30] and [41]).

Proposition 7.10 *The following are equivalent for a group Γ.*

- *For any finite Γ-module M, the induced map $H^2(\widehat{\Gamma}; M) \to H^2(\Gamma; M)$ is an isomorphism;*
- *For every group extension $1 \to N \to E \to \Gamma \to 1$ with N finitely generated, the map $\widehat{N} \to \widehat{E}$ is injective.*

Before discussing this we deduce the following.

Corollary 7.11 *Suppose that Γ is residually finite and for any finite Γ-module M, the induced map $H^2(\widehat{\Gamma}; M) \to H^2(\Gamma; M)$ is an isomorphism. Then any extension E (as above) by a finitely generated residually finite group N is residually finite.*

Groups as in Corollary 7.11 are called *highly residually finite* in [41], and *super residually finite* in [22].

Proof By Proposition 7.10, and referring to the diagram below, we have exact sequences with vertical homomorphisms i_N and i_Γ being injective by residual finiteness. Now the squares commute, and so a 5-Lemma argument implies that i_E is injective; i.e., E is residually finite.

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \longrightarrow & E & \longrightarrow & \Gamma & \longrightarrow & 1 \\ & & \downarrow i_N & & \downarrow i_E & & \downarrow i_\Gamma & & \\ 1 & \longrightarrow & \widehat{N} & \longrightarrow & \widehat{E} & \longrightarrow & \widehat{\Gamma} & \longrightarrow & 1 \end{array}$$

□

Proof We discuss the "if" direction below, and refer the reader to [41] for the "only if". We will show that it suffices to prove the result with N finite. For then the case of N finite is dealt with by Proposition 6.1 of [30].

Thus assume that N is finitely generated and J a finite index subgroup of N. Recall that from §4.5 we need to show that there exists a finite index subgroup $E_1 < E$ such that $E_1 \cap N < J$.

To that end, since N is finitely generated we can find a characteristic subgroup $H < J$ of finite index in N that is normal in E. Thus we have:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \longrightarrow & E & \longrightarrow & \Gamma & \longrightarrow & 1 \\ & & \downarrow & & \downarrow \pi & & & & \\ 1 & \longrightarrow & N/H & \longrightarrow & E/H & \longrightarrow & \Gamma & \longrightarrow & 1 \end{array}$$

Assuming that the result holds for the case of N finite we can appy this to N/H. That is to say we can find $E_0' < E/H$ such that $E_0' \cap (N/H) = 1$. Set $E_0 = \pi^{-1}(E_0')$, then $E_0 \cap N < H < J$ as required. □

Given Corollary 7.11 and Theorem 7.3 we have:

Corollary 7.12 *Let Γ be a group as in Theorem 7.3. Then Γ is highly residually finite.*

Examples of groups that are not highly residually finite are $\mathrm{SL}(3, \mathbf{Z})$ (see [33]) and $\mathrm{Sp}(2g, \mathbf{Z})$ ([24]). In particular in [24] lattices in a connected Lie group are constructed that are not residually finite. These arise as extensions of $\mathrm{Sp}(2g, \mathbf{Z})$.

7.6. We now return to Question 1, and in particular deduce some consequences about a group Γ in the same genus as a finitely generated free group. To that end, the following simple observation will prove useful.

Corollary 7.13 *Let Γ_1 and Γ_2 be finitely-generated (abstract) residually finite groups with $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$. Assume that Γ_1 is good and $\mathrm{cd}(\Gamma_1) = n < \infty$. Furthermore, assume that H is a good subgroup of Γ_2 for which the natural mapping $\widehat{H} \to \widehat{\Gamma}_2$ is injective. Then $H^q(H; \mathbf{F}_p) = 0$ for all $q > n$.*

Proof If $H^q(H; \mathbf{F}_p)$ were non-zero for some $q > n$, then by goodness we would have $H^q(\widehat{H}; \mathbf{F}_p) \neq 0$, so $\mathrm{cd}_p(\widehat{H}) \geq q > n$. Now $\widehat{H} \to \widehat{\Gamma}_2$ is injective and so $\widehat{H} \cong \overline{H}$. Hence $\widehat{\Gamma}_1$ contains a closed subgroup of p-cohomological dimension greater than n, a contradiction. □

Corollary 7.14 *If Γ contains a surface group S, and $\widehat{S} \to \widehat{\Gamma}$ is injective, then $\widehat{\Gamma}$ is not isomorphic to the profinite completion of any free group.*

In particular, this also shows the following:

Corollary 7.15 *If L is a non-abelian free group, then $\widehat{L}$ does not contain the profinite completion of any surface group, nor that of any free abelian group of rank greater than 1.*

Remark 7.16 Note that $\widehat{L}$ does contain surface subgroups S of arbitrary large genus (as shown in [12] for example) and free abelian subgroups of arbitrary rank, but the natural map $\widehat{S} \to \widehat{L}$ is never injective. The surface subgroup examples of [12] are in fact dense in $\widehat{L}$.

Next we single out a particular case of an application of the above discussion that connects to two well-known open problems about word hyperbolic groups, namely:

(A) *Does every* 1*-ended word-hyperbolic group contain a surface subgroup?*

(B) *Is every word-hyperbolic group residually finite?*

The first question, due to Gromov, was motivated by the case of hyperbolic 3-manifolds, and in this special case the question was settled recently by Kahn and Markovic [35]. Indeed, given [35], a natural strengthening of (A) above is to ask:

(A′) *Does every* 1*-ended word-hyperbolic group contain a quasi-convex surface subgroup?*

Theorem 7.17 *Suppose that every* 1*-ended word-hyperbolic group is residually finite and contains a quasi-convex surface subgroup. Then there exist no 1-ended word-hyperbolic group* Γ *and free group* F *such that* $\widehat{\Gamma} \cong \widehat{F}$.

Proof Assume the contrary, and let Γ be a counter-example, with $\widehat{\Gamma} \cong \widehat{F}$ for some free group F. Let H be a quasi-convex surface subgroup of Γ. Note that the finite-index subgroups of H are also quasi-convex in Γ. Under the assumption that all 1-ended hyperbolic groups are residually finite, it is proved in [3] that H and all its subgroups of finite index must be separable in Γ. Hence by Lemma 4.6, the natural map $\widehat{H} \to \overline{H} < \widehat{\Gamma}$ is an isomorphism. But as above this yields a contradiction. □

Corollary 7.18 *Suppose that there exists a* 1*-ended word hyperbolic group* Γ *with* $\widehat{\Gamma} \cong \widehat{F}$ *for some free group* F*. Then either there exists a word-hyperbolic group that is not residually finite, or there exists a word-hyperbolic group that does not contain a quasi-convex surface subgroup.*

8 Fuchsian groups, 3-manifold groups and related groups

In this section we prove several results in connection with distinguishing free groups within certain classes of groups. In addition we also prove some results distinguishing 3-manifold groups.

In what follows we denote by $\mathcal{F}$ the collection of Fuchsian groups, and $\mathcal{L}$ the collection of lattices in connected Lie groups.

8.1. In this section we sketch the proof of the following result from [19].

Theorem 8.1 *Let* $\Gamma \in \mathcal{F}$*, then* $\mathcal{G}(\Gamma, \mathcal{L}) = \{\Gamma\}$.

Before commencing with a sketch of the proof, we remark that there exist lattices in connected Lie groups that are not residually finite (recall the discussion at the end of §7.5). For simplicity, in the sketch below we will simply assume all lattices considered are residually finite. This can be bypassed, and we refer the reader to [19] for details on how this is done.

Proof Suppose that $\Delta \in \mathcal{G}(\Gamma, \mathcal{L})$ is residually finite. Then, Corollary 6.7 shows that $b_1^{(2)}(\Delta) \neq 0$. It now follows (see [43] Lemma 1 for example) that Δ fits into a a short exact sequence

$$1 \to N \to \Delta \to F \to 1$$

such that N is finite and F is a lattice in $\mathrm{PSL}(2, \mathbf{R})$.

We next claim that this forces N to be trivial and so Δ is Fuchsian. To see this, suppose that $N \neq 1$. Since N is finite, and Δ is residually finite, it follows that the short exact sequence above can be promoted to a short exact sequence of profinite groups (recall §7.5). Hence the full profinite topology is induced on N by Δ, and we deduce that $\widehat{\Delta}$ contains a non-trivial finite normal subgroup. But, $\widehat{\Delta} \cong \widehat{\Gamma}$, where Γ is a Fuchsian group. This is excluded by the following result proved in [19]. We will not comment on the proof of this result other than to say that it uses profinite group actions on profinite trees. We recall some notation. Write $\mathfrak{cf}(\Gamma)$ to denote the set of conjugacy classes of *maximal finite subgroups* of a group Γ.

Theorem 8.2 *If Γ is a finitely generated Fuchsian group, then the natural inclusion $\Gamma \to \widehat{\Gamma}$ induces a bijection $\mathfrak{cf}(\Gamma) \to \mathfrak{cf}(\widehat{\Gamma})$. More precisely, every finite subgroup of $\widehat{\Gamma}$ is conjugate to a subgroup of Γ, and if two maximal finite subgroup of Γ are conjugate in $\widehat{\Gamma}$ then they are already conjugate in Γ.*

It follows from this that if Γ is a finitely generated non-elementary Fuchsian group, then $\widehat{\Gamma}$ cannot contain a finite non-trivial normal subgroup, since Γ does not.

Given this discussion, to prove Theorem 8.1, it suffices to prove:

Claim: $\mathcal{G}(\Gamma, \mathcal{F}) = \{\Gamma\}$.

Proof of Claim: Suppose that $\Delta \in \mathcal{G}(\Gamma, \mathcal{F})$. If Γ is torsion-free then Δ is torsion-free by Corollary 7.6. Still assuming that Γ is torsion-free, if Γ is a cocompact surface group of genus g then so is Δ. That is to say, Δ cannot be free—this was ruled out by the discussion in §4.7 or Corollary 7.15. In addition, it also cannot be the case that Γ is cocompact and Δ is not (or vice versa). For if this were so, then we could pass to torsion-free subgroups of common finite index that would still have isomorphic profinite completions and this is ruled out by the previous sentence.

If neither Γ_1 nor Γ_2 is cocompact, then each is a free product of cyclic groups. We know that $b_1(\Gamma) = b_1(\Delta)$, and so by Proposition 3.2 the number of infinite cyclic factors in each product is the same. By Theorem 8.2, the finite cyclic factors, being in bijection with the conjugacy classes of maximal finite subgroups, are also the same. Hence the claim follows in this case too.

It only remains to consider the case where both Γ and Δ are cocompact groups with torsion. The genus of Γ is determined by $b_1(\Gamma)$, and so, by Proposition 3.2, Γ and Δ are of the same genus. The periods of Γ and Δ are the orders of representatives of the conjugacy classes of maximal finite subgroups of Γ_i, and so by Theorem 8.2 these must also be the same for Γ and Δ. Thus Γ and Δ have the same signature, and are therefore isomorphic.

This completes the proof of the claim and also Theorem 8.1. □

8.2. In this subsection we focus on proving results distinguishing 3-manifold groups. We summarize this in the following theorem.

Theorem 8.3 *1. Let M be a prime 3-manifold. Then $\pi_1(M)$ is Grothendieck Rigid.*

2. Let Γ be a finitely generated free group of rank $r \geq 2$, and let M be a closed 3-manifold with $\pi_1(M) \in \mathcal{G}(\Gamma)$. Then M is a connect sum of r copies of $S^2 \times S^1$.

3. For $i = 1, 2$, let $M_i = \mathbf{H}^3/\Gamma_i$ where M_1 closed and M_2 non-compact. Then $\widehat{\Gamma}_1$ is not isomorphic to $\widehat{\Gamma}_2$.

4. Let M be a closed hyperbolic 3-manifold and N a geometric 3-manifold. Then if $\pi_1(N) \in \mathcal{G}(\pi_1(M))$, N is a closed hyperbolic 3-manifold.

Proof 1. We have already seen that this holds if M is geometric. Thus we can assume that M is not geometric. Since M is prime, it must therefore have a non-trivial JSJ decomposition. By Theorem 7.3 $\pi_1(M)$ is good. Since M is prime it is aspherical and so we have $H^3(M; \mathbf{F}_p) = H^3(\pi_1(M); \mathbf{F}_p) = \mathbf{F}_p$ for all primes p. On the other hand, if $(\pi_1(M), H)$ is a Grothendieck Pair, where H is finitely generated subgroup of $\pi_1(M)$, then by the discussion in §5.2, H is of infinite index. Moreover, H is also good by Theorem 7.3, and the cover of M corresponding to H, denoted by M_H is still aspherical. However, since this is an infinite sheeted cover, $0 = H^3(M_H; \mathbf{F}_p) = H^3(H; \mathbf{F}_p)$, and hence a contradiction.

2. First, it is clear that $\pi_1(M)$ is infinite. If M is prime, then using the remark at the end of §6, $b_1^{(2)}(\pi_1(M)) = 0$ and the result follows from Corollary 6.4. Thus we can assume that M decomposes as a connect sum $X_1 \# X_2 \# \ldots \# X_s$. Again using the remark in §6 and Example 6.2, we have $s = r$. Also, each X_i has infinite fundamental group since free groups are good and so Lemma 7.6 excludes torsion in the profinite completion.

Now $\pi_1(M)$ has the structure of a free product and so by Lemma 4.8, the profinite topology is efficient. In particular each $\pi_1(X_i)$ is a closed subgroup of $\pi_1(M)$. Suppose that some X_i is not homeomorphic to $S^2 \times S^1$. Then X_i is aspherical, and then either there exists a subgroup $A \cong \mathbf{Z} \oplus \mathbf{Z}$ which is closed in the profinite topology on $\pi_1(X_i)$ and for which the full profinite topology is induced on A (by [57] in the case of Seifert manifolds and [61] for the case where X_i has a non-trivial JSJ decomposition), or there exists a closed surface subgroup of genus > 1 (by [35]) which is closed in the profinite topology and for which the full profinite topology is induced (by [2]). In either case we deduce that $\widehat{\pi_1(M)}$ contains a closed subgroup to which we can apply Corollary 7.13 and deduce a contradiction (by Corollary 7.15).

3. This follows easily from Theorem 7.3 since for all primes p, $H^3(M_2; \mathbf{F}_p) = H^3(\pi_1(M_2); \mathbf{F}_p) = 0$ and $H^3(M_2; \mathbf{F}_p) = H^3(\pi_1(M_2); \mathbf{F}_p) \neq 0$.

4. Since M is closed and hyperbolic, as above, by Theorem 7.3, we can assume that N is closed. It is well known that $\pi_1(M)$ has infinitely many non-abelian finite simple quotients (see [39] for example). Thus we quickly eliminate all possibilities for N apart from those modelled on $\mathbf{H}^2 \times \mathbf{R}$ and $\widetilde{\mathrm{SL}}_2$. In this case, $\pi_1(N)$ has a

description as:

$$1 \to Z \to \pi_1(N) \to F \to 1$$

where Z is infinite cyclic, and F is a cocompact Fuchsian group (we can pass to a subgroup of finite index if necessary so as to arrange the base to be orientable). Since $\pi_1(N)$ is LERF, this short exact sequence can be promoted to (recall the discussion in §4.5):

$$1 \to \widehat{Z} \to \widehat{\pi_1(N)} \to \widehat{F} \to 1.$$

Setting $G = \widehat{\pi_1(N)}$ we have that $G \cong \widehat{\pi_1(M)}$ and so $\pi_1(M)$ is a dense subgroup of G. If $\pi_1(M) \cap \widehat{Z} \neq 1$, then it follows that $\pi_1(M)$ contains an abelian normal subgroup, and this is impossible (as M is a closed hyperbolic 3-manifold). Thus $\pi_1(M) \cap \widehat{Z} = 1$ and therefore $\pi_1(M)$ projects injectively to a dense subgroup of $\widehat{F}$. However, this then contradicts Proposition 6.3. This completes the proof. □

Remarks:

1. Part 1. of Theorem 8.3 was proved in the PhD thesis of W. Cavendish [23] (assuming the then open virtual fibration conjecture for finite volume hyperbolic 3-manifolds).
2. There appears to be no direct proof that distinguishes closed hyperbolic 3-manifolds from finite volume non-compact hyperbolic 3-manifolds by the profinite completions of their fundamental groups. In particular the issue of detecting a peripheral $\mathbf{Z} \oplus \mathbf{Z}$ seems rather delicate.
3. In a similar vein, at present it also seems hard to distinguish a closed prime 3-manifold with a non-trivial JSJ decomposition from a closed hyperbolic 3-manifold by the profinite completions of their fundamental groups. As above the issue of detecting a $\mathbf{Z} \oplus \mathbf{Z}$ is rather subtle.
   However, the author has recently been informed that Wilton and Zalesskii claimed to have now shown that a closed prime 3-manifold with a non-trivial JSJ decomposition from a closed hyperbolic 3-manifold by the profinite completions of their fundamental groups.
4. Funar [26] has shown that there are non-homeomorphic geometric 3-manifolds whose fundamental groups have isomorphic profinite completions. The known examples are torus bundles with SOL geometry. At present, we do not know whether other torus bundles modelled on NIL geometry (which are Seifert fibered), or more generally other Seifert fibered spaces can be distinguished by their finite quotients (even amongst Seifert fibered spaces).

8.3. In this subsection we discuss further properties of a group that is in the same genus as a finitely generated free group. The starting point for this discussion is Section 4 of Peterson and Thom [51] which contains a number of results concerning the structure of finitely presented groups that satisfy their condition ($\star$) and have non-zero $b_1^{(2)}$. We will not state their condition ($\star$) here, but rather remark that the condition is known to hold for left orderable groups and groups that are residually torsion-free nilpotent. We prove the following:

Theorem 8.4 *Let Γ be a finitely presented group in the same genus as a free group F of rank $r \geq 2$. Then:*

(1) *the reduced group C^*-algebra $C^*_\lambda(\Gamma)$ is simple and carries a unique normalised trace.*

(2) *Γ satisfies a the following Freiheitssatz; every generating set $S \subset \Gamma$ has an r-element subset $T \subset S$ such that the subgroup of G generated by T is free of rank r.*

Recall that the *reduced C^*-algebra* $C^*_\lambda(\Gamma)$ is the norm closure of the image of the complex group algebra $\mathbf{C}[\Gamma]$ under the left-regular representation $\lambda_\Gamma : \mathbf{C}[\Gamma] \to \mathcal{L}(\ell^2(\Gamma))$ defined for $\gamma \in \Gamma$ by $(\lambda_\Gamma(\gamma)\xi)(x) = \xi(\gamma^{-1}x)$ for all $x \in \Gamma$ and $\xi \in \ell^2(\Gamma)$. A group Γ is C^*-simple if its reduced C^*-algebra is simple as a complex algebra (i.e., has no proper two-sided ideals). This is equivalent to the statement that any unitary representation of Γ which is weakly contained in λ_Γ is weakly equivalent to λ_Γ. We refer the reader to [32] for a thorough account of the groups that were known to be *C^*-simple* by 2006. The subsequent work of Peterson and Thom [51] augments this knowledge.

An important early result in the field is the proof by Powers [55] that non-abelian free-groups are C^*-simple. In contexts where one is able to adapt the Powers argument, one also expects the canonical trace to be the only normalized trace on $C^*_\lambda(\Gamma)$ (cf. Appendix to [16]). By definition, a linear form τ on $C^*_\lambda(\Gamma)$ is a *normalised trace* if $\tau(1) = 1$ and $\tau(U^*U) \geq 0,\ \tau(UV) = \tau(VU)$ for all $U, V \in C^*_\lambda(\Gamma)$. The *canonical trace* is uniquely defined by

$$\tau_{\text{can}}\left(\sum_{f\in F} z_f \lambda_\Gamma(f)\right) = z_e$$

for every *finite* sum $\sum_{f\in F} z_f\lambda_\Gamma(f)$ where $z_f \in \mathbf{C}$ and $F \subset \Gamma$ contains 1.

Proof Note that Γ is necessarily torsion free since $\widehat{\Gamma} \cong \widehat{F}$. By assumption, we have from Corollary 6.7 that $b_1^{(2)}(\Gamma) = r - 1 \neq 0$ and so both parts of the theorem will follow from [51] once we establish that Γ is left orderable (see Corollary 4.6 and 4.7 of [51]). For this we will make use of a result of Burns and Hale [20] that states that if a group Γ is *locally indicable* (i.e., every finitely generated non-trivial subgroup A admits an epimorphism to $\mathbf{Z}$), then Γ is left orderable. Thus the result will follow from the next theorem. Details of the proof will appear elsewhere, we sketch some of the ideas.

Theorem 8.5 *Γ as in Theorem 8.4 is locally indicable.*

Sketch Proof: Let $A < \Gamma$ be a finitely generated non-trivial subgroup. Since Γ is residually finite, A injects in $\widehat{\Gamma} \cong \widehat{F}$ for a finitely generated free group F of rank ≥ 2. Consider the closure $\overline{A} < \widehat{\Gamma}$ which by a slight abuse of notation we view as sitting in $\widehat{F}$. As a closed subgroup we have from Proposition 7.5 that $\text{cd}(\overline{A}) \leq \text{cd}(\widehat{F}) = 1$ (recall Example 7.9). Since $\overline{A} \neq 1$, and $\text{cd}(\widehat{F}) = 1$ we must have that $\text{cd}_p(\overline{A}) = 1$ for some prime p (see Lemma 7.8). The proof is completed by establishing the following claims:

Claim

(1) *There is an epimorphism* $\overline{A} \to \mathbf{Z}/p\mathbf{Z}$.

(2) *The epimorphism* $\overline{A} \to \mathbf{Z}/p\mathbf{Z}$ *in (1) lifts to an epimorphism* $\overline{A} \to \mathbf{Z}_p$.

Given these claims we can now complete the proof that A surjects onto $\mathbf{Z}$. For A being a dense subgroup of $\overline{A}$ must surject all the finite quotients arising from $\overline{A} \to \mathbf{Z}_p \to \mathbf{Z}/p^n\mathbf{Z}$. That is to say A must surject onto $\mathbf{Z}$.

To prove (1) we exploit the fact that $\mathrm{cd}_p(\overline{A}) = 1$ for some prime p, which allows us to conclude that $H^1(\overline{A}; M) \neq 0$ for some finite $\overline{A}$-module M which is p-primary. To prove (2) we use the fact that since $\mathrm{cd}(\overline{A}) = 1$, $\overline{A}$ is a *projective profinite group* (see [56] Chapter 7.6). In particular this allows for lifting problems to be solved, which is needed to pass from (1) to (2). □

Note that fully residually free groups are residually torsion-free nilpotent and non-abelian fully residually free groups have $b_1^{(2)} \neq 0$ (by [17]). As noted above, ($\star$) of [51] applies, and so these groups also satisfy a similar Freiheitssatz.

9 Parafree groups

Recall that a residually nilpotent group with the same nilpotent genus as a free group is called *parafree*. Many examples of such groups are known (see [7], [8] and [10]). Although much is known about finitely generated parafree groups, a good structure theory for these groups is as yet out of reach. Being in the same nilpotent genus as a parafree group, one might wonder about what properties of a free group are shared by a parafree group. For example, in [10], Baumslag asks:

Question 6: *Let G be a finitely generated parafree group and let $N < G$ be a finitely generated, non-trivial, normal subgroup. Must N be of finite index in G?*

This was answered affirmatively in Corollary 6.8 for groups in the same *genus* as a free group, and using similar methods, in [18] we showed this also holds for the nilpotent genus. In particular we showed that if Γ is a finitely generated parafree group in the same nilpotent genus as a free group of rank $r \geq 2$, then $b_1^{(2)}(\Gamma) \geq r - 1$ and in particular is non-zero. Hence the argument given for proving Corollary 6.8 can still be applied. The argument in [18] uses the following variation of Proposition 6.3.

Proposition 9.1 *Let Γ be a finitely generated group and let F be a finitely presented group that is residually-p for some prime p. Suppose that there is an injection $\Gamma \hookrightarrow \widehat{F}_p$ and that $\overline{\Gamma} = \widehat{F}_p$. Then $b_1^{(2)}(\Gamma) \geq b_1^{(2)}(F)$.*

This has various other consequences for parafree groups; for example the reduced group C^*-algebra is simple and carries a unique normalised trace, and recovers Baumslag's result ([8] Theorem 4.1) that parafree groups also satisfy a Freiheitssatz.

We now discuss some other properties of finitely generated parafree groups. In [37], it was shown that a non-abelian finitely presented parafree group is large (i.e., it contains a finite index subgroup that surjects a non-abelian free group). Another property of free groups (which has come to prominence of late through its connections

to 3-manifold topology) is Agol's RFRS condition (see [1]). To define this recall that the rational derived series of a group Γ is defined inductively as follows. If $\Gamma^{(1)} = [\Gamma,\Gamma]$, then $\Gamma_r^{(1)} = \{\, x \in \Gamma : \text{there exists } k \neq 0, \text{ such that } x^k \in \Gamma^{(1)} \,\}$. If $\Gamma_r^{(n)}$ has been defined then define $\Gamma_r^{(n+1)} = (\Gamma_r^{(n)})_r^{(1)}$.

A group Γ is called *residually finite rationally solvable* (RFRS for short) if there is a sequence of subgroups:

$$\Gamma = \Gamma_0 > \Gamma_1 > G_2 \dots$$

such that $\bigcap_i \Gamma_i = 1$, $[\Gamma : \Gamma_i] < \infty$ and $(\Gamma_i)_r^{(1)} < G_{i+1}$.

Theorem 9.2 *Let Γ be a finitely generated parafree group with the same nilpotent genus of a free group of rank $r \geq 2$. Then Γ is RFRS.*

Proof Fix a prime p, and let G denote the pro-p completion of Γ (which by assumption is the free pro-p group of rank $r \geq 2$). Consider the tower of finite index subgroups defined as $P_1(G) = G$, and $P_{i+1}(G) = (P_i(G))^p[G, P_i(G)]$. Note that each $P_i(G)$ is a closed normal subgroup of G, that $\{P_i(G)\}$ forms a basis of open neighbourhoods of the identity, $\bigcap P_i(G) = 1$ and $P_i(G)/P_{i+1}(G)$ is an elementary abelian p-group.

Since $\Gamma \to G$ is injective, we will consider the subgroups $\{\Delta_i = P_i(G) \cap \Gamma\}$. These are then normal subgroups of finite index in Γ that intersect in the identity. RFRS will follow once we show that $(\Delta_i)_r^{(1)} < \Delta_{i+1}$.

To see this, first note that since each quotient Δ_i/Δ_{i+1} is an elementary abelian p-group, then each Δ_i is normal of p-power index in Γ. Hence $\widehat{\Delta}_{i,p} \to \overline{\Delta}_i < G$ is an isomorphism (since Γ is residually p and Δ_i is normal and of p-power index, the full pro-p topology is induced). Hence $\widehat{\Delta}_{i,p}$ is a free pro-p group of rank l say. It follows that Δ_i has first Betti number equal to l (see [18] Corollary 2.9 for example). Moreover, Δ_i and the free group of rank l have the same p-group quotients, and so it follows that $|\mathrm{Tor}(H_1(\Delta_i;\mathbf{Z})|$ is not divisible by p. The proof is completed by the following lemma. □

Before stating and proving this, we make a preliminary remark. If H is a finitely generated group, then trivially $[H,H] < H_r^{(1)}$, and if $H_1(H;\mathbf{Z})$ is torsion-free, then $H_r^{(1)} = [H,H]$. The next lemma is a variation of this.

Lemma 9.3 *Let p be a prime, H be a finitely generated group and K a normal subgroup of H satisfying:*

- *H/K is an elementary abelian p-group.*
- *$|\mathrm{Tor}(H_1(H;\mathbf{Z})|$ is not divisible by p.*

Then $H_r^{(1)} < K$.

Proof As noted above, if $\mathrm{Tor}(H_1(H;\mathbf{Z})) = 1$ then we are done since $H_r^{(1)} = [H,H] < K$. Thus we may suppose that $\mathrm{Tor}(H_1(H;\mathbf{Z})) \neq 1$. Let $x \in H_r^{(1)}$, so that $x^d \in [H,H]$ for some $d \geq 1$. We will assume that $x \notin K$, otherwise we are done. In particular, $d \geq 2$ since $[H,H] < K$ by the first assumption. Hence x projects to a non-trivial element in $H/[H,H]$ and H/K.

Since H/K is an elementary abelian p-group, it follows from the previous discussion that d is divisible by p. On the other hand, the second assumption is that $|\mathrm{Tor}(H_1(H;\mathbf{Z})|$ is not divisible by p. Putting these statements together, it follows that the image of x must have infinite order in $H/[H,H]$, and this is false. In particular we conclude that d cannot be greater than or equal to 2; i.e., $x \in [H,H] < K$ and the lemma is proved. □

Perhaps the most famous open problem about parafree groups is the *Parafreee Conjecture*. This asserts that if Γ is a finitely generated parafree group, then $H_2(G;\mathbf{Z}) = 0$. Although goodness seems like it may be relevant here, it is not quite the right thing—since the nilpotent genus is only concerned with nilpotent quotients. However, a variation is relevant.

One says that a group Γ is *pro-p good* if for each $q \geq 0$, the homomorphism of cohomology groups

$$H^q(\widehat{\Gamma}_p;\mathbf{F}_p) \to H^q(\Gamma;\mathbf{F}_p)$$

induced by the natural map $\Gamma \to \widehat{\Gamma}_p$ is an isomorphism. One says that the group Γ is *cohomologically complete* if Γ is pro-p good for all primes p. Many groups are known to be cohomologically complete. For example finitely generated free groups, RAAG's [42], and the fundamental group of certain link complements in S^3 (see [11]). However, as is pointed out in [18], there are link complements (even hyperbolic) for which the fundamental group is not cohomologically complete. Note that such an example is good by Theorem 7.3.

The connection with the Parafree Conjecture is the following.

Proposition 9.4 *If finitely generated parafree groups are cohomologically complete, then the Parafree Conjecture is true.*

Proof Suppose that Γ is a finitely generated parafree group. Since Γ is parafree, $\widehat{\Gamma}_p$ is a free pro-p group for all primes p. If we now assume that $H_2(\Gamma;\mathbf{Z}) \neq 0$, then for some prime p we must have $H_2(\Gamma;\mathbf{F}_p) \neq 0$. But then the Universal Coefficient Theorem implies that $H^2(\Gamma;\mathbf{F}_p) \neq 0$. If Γ is pro-p good a contradiction is obtained. □

10 Questions and comments

We close with a list of problems and comments motivated by these notes. First, call a finitely generated discrete group *profinitely rigid* if $\mathcal{G}(\Gamma) = \{\Gamma\}$. We begin with various strengthenings of Question 1.

Question 7: *Are finitely generated Fuchsian groups profinitely rigid?*

Question 8: *Are finitely generated Kleinian groups profinitely rigid?*

Restricting to lattices in $\mathrm{PSL}(2,\mathbf{C})$ we can ask by analogy with the hard part of Theorem 8.1:

Question 9: *Let Γ_1 and Γ_2 be lattices in $\mathrm{PSL}(2,\mathbf{C})$. If $\widehat{\Gamma}_1 \cong \widehat{\Gamma}_2$ is $\Gamma_1 \cong \Gamma_2$?*

Much more ambitious is the following:

Question 10: *Are lattices in rank* 1 *semisimple Lie groups profinitely rigid?*

There is some chance this may be false. In particular, an answer to this question seems closely related to the status of CSP for lattices in $\mathrm{Sp}(n,1)$ $(n \geq 2)$. This is also related to the next three questions.

Question 11: *Does there exist a residually finite word hyperbolic group that is not good?*

Question 12: *Does there exist a residually finite torsion free word hyperbolic group* Γ *for which* $\widehat{\Gamma}$ *contains non-trivial elements of finite order?*

Question 13: *Does there exist a residually finite word hyperbolic group that is not highly residually finite?*

Question 14: *Does there exist a word hyperbolic* Γ *for which* $\mathcal{G}(\Gamma)$ *contains another word hyperbolic group?*

Using Proposition 5.1(2) Grothendieck Pairs (Γ, N) can be constructed so that Γ is word hyperbolic. However, in the known examples, N is not word hyperbolic.

As discussed in §3.4, there are lattices of higher rank for which the genus contains more than one element. However, some interesting special cases seem worth considering.

Question 15: *Is* $\mathrm{SL}(n,\mathbf{Z})$ *profinitely rigid for all* $n \geq 3$*? Is* $\mathrm{SL}(n,\mathbf{Z})$ *Grothendieck Rigid for all* $n \geq 3$*?*

Note that using [15] and [54], for large enough n examples of subgroups $H < \Gamma < \mathrm{SL}(n,\mathbf{Z})$ can be constructed so that (Γ, H) is a Grothendieck Pair.

Motivated by the Parafree Conjecture and a desire to have some type of structure theory for finitely generated parafree groups we raise:

Question 16: *Are finitely generated parafree groups cohomologically complete? How about good?*

We saw in Theorem 9.2 that finitely generated parafree groups are RFRS. The RFRS property holds for groups that are *special* (see [1]). That parafree groups are special seems too much to ask, however, the following seems plausible:

Question 17: *Are finitely generated parafree groups virtually special?*

Note that a positive answer to Question 17 would also imply that finitely generated parafree groups are linear. This is still an open question (see [10] Question 8).

On a slightly different topic. Let Γ_g denote the Mapping Class Group of a closed orientable surface of genus $g \geq 2$.

Question 18: *Is Γ_g profinitely rigid?*

Question 19: *Is Γ_g good?*

This question was raised in [38] in connection with the geometry of moduli spaces of curves of genus g. As pointed out in [38], the answer is known for $g \leq 2$ (the case $g = 1$ follows from Theorem 7.3).

References

[1] I. Agol, *Criteria for virtual fibering*, J. Topology **1** (2008), 269–284.

[2] I. Agol, *The virtual Haken conjecture*, with an appendix by I. Agol, D. Groves, and J. Manning, Documenta Math. **18** (2013), 1045–1087.

[3] I. Agol, D. Groves and J. Manning, *Residual finiteness, QCERF and fillings of hyperbolic groups*, Geometry and Topology **13** (2009), 1043–1073.

[4] M. Aka, *Profinite completions and Kazhdan's Property T*, Groups, Geometry and Dynamics **6** (2012), 221–229.

[5] M. Aka, *Arithmetic groups with isomorphic finite quotients*, J. Algebra **352**, (2012), 322–340.

[6] M.P. Anderson, *Subgroups of finite index in profinite groups*, Pacific J. Math. **62** (1976), 19–28.

[7] G. Baumslag, *Groups with the same lower central sequence as a relatively free group. I*, Trans. Amer. Math. Soc. **129** (1967), 308–321.

[8] G. Baumslag, *Groups with the same lower central sequence as a relatively free group. II*, Trans. Amer. Math. Soc. **142** (1969), 507–538.

[9] G. Baumslag, *Residually finite groups with the same finite images*, Compositio Math. **29** (1974), 249–252.

[10] G. Baumslag, *Parafree groups*, in Infinite groups: geometric, combinatorial and dynamical aspects, 1–14, Progr. Math. **248**, Birkhäuser, (2005).

[11] I. Blomer, P.A. Linnell and T. Schick, *Galois cohomology of completed link groups*, Proc. Amer. Math. Soc. **136** (2008), 3449–3459.

[12] E. Breuillard, T. Gelander, J. Souto and P. Storm, *Dense embeddings of surface groups*, Geometry and Topology **10** (2006), 1373–1389.

[13] M.R. Bridson, *The Schur multiplier, profinite completions and decidability*, Bull. London Math. Soc. **42** (2010), 412–416.

[14] M.R. Bridson, *The strong profinite genus of a finitely presented group can be infinite*, ArXiv: 1401.4084.

[15] M.R. Bridson and F. Grunewald, *Grothendieck's problems concerning profinite completions and representations of groups*, Annals of Math. **160** (2004), 359–373.

[16] M.R. Bridson and P. de la Harpe, *Mapping class groups and outer automorphism groups of free groups are C^*-simple*, J. Funct. Anal. **212** (2004), 195–205.

[17] M.R. Bridson and D.H. Kochloukova, *Volume gradients and homology in towers of residually-free groups*, ArXiv: 1309.1877.

[18] M.R Bridson and A.W Reid, *Nilpotent completions of groups, Grothendieck Pairs and four problems of Baumslag*, to appear I.M.R.N.

[19] M.R. Bridson, M. Conder and A.W. Reid, *Determining Fuchsian groups by their finite quotients*, Israel J. Math., to appear.

[20] R. Burns and V. Hale, *A note on group rings of certain torsion-free groups*, Canadian Math. Bull. **15** (1972), 441–445.

[21] J. Cheeger and M. Gromov, *Bounds on the von Neumann dimension of L^2-cohomology and the Gauss-Bonnet theorem for open manifolds*, J. Diff. Geom. **21** (1985), 1–34.

[22] J. Corson and T. Ratkovich, *A strong form of residual finiteness for groups*, J. Group Theory **9** (2006), 497–505.

[23] W. Cavendish, *Finite-sheeted covering spaces and solenoids over 3-manifolds*, PhD Thesis, Princeton University (2012).
[24] P. Deligne, *Extensions centrales non résiduellment finies de groupes arithmétqiues*, C. R. Acad. Sci. Paris Sér. A **287** (1978), 203–208.
[25] J.D. Dixon, E.W. Formanek, J.C. Poland, L. Ribes, *Profinite completions and isomorphic finite quotients*, J. Pure Appl. Algebra **23** (1982), 227–231.
[26] L. Funar, *Torus bundles not distinguished by TQFT invariants*, Geometry and Topology **17** (2013), 2289–2344.
[27] D. Gaboriau, *Invariants ℓ^2 de relations d'eq́uivalence et de groupes*, Publ. Math. I.H.E.S. **95** (2002), 93–150.
[28] A. Grothendieck, *Representations lineaires et compatifications profinie des groupes discrets*, Manuscripta Math. **2** (1970), 375–396.
[29] F. Grunewald and D. Segal, *On polycyclic groups with the isomorphic finite quotients*, Math. Proc. Camb. Phil. Soc. **84** (1978), 235–246.
[30] F. Grunewald, A. Jaikin-Zapirain, and P.A. Zalesskii, *Cohomological goodness and the profinite completion of Bianchi groups*, Duke Math. J. **144** (2008), 53–72.
[31] F. Grunewald and P. Zalesskii, *Genus for groups*, J. Algebra **326** (2011), 130–168.
[32] P. de la Harpe, *On simplicity of reduced C^*-algebras of groups*, Bull. London Math. Soc. **39** (2007), 1–26.
[33] P.R. Hewitt, *Extensions of residually finite groups*, J. Algebra **163** (1994), 757–772.
[34] R. Hirshon, *Some cancellation theorems with applications to nilpotent groups*, J. Austral. Math. Soc. **23** (1977), 147–165.
[35] J. Kahn and V. Markovic, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, Annals of Math. **175** (2012), 1127–1190.
[36] K. Komatsu, *On the Adele rings of algebraic number fields*, Kodai Math. Sem. Rep. **28** (1976), 78–84.
[37] M. Lackenby, *Detecting large groups*, J. Algebra **324** (2010), 2636–2657.
[38] P. Lochak and L. Schneps, *Open problems in Grothendieck-Teichmuller theory*, in Proc. Symp. Pure Math. **74** 165–186, ed. B. Farb, Amer. Math. Soc. Publications (2006).
[39] D.D. Long and A.W. Reid, *Simple quotients of hyperbolic* 3*-manifold groups*, Proc. Amer. Math. Soc. **126** (1998), 877–880.
[40] D.D. Long and A.W. Reid, *Grothendieck's problem for 3-manifold groups*, Groups, Geometry and Dynamics **5** (2011), 479–499.
[41] K. Lorensen, *Groups with the same cohomology as their profinite completions*, J. Algebra **320** (2008), 1704–1722.
[42] K. Lorensen, *Groups with the same cohomology as their pro-p completions*, J. Pure and Applied Algebra **214** (2010), 6–14.
[43] J. Lott, *Deficiencies of lattice subgroups of Lie groups*, Bull. London Math. Soc. **31** (1999), 191–195.
[44] J. Lott and W. Lück, *L^2-topological invariants of 3-manifolds*, Invent. Math. **120** (1995), 15–60.
[45] A. Lubotzky, *Torsion in the profinite completions of torsion-free groups*, Quart. J. Math. **44** (1993), 327–332.
[46] W. Lück, *Approximating L^2-invariants by their finite-dimensional analogues*, Geom. Funct. Anal. **4** (1994), 455–481.
[47] W. Lück, *L^2-invariants of regular coverings of compact manifolds and CW-complexes*, Handbook of Geometric Topology, North-Holland, Amsterdam (2002), 735–817.
[48] W. Lück and D. Osin, *Approximating the first L^2-Betti number of residually finite groups*, J. of Topology and Analysis, **3** (2011), 153–160.
[49] O.V. Melnikov and P.A. Zalesskii, *Subgroups of profinite groups acting on trees*, Math. USSR Sb. **63** (1989), 405–424.
[50] N. Nikolov and D. Segal, *On finitely generated profinite groups. I. Strong completeness*

and uniform bounds, Annals of Math. **165** (2007), 171–238.
[51] J. Peterson and A. Thom, *Group cocycles and the ring of affiliated operators*, Invent. Math. **185** (2011), 561–592.
[52] P.F. Pickel, *Finitely generated nilpotent groups with isomorphic quotients*, Trans. Amer. Math. Soc. **160** (1971), 327–341.
[53] V.P. Platonov and A. Rapinchuk, *Algebraic groups and Number Theory*, Pure and Applied Mathematics **139** Academic Press (1994).
[54] V.P. Platonov and O.I. Tavgen, *Grothendieck's problem on profinite completions and representations of groups*, K-Theory **4** (1990), 89–101.
[55] R.T. Powers, *Simplicity of the C^*-algebra associated with the free group on two generators*, Duke Math. J. **42** (1975), 151–156.
[56] L. Ribes and P.A. Zalesskii, *Profinite Groups*, Ergeb. der Math. **40**, Springer-Verlag (2000).
[57] G.P. Scott, *Subgroups of surface groups are almost geometric*, J. London Math. Soc. **17** (1978), 555–565. See also *ibid Correction:* J. London Math. Soc. **32** (1985), 217–220.
[58] D. Segal, *Polycyclic Groups*, Cambridge University Press (1983).
[59] J-P. Serre, *Galois Cohomology*, Springer-Verlag (1997).
[60] H. Wilton, *Hall's theorem for limit groups*, Geom. Funct. Anal. **18** (2008), 271–303.
[61] H. Wilton and P.A. Zalesskii, *Profinite properties of graph manifolds*, Geometriae Dedicata **147**, (2010), 29–45.
[62] D.T. Wise, *The structure of groups with a quasi-convex hierarchy*, preprint 2012.

$GL(n,\mathbb{Z})$, $Out(F_n)$ AND EVERYTHING IN BETWEEN: AUTOMORPHISM GROUPS OF RAAGs

KAREN VOGTMANN

University of Warwick, Coventry, UK and Cornell University, Ithaca NY, USA
Email: kvogtmann@gmail.com

Abstract

A right-angled Artin group (RAAG) is a group given by a finite presentation in which the only relations are that some of the generators commute. Free groups and free abelian groups are the extreme examples of RAAGs. Their automorphism groups $GL(n,\mathbb{Z})$ and $Out(F_n)$ are complicated and fascinating groups which have been extensively studied. In these lectures I will explain how to use what we know about $GL(n,\mathbb{Z})$ and $Out(F_n)$ to study the structure of the (outer) automorphism group of a general RAAG. This will involve both inductive local-to-global methods and the construction of contractible spaces on which these automorphism groups act properly. For the automorphism group of a general RAAG the space we construct is a hybrid of the classical symmetric space on which $GL(n,\mathbb{Z})$ acts and Outer space with its action of $Out(F_n)$.

1 Introduction

In these lectures we will study the group of (outer) automorphism groups of a right-angled Artin group. Most of the material can be found in the papers [5, 7, 8, 6] which are all joint with Ruth Charney, some with additional authors. I will first go over some basic facts about right-angled Artin groups, then introduce inductive algebraic methods for studying these groups, then turn to more recent work on geometric methods. I will concentrate on describing and motivating the constructions but avoid proofs, however I will give explicit references to sources where the interested reader can find detailed proofs.

The Groups St Andrews conference was run seamlessly by Colin Campbell, Edmund Robertson, Max Neunhoeffer, Colva Roney-Dougal and Martyn Quick, and I would like to thank them all warmly for inviting me to give these lectures.

2 Lecture 1

2.1 Definition of a RAAG

A right-angled Artin group, or RAAG for short, is a finitely-presented group whose relators (if any) are all simple commutators of generators. The extreme examples of RAAGs are free groups (with no relators), and free abelian groups (with all possible commutators of generators as relators). A RAAG is often specified by drawing a graph Γ with one vertex for each generator and one edge between two vertices if the corresponding generators commute (see Figure 1). Note that Γ is a simplicial complex, i.e., it has no loops or multiple edges.

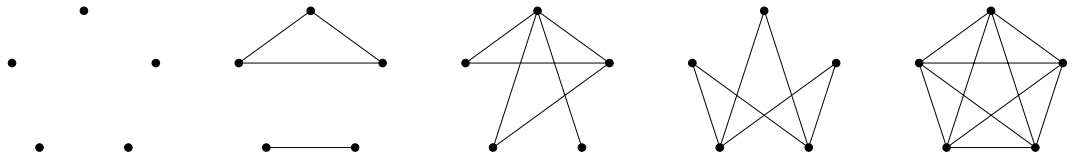


Figure 1. Graphs Γ and the associated RAAGs A_Γ

If we start with any simplicial graph Γ the corresponding RAAG is denoted A_Γ. If Γ is disconnected with components $C_1, \ldots, C_k$, then A_Γ is the free product $A_{C_1} * \ldots * A_{C_k}$ and (just to maximize notational confusion) if Γ is a simplicial join $\Gamma = \Gamma_1 * \Gamma_2$, then A_Γ is the direct product $A_{\Gamma_1} \times A_{\Gamma_2}$.

2.2 Cell complexes with fundamental group A_Γ

Given a presentation $G = \langle X \mid R \rangle$ of a group there is a standard way of constructing a cell complex with fundamental group G, called the *presentation 2-complex*. This has one vertex, an edge for each generator $x \in X$, and a 2-cell for each relator $r \in R$. For a RAAG A_Γ, the 2-cells are all squares (see Figure 2).

The universal cover of the presentation 2-complex for A_Γ is not necessarily contractible, but it can be made contractible by attaching a few more cells. Recall that a *k-clique* in a graph is a complete subgraph with k vertices. If $\Delta \subset \Gamma$ is a k-clique, then the presentation 2-complex for A_Δ is a subcomplex of the presentation 2-complex for A_Γ, and is easily seen to be the 2-skeleton of a k-torus (constructed by gluing opposite sides of a k-dimensional cube). If we fill in this 2-skeleton with the entire k-torus for every k-clique in Γ, the result is called the *Salvetti complex* for A_Γ and is denoted S_Γ.

The Salvetti complex S_Γ is a cube complex which by construction satisfies Gromov's *link condition* and therefore supports a non-positively curved (*locally CAT(0)*) metric. In particular its universal cover is CAT(0) and therefore contractible. The Salvetti complex of a RAAG is in fact a particular kind of non-positively curved cube complex in which hyperplanes are well separated, called a *special cube complex* by Haglund and Wise [16]. We will say a little more about CAT(0) geometry and Gromov's link condition in Lecture 4, but for a thorough introduction to these concepts we refer to [2].

2.3 RAAGs and geometric group theory

RAAGs are important in geometric group theory for many reasons, including the fact that they have very interesting subgroups. They have been in the news lately because of Ian Agol's proof of Thurston's virtual fibering and virtual Haaken conjectures. A key step in those proofs is showing that the fundamental groups of closed hyperbolic 3-manifolds have finite-index subgroups which embed into RAAGs.

The extreme examples of RAAGs do not have such interesting subgroups. A subgroup of $\mathbb{Z}^n$ is a free abelian group of rank at most n. Things get slightly more interesting for F_n, where a subgroup is still a free group but can be of any rank,

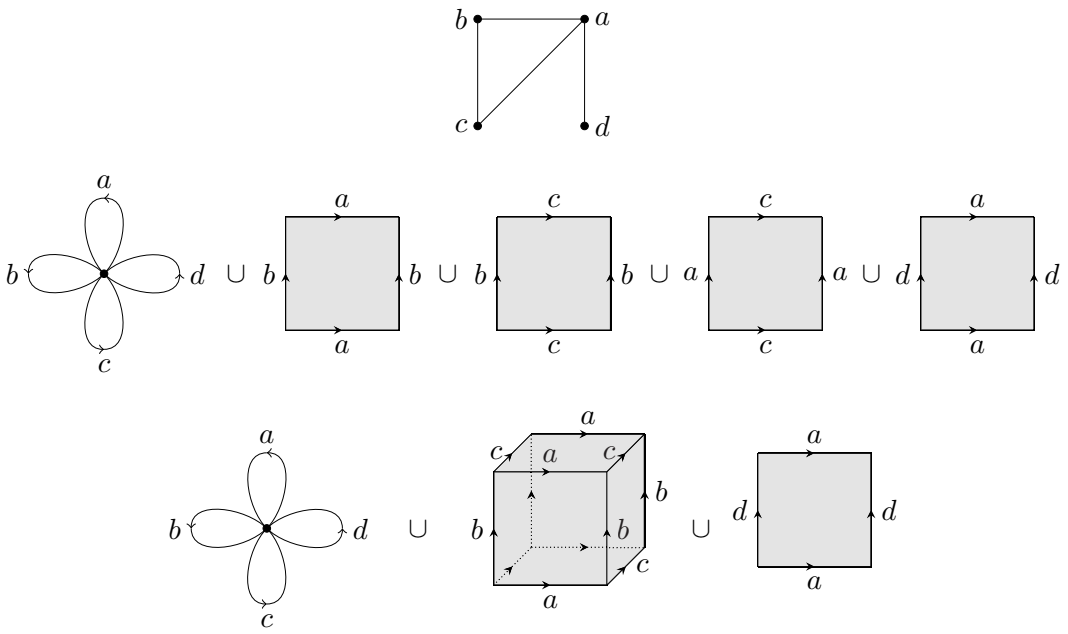


Figure 2. A graph, and kits for making its presentation 2-complex and its Salvetti complex

including infinity. Things got much more interesting when Stallings showed that the RAAG $F_2 \times F_2$ contains finitely generated subgroups which are not finitely presentable. In fact $F_2 \times F_2 \times \ldots \times F_2$ contains subgroups which are FP_{n-1} but not FP_n for all n, where FP_k is a k-dimensional algebraic finiteness property. This shows in particular that finitely-generated subgroups of RAAGs are not necessarily RAAGs. Droms clarified the situation by characterizing exactly which RAAGs have the property that all of their finitely-generated subgroups are RAAGs: they are those for which the subgraph spanned by four vertices is never a square or a straight line [13]. Servatius, Droms and Servatius showed that if Γ is a pentagon, then A_Γ contains the fundamental group of a closed surface [26], and there is a great deal of recent work on surface subgroups of RAAGs by authors including S. Kim, T. Koberda, A. Duncan, I. Kazachkov, M. Cassals-Ruiz, R. Weidman, I. Kapovich and A. Minasyan.

2.4 Automorphism groups of RAAGs

The emphasis of the present lectures is on automorphism groups of RAAGs. We will address the following three natural questions:

- How does the shape of Γ affect properties of $Out(A_\Gamma)$?
- $Aut(F_n), Out(F_n)$ and $GL(n,\mathbb{Z})$ share many basic properties. Which are in fact properties of $Out(A_\Gamma)$ for *any* Γ?
- How can we leverage information about $Out(F_n)$ and $GL(n,\mathbb{Z})$ to gain information about $Out(A_\Gamma)$?
- What techniques classically used to study $Out(F_n)$ and $GL(n,\mathbb{Z})$ can be adapted

to $Out(A_\Gamma)$? We are especially interested in geometric techniques.

For the most part we will concentrate on the outer automorphism group $Out(A_\Gamma)$ instead of the full automorphism group $Aut(A_\Gamma)$. For $A_\Gamma = \mathbb{Z}^n$ there is no difference. For any Γ the abelianization map $A_\Gamma \to \mathbb{Z}^n$ induces a map on automorphism groups since the commutator subgroup is characteristic. But this map factors through $Out(A_\Gamma)$, and for $A_\Gamma = F_2$, the induced map on $Out(F_2)$ is an isomorphism, so Out seems the more natural comparison group. Further motivation is provided by the fact that we want to model automorphisms by maps on spaces with fundamental group A_Γ, and passing to $Out(A_\Gamma)$ means that we do not have to endow these spaces with basepoints and keep track of where the basepoint goes under the maps.

2.5 Generators for $Aut(A_\Gamma)$

For $A_\Gamma = \mathbb{Z}^n$ it is an easy consequence of the Euclidean algorithm that $Out(A_\Gamma) = GL(n,\mathbb{Z})$ is generated by the elementary matrices $A_{ij} = I_n + E_{ij}$ for $i \neq j \in \{1,\dots,n\}$ (where the only non-zero entry of E_{ij} is a 1 in the (i,j)-position) and the matrix

$$T = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

If we let $GL(n,\mathbb{Z})$ act on $\mathbb{Z}^n$ on the right then A_{ij} sends $e_i \mapsto e_i + e_j$ and fixes e_k for $k \neq i$. This is called a *transvection*. In multiplicative notation for the free abelian group with generators $\{a_1,\dots,a_n\}$ these generators become $T\colon a_1 \mapsto a_1^{-1}$ and $A_{ij}\colon a_i \mapsto a_i a_j = a_j a_i$.

For $A_\Gamma = F_n$, the group $Out(A_\Gamma) = Out(F_n)$ is also generated by $T\colon a_1 \mapsto a_1^{-1}$ and by transvections, but right transvections $\rho_{ij}\colon a_i \mapsto a_i a_j$ are now different from left transvections $\lambda_{ij}\colon a_i \mapsto a_j a_i$ and the most natural presentation of $Out(F_n)$ (due to Gersten [14]) uses both. The fact that these generate $Out(F_n)$ was originally proved by Magnus [20], but the slickest proof is the one by Stallings using foldings of graphs [27].

For a general RAAG, not every transvection gives an automorphism: if a commutes with c but b does not, then the transvection $a \mapsto ab$ is not a homomorphism. This is the only thing that can go wrong, though: one just needs to check that everything that commutes with a also commutes with b; in this case we say the transvection $a \mapsto ab$ is Γ-*legal*. It is convenient to express this in terms of the defining graph Γ using the following standard terminology, which will be used throughout these lectures:

Definition 2.1 Let a be a vertex of Γ. The *link* of a is the full subgraph $lk(a)$ spanned by all vertices adjacent to a, and the *star* of a is the full subgraph $st(a)$ spanned by $lk(a)$ and a, i.e., $st(a)$ is the simplicial join $a * lk(a)$.

If Θ is a full subgraph of Γ, then the *link of* Θ is the intersection of the links of vertices in Θ

$$lk(\Theta) = \bigcap_{b\in\Theta} lk(b),$$

and the *star of* Θ is the simplicial join of Θ and $lk(\Theta)$

$$st(\Theta) = \Theta * lk(\Theta).$$

Twists and folds. Using the above notation, the condition for a transvection to be Γ-legal is: transitions $a \mapsto ab$ and $a \mapsto ba$ are Γ-legal if and only if

- $ab \neq ba$ and $lk(a) \subseteq lk(b)$, or
- $ab = ba$ and $st(a) \subseteq st(b)$.

This can be said more economically by the single condition $lk(a) \subseteq st(b)$, but it is often important to retain the distinction (commuting versus non-commuting is a critical difference here!) so we also introduce different terminology for the two types of transvections.

Definition 2.2 If $ab = ba$, then a Γ-legal transvection $a \mapsto ab$ is a *twist*. If $ab \neq ba$ then a Γ-legal transvections $a \mapsto ab$ and $a \mapsto ba$ are called *(right and left) folds*.

The reason for this terminology will become clear when we discuss geometric models for these automorphisms.

Partial conjugations. Even if we can't transvect b onto a we can still try to conjugate a by b. If we do that, we must also conjugate everything which commutes with a, and everything that commutes with things that commute with a, etc. However, the vertices in $lk(b)$ don't know whether they've been conjugated by b or not, so if a and a' are separated by $lk(b)$, we could conjugate a by b but not a'. In other words, conjugating an entire component of $\Gamma - lk(b)$ by b gives an automorphism; this is called a (Γ-*legal*) *partial conjugation*.

Inversions and graph automorphisms. It is clear that a permutation of the generators will be an automorphism if and only if it extends to an automorphism of Γ, since Γ encodes the commuting relations. Since $Aut(A_\Gamma)$ does not contain all transvections, we can't assume that all of these permutations are products of transvections, so we include graph automorphisms in the generating set. Similarly, we add all inversions $a_i \mapsto a_i^{-1}$ instead of just $a_1 \mapsto a_1^{-1}$.

The types of automorphisms described in the last three paragraphs now do generate $Aut(A_\Gamma)$, by a theorem of Laurence and Servatius.

Theorem 2.3 ([19, 25]) *$Aut(A_\Gamma)$ is generated by graph automorphisms, inversions, and Γ-legal twists, folds and partial conjugations.*

In particular this shows that $Aut(A_\Gamma)$ (and therefore $Out(A_\Gamma)$) is finitely generated. Both $Aut(A_\Gamma)$ and $Out(A_\Gamma)$ are also finitely presented. This was proved in some special cases in [3], and an explicit finite presentation in all cases was given by Matt Day [11]. Day's proof closely follows McCool's proof that $Aut(F_n)$ is finitely presented using a "Peak reduction" algorithm (see [21]).

3 Lecture 2

In the last lecture we introduced right-angled Artin groups and their automorphisms. In this lecture we will show how to infer information about general $Out(A_\Gamma)$ from known facts about $Out(F_n)$ and $GL(n, \mathbb{Z})$.

We already mentioned that $Out(A_\Gamma)$ is finitely generated and finitely presented for all A_Γ. We claim that it also has higher-dimensional homological finiteness properties:

- $Out(A_\Gamma)$ has subgroups of finite index which are torsion-free, i.e., $Out(A_\Gamma)$ is *virtually torsion-free*.
- In fact, $Out(A_\Gamma)$ has lots of torsion-free finite index subgroups: for any given element $\phi \in Out(A_\Gamma)$ there is a torsion-free finite index subgroup which does not contain ϕ, i.e., $Out(A_\Gamma)$ is *residually finite*.
- The homology of any torsion-free finite index subgroup is finitely generated. In particular its homology vanishes above some point, i.e., $Out(A_\Gamma)$ has *finite virtual cohomological dimension*.

We will introduce our method for bootstrapping information about $Out(F_n)$ and $GL(n,\mathbb{Z})$ to $Out(A_\Gamma)$ by giving a proof that $Out(A_\Gamma)$ contains a torsion-free subgroup of finite index.

The proof that $GL(n,\mathbb{Z})$ is virtually torsion-free is quite easy, one just checks that the kernel of the natural map $GL(n,\mathbb{Z}) \to GL(n,\mathbb{Z}/3)$ has no torsion; the proof uses only the binomial theorem.

The proof that $Out(F_n)$ is virtually torsion-free relies on this calculation plus the non-trivial fact, due to Baumslag and Taylor [1], that the kernel of the natural map $Out(F_n) \to GL(n,\mathbb{Z})$ is torsion-free.

The kernel of the map $Out(A_\Gamma) \to GL(n,\mathbb{Z})$ for general A_Γ is also torsion-free; this is one consequence of a recent paper by Toinet [29]. We will avoid appealing to this, however, since our point is to illustrate the general bootstrapping method.

Since we are interested in a virtual notion, it suffices to pass to a subgroup of finite index. Let $Aut^0(A_\Gamma)$ denote the subgroup of $Aut(A_\Gamma)$ generated by inversions, transvections and partial conjugations (we are leaving out only the graph automorphisms), and let $Out^0(A_\Gamma)$ be the image of $Aut^0(A_\Gamma)$ in $Out(A_\Gamma)$.

Exercise 3.1 Show that $Aut^0(A_\Gamma)$ and $Out^0(A_\Gamma)$ are normal subgroups of $Aut(A_\Gamma)$ and $Out(A_\Gamma)$ respectively, and then that they have finite index.

To show that $Out^0(A_\Gamma)$ has a torsion-free subgroup of finite index the key idea we will exploit is that there are lots of subgroups in A_Γ which must be sent to conjugates of themselves by any automorphism. To describe these subgroups, we introduce some basic facts and some new terminology.

Lemma 3.2 *Let V be a set of vertices in Γ and Θ the full subgraph spanned by V. Then the subgroup generated by V is isomorphic to A_Θ.*

Such a subgroup is called a *special subgroup*. By convention, we set $A_\emptyset = 1$.

Now recall that a transvection $a \mapsto ab$ (or $a \mapsto ba$) is an automorphism of A_Γ if and only if $st(a) \subseteq lk(b)$. In this case we write $a \preceq b$. If $a \preceq b$ and $b \preceq a$ we say $a \sim b$; this defines an equivalence relation on the vertices of Γ. The notation is justified by the following observation.

Exercise 3.3 The set of equivalence classes of vertices of Γ is a partially ordered set, with partial order induced by $\preceq$.

A vertex is called *maximal* if its equivalence class is maximal in this partial order.

Let $[a]$ denote the full subgraph of Γ spanned by vertices equivalent to a.

Proposition 3.4 ([7], Proposition 3.2) *Let Γ be a connected graph and a a maximal vertex in Γ. Then any $\phi \in Out^0(A_\Gamma)$ is represented by some $f_a \in Aut^0(A_\Gamma)$ with*

$$f_a(A_{st(a)}) = A_{st(a)} \text{ and } f_a(A_{[a]}) = A_{[a]}.$$

Proof The proof is accomplished by checking that the statement is true for all of the generators of $Out^0(A_\Gamma)$. □

Proposition 3.5 ([8], Section 3) *For Γ connected and a maximal in Γ there are homomorphisms*

- **[Restriction]** $R_a \colon Out^0(A_\Gamma) \to Out^0(A_{st[a]})$
- **[Exclusion]** $E_a \colon Out^0(A_\Gamma) \to Out^0(A_{\Gamma - [a]})$
- **[Projection]** $P_a \colon Out^0(A_\Gamma) \to Out^0(A_{lk[a]})$

Sketch of proof If f_a is the map representing $\phi \in Out^0(A_\Gamma)$ described in Proposition 3.4, then the restriction map sends ϕ to the class of the restriction of f_a to $A_{st[a]}$. This is well-defined because $A_{st[a]}$ is its own normalizer (see, e.g., [8], Proposition 2.2).

Exclusion is induced by the map $A_\Gamma \to A_{\Gamma-[a]}$ sending $v \mapsto 1$ if $v \in [a]$ and $v \mapsto v$ if $v \notin [a]$. This is well-defined because the normal subgroup generated by a maximal equivalence class $[a]$ is characteristic, by Proposition 3.4.

Projection is the composition $P_a = E_a \circ R_a$. This is well-defined because $[a]$ is maximal in $st[a]$, so E_a is defined on the image of R_a. □

We can put all of the projection homomorphisms together to get a single homomorphism $P = \prod P_a$. The following theorem is the basic result which enables our bootstrapping technique.

Theorem 3.6 ([7, Theorem 4.1], [8, Section 3]) *Let Γ be connected, and set*

$$P = \prod_{[a]\, maximal} P_a \colon Out^0(A_\Gamma) \to \prod_{[a]\, maximal} Out^0(A_{lk[a]}).$$

Then $ker(P)$ is finitely generated and free abelian. The rank of $ker(P)$ is computable in terms of Γ.

Notice that $lk[a]$ is smaller than Γ. We would like to use this fact to do induction. There is, however a problem: all of the above results have the hypothesis that Γ is connected, but $lk[a]$ need not be connected, even if Γ is. There are two ways to get around this. First, since disconnected graphs give rise to free products of RAAGs we can sometimes take advantage of known results about free products. If these are not available, we can simply assume that all non-empty links are either connected or discrete (reducing us by induction to the general linear and free group cases).

We illustrate the first option by proving that $Out(A_\Gamma)$ has torsion-free subgroups of finite index. We take advantage of the following theorem of Guirardel and Levitt:

Theorem 3.7 ([15]) *Let $G = G_1 * G_2$ with G_i and $G_i/Z(G_i)$ torsion-free. If $Out(G_i)$ is virtually torsion-free for $i = 1, 2$ then so is $Out(G)$.*

With this in our repertoire we can now prove our theorem.

Theorem 3.8 ([7], Theorem 5.2) *$Out(A_\Gamma)$ has torsion-free subgroups of finite index, for any Γ.*

Proof If Γ is a disjoint union $\Gamma = \Gamma_1 \sqcup \Gamma_2$, then $A_\Gamma = A_{\Gamma_1} * A_{\Gamma_2}$, so by Theorem 3.7 it suffices to consider connected graphs Γ.

If Γ is a complete graph then $A_\Gamma = \mathbb{Z}^n$ and the theorem is true, as noted above. If Γ is not complete, then $lk[a]$ is non-empty for some maximal vertex a. For any vertex a of Γ the maximal size of a clique in $lk[a]$ is strictly less than the maximal size of a clique in Γ. Therefore we can use this number, which we denote $m(\Gamma)$, to do induction.

If $m(\Gamma) = 1$ then Γ is discrete and $A_\Gamma = F_n$, in which case the theorem is true by the theorem of Baumslag and Taylor.

If $m(\Gamma) = 2$ then $lk[a]$ is discrete for all a, so we can use the map P defined in Theorem 3.6 to pull back a product of torsion-free finite index subgroups of $Out^0(A_{lk[a]})$ to obtain a torsion-free finite index subgroup of $Out(A_\Gamma)$.

Now induction on $m(\Gamma)$ together with Theorem 3.7 completes the proof. □

The groups A_Γ are residually finite; this follows from the fact that they are linear, which was proved by Davis and Januskiewicz [10]. Residual finiteness for $Aut(A_\Gamma)$ then follows from Baumslag's theorem that the automorphism group of any residually finite group is itself residually finite. Residual finiteness for $Out(A_\Gamma)$ is more subtle, but using the homomorphisms P, R and E above together with various inductive schemes we can also settle this question.

Theorem 3.9 ([8], Theorem 4.2) *For any RAAG A_Γ, $Out(A_\Gamma)$ is residually finite.*

This was also proved by Minasyan [22]. Both proofs rely on the result of Minasyan and Osin that if G_1 and G_2 are finitely generated groups with $Out(G_1)$ and $Out(G_2)$ residually finite, then $Out(G_1 * G_2)$ is residually finite [23].

Another result which can be proved using the maps P, R and E is:

Theorem 3.10 ([7], Theorem 5.2) *$Out(A_\Gamma)$ has finite virtual cohomological dimension.*

Here again we rely on a result of Guirardel and Levitt about free products, namely:

Theorem 3.11 ([15]) *Let $G = G_1 * G_2$ with G_i and $G_i/Z(G_i)$ torsion-free. If $Out(G_i)$ has finite virtual cohomological dimension for $i = 1, 2$ then so does $Out(G)$.*

If we do not have a suitable result in the wings for free products, we need to hypothesize that Γ is connected and the link of every non-maximal clique is either connected or discrete; such a graph Γ is called *homogeneous*. This is automatically true if Γ has no triangles (in which case the Salvetti complex is 2-dimensional, so we say A_Γ is two-dimensional). It is also true, e.g., if Γ is the 1-skeleton of a triangulated manifold. As an example, we can prove:

Theorem 3.12 ([8], Theorem 5.5) *If Γ is homogeneous then every subgroup of $Out(A_\Gamma)$ is either virtually solvable or contains a free group of rank 2.*

We can also bound the maximum derived length of a solvable subgroup in terms of the shape of Γ (see [8], Section 6). The crudest such estimate is that this length is always less than or equal to the number of vertices in Γ.

4 Lecture 3

In the last lecture we studied $Out(A_\Gamma)$ via the projection map

$$P = \prod_{[a]\, maximal} P_a \colon Out^0(A_\Gamma) \to \prod_{[a]\, maximal} Out^0(A_{lk[a]})$$

and its free abelian kernel.

We remarked that we can use this to bound the virtual cohomological dimension of $Out(A_\Gamma)$. However, P is far from surjective, and $ker(P)$ is far from being maximal rank among abelian subgroups, so the upper and lower bounds this gives are not very good.

In this lecture we take a more geometric approach to the study of $Out(A_\Gamma)$ by attempting to realize $Out(A_\Gamma)$ as symmetries of an "outer space." As before the classical theory of $GL(n,\mathbb{Z})$ and $Out(F_n)$ provide guidance.

$GL(n,\mathbb{Z})$ acts on the symmetric space $SO(n)\backslash SL(n,\mathbb{R})$, and $Out(F_n)$ acts on Outer space. Useful features of these actions include:

- the spaces are contractible
- the actions are proper

These two properties imply that algebraic invariants of the groups can be computed by computing topological invariants of the quotient spaces; in particular the cohomology $H^*(\Gamma) \cong H^*(X/\Gamma)$. Further properties of the classical actions include

- the spaces are finite-dimensional

from which we can immediately conclude that the virtual cohomological dimension of the groups are finite, and

- there is a cocompact equivariant deformation retract

which implies that the group cohomology is finitely generated in all dimensions. Furthermore, the quotient of the retract by the action can be described combinatorially, making it possible to do explicit cohomology calculations, at least in small dimensions.

More sophisticated features include

- the spaces have *bordifications*, i.e., they can be enlarged to spaces with proper cocompact actions, whose cohomology at infinity is concentrated in one dimension.

By work of Bieri and Eckmann this implies that the groups are *virtual duality groups*, i.e., there is a dualizing module D and isomorphisms

$$H^*(G;A) \cong H_{d-*}(G, D \otimes A)$$

between cohomology with any coefficients A and homology with coefficients in $D \otimes A$.

Outer space for a general RAAG will be a hybrid species, combining features of both symmetric spaces and Outer space. So let us now review these spaces.

4.1 Symmetric space

The symmetric space $\mathbb{D}_n = SO(n)\backslash SL(n,\mathbb{R})$ has several useful alternate descriptions. A coset $SO(n)A$ gives a well-defined positive definite symmetric matrix $Q = A^tA$, identifying $\mathbb{D}_n$ with

- the space of positive definite quadratic forms Q on $\mathbb{R}^n$, modulo homothety.

If we fix the standard lattice $\mathbb{Z}^n \subset \mathbb{R}^n$, then any linear map $A\colon \mathbb{Z}^n \to \mathbb{R}^n$ defines a *marked lattice*, modulo homothety. If we also mod out by rotations, then $\mathbb{D}_n$ can also be described as

- the space of marked lattices $A\colon \mathbb{Z}^n \to \mathbb{R}^n$ modulo homothety and rotation.

Finally, the map $A\colon \mathbb{R}^n \to \mathbb{R}^n$ induces a map $\bar{A}\colon \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{R}^n/A(\mathbb{Z}^n)$. We think of $\mathbb{R}^n/\mathbb{Z}^n$ as a standard torus T^n, $Y = \mathbb{R}^n/A(\mathbb{Z}^n)$ as a torus with a flat metric, and $\bar{A}$ as an isotopy class of homeomorphisms; then $\mathbb{D}_n$ is identified with

- the space of marked flat tori $\bar{A}\colon T^n \to Y$, modulo homothety.

In each case, the group $GL(n,\mathbb{Z})$ acts on the right. If $g \in GL(n,\mathbb{Z})$, then

- $SO(n)A \cdot g = SO(n)Ag$
- $Q \cdot g = g^tQg$
- $(A\colon \mathbb{Z}^n \to \mathbb{R}^n) \cdot g = Ag\colon \mathbb{Z}^n \to \mathbb{R}^n$
- $(\bar{A}\colon T^n \to Y) \cdot g = \overline{Ag}\colon T^n \to Y$.

Each of these descriptions of the symmetric space has its advantages. For example, the description as the space of positive definite quadratic forms makes it easy to see that $\mathbb{D}_n$ is contractible, since the set of positive definite quadratic forms is a convex cone in the space of $n \times n$ matrices. The description that will be most relevant for us is the last, as a space of marked flat tori. Note that the action of $GL(n,\mathbb{Z})$ changes the marking, but does *not* change the flat metric.

4.2 Outer space

Outer space also has several useful descriptions. We can mimic the description of the symmetric space as the space of marked flat tori by defining Outer space as a space of *marked metric graphs*. To do this, we fix a *rose* R_n, i.e., a graph with one vertex and n directed edges, as a "model space" to play the role of the torus T^n. A metric on a graph X is simply an assignment of positive real lengths to its edges, making X into a metric space with the path metric. A *marking* is a homotopy equivalence $h\colon R_n \to X$. For technical reasons we don't allow our graphs to have univalent or bivalent vertices, and they must be finite. Marked graphs (X,h) and (X',h') are *equivalent* if there is an isometry or a homothety $f\colon X \to X'$ with $f \circ h$ homotopic to h'.

Definition 4.1 *Outer space* CV_n is the space of equivalence classes of marked metric graphs with fundamental group F_n.

$Out(F_n)$ acts on CV_n on the right by changing the marking, i.e., for $\phi \in Out(F_n)$ take a map $f\colon R_n \to R_n$ that induces ϕ on $\pi_1(R_n) \equiv F_n$ and set $(X,h)\cdot\phi = (X, h\circ f)$.

There is an obvious equivariant deformation retraction of CV_n onto the subspace consisting of marked metric graphs with no separating edges (simply shrink each

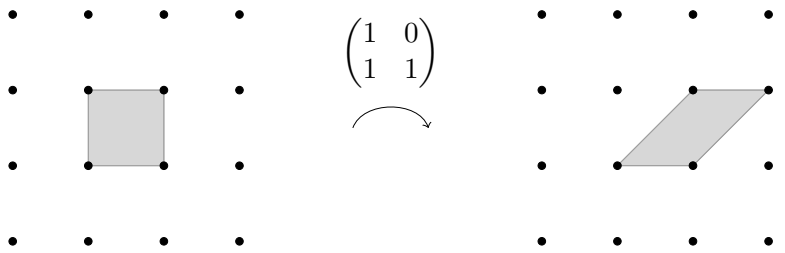


Figure 3. Action of A_{21} on Λ

separating edge to a point). This subspace, called *reduced Outer space* is sometimes more convenient to work with.

Given a marked metric graph X, the marking serves to identify F_n with the fundamental group of X. By looking at the universal cover $\widetilde{X}$ we thus obtain a simplicial tree with a free action of F_n. The fact that we don't allow X to be infinite or have univalent or bivalent vertices translates into the condition that the action is *minimal*, i.e., there are no F_n-invariant subtrees. Therefore an alternate description of CV_n is as the space of *free minimal actions of F_n on metric simplicial trees*. This is analogous to the description of the symmetric space as a space of lattices instead of as a space of flat tori. (There is a third definition of CV_n in terms of isotopy classes of spheres in a doubled handlebody which is extremely useful, but will not be relevant for these lectures.)

4.3 Lattices, tori and graphs in rank 2

To motivate our definition of Outer space for a general RAAG, we first compare the symmetric space and Outer space in rank 2. In rank 2 the natural map $Out(F_2) \to GL(2,\mathbb{Z})$ induced by abelianization $F_2 \to \mathbb{Z}^2$ is an isomorphism and both (reduced) Outer space and the symmetric space can be identified with the hyperbolic plane. The spaces diverge dramatically in higher ranks, but the rank 2 picture gives us some insight into the general situation because we can look at the same space from two different points of view.

$GL(n,\mathbb{Z})$ is generated by elementary matrices, so consider the action of $A_{21} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ on the standard lattice $\Lambda = \mathbb{Z}^2 \subset \mathbb{R}^2$, marked by the identity. This sends $e_1 \mapsto e_1$ and $e_2 \mapsto e_1 + e_2$ (remember we are acting on the *right*). This action is illustrated in Figure 3.

Note that the action of A_{21} does not change the lattice, it just changes the marking. Thus the orbit of $GL(n,\mathbb{Z})$ is a discrete subset of the space of all marked lattices. To get a *path* from Λ to $\Lambda \cdot A_{21}$ we must gradually shear the marked lattice, as in Figure 4. In the figure we have drawn the original fundamental domain for reference.

The path in the space of marked flat tori is obtained by identifying opposite sides of the fundamental domain for the lattices, as in Figure 5. In this figure, too, we have marked the original fundamental domain for reference.

This gives us a clue for what this path looks like if we think of it as a path in CV_2: if we puncture the torus its fundamental group is F_2 instead of $\mathbb{Z}^2$, and it deformation

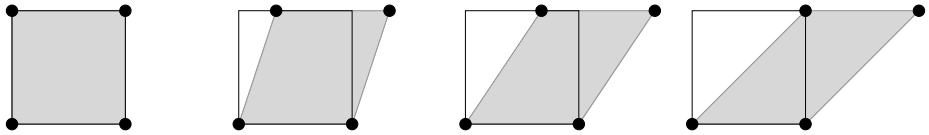


Figure 4. A path from Λ to $\Lambda \cdot A_{21}$.

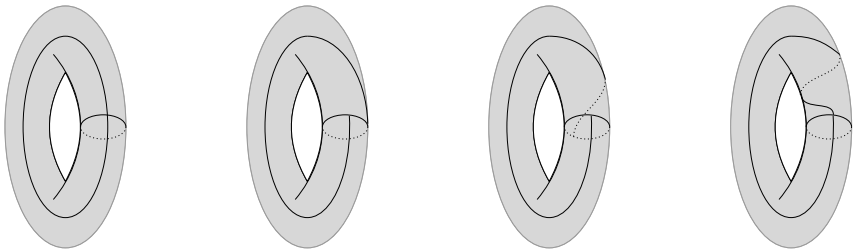


Figure 5. The same path, as a path of tori

retracts onto the dark graph. The path in CV_2 is then given by the graphs in Figure 6. Note that the total length of the graph is constant (equal to 2) in this picture. Under the deformation retraction of the punctured torus onto the graph, the action of A_{21} on $\mathbb{Z}^2$ becomes the action of $\rho_{21}\colon x_2 \mapsto x_2x_1$ on F_2. We indicate the loop representing x_2 by the thicker curve in the picture.

4.4 A simple example

We have described CV_n as a space of marked metric graphs $R_n \to X$ and the symmetric space for $\mathbb{Z}^n$ as a space of marked flat tori $T^n \to X$. In each case we needed a model space and a homotopy equivalence to a metric space. For a general *RAAG* A_Γ we have a model space, namely the Salvetti complex S_Γ, so we would like to have an outer space of marked metric spaces $S_\Gamma \to X$. We now need to decide:

- What homeomorphism types X should we allow?
- What metric structures should we allow on these spaces X?

We begin by looking at a very simple RAAG, i.e.,

$$\langle\, a, b, c \mid [a, c] = [b, c] = 1 \,\rangle.$$

Figure 6. The same path, as a path in Outer space CV_2

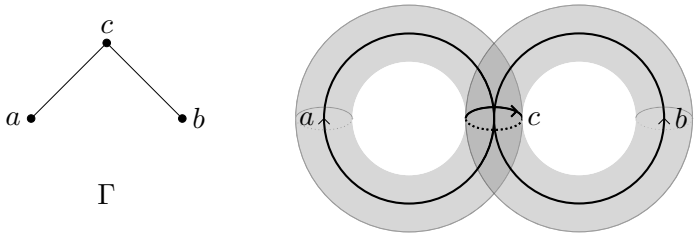


Figure 7. A simple graph Γ and its Salvetti complex S_Γ

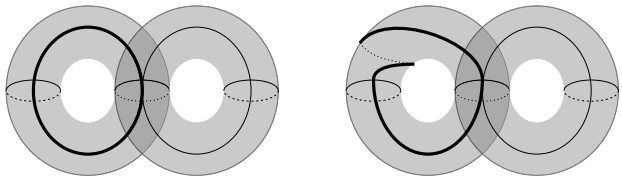


Figure 8. Realizing the twist τ_{ac} on S_Γ

This is the RAAG associated to the graph Γ with three vertices a, b and c and two edges, one from a to c and one from c to b. The Salvetti complex S_Γ is the union of two tori, glued along a common meridian curve labeled c (see Figure 7).

Generators for $Out(A_\Gamma)$ consist of the graph automorphism interchanging a and b, inversion in a, inversion in c, the twist $\tau_{ac}\colon a \mapsto ac = ca$, and the folds $\rho_{ab}\colon a \mapsto ab$ and $\lambda_{ab}\colon a \mapsto ba$.

The group A_Γ is the product of the cyclic group generated by c with the free group generated by a and b, and the Salvetti complex S_Γ is the product of the loop labeled c by the rose formed by the two longitudinal curves. To realize the twist τ_{ac} on S_Γ, we can perform a Dehn twist of the left-hand torus around a curve parallel to c but disjoint from c (see Figure 8).

To realize the transvection ρ_{ab} we *fold* the left-hand torus around the right-hand torus (see Figure 9). This is can also be described as first expanding the intersection circle into a cylinder, then collapsing a different cylinder (the bottom shaded cylinder in the figure).

Thus to make a path from S_Γ to $S_\Gamma \cdot \tau_{ac}$ we need to gradually shear the metric on the left-hand torus, and to make a path from S_Γ to $S_\Gamma \cdot \rho_{ab}$ we need to pass through spaces X which are not homeomorphic to S_Γ. In our outer space for A_Γ we need to be able to vary both the homeomorphism type of spaces and the metrics on the spaces. But we want to restrict both as much as possible so that we can control the topology and geometry of the space.

Note that both S_Γ and the intermediate complexes X are combinatorially cube complexes. With standard Euclidean metrics on the cubes they are non-positively curved cube complexes (i.e., their universal covers are CAT(0)), in fact they are *special* cube complexes, in the sense of Haglund and Wise [16]. However, we want to vary the (projective classes of) the metrics to allow shearing of the tori. This can be accomplished in the intermediate spaces X by giving all three cylinders flat

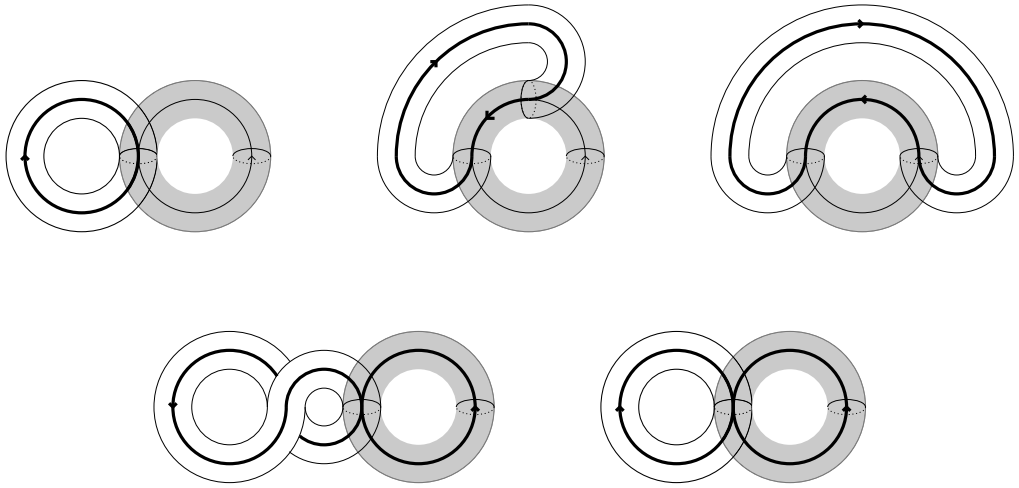


Figure 9. Realizing the fold ρ_{ab} on S_Γ

right-angled metrics with the same circumference, then specifying the attaching maps to the two circles by shear parameters.

There are *a priori* two shear parameters for each cylinder, but shearing both ends of a cylinder by the same amount does not change the metric on X, so there are actually only three total parameters. These are not independent either, since shearing all three by the same amount simply twists the circle without affecting the metric; thus in the end we have only 2 independent shear parameters.

Shearing the top or bottom cylinder by an entire rotation changes the marking by a twist, and shearing in the opposite direction changes the marking by its inverse. Expanding and collapsing cylinders without shearing varies the space independently of the c direction, so may be thought of as moving around the space of metric graphs marked by the free subgroup generated by a and b, i.e., around reduced Outer space in rank 2. Since reduced Outer space in rank 2 is homeomorphic to $\mathbb{R}^2$, the entire moduli space of marked metric blowups is homeomorphic to the product $\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}$.

5 Lecture 4

In this lecture we show how to construct an Outer space for any RAAG A_Γ; this will be a space of marked metric cube complexes. We then outline very briefly how to prove the space we have constructed is contractible and that the action is proper.

We recall the example we studied in the last lecture, where Γ has three vertices and two edges. To get a path from the standard Salvetti $id\colon S_\Gamma \to S_\Gamma$ to its image under the twist $\tau_{ac}\colon a \mapsto ac = ca$ we needed to shear the metric on the left-hand torus, while to get a path to its image under the fold $\rho_{ab}\colon a \mapsto ab$ we needed to expand a circle of the Salvetti into a cylinder, then collapse a different cylinder. The first operation involves changing the metrics on the spaces without changing their homeomorphism type, while the second operation can be described combinatorially in terms of "blowing up" and collapsing subcomplexes.

We begin our construction by determining which cube complexes we need, tem-

porarily ignoring their metric structure. For $A_\Gamma = F_n$, this amounts to describing the (vertices of the) *spine* K_n instead of the full space CV_n, so we briefly recall that construction.

Outer space CV_n for a free group is the union of open simplices, one for each equivalence class of marked (combinatorial) graphs (X, h), where X is a finite graph with no bivalent vertices or separating edges (and hence no univalent vertices, either). The open simplex associated to (X, h) is obtained by assigning all possible positive real lengths to the edges of X, then either projectivising or (equivalently) normalizing so that the sum of the lengths is one. If we take small neighborhoods of these simplices we obtain an open cover of CV_n by contractible sets, such that the intersection of any two elements is either empty or is in the cover. The nerve of this cover is known as the *spine of Outer space*, and is an equivariant deformation retract of all of CV_n. The spine can be described combinatorially as the geometric realization of the partially ordered set of marked graphs, where the poset relation is given by forest collapse: $(X, h) > (X', h')$ if there is a forest $\Phi \subset X$ such that X' is obtained from X by collapsing each edge of Φ to a point, and h' is (homotopic to) the composition of h with this collapse. The full space CV_n can be recovered from the spine by putting the metric information back into the graphs.

Motivated by this, we will now construct a similar spine for any A_Γ.

5.1 The spine of outer space for A_Γ

For general A_Γ we need analogs of graphs, forests and forest collapses. The analog of a graph will be a particular type of non-positively curved cube complex adapted to Γ, which we call a Γ-complex. We first recall some standard background about cube complexes.

5.1.1 NPC cube complexes and hyperplanes

A *cube complex* is a CW complex X in which every cell is homeomorphic to a Euclidean cube (of some dimension) and the attaching maps identify faces with lower-dimensional cells by homeomorphisms.

If v is a vertex of a cube complex X there is an associated simplicial complex called the *link* of v and denoted $lk(v)$. This has one vertex for each half-edge terminating at v, and a set of half-edges spans a k-simplex if they belong to distinct edges of the same cube. Gromov gave a simple condition on links which guarantees that X can be given a metric of non-positive curvature. This says that if the 1-skeleton of a simplex appears in $lk(v)$, then the entire simplex must be in $lk(v)$. This is called the *flag condition* on links, and a cube complex X whose links satisfy the flag condition is said to be *NPC*.

Cubes in a cube complex are cut by *hyperplanes*. A hyperplane is dual to an equivalence class of edges, where the equivalence relation is generated by saying two edges are equivalent if they are parallel in some cube. If H is the hyperplane dual to $[e]$, then the intersection of H with a cube C is spanned by midpoints of edges of C which are in $[e]$; thus $H \cap C$ is either empty or is a codimension one linear subspace cutting C in half.

Example 5.1 The space X from the last lecture is an NPC cube complex. There are four hyperplanes. Three of them are circles midway up the three cylinders, and the other is a theta graph, with one edge running the length of each cylinder.

Example 5.2 S_Γ is an NPC cube complex with one k-cube for each k-clique in Γ. There is one hyperplane for each generator a of A_Γ (i.e., each vertex of Γ), and the associated hyperplane is isomorphic to the Salvetti complex $S_{lk[a]}$.

The *carrier* of a hyperplane H is the closure of the union of all cubes which intersect H. If the carrier of H is an embedded copy of $H \times [0,1]$, then collapsing each cube in the carrier to its intersection with H is called a *hyperplane collapse*, though maybe it should be called a carrier collapse. A hyperplane collapse is *trivial* if the resulting complex is still homeomorphic to X.

A set of hyperplanes $\{H_1, \dots, H_k\}$ in X is called a *hyperplane forest* if any cycle formed by edges dual to the H_i is null-homotopic. In this case each H_i determines a hyperplane collapse in which the images of the remaining H_j form a new hyperplane forest.

5.1.2 Marked Γ-complexes

Definition 5.3 A compact NPC cube complex X is called a Γ-*complex* if there is a hyperplane forest $\{H_1, \dots, H_k\}$ in X such that performing the associated hyperplane collapses (in any order) gives a cube complex isomorphic to S_Γ and

1. The hyperplane collapse associated to H_i is non-trivial for each i.
2. After each collapse, the image of any hyperplane in X is either a hyperplane or a subcomplex parallel to a hyperplane.

A *marking* of a Γ-complex X is a homotopy equivalence $h\colon S_\Gamma \to X$. Two marked Γ-complexes (X,h) and (X',h') are *equivalent* if there is an isomorphism of cube complexes $f : X \to X'$ with $h' \simeq f \circ h$.

The Salvetti complex S_Γ is of course an example of a Γ-complex. If we mark it with the identity map $id\colon S_\Gamma \to S_\Gamma$, the result is called the *standard Salvetti*.

The group $Out(A_\Gamma)$ acts on the set of marked Γ-complexes by changing the marking: any $\phi \in Out(A_\Gamma)$ can be induced by a homotopy equivalence $f\colon S_\Gamma \to S_\Gamma$, and we define $(X,h)\phi = (X, h \circ f)$.

The set of equivalence classes of marked Γ-complexes forms a partially ordered set under the relation of hyperplane collapse, and we define $\mathcal{M}_\Gamma$ to be the geometric realization of this poset.

5.1.3 The untwisted subgroup

We expect to have to shear the metric to find a path from the standard Salvetti to its image under a twist, as in the example from the last lecture. Since blowups, collapses and isometries don't do any shearing, we shouldn't even expect $\mathcal{M}_\Gamma$ to be connected, much less contractible. But it turns out that we can find a large contractible piece of $\mathcal{M}_\Gamma$ by ignoring the twists at first, i.e., we consider an orbit of the subgroup of $Out(A_\Gamma)$ generated by all other types of generators.

Definition 5.4 The *untwisted subgroup* $U(A_\Gamma)$ of $Out(A_\Gamma)$ is the subgroup generated by inversions, graph automorphisms, partial conjugations and folds.

The untwisted subgroup can be all of $Out(A_\Gamma)$ (e.g., if Γ is discrete or is without triangles and univalent vertices) or it can be finite (e.g., if Γ is complete or is an n-gon with $n \geq 5$), and is generally somewhere in between. It is not usually normal (e.g., conjugating an inversion $v \mapsto v^{-1}$ by a twist $u \mapsto uv$ results in the square of the twist times the inversion).

Definition 5.5 Let $\mathcal{M}_\Gamma$ be the geometric realization of the poset of equivalence classes of marked Γ-complexes. Let $\sigma_0 = (S_\Gamma, id)$ be the standard Salvetti and $st(\sigma_0)$ the star of σ_0 in $\mathcal{M}_\Gamma$. The *spine of outer space* K_Γ is the orbit of $st(\sigma_0)$ under $U(A_\Gamma)$.

The following theorem is joint with Charney and Stambaugh.

Theorem 5.6 ([6]) *The spine K_Γ is contractible. The untwisted subgroup $U(A_\Gamma)$ acts cocompactly with finite stabilizers on K_Γ.*

The following subsections give a brief indication of the proof.

5.1.4 Blowups and Γ-Whitehead partitions

In order to prove Theorem 5.6 we need to understand exactly which complexes occur in the star of σ_0, i.e., which marked Γ-complexes collapse to (S_Γ, id). The opposite of a hyperplane collapse is called a *blowup*, and to explain how to find all blowups of σ_0 we first recall the situation for $A_\Gamma = F_n$.

If we can obtain a rose R_n by collapsing a single edge e of a graph G, then the other edges of G can be identified with the petals a_i of R_n. Each petal is an edge with two ends, a_i^+ and a_i^-. We can reconstruct G by saying which a_i^ϵ get attached to which end of e, i.e., by giving a partition of the set $\{a_1^+, a_1^-, \ldots, a_n^+, a_n^-\}$ into two subsets, called the *sides* of the partition. In fact, any marked graph in CV_n can be collapsed to a rose R_n by collapsing a maximal tree, and each edge in the maximal tree gives a partition of the ends $\{a_1^+, a_1^-, \ldots, a_n^+, a_n^-\}$ of the petals of R_n. A collection of partitions corresponds to a graph if and only if the partitions are pairwise compatible (P and Q are compatible if some side of P is disjoint from some side of Q).

Now let H be a hyperplane in a Γ-complex X, and suppose the corresponding hyperplane collapse is defined and gives the standard Salvetti (S_Γ, id). We will partition the edges in the 1-skeleton of S_Γ by looking at their pre-images in the 1-skeleton of $Y = X - (H \times (0,1))$. By condition (2) the image of H must be parallel to a hyperplane in S_Γ, say the hyperplane $S_{lk(v)}$ dual to v. If a_i is an edge in this image, then a_i has two pre-image loops in the 1-skeleton of Y, at the top and bottom of the hyperplane carrier $H \times [0,1]$. All other edges in S_Γ have one pre-image, and we will partition their ends according to whether they are attached at the top or bottom of the carrier. We cannot partition these arbitrarily, however, since there are constraints imposed by existence of the higher-dimensional cubes. For example, we cannot put the meridian of a torus at one vertex and the longitude at the other. Careful consideration of these constraints leads to the following definition.

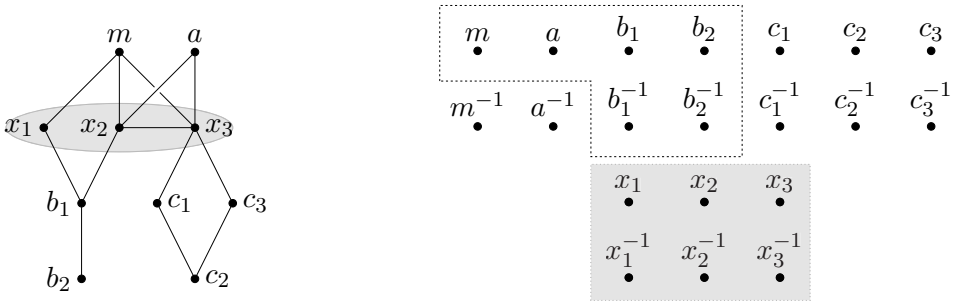


Figure 10. A graph Γ and a Γ-Whitehead partition

Definition 5.7 Let m be a vertex of Γ, L_m the vertices in $lk(m)$ and $V_m^{\pm}$ the set of vertices in $\Gamma - lk(m)$ and their inverses. A partition P of $V_m^{\pm}$ is a Γ-*Whitehead partition* if

1. Each side of P has at least two elements.
2. Each side of P is a union of (vertices of) components of $\Gamma - lk(m)$ and their inverses *except*
   - P separates m from m^{-1} and
   - if $lk(a) \subseteq lk(m)$ then P may separate a from a^{-1}.

The vertex m is called a *maximal vertex for* P. Note that a Γ-Whitehead partition may have more than one maximal vertex but any two maximal vertices have the same link, which we therefore call $lk(P)$.

An example of a Γ-Whitehead partition is shown in Figure 10.

The terminology "Γ-Whitehead" has a historical basis. Suppose the rose R_n is blown up by inserting a single edge e which partitions the half-edges of R_n. Then collapsing a different edge of the blowup gives a homotopy equivalence $R_n \to R_n$ which induces a *Whitehead automorphism* of F_n. If the newly collapsed edge was labeled with the generator a, this automorphism multiplies some generators by a (or a^{-1}) and conjugates some others by a (or a^{-1}).

Not every Whitehead automorphism of the free group on the generators of A_Γ induces an automorphism of A_Γ, but we can tell exactly which ones do. If P is a Γ-Whitehead partition as defined above, we can complete P to a partition $\widehat{P}$ of all of the generators of A_Γ and their inverses by putting $L^{\pm} = L_m \cup L_m^{-1}$ on one side of P (it doesn't matter which side). Then the induced Whitehead automorphism of the generators does give an automorphism of A_Γ.

We have seen that a two-vertex Γ-complex in $st(\sigma)$ gives a Γ-Whitehead partition. Conversely, given a Γ-Whitehead partition we can construct a two-vertex Γ-complex, which we call S^P. Here are instructions for its construction.

- Start with a copy of $S_{lk(P)} \times [0,1]$.
- For each a which is separated from a^{-1} by P (including each maximal vertex), glue in a copy of $S_{lk(a)} \times [0,1]$, attaching $S_{lk(a)} \times \{i\}$ by its inclusion into $S_{lk(P)} \times \{i\}$ for $i = 0, 1$.
- For each remaining component C of $\Gamma - lk(P)$ attach a copy of $S_{lk(P)\cup C}$ via

the inclusion of $lk(P)$, where components on opposite sides of P are attached at opposite ends of $S_{lk(P)} \times [0,1]$.

Collapsing the initial subcomplex $S_{lk(P)} \times [0,1]$ to its hyperplane $S_{lk(P)} \times \{\frac{1}{2}\}$ recovers the Salvetti complex S_Γ, so S^P is a Γ-complex in $st(\sigma)$.

Remark 5.8 For any vertex a of Γ the subcomplex $S_{st(a)}$ of S_Γ is a product $S_{lk(a)} \times S^1$. Another way of describing S^P is as the union of these subcomplexes, some of which have been subdivided by hyperplanes, modulo appropriate identifications. In particular, $S_{st(a)}$ embeds into the blowup S^P.

5.1.5 Compatible partitions and iterated blowups

The Γ-complexes in $st(\sigma)$ which have exactly two vertices are those which can be obtained from σ by blowing up a single Γ-Whitehead partition. In order to obtain any Γ-complex in $st(\sigma)$ we may need to blow up several times. This is possible if we are given a collection of Γ-Whitehead partitions which are compatible, in the following sense.

Definition 5.9 Two Γ-Whitehead partitions P and Q are *compatible* if either

1. maximal elements of P and Q are distinct and commute, or
2. some side of P is disjoint from some side of Q.

A precise recipe for constructing a Γ-complex from a collection of pairwise compatible Γ-Whitehead partitions is given in [6]. We omit the details here.

5.1.6 Contractibility of the spine

The proof that the spine K_Γ is contractible follows the general outline of the original proof that CV_n is contractible [9]. A vertex (S,h) of K_Γ is called a *Salvetti vertex* if S is homeomorphic to S_Γ. We define a total order on Salvetti vertices (S,h) by measuring the lengths of conjugacy classes of elements of A_Γ in S. More precisely, for each conjugacy class $w \in \pi_1(S_\Gamma) = A_\Gamma$ we record the length of the minimal loop in the 1-skeleton of S that represents $h(w)$, then list all these lengths in an infinite sequence. We then build K_Γ by gluing on stars of Salvetti vertices according to the lexicographical order of these sequences. We need to prove that a Salvetti vertex is determined by its length sequence, that there is a unique smallest Salvetti vertex, and that at each stage of the construction we are attaching the next star along a contractible subcomplex of its link. The proof of this last fact uses a variation of the classical Peak Reduction algorithm for free group automorphisms.

5.2 The full outer space

In order to get a contractible space on which all of $Out(A_\Gamma)$ acts properly, we need to add metric information to the marked Γ-complexes used to define K_Γ.

To explain the idea, we again recall the relation of the spine K_n to the full Outer space CV_n, from a slightly different point of view. The full space CV_n decomposes as a disjoint union of open simplices of various dimensions, where the simplex containing

(X, h) is obtained by varying the (positive) lengths of the edges of X. If we allow an edge length to shrink to zero, we pass to a face of the simplex. Thus the closure of $\sigma(X, h)$ in CV_n is a simplex together with some of its faces, but some faces are missing: if we try to shrink a set of edges containing a loop to zero, we leave CV_n. If we formally add all of the missing faces to each $\sigma(X, h)$ we obtain a simplicial complex $\overline{CV}_n$ called the *simplicial closure of Outer space*. The spine K_n is a subcomplex of the barycentric subdivision $\overline{CV}_n'$, namely K_n is the subcomplex spanned by vertices of $\overline{CV}_n$ (i.e., faces of $\overline{CV}_n$) which are actually in CV_n.

We can verify that CV_n is homotopy equivalent to K_n using a suitable open cover of CV_n. For each vertex $v \in K_n$, let $U_v = st^o(v) \subset CV_n$ be the open star of v in $\overline{CV}_n'$. Each U_v is a contractible subset of CV_n and contains no other vertices of K_n. An intersection $U_{v_0} \cap \ldots \cap U_{v_k}$ is non-empty if and only if $v_0, \ldots, v_k$ are the vertices of a simplex of K_n, in which case the intersection is contractible. Thus the nerve of the cover $\{U_v\}$ is isomorphic to K_n and is homotopy equivalent to CV_n.

For general A_Γ we would like to do something similar, i.e., add metric information to marked Γ-complexes to produce a space of marked metric Γ-complexes and an open cover by sets $\{U_v\}$ corresponding to vertices $v \in K_\Gamma$. We want the nerve of this cover to be isomorphic to K_Γ and the cover to be by contractible sets with contractible intersections, so that the whole space is homotopy equivalent to K_Γ (and hence contractible).

5.2.1 Untwisted metrics

As in the free group case we can assign positive lengths to the edges of the (rectilinear) cubes of a Γ-complex X to obtain a set of metrics on X which forms an open simplex $\sigma(X, h)$ of marked metric Γ-complexes, one for each vertex $v = (X, h)$ of K_Γ. Let Σ_G denote the union of these open simplices, modulo the natural face relations. Formally adding missing faces to the simplices $\sigma(X, h)$ completes Σ_Γ to a simplicial complex $\overline{\Sigma}_\Gamma$. The spine K_Γ is a subcomplex of the barycentric subdivision $\overline{\Sigma}_\Gamma'$, and the space Σ_Γ is covered by open stars $st^o(v)$ in $\overline{\Sigma}_\Gamma'$ of vertices $v \in K_n$. The action of $U(A_\Gamma)$ on K_Γ extends naturally to a proper action on Σ_Γ and $\overline{\Sigma}_G$.

In the free group case this is all we needed to do, but for general A_Γ this is not enough...we also need to allow the metrics on some cubes to be sheared in order to get a space on which all of $Out(A_\Gamma)$ acts properly. This shearing is governed by the subgroup $T(A_\Gamma)$ generated by twists, so we next investigate this subgroup.

5.2.2 Twisted metrics

If we order the generators $\{a_1, \ldots, a_n\}$ of A_Γ we obtain a map $Out(A_\Gamma) \to GL(n, \mathbb{Z})$ induced by abelianization $A_\Gamma \to \mathbb{Z}^n$. This map sends the twist subgroup $T(A_\Gamma)$ injectively into $SL(n, \mathbb{Z})$. The image of $T(A_\Gamma)$ is generated by the matrices $A_{ij} = I_n + E_{ij}$ for i, j such that $st(a_i) \subseteq st(a_j)$. If the ordering of the generators is subordinate to the partial ordering on the vertices of Γ, the image of $T(A_\Gamma)$ is block upper triangular. The diagonal blocks correspond to equivalence classes of vertices and a non-zero upper block corresponds to an inclusion of stars.

Now let $T_\mathbb{R}(A_\Gamma) \subset SL(n, \mathbb{R})$ be the subgroup generated by matrices $A_{ij}(r) = I_n + E_{ij}(r)$ with r real, for i, j such that $st(a_i) \subseteq st(a_j)$. $T_\mathbb{R}(A_\Gamma)$ is contained in a

parabolic subgroup of $SL(n,\mathbb{R})$ and $T(A_\Gamma)$ is a lattice in $T_{\mathbb{R}}(A_\Gamma)$. The quotient space

$$\mathbb{D}_\Gamma = (T_{\mathbb{R}}(A_\Gamma) \cap SO(n))\backslash T_{\mathbb{R}}(A_\Gamma)$$

is a subspace of the symmetric space $\mathbb{D}_n = SO(n)\backslash SL(n,\mathbb{R})$. It is homeomorphic to a product of one symmetric space for each diagonal block and one copy of $\mathbb{R}$ for each pair $i > j$ with $st(a_i) \subsetneq st(a_j)$. In particular, it is a contractible subspace of $\mathbb{D}_n$.

Each point of $\mathbb{D}_\Gamma$ corresponds to a marked flat metric on T^n, but we will ignore the marking for a moment. If we regard the Salvetti complex S_Γ as a subcomplex of T^n, then the flat metric on T^n induces a metric on S_Γ, where the distance between two points is the length of the shortest path in S_Γ joining them. Note that the metric on each subtorus S_Δ corresponding to a clique $\Delta \subset \Gamma$ is flat. A metric on S_Γ induced in this way by a point in $\mathbb{D}_\Gamma$ is said to be Γ-*adapted*.

If X is a blowup of S_Γ then X contains a subcomplex X_Δ for each clique Δ which is a (possibly subdivided) copy of the cube complex S_Δ. A CAT(0) metric on X is Γ-*adapted* if the metric restricted to each X_Δ is equal to the flat metric on S_Δ obtained from some Γ-adapted metric on S_Γ.

Definition 5.10 A *marked metric* Γ-*complex* is a triple (X, h, d), where

1. X is a Γ-complex,
2. $h\colon S_\Gamma \to X$ is a homotopy equivalence, and
3. d is a Γ-adapted CAT(0) metric on X.

Two marked metric Γ-complexes (X, h, d) and (X', h', d') are *equivalent* if there is an isometry or a homothety $\iota : X \to X'$ with $\iota \circ h' \simeq h$.

5.2.3 Outer space for A_Γ

We now define outer space $\mathcal{O}_\Gamma$ to be the space of equivalence classes of marked metric Γ-complexes. The group $Out(A_\Gamma)$ acts on $\mathcal{O}_\Gamma$ on the right by changing the marking, i.e., given $\phi \in Out(A_\Gamma)$, choose a homotopy equivalence $f : S_\Gamma \to S_\Gamma$ inducing ϕ on $\pi_1(S_\Gamma)$; then $(X, h, d) \cdot \phi = (X, h \circ f)$.

Claim *Outer space $\mathcal{O}_\Gamma$ is contractible, and $Out(A_\Gamma)$ acts properly.*

Caveat. I have refrained from calling this a Theorem since the details have not yet been posted on the arXiv.

Sketch of proof We cover $\mathcal{O}_\Gamma$ by open sets U_v corresponding to vertices $v = (X, h)$ of K_Γ. Each U_v is homeomorphic to the product of $\mathbb{D}_\Gamma$ with the open star $st^o(v)$ of v in the barycentric subdivision $\overline{\Sigma}'_\Gamma$, and is hence contractible. The nerve of the cover $\{U_v\}$ is isomorphic to K_Γ and intersections $U_{v_1} \cap \ldots \cap U_{v_k}$ are either empty or contractible. Thus $\mathcal{O}_\Gamma$ is homotopy equivalent to K_Γ, which is contractible by Theorem 5.6. Finally, one must check the stabilizer of a point (X, h, d) in $\mathcal{O}_\Gamma$ under the action of $Out(A_\Gamma)$. The action of a twist moves (X, h, d) "up the $\mathbb{D}_\Gamma$-direction," and the stabilizer of (X, h, d) is isomorphic to the group of isometries of (X, d), which is finite.

5.3 Questions

1. The action of $Out(A_\Gamma)$ on O_Γ is not cocompact, since we're using *all* Γ-adapted metrics on the Γ-complexes X. Inside this space of metrics there should be an analog of Ash's *well-rounded retract* of $SO(n)\backslash SL(n,\mathbb{R})$, which is a cocompact deformation retract, equivariant with respect to the action of $SL(n,\mathbb{Z})$. Incorporating this idea should result in an Outer space with a cocompact action.
2. Is the fixed point set of a finite subgroup of $Out(A_\Gamma)$ contractible (i.e., is $\mathcal{O}_\Gamma$ an $\underline{EG}$?)? Is it even non-empty, i.e., can every finite subgroup of $Out(A_\Gamma)$ be realized as isometries of a marked Γ-complex?
3. Is $Out(A_\Gamma)$ a virtual duality group? Is there a *bordification* of $\mathcal{O}_\Gamma$ which is a hybrid of the Borel-Serre bordification of the symmetric space $\mathbb{D}_n$ and the Bestvina-Feighn bordification of Outer space CV_n? If so, is bordified $\mathcal{O}_\Gamma$ highly connected at infinity?
4. The *metric theory* of symmetric spaces is classical and highly developed. There has also been a lot of activity recently on the metric theory of Outer space, using the asymmetric Lipschitz metric. Is there a good metric theory of $\mathcal{O}_\Gamma$? What are the geodesics? Can they be used to help classify elements of $Out(A_\Gamma)$?
5. Handel and Mosher recently proved that the 1-skeleton of the simplicial closure $\overline{CV}_n$ is a Gromov hyperbolic graph [17]. Is the 1-skeleton of $\overline{\mathcal{K}}_\Gamma$ Gromov hyperbolic? If so, is there an associated Gromov hyperbolic space on which all of $Out(A_\Gamma)$ acts?

References

[1] Gilbert Baumslag and Tekla Taylor, The centre of groups with one defining relator, *Math. Ann.* **175** (1968) 315–319.

[2] Martin Bridson and Andre Haefliger, *Metric spaces of non-positive curvature*, Springer-Verlag, New York, (1999).

[3] Kai-Uwe Bux, Ruth Charney and Karen Vogtmann, Automorphism groups of RAAGs and partially symmetric automorphisms of free groups, *Groups Geom. Dyn.* **3** (2009), no. 4, 541–554.

[4] Ruth Charney, An introduction to right-angled Artin groups, *Geom. Dedicata* **125** (2007), 141–158.

[5] Ruth Charney, John Crisp and Karen Vogtmann, Automorphisms of two-dimensional right-angled Artin groups, *Geom. Top.* **11** (2007), 2227–2264

[6] Ruth Charney, Nate Stambaugh and Karen Vogtmann, *Outer space for right-angled Artin groups, I*, arXiv:1212.4791.

[7] Ruth Charney and Karen Vogtmann, Finiteness properties of automorphism groups of right-angled Artin groups, *Bull. London Math. Soc.* **41** (2009), no. 1, 94–102.

[8] Ruth Charney and Karen Vogtmann, *Subgroups and quotients of automorphism groups of RAAGs Low-dimensional and symplectic topology*, 9–27, Proc. Sympos. Pure Math., **82**, Amer. Math. Soc., Providence, RI, 2011.

[9] Marc Culler and Karen Vogtmann, Moduli of graphs and automorphisms of free groups, *Invent. Math.* **84** (1986), no. 1, 91–119.

[10] Michael Davis and Tadeusz Januszkiewicz, Right-angled Artin groups are commensurable with right-angled Coxeter groups, *J. Pure Appl. Algebra* **153** (2000), no. 3, 229–235.

[11] Matthew B. Day, Peak reduction and finite presentations for automorphism groups of right-angled Artin groups, *Geom. Topol.* **13** (2009), no. 2, 817–855.

[12] Matthew B. Day, Full-featured peak reduction in right-angled Artin groups, arXiv:1211.0078.

[13] Carl Droms, Subgroups of graph groups, *J. Algebra* **110** (1987), no. 2, 519–522.
[14] S. M. Gersten, A presentation for the special automorphism group of a free group, *J. Pure Appl. Algebra* **33** (1984), no. 3, 269–279.
[15] Vincent Guirardel and Gilbert Levitt, The outer space of a free product, *Proc. London Math. Soc.* (3) **94** (2007), no. 3, 695–714.
[16] Frederic Haglund and Daniel T. Wise, Special cube complexes, *Geom. Funct. Anal.* **17** (2008), no. 5, 1551–1620.
[17] Michael Handel and Lee Mosher, The free splitting complex of a free group I: Hyperbolicity, arXiv:1111.1994.
[18] A. H. M. Hoare, Coinitial graphs and Whitehead automorphisms, *Canad. J. Math.* **21** (1979), no. 1, 112–123.
[19] Michael R. Laurence, A generating set for the automorphism group of a graph group, *J. London Math. Soc. (2)* **52** (1995), no. 2, 318–334.
[20] Wilhelm Magnus, Über n-dimensionale Gittertransformationen, *Acta Math.* **64** (1935), no. 1, 353–367.
[21] James McCool, Some finitely presented subgroups of the automorphism group of a free group, *J. Algebra* **35** (1975), 205–213.
[22] Ashot Minasyan, Hereditary conjugacy separability of right-angled Artin groups and its applications, *Groups Geom. Dyn.* **6** (2012), no. 2, 335–388.
[23] Ashot Minasyan and Denis Osin, Normal automorphisms of relatively hyperbolic groups, *Trans. Amer. Math. Soc.* **362** (2010), no. 11, 6079–6103.
[24] Daniel Quillen, *Higher Algebraic K-Theory, I*, Lect. Notes Math. **341** (1974), 85–147.
[25] Herman Servatius, Automorphisms of graph groups, *J. Algebra* **126** (1989), no. 1, 34–60.
[26] Herman Servatius, Carl Droms and Brigitte Servatius, Surface subgroups of graph groups, *Proc. Amer. Math. Soc.* **106** (1989), no. 3, 573–578.
[27] John R. Stallings, Topology of finite graphs, *Invent. Math.* **71** (1983), no. 3, 551–565.
[28] Nate Stambaugh, Toward an outer space for right-angled Artin groups, Ph.D. Dissertation, Brandeis University, August 2011.
[29] Emmanuel Toinet, Conjugacy p-separability of right-angled Artin groups and applications, *Groups Geom. Dyn.*, to appear.
[30] Anna Vijayan, Compactifying the space of length functions of a right-angled Artin group, Ph.D. Dissertation, Brandeis University, December 2012.

PERMUTATION GROUPS AND TRANSFORMATION SEMIGROUPS: RESULTS AND PROBLEMS

JOÃO ARAÚJO* and PETER J. CAMERON†

*Universidade Aberta and Centro de Algebra, Universidade de Lisboa, Av. Gama Pinto 2, 1649-003 Lisboa, Portugal
Email: jaraujo@ptmat.fc.ul.pt

†Mathematical Institute, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SS, U.K.
Email: pjc@mcs.st-andrews.ac.uk

Abstract

J.M. Howie, the influential St Andrews semigroupist, claimed that we value an area of pure mathematics to the extent that (a) it gives rise to arguments that are deep and elegant, and (b) it has interesting interconnections with other parts of pure mathematics.

This paper surveys some recent results on the transformation semigroup generated by a permutation group G and a single non-permutation a. Our particular concern is the influence that properties of G (related to homogeneity, transitivity and primitivity) have on the structure of the semigroup. In the first part of the paper, we consider properties of $S = \langle G, a\rangle$ such as regularity and idempotent generation. The second is a brief report on the synchronization project, which aims to decide in what circumstances S contains an element of rank 1. The paper closes with a list of open problems on permutation groups and linear groups, and some comments about the impact on semigroups are provided.

These two research directions outlined above lead to very interesting and challenging problems on primitive permutation groups whose solutions require combining results from several different areas of mathematics, certainly fulfilling both of Howie's elegance and value tests in a new and fascinating way.

1 Regularity and generation

1.1 Introduction

How can group theory help the study of semigroups?

If a semigroup has a large group of units, we can apply group theory to it. But there may not be any units at all! According to a widespread belief, almost all finite semigroups have only one idempotent, which is a zero, not an identity (see [25] and [18]). This conjecture, however, should not deter us from the general goal of investigating how the group of units shapes the structure of the semigroup. Infinitely many families of finite semigroups, and the most interesting, are composed by semigroups with a group of units. Some of those families are interesting enough to keep many mathematicians busy their entire lives; in fact a unique family of finite semigroups, the endomorphism semigroups of vector spaces over finite fields, has been keeping experts in linear algebra busy for more than a century.

Regarding the general question of how the group of units can shape the structure of the semigroup, an especially promising area is the theory of *transformation semigroups*, that is, semigroups of mappings $\Omega \to \Omega$ (subsemigroups of the *full transformation semigroup* $T(\Omega)$, where $\Omega := \{1, \ldots, n\}$). This area is especially promising for two reasons. First, in a transformation semigroup S, the units are the permutations; if there are any, they form a *permutation group* G and we can take advantage of the very deep recent results on them, chiefly the classification of finite simple groups (CFSG). Secondly, even if there are no units, we still have a group to play with, the *normaliser* of S in $\mathrm{Sym}(\Omega)$, the set of all permutations g such that $g^{-1}Sg = S$.

The following result of Levi and McFadden [28] is the prototype for results of this kind. Let S_n and T_n denote the symmetric group and full transformation semigroup on $\Omega := \{1, 2, \ldots, n\}$.

Theorem 1.1 *Let $a \in T_n \setminus S_n$, and let S be the semigroup generated by the conjugates $g^{-1}ag$ for $g \in S_n$. Then*

(a) S is idempotent-generated;

(b) S is regular;

(c) $S = \langle a, S_n \rangle \setminus S_n$.

In other words, semigroups of this form, with normaliser S_n, have *very nice* properties!

Inspired by this result, we could formulate a general problem:

Problem 1.2 (a) Given a semigroup property P, for which pairs (a, G), with $a \in T_n \setminus S_n$ and $G \leq S_n$, does the semigroup $\langle g^{-1}ag : g \in G \rangle$ have property P?

(b) Given a semigroup property P, for which pairs (a, G) as above does the semigroup $\langle a, G \rangle \setminus G$ have property P?

(c) For which pairs (a, G) are the semigroups of the preceding parts equal?

The following portmanteau theorem lists some previously known results on this problem. The first part is due to Levi [26], the other two to Araújo, Mitchell and Schneider [8].

Theorem 1.3 *(a) For any $a \in T_n \setminus S_n$ the semigroups $\langle g^{-1}ag : g \in S_n \rangle$ and $\langle g^{-1}ag : g \in A_n \rangle$ are equal.*

(b) $\langle g^{-1}ag : g \in G \rangle$ is idempotent-generated for all $a \in T_n \setminus S_n$ if and only if $G = S_n$ or $G = A_n$ or G is one of three specific groups of low degrees.

(c) $\langle g^{-1}ag : g \in G \rangle$ is regular for all $a \in T_n \setminus S_n$ if and only if $G = S_n$ or $G = A_n$ or G is one of eight specific groups of low degrees.

Recently, we have obtained several extensions of these results. The first theorem is proved in [3].

Theorem 1.4 *Given k with $1 \leq k \leq n/2$, the following are equivalent for a subgroup G of S_n:*

(a) for all rank k transformations a, a is regular in $\langle a, G \rangle$;

(b) for all rank k transformations a, $\langle a, G \rangle$ is regular;

(c) for all rank k transformations a, a is regular in $\langle g^{-1}ag : g \in G\rangle$;

(d) for all rank k transformations a, $\langle g^{-1}ag : g \in G\rangle$ is regular.

Moreover, we have a complete list of the possible groups G with these properties for $k \geq 5$, and partial results for smaller values.

It is worth pointing out that in the previous theorem the equivalence between (a) and (c) is not new (it appears in [27]). Really surprising, and a great result that semigroups owe to the classification of finite simple groups, are the equivalences between (a) and (b), and between (c) and (d).

The four equivalent properties above translate into a transitivity property of G which we call the *k-universal transversal property*, which we will describe in the Subsection 1.3.

In the framework of Problem 1.2, let P be the following property: the pair (a, G), with $a \in T_n \setminus S_n$ and $G \leq S_n$, satisfies $\langle a, G\rangle \setminus G = \langle a, S_n\rangle \setminus S_n$.

The classification of the pairs (a, G) with this property poses a very interesting group theoretical problem. Recall that the rank of a map $a \in T_n$ is $|\Omega a|$ and the kernel of a is $\ker(a) := \{(x, y) \in \Omega^2 \mid xa = ya\}$; by the usual correspondence between equivalences and partitions, we can identify $\ker(a)$ with a partition $\{A_1, \ldots, A_k\}$. Suppose $|\Omega| > 2$ and we have a rank 2 map $a \in T_n$. It is clear that $g^{-1}a \in \langle a, S_n\rangle$, for all $g \in S_n$. In addition, if $\ker(a) = \{A_1, A_2\}$, then $\ker(g^{-1}a) = \{A_1g, A_2g\}$. Therefore, in order to classify the groups with property P above we need to find the groups G such that

$$\{\{A_1, A_2\}g \mid g \in G\} \quad = \quad \{\{A_1, A_2\}g \mid g \in S_n\}. \tag{1}$$

If $|A_1| < |A_2|$, this is just $|A_1|$-homogeneity; but if these two sets have the same size, the property is a little more subtle.

Extending this analysis to partitions with more than two parts, we see that the group-theoretic properties we need to investigate are transitivity on ordered partitions of given shape (this notion was introduced by Martin and Sagan [32] under the name *partition-transitivity*) and the weaker notion of transitivity on unordered partitions of given shape. This is done in Section 1.4, where we indicate the proof of the following theorem from [1].

Theorem 1.5 *We have a complete list (in terms of the rank and kernel type of a) for pairs (a, G) for which $\langle a, G\rangle \setminus G = \langle a, S_n\rangle \setminus S_n$.*

As we saw, the semigroups $\langle a, S_n\rangle \setminus S_n$ have very nice properties. In particular, the questions of calculating their automorphisms and congruences, checking for regularity, idempotent generation, etc., are all settled. Therefore the same happens for the groups G and maps $a \in T_n \setminus S_n$ such that $\langle a, G\rangle \setminus G = \langle a, S_n\rangle \setminus S_n$, and all these pairs (a, G) have been classified.

Another long-standing open question was settled by the following theorem, from [5].

Theorem 1.6 *The semigroups $\langle a, G\rangle \setminus G$ and $\langle g^{-1}ag : g \in G\rangle$ are equal for all $a \in T_n \setminus S_n$ if and only if $G = S_n$, or $G = A_n$, or G is the trivial group, or G is one of five specific groups.*

Problem 1.7 It would be good to have a more refined version of this where the hypothesis refers only to all maps of rank k, or just a single map a.

1.2 Homogeneity and related properties

A permutation group G on Ω is *k-homogeneous* if it acts transitively on the set of k-element subsets of Ω, and is *k-transitive* if it acts transitively on the set of k-tuples of distinct elements of Ω.

It is clear that k-homogeneity is equivalent to $(n-k)$-homogeneity, where $|\Omega| = n$; so we may assume that $k \leq n/2$. It is also clear that k-transitivity implies k-homogeneity.

We say that G is *set-transitive* if it is k-homogeneous for all k with $0 \leq k \leq n$. The problem of determining the set-transitive groups was posed by von Neumann and Morgenstern [33] in the first edition of their influential book on game theory. In the second edition, they refer to an unpublished solution by Chevalley, but the first published solution was by Beaumont and Peterson [10]. The set-transitive groups are the symmetric and alternating groups, and four small exceptions with degrees $5, 6, 9, 9$.

In an elegant paper in 1965, Livingstone and Wagner [30] showed:

Theorem 1.8 *Let G be k-homogeneous, where $2 \leq k \leq n/2$. Then*

(a) G is $(k-1)$-homogeneous;

(b) G is $(k-1)$-transitive;

(c) if $k \geq 5$, then G is k-transitive.

In particular, part (a) of this theorem is proved by a short argument using character theory of the symmetric group. This can be translated into combinatorics, and generalised to linear and affine groups: see Kantor [23].

The k-homogeneous but not k-transitive groups for $k = 2, 3, 4$ were determined by Kantor [21, 22]. All this was pre-CFSG.

The k-transitive groups for $k > 1$ are known, but the classification uses CFSG. Lists can be found in various references such as [11, 15].

1.3 The *k*-universal transversal property

Let $G \leq S_n$, and k an integer smaller than n.

The group G has the *k-universal transversal property*, or *k-ut* for short, if for every k-element subset S of $\{1, \ldots, n\}$ and every k-part partition P of $\{1, \ldots, n\}$, there exists $g \in G$ such that Sg is a *transversal* or *section* for P: that is, each part of P intersects Sg in a single point.

Theorem 1.9 *For $k \leq n/2$, the following are equivalent for a permutation group $G \leq S_n$:*

(a) for all $a \in T_n \setminus S_n$ with rank k, a is regular in $\langle a, G\rangle$;

(b) G has the k-universal transversal property.

In order to get the surprising equivalence (noted after Theorem 1.4) of "a is regular in $\langle a, G\rangle$" and "$\langle a, G\rangle$ is regular", we need to know that, for $k \leq n/2$, a group with the k-ut property also has the $(k-1)$-ut property. This fact, the analogue of Theorem 1.8(a), is not at all obvious.

We go by way of a related property: G is $(k-1,k)$-*homogeneous* if, given any two subsets A and B of $\{1,\ldots,n\}$ with $|A| = k-1$ and $|B| = k$, there exists $g \in G$ with $Ag \subseteq B$.

Now the k-ut property implies $(k-1,k)$-homogeneity. (Take a partition with k parts, the singletons contained in A and all the rest. If Bg is a transversal for this partition, then $Bg \supseteq A$, so $Ag^{-1} \subseteq B$.)

The bulk of the argument involves these groups. We show that, if $3 \leq k \leq (n-1)/2$ and G is $(k-1,k)$-homogeneous, then either G is $(k-1)$-homogeneous, or G is one of four small exceptions (with $k = 3,4,5$ and $n = 2k-1$).

It is not too hard to show that such a group G must be transitive, and then primitive. Now careful consideration of the orbital graphs shows that G must be 2-homogeneous, at which point we invoke the classification of 2-homogeneous groups (a consequence of CFSG).

One simple observation: if G is $(k-1,k)$-homogeneous but not $(k-1)$-homogeneous of degree n, then colour one G-orbit of $(k-1)$-sets red and the others blue; by assumption, there is no monochromatic k-set, so n is bounded by the Ramsey number $R(k-1,k,2)$. The values $R(2,3,2) = 6$ and $R(3,4,2) = 13$ are useful here; $R(4,5,2)$ is unknown, and in any case too large for our purposes.

Now we return to considering the k-ut property.

First, we note that the 2-ut property says that every orbit on pairs contains a pair crossing between parts of every 2-partition; that is, every orbital graph is connected. By Higman's Theorem, this is equivalent to primitivity.

For $2 < k < n/2$, we know that the k-ut property lies between $(k-1)$-homogeneity and k-homogeneity, with a few small exceptions. In fact k-ut is equivalent to k-homogeneous for $k \geq 6$; we classify all the exceptions for $k = 5$, but for $k = 3$ and $k = 4$ there are some groups we are unable to resolve (affine, projective and Suzuki groups), which pose interesting problems (see Problems 3.1 and 3.2).

For large k we have:

Theorem 1.10 *For $n/2 < k < n$, the following are equivalent:*

(a) G has the k-universal transversal property;

(b) G is $(k-1,k)$-homogeneous;

(c) G is k-homogeneous.

In the spirit of Livingstone and Wagner, we could ask:

Problem 1.11 Without using CFSG, show any or all of the following implications:

(a) k-ut implies $(k-1)$-ut for $k \leq n/2$;

(b) $(k-1,k)$-homogeneous implies $(k-2,k-1)$-homogeneous for $k \leq n/2$;

(c) k-ut (or $(k-1,k)$-homogeneous) implies $(k-1)$-homogeneous for $k \leq n/2$.

1.4 Partition transitivity and homogeneity

Let λ be a partition of n (a non-increasing sequence of positive integers with sum n). A partition of $\{1,\ldots,n\}$ is said to have *shape* λ if the size of the ith part is the ith part of λ.

The group G is λ-*transitive* if, given any two (ordered) partitions of shape λ, there is an element of G mapping each part of the first to the corresponding part of the second. (This notion is due to Martin and Sagan [32].) Moreover, G is λ-*homogeneous* if there is an element of G mapping the first partition to the second (but not necessarily respecting the order of the parts).

Of course λ-transitivity implies λ-homogeneity, and the converse is true if all parts of λ are distinct. If $\lambda = (n-t,1,\ldots,1)$, then λ-transitivity and λ-homogeneity are equivalent to t-transitivity and t-homogeneity.

The connection with semigroups is given by the next result, from [1]. Let G be a permutation group, and $a \in T_n \setminus S_n$, where r is the rank of a, and λ the shape of the kernel partition.

Theorem 1.12 *For $G \leq S_n$ and $a \in T_n \setminus S_n$, the following are equivalent:*

(a) $\langle a, G\rangle \setminus G = \langle a, S_n\rangle \setminus S_n$*;*

(b) G is r-homogeneous and λ-homogeneous.

So we need to know the λ-homogeneous groups. First, we consider λ-transitive groups.

If the largest part of λ is greater than $n/2$ (say $n-t$, where $t < n/2$), then G is λ-transitive if and only if it is t-homogeneous and the group H induced on a t-set by its setwise stabiliser is λ'-transitive, where λ' is λ with the part $n-t$ removed.

So if G is t-transitive, then it is λ-transitive for all such λ.

If G is t-homogeneous but not t-transitive, then $t \leq 4$, and examination of the groups in Kantor's list gives the possible λ' in each case.

So what remains is to show that, if G is λ-transitive but not S_n or A_n, then λ must have a part greater than $n/2$.

If $\lambda \neq (n), (n-1,1)$, then G is primitive.

If $n \geq 8$, then by *Bertrand's Postulate*, there is a prime p with $n/2 < p \leq n-3$. If there is no part of λ which is at least p, then the number of partitions of shape λ (and hence the order of G) is divisible by p. A theorem of Jordan (see Wielandt [40], Theorem 13.9) now shows that G is symmetric or alternating.

The classification of λ-homogeneous but not λ-transitive groups is a bit harder. We have to use

(a) a little character theory to show that either G fixes a point and is transitive on the rest, or G is transitive;

(b) the argument using Bertrand's postulate and Jordan's theorem as before;

(c) CFSG (to show that G cannot be more than 5-homogeneous if it is not S_n or A_n).

The outcome is a complete list of such groups.

1.5 Normalising groups

We define a permutation group G to be *normalising* if $\langle g^{-1}ag : g \in G\rangle = \langle a, G\rangle \setminus G$ for all $a \in T_n \setminus S_n$.

The classification of normalising groups given by Theorem 1.6 is a little different; although permutation group techniques are essential in the proof, we didn't find a simple combinatorial condition on G which is equivalent to this property. We will not discuss it further here.

2 Synchronization

2.1 Introduction

In this section, we give a brief report on synchronization.

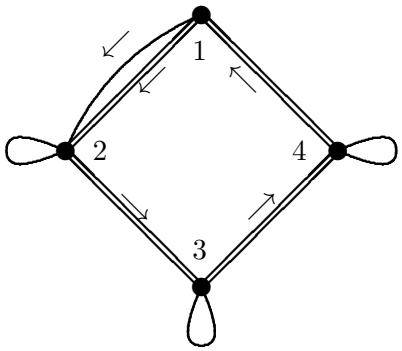
A (finite deterministic) *automaton* consists of a finite set Ω of *states* and a finite set of maps from Ω to Ω called *transitions*, which may be composed freely.

In other words, it is a transformation semigroup with a distinguished set of generators.

An automaton is *synchronizing* if there is a map of rank 1 (image of size 1) in the semigroup. A word in the generators expressing such a map is called a *reset word*.

We will also call a transformation semigroup *synchronizing* if it contains an element of rank 1.

Example 2.1 This example has four (numbered) states, and two transitions A and B, shown as double and single lines respectively.



The reader can check easily that, irrespective of the starting state, following the path $BAAABAAAB$ always ends in state 2, and hence this is a reset word of length 9. In fact, this is the shortest reset word.

The *Černý Conjecture* asserts that if an n-state automaton is synchronizing, then it has a reset word of length at most $(n-1)^2$. The above example, with the square replaced by an n-gon, shows that this would be best possible. The problem has been open for about 45 years. The best known bound is cubic.

It is known that testing whether an automaton is synchronizing is in P, but finding the length of the shortest reset word is NP-hard.

2.2 Graph homomorphisms and transformation semigroups

All graphs here are undirected simple graphs (no loops or multiple edges).

A *homomorphism* from a graph X to a graph Y is a map f from the vertex set of X to the vertex set of Y which carries edges to edges. (We don't specify what happens to a non-edge; it may map to a non-edge, or to an edge, or collapse to a vertex.) An *endomorphism* of a graph X is a homomorphism from X to itself.

Let K_r be the complete graph with r vertices. The *clique number* $\omega(X)$ of X is the size of the largest complete subgraph, and the *chromatic number* $\chi(X)$ is the least number of colours required for a proper colouring of the vertices (adjacent vertices getting different colours).

(a) There is a homomorphism from K_r to X if and only if $\omega(X) \geq r$.

(b) There is a homomorphism from X to K_r if and only if $\chi(X) \leq r$.

There are correspondences in both directions between graphs and transformation semigroups (not quite functorial, or a Galois correspondence, sadly!).

First, any graph X has an *endomorphism semigroup* $\mathrm{End}(X)$.

In the other direction, given a transformation semigroup S on Ω, its *graph* $\mathrm{Gr}(S)$ has Ω as vertex set, two vertices v and w being joined if and only if there is no element of S which maps v and w to the same place.

(a) $\mathrm{Gr}(S)$ is complete if and only if $S \leq S_n$;

(b) $\mathrm{Gr}(S)$ is null if and only if S is synchronizing;

(c) $S \leq \mathrm{End}(\mathrm{Gr}(S))$ for any $S \leq T_n$;

(d) $\omega(\mathrm{Gr}(S)) = \chi(\mathrm{Gr}(S))$; this is equal to the minimum rank of an element of S.

Now the main theorem of this section describes the unique obstruction to synchronization for a transformation semigroup.

Theorem 2.2 *A transformation semigroup S on Ω is non-synchronizing if and only if there is a non-null graph X on the vertex set Ω with $\omega(X) = \chi(X)$ and $S \leq \mathrm{End}(X)$.*

In the reverse direction, the endomorphism semigroup of a non-null graph cannot be synchronizing, since edges can't be collapsed. In the forward direction, take $X = \mathrm{Gr}(S)$; there is some straightforward verification to do. (For details see [4].)

2.3 Maps synchronized by groups

Let $G \leq S_n$ and $a \in T_n \setminus S_n$. We say that G *synchronizes* a if $\langle a, G\rangle$ is synchronizing.

By abuse of language, we say that G is *synchronizing* if it synchronizes every element of $T_n \setminus S_n$.

Our main problem is to determine the synchronizing groups. From the theorem, we see that G is non-synchronizing if and only if there is a G-invariant graph whose clique number and chromatic number are equal.

Rystsov [36] showed the following result, which implies that synchronizing groups are necessarily primitive.

Theorem 2.3 *A permutation group G of degree n is primitive if and only if it synchronizes every map of rank $n-1$.*

We give a brief sketch of the proof, to illustrate the graph endomorphism technique. The backward implication is trivial; so suppose, for a contradiction, that G is primitive but fails to synchronize the map a of rank $n-1$. Then there are two points x, y with $xa = ya$, and a is bijective on the remaining points. Choose a graph X with $\langle G, a\rangle \leq \mathrm{End}(X)$. Note that X is a regular graph. Since a is an endomorphism, x and y are non-adjacent; so a maps the neighbours of x bijectively to the neighbours of xa, and similarly the neighbours of y to those of ya. Since $xa = ya$, we see that x and y have the same neighbour set. Now "same neighbour set" is an equivalence relation preserved by G, contradicting primitivity.

So a synchronizing group must be primitive.

We have recently improved this: a primitive group synchronizes every map of rank $n-2$. The key tool in the proof is graph endomorphisms. Also, a primitive group synchronizes every map of kernel type $(k, 1, \ldots, 1)$. For both results, and further information, see [4].

Also, G is synchronizing if and only if there is no G-invariant graph, not complete or null, with clique number equal to chromatic number. For more on this see [9, 13, 34, 35, 36, 38, 39]. Thus, a 2-homogeneous group is synchronizing, and a synchronizing group is primitive. For if G is 2-transitive, the only G-invariant graphs are complete or null; and if G is imprimitive, then it preserves a complete multipartite graph.

Furthermore, a synchronizing group is *basic* in the O'Nan–Scott classification, that is, not contained in a wreath product with the product action. (For non-basic primitive groups preserve Hamming graphs, which have clique number equal to chromatic number.) By the O'Nan–Scott Theorem, such a group is affine, diagonal or almost simple.

None of the above implications reverses. Indeed, there are non-synchronizing basic groups of all three O'Nan–Scott types.

We are a long way from a classification of synchronizing groups. The attempts to classify them lead to some interesting and difficult problems in extremal combinatorics, finite geometry, computation, etc. But that is another survey paper! We content ourselves here with a single result about an important class of primitive groups, namely the classical symplectic, orthogonal and unitary groups, acting on their associated polar spaces. The implicit geometric problem has not been completely solved, despite decades of work by finite geometers. We refer to Thas [37] for a survey.

Theorem 2.4 *A classical group, acting on the points of its associated polar space, is non-synchronizing if and only if the polar space possesses either an ovoid and a spread, or a partition into ovoids.*

2.4 A conjecture

We regard the following as the biggest open problem in the area. A map $a \in T_n$ is *non-uniform* if its kernel classes are not all of the same size.

Conjecture 2.5 *A primitive permutation group synchronizes every non-uniform map.*

We have some partial results about this (see [2, 4]) but are far from a proof!

3 Problems

One of the goals of this paper is to provide a list of problems that might help the interested reader involve himself in this fascinating topic. In addition to the problems included above, we collect here a number of problems on the general interplay between properties of the group of units and properties of the semigroup containing it.

We start by proposing a problem to experts in number theory. If this problem can be solved, the results on $\mathrm{AGL}(1,p)$, in [3], will be dramatically sharpened.

Problem 3.1 Classify the prime numbers p congruent to 11 (mod 12) such that for some $c \in \mathrm{GF}(p)^*$ we have $|\langle -1, c, c-1\rangle| < p-1$.

The primes less than 500 with this property are 131, 191, 239, 251, 311, 419, 431, and 491.

Problem 3.2 Do the Suzuki groups $\mathrm{Sz}(q)$ have the 3-universal transversal property?

Classify the groups G that have the 4-ut property, when $\mathrm{PSL}(2,q) \leq G \leq \mathrm{P\Gamma L}(2,q)$, with either q prime (except $\mathrm{PSL}(2,q)$ for $q \equiv 1 \pmod 4$, which is not 3-homogeneous), or $q = 2^p$ for p prime.

A group $G \leq S_n$ has the $(n-1)$-universal transversal property if and only if it is transitive. And $\langle a, G\rangle$ (for a rank $n-1$ map a) contains all the rank $n-1$ maps of T_n if and only if G is 2-homogeneous. In this last case $\langle a, G\rangle$ is regular for all $a \in T_n$, because $\langle a, G\rangle = \{b \in T_n \mid |\Omega b| \leq n-1\} \cup G$, and this semigroup is well known to be regular.

Problem 3.3 Classify the groups $G \leq S_n$ such that G together with any rank $n-k$ map, where $k \leq 5$, generate a regular semigroup. We already know that such G must be k-homogeneous; so we know which groups to look at (see Theorem 1.10).

The difficulty here (when rank $k > \lfloor (n+1)/2 \rfloor$) is that a k-homogenous group is not necessarily $(k-1)$-homogenous. Therefore a rank k map $a \in T_n$ might be regular in $\langle a, G\rangle$, but we are not sure that there exists $g \in G$ such that $\mathrm{rank}(bgb) = \mathrm{rank}(b)$, for $b \in \langle a, G\rangle$ such that $\mathrm{rank}(b) < \mathrm{rank}(a)$.

It is clear that if $\langle a, G\rangle \setminus G$ is idempotent generated, for all rank k transformation $a \in T_n \setminus S_n$, then G has the k-ut property (see [8]).

Problem 3.4 Classify the groups $G \leq S_n$ such that $\langle a, G\rangle \setminus G$ is idempotent generated, for all rank k maps, where $k \leq n/2$. Even if the classification of the groups with the k-ut property is not quite finished (Problem 3.2 is the missing part), it might be possible to settle the idempotent generation problem.

Problem 3.5 The most general problem that has to be handled is the classification of pairs (a, G), where $a \in T_n$ and $G \leq S_n$, such that $\langle a, G\rangle$ is a regular semigroup.

When investigating $(k-1)$-homogenous groups without the k-universal transversal property (k-ut property), it was common that some of the orbits on the k-sets have transversals for all the partitions. Therefore the following definition is natural.

A group $G \leq S_n$ is said to have the weak k-ut property if there exists a k-set $S \subseteq \Omega$ such that the orbit of S under G contains a section for all k-partitions. Such a set is called a G-universal transversal set. A solution to the following problem would have important consequences in semigroup theory.

Problem 3.6 Classify the groups with the weak k-ut property; in addition, for each of them, classify their G-universal transversal sets.

In McAlister's celebrated paper [31] it is proved that, if $e^2 = e \in T_n$ is a rank $n-1$ idempotent, then $\langle G, e\rangle$ is regular for all groups $G \leq S_n$. In addition, assuming that $\{\alpha, \beta\}$ is the non-singleton kernel class of e and $\alpha e = \beta$, if α and β are not in the same orbit under G, then $\langle e, G\rangle$ is an orthodox semigroup (that is, the idempotents form a subsemigroup); and $\langle e, G\rangle$ is inverse if and only if α and β are not in the same orbit under G and the stabilizer of α is contained in the stabilizer of β.

Problem 3.7 Classify the groups $G \leq S_n$ that together with any idempotent [rank k idempotent] generate a regular [orthodox, inverse] semigroup.

Classify the pairs (G, a), with $a \in T_n$ and $G \leq S_n$, such that $\langle e, G\rangle$ is inverse [orthodox].

The theorems and problems in this paper admit linear versions that are interesting for experts in groups and semigroups, but also to experts in linear algebra and matrix theory.

Problem 3.8 Prove (or disprove) that if $G \leq \mathrm{GL}(n, q)$ such that for all singular matrices a there exists $g \in G$ with $\mathrm{rank}(a) = \mathrm{rank}(aga)$, then G contains the special linear group.

For $n = 2$ and for $n = 3$, this condition is equivalent to irreducibility of G. But we conjecture that, for sufficiently large n, it implies that G contains the special linear group.

Problem 3.9 Classify the groups $G \leq \mathrm{GL}(n, q)$ such that for all rank k (for a given k) singular matrices a we have that a is regular in $\langle G, a\rangle$ [the semigroup $\langle G, a\rangle$ is regular].

To handle this problem it is useful to keep in mind the following results. Kantor [23] proved that if a subgroup of $\mathrm{P\Gamma L}(d, q)$ acts transitively on k-dimensional subspaces, then it acts transitively on l-dimensional subspaces for all $l \leq k$ such that $k + l \leq n$; in [24], he showed that subgroups transitive on 2-dimensional subspaces are 2-transitive on the 1-dimensional subspaces with the single exception of a subgroup of $\mathrm{PGL}(5, 2)$ of order $31 \cdot 5$; and, with the second author [12], he showed that such groups must contain $\mathrm{PSL}(d, q)$ with the single exception of the alternating group A_7 inside $\mathrm{PGL}(4, 2) \cong A_8$. Also Hering [19, 20] and Liebeck [29], using CFSG, classified the subgroups of $\mathrm{PGL}(d, p)$ which are transitive on 1-spaces.

Regarding synchronization, the most important question (in our opinion) is the following conjecture, stated earlier.

Problem 3.10 Is it true that every primitive group of permutations of a finite set Ω synchronizes every non-uniform transformation on Ω?

Assuming the previous question has an affirmative answer (as we believe), an intermediate step in order to prove it would be to solve the following set of connected problems:

Problem 3.11 (a) Prove that every map of rank $n-3$, with non-uniform kernel, is synchronized by a primitive group. This is known for idempotent maps (see [4]).

(b) Prove that a primitive group synchronizes every non-uniform map of rank 5.

(c) Prove that if in $S = \langle f, G\rangle$ there is a map of minimal rank $r > 1$, there can be no map in S with rank $r + 2$.

The next class of groups lies strictly between primitive and synchronizing.

Problem 3.12 Is it possible to classify the primitive groups which synchronize every rank 3 map?

Note that there are primitive groups that do not synchronize a rank 3 map (see [34]). And there are non-synchronizing groups which synchronize every rank 3 map. Take for example $\mathrm{PGL}(2, 7)$ of degree 28; this group is non-synchronizing, but synchronizes every rank 3 map, since 28 is not divisible by 3.

There are very fast polynomial-time algorithms to decide if a given set of permutations generates a primitive group, or a 2-transitive group.

Problem 3.13 Find an efficient algorithm to decide if a given set of permutations generates a synchronizing group.

It would be quite remarkable if such an algorithm exists; as we saw, it would in particular resolve questions about ovoids and spreads in certain polar spaces (among other things).

There are a number of natural problems related to λ-homogeneity.

Problem 3.14 Let $H \leq S_n$ be a 2-transitive group. Classify the pairs (a, G), where $a \in S_n$ and $G \leq S_n$, such that $\langle a, G\rangle = H$.

Problem 3.15 Let $G \leq S_n$ be a 2-transitive group. (The list of those groups is available in [11, 15].) For every $a \in T_n$ describe the structure of $\langle G, a\rangle \setminus G$. In particular (where G is a 2-transitive group and $a \in T_n$):

(a) classify all the pairs (a, G) such that $\langle a, G\rangle$ is a regular semigroup (that is, for all $x \in \langle a, G\rangle$ there exists $y \in \langle a, G\rangle$ such that $x = xyx$);

(b) classify all the pairs (a, G) such that $\langle a, G\rangle \setminus G$ is generated by its idempotents;

(c) classify all the pairs (a, G) such that $\langle a, G\rangle \setminus G = \langle g^{-1}ag \mid g \in G\rangle$;

(d) describe the automorphisms, congruences, principal right, left and two-sided ideals of the semigroups $\langle a, G\rangle$ (when G is a 2-transitive group).

Problem 3.16 For each 2-transitive group G classify the G-pairs, that is, the pairs (a, H) such that $H \leq S_n$, $a \in T_n$ and $\langle a, G\rangle \setminus G = \langle a, H\rangle \setminus H$.

Problem 3.17 Let V be a finite dimensional vector space. A pair (a, G), where a is a singular endomorphism of V and $G \leq \mathrm{Aut}(V)$, is said to be an $\mathrm{Aut}(V)$-pair if

$$\langle a, G\rangle \setminus G = \langle a, \mathrm{Aut}(V)\rangle \setminus \mathrm{Aut}(V).$$

Classify the $\mathrm{Aut}(V)$-pairs.

Problem 3.18 Formulate and prove analogues of the results in this paper, but for semigroups of linear maps on a vector space.

Problem 3.19 Solve the analogue of Problem 3.18 for independence algebras (for definitions and fundamental results see [6, 7, 14, 16, 17]).

Acknowledgment The first author was partially supported by Pest-OE/MAT/UI0143/2011 of Centro de Algebra da Universidade de Lisboa, and by FCT and PIDDAC through the project PTDC/MAT/101993/2008.

References

[1] J. André, J. Araújo and P.J. Cameron, The classification of partition homogeneous groups with applications to semigroup theory, http://arxiv.org/abs/1304.7391

[2] J. Araújo, W. Bentz and P.J. Cameron, Groups synchronizing a transformation of non-uniform kernel, *Theoret. Comput. Sci.* **498** (2013), 1–9.

[3] J. Araújo and P.J. Cameron, Two generalizations of homogeneity in groups with applications to regular semigroups, http://arxiv.org/abs/1204.2195

[4] J. Araújo and P.J. Cameron, Primitive groups synchronize non-uniform maps of extreme ranks, http://arxiv.org/abs/1306.4827

[5] J. Araújo, P.J. Cameron, J.D. Mitchell and M. Neunhöffer, The classification of normalizing groups, *J. Algebra* **373** (2013), 481–490.

[6] J. Araújo, M. Edmundo and S. Givant, v^*-algebras, independence algebras and logic, *Internat. J. Algebra Comput.* **21** (2011), 1237–1257.

[7] J. Araújo and J. Fountain, The origins of independence algebras, *Proceedings of the Workshop on Semigroups and Languages* (Lisbon 2002), 54–67, World Scientific, 2004.

[8] J. Araújo, J.D. Mitchell and C. Schneider, Groups that together with any transformation generate regular semigroup or idempotent generated semigroups, *J. Algebra* **343** (2011), 93–106.

[9] F. Arnold and B. Steinberg, Synchronizing groups and automata, *Theoret. Comput. Sci.* **359** (2006), 101–110.

[10] R.A. Beaumont and R.P. Peterson, Set-transitive permutation groups, *Canad. J. Math.* **7** (1955), 35–42.

[11] P.J. Cameron, *Permutation groups*, London Math. Soc. Student Texts **45**, Cambridge Univ. Press, Cambridge, 1999.

[12] P.J. Cameron and W.M. Kantor, 2-transitive and antiflag transitive collineation groups of finite projective spaces, *J. Algebra* **60** (1979), 384–422.

[13] P.J. Cameron and A. Kazanidis, Cores of symmetric graphs, *J. Austral. Math. Soc.* **85** (2008), 145–154.

[14] P.J. Cameron and C. Szabó, Independence algebras, *J. London Math. Soc.* **61** (2000), 321–334.

[15] J.D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts Math. **163**, Springer-Verlag, New York, 1996.

[16] J. Fountain and A. Lewin, Products of idempotent endomorphisms of an independence algebra of finite rank, *Proc. Edinburgh Math. Soc.* **35** (1992), 493–500.

[17] V. Gould, Independence algebras, *Algebra Universalis* **33** (1995), 294–318.
[18] H. Jürgensen, F. Migliorini, and J. Szép, *Semigroups*, Akadémiai Kiadó, Budapest, 1991.
[19] C. Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order, *Geom. Dedicata* **2** (1974), 425–460.
[20] C. Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order, II, *J. Algebra* **93** (1985), 151–164.
[21] W.M. Kantor, 4-homogeneous groups, *Math. Z.* **103** (1968), 67-68; correction *ibid.* **109** (1969), 86.
[22] W.M. Kantor, k-homogeneous groups, *Math. Z.* **124** (1972), 261–265.
[23] W.M. Kantor, On incidence matrices of projective and affine spaces, *Math. Z.* **124** (1972), 315–318.
[24] W.M. Kantor, Line-transitive collineation groups of finite projective spaces, *Israel J. Math.* **14** (1973), 229–235.
[25] D.J. Kleitman, B.R. Rothschild and J.H. Spencer, The number of semigroups of order n, *Proc. Amer. Math. Soc.* **55** (1976), 227–232.
[26] I. Levi, On the inner automorphisms of finite transformation semigroups, *Proc. Edinburgh Math. Soc.* **39** (1996), 27–30.
[27] I. Levi, D.B. McAlister and R.B. McFadden, Groups associated with finite transformation semigroups, *Semigroup Forum* **61** (2000), 453–467.
[28] I. Levi and R.B. McFadden, S_n-normal semigroups, *Proc. Edinburgh Math. Soc.* **37** (1994), 471–476.
[29] M.W. Liebeck, The affine permutation groups of rank 3, *Bull. London Math. Soc.* **18** (1986), 165–172.
[30] D. Livingstone and A. Wagner, Transitivity of finite permutation groups on unordered sets, *Math. Z.* **90** (1965), 393–403.
[31] D.B. McAlister, Semigroups generated by a group and an idempotent, *Comm. Algebra* **26** (1998), 515–547.
[32] W.J. Martin and B.E. Sagan, A new notion of transitivity for groups and sets of permutations, *J. London Math. Soc.* **73** (2006), 1–13.
[33] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1944.
[34] P.M. Neumann, Primitive permutation groups and their section-regular partitions, *Michigan Math. J.* **58** (2009), 309–322.
[35] J.-E. Pin, Černý's conjecture, http://www.liafa.jussieu.fr/~jep/Problemes/Cerny.html
[36] I. Rystsov, Quasioptimal bound for the length of reset words for regular automata, *Acta Cybernet.* **12** (1995), 145–152.
[37] J.A. Thas, Projective geometry over a finite field, in *Handbook of Incidence Geometry* (ed. F. Buekenhout), 295–347, Elsevier, Amsterdam, 1995.
[38] A.N. Trahtman, Bibliography, synchronization @ TESTAS, http://www.cs.biu.ac.il/~trakht/syn.html
[39] A.N. Trahtman, The Černý conjecture for aperiodic automata, *Discrete Math. Theor. Comput. Sci.* **9** (2007), 3–10.
[40] H. Wielandt, *Finite Permutation Groups*, Academic Press, New York, 1964.

NEW PROGRESS ON FACTORIZED GROUPS AND SUBGROUP PERMUTABILITY

MILAGROS ARROYO-JORDÁ*, PAZ ARROYO-JORDÁ*,
ANA MARTÍNEZ-PASTOR† and M. DOLORES PÉREZ-RAMOS§

*Escuela Técnica Superior de Ingenieros Industriales, Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain
E-mail: marroyo@mat.upv.es, parroyo@mat.upv.es

†Escuela Técnica Superior de Ingeniería Informática, Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain
E-mail: anamarti@mat.upv.es

§Departament d'Àlgebra, Universitat de València, Doctor Moliner 50, 46100 Burjassot (València), Spain
E-mail: Dolores.Perez@uv.es

Abstract

The study of products of groups whose factors are linked by certain permutability conditions has been the subject of fruitful investigations by a good number of authors. A particular starting point was the interest in providing criteria for products of supersoluble groups to be supersoluble. We take further previous research on total and mutual permutability by considering significant weaker permutability hypotheses. The aim of this note is to report about new progress on structural properties of factorized groups within the considered topic. As a consequence, we discuss new attainments in the framework of formation theory.

1 Introduction

In this survey only finite groups are considered.

The study of groups factorized as the product of two subgroups has been the subject of considerable interest in recent years. One of the important questions dealing with this study is how the structure of the factors affects the structure of the whole group and vice versa. A natural approach to this problem is provided by the theory of classes of groups. In this context, the above question can be reformulated as when the belonging of the factors of a factorized group to a class of groups is transferred to the whole group and reciprocally. It is well known that the product of two normal supersoluble subgroups is not supersoluble, in general. Nevertheless, the class of all supersoluble groups $\mathcal{U}$ is closed under forming direct and central products. It seems then natural to consider factorized groups in which certain subgroups of the corresponding factors permute, in order to obtain new criteria of supersolubility. A starting point of this research can be located at M. Asaad and A. Shaalan's paper [6]. They considered factorized groups $G = AB$ where A and B are supersoluble

subgroups and, in particular, they proved that G is supersoluble under any of the following conditions:

(i) Every subgroup of A permutes with every subgroup of B.

(ii) A permutes with every subgroup of B, B permutes with every subgroup of A and, moreover, the derived subgroup G' of G is a nilpotent group.

Products of groups whose factors satisfy condition (i) were called *totally permutable products* by R. Maier in [33], where he proved that a corresponding result remains valid when the saturated formation $\mathcal{U}$ of all supersoluble groups is replaced by any saturated formation containing $\mathcal{U}$. Later on this result was also extended to non-saturated formations which contain $\mathcal{U}$ (see [14]). On the other hand, factorized groups whose factors satisfy the permutability property stated in (ii) were called *mutually permutable products* by A. Carocca in [21]. In [9] it was proved that a corresponding result under condition (ii) is true when supersolubility is again replaced by containment in any saturated formation containing $\mathcal{U}$. Totally and mutually permutable products of groups have since been subject of an in-depth study, both in the frameworks of formation theory (see [9, 10, 11, 13, 14, 15, 19, 22, 23, 33]) as well as in the theory of Fitting classes (see [16, 18, 20, 27, 28, 29]), and its structure is currently very well understood. An exhaustive report on this matter appears in [8].

More recently, this study has been spread by introducing a weaker condition of subgroup permutability, namely *conditional permutability*, which requires permutability for some conjugates of the considered subgroups. Using this permutability property new criteria for a product of supersoluble groups to be supersoluble are obtained in [26, 31, 32] and by the authors in [2], extending known results. A generalization in the framework of formation theory has been initiated in [5]. In [3] recent developments involving conditional permutability and other close permutability properties, in relation with products of groups, supersolubility and formation theory, have been collected.

This survey reports about new progress on the current research on the topic. We aim to contribute a better understanding of structural properties of products of groups under consideration as well as to analyze to which extent permutability hypotheses can be weakened. This has led to the interest in a stronger variant of conditional permutability, namely complete conditional permutability. This information will be used to obtain new achievements in the context of formation theory. The results presented here can be found mainly in [2, 4, 5].

We recall that a *formation* is a class $\mathcal{F}$ of groups closed under homomorphic images, such that $G/(N \cap M) \in \mathcal{F}$ whenever G is a group and M, N are normal subgroups of G with $G/N, G/M \in \mathcal{F}$. In this case the *$\mathcal{F}$-residual* $G^{\mathcal{F}}$ of G is the smallest normal subgroup of G such that $G/G^{\mathcal{F}} \in \mathcal{F}$. The formation $\mathcal{F}$ is *saturated* if $G \in \mathcal{F}$ whenever $G/\Phi(G) \in \mathcal{F}$, where $\Phi(G)$ denotes the Frattini subgroup of G.

For notation we refer to [24].

2 Conditional permutability

We follow W. Guo et al. in [25, 26], and collect the following concepts:

Definition 2.1 Let G be a group. Two subgroups X and Y of G are called *condi-*

tionally permutable (*c-permutable*, for brevity) in G if X permutes with Y^g for some element $g \in G$.

The subgroups X and Y are called *completely conditionally permutable* (or *completely c-permutable*) in G if X permutes with Y^g for some element $g \in \langle X, Y \rangle$, the subgroup generated by X and Y.

Two subgroups A and B of G are said to be *totally conditionally permutable* (or *totally c-permutable*) if every subgroup of A is c-permutable in G with every subgroup of B. Moreover, if $G = AB$, we say that G is the *totally c-permutable product* of the subgroups A and B.

Two subgroups A and B of G are said to be *totally completely conditionally permutable* (or *totally completely c-permutable*) in G if every subgroup of A is completely c-permutable in G with every subgroup of B. If $G = AB$, we say that G is the *totally completely c-permutable product* of the subgroups A and B.

Conditional permutability has been considered by several authors in extending classical results about the influence of permutability properties of certain families of subgroups on the structure of groups (see [25, 30, 34]).

We mention that c-permutability fails to satisfy the property of persistence in intermediate subgroups; i.e., if X and Y are c-permutable subgroups in a group G, then X and Y are not necessarily c-permutable in any subgroup M of G such that $X, Y \leq M \leq G$, as the next example shows (see [2, 5]). This makes a relevant difference between c-permutability and complete c-permutability. In fact, complete c-permutability appears when requiring c-permutability to satisfy this persistence property and becomes a stronger hypothesis.

Example 2.2 Let $G = \mathrm{Sym}(4)$ be the symmetric group of degree 4, Y a subgroup of G of order 2 generated by a transposition, V the normal subgroup of G of order 4 and X a subgroup of V of order 2, $X \neq Z(VY)$. Then X and Y are c-permutable in G but they are not c-permutable in $\langle Y, X \rangle$.

We point out that totally permutable products are examples of totally (completely) c-permutable products, but the converse is not true in general. Also mutually permutable products are not necessarily totally (completely) c-permutable products and vice versa. These facts will appear clear along this survey.

We remark however that a totally completely c-permutable product of two subgroups is indeed a mutually permutable product whenever the factors are nilpotent groups, and it is totally permutable if one of the factors is a nilpotent normal subgroup (see [4]).

3 Structural properties

In this section we report on structural properties of the factorized groups under consideration. It is nice to think of totally permutable products as extension of central products as well as of mutually permutable products as extension of normal products. In this sense, the goal of finding either centralizing or subnormality properties of relevant subgroups of the factors have become a key point in this research. Many authors have contributed to the knowledge of the structure of this kind of products

(for instance, [1, 7, 14, 15, 16, 17, 20, 21, 33]). We wonder whether relevant structural properties known for totally and mutually permutable product are satisfied under weaker permutability conditions as complete conditional permutability.

We begin by stating that one cannot find totally completely c-permutable products with all factors being core-free subgroups:

Lemma 3.1 ([5]) *Let the group $1 \neq G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Then there exists $1 \neq N \trianglelefteq G$ such that $N \leq G_i$ for some $i \in \{1, \ldots, r\}$.*

The corresponding result for totally permutable products was proved in [33] (for $r = 2$) and in [22] (for r arbitrary). In [17] it was shown that in a product of two mutually permutable subgroups the product of the cores of the factors is non-trivial.

Within the study of mutually and totally permutable products, the intersection of the factors plays an important role (see, for instance, [1, 7, 16, 19, 21]). In particular, a mutually permutable product of subgroups with trivial intersection is totally permutable ([21]). More in general, in a mutually permutable product, this intersection is a subnormal subgroup in the whole group (see [16, 21]).

The following example shows that this result is not further true in general in a totally completely c-permutable product:

Example 3.2 ([2, 5]) We consider the symmetric group $G = \mathrm{Sym}(3)$ of degree 3 and the trivial factorization $G = AB$ being $A = G$ and $B = X$ a 2-Sylow subgroup of G. Then every subgroup of A is completely c-permutable in G with every subgroup of B, but $B = A \cap B$ is not subnormal in G.

However we will see next that it is possible to find a variety of relevant subnormal subgroups in the factors of a product of totally completely c-permutable subgroups.

In [16] it was shown that the derived subgroup of each factor in a mutually permutable product is a subnormal subgroup in the product. The same holds true for products of totally completely c-permutable subgroups.

Proposition 3.3 ([4]) *Let the group $G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Then G_i' is a subnormal subgroup of G, for all $i \in \{1, \ldots, r\}$.*

Consequently, if $\mathcal{F}$ is a formation containing $\mathcal{A}$, the class of all abelian groups, we can deduce that the $\mathcal{F}$-residual of each factor in a product of totally completely c-permutable subgroups is a subnormal subgroup of the product.

Totally permutable products are also close to central products since in such products the nilpotent residual of each factor centralizes the other one, that is, if $G = AB$ is the totally permutable product of the subgroups A and B, then $[A^{\mathcal{N}}, B] = [B^{\mathcal{N}}, A] = 1$, where $\mathcal{N}$ is the class of all nilpotent groups (see [15]).

In particular, if $\mathcal{F}$ is a formation such that $\mathcal{N} \subseteq \mathcal{F}$, then $[A^{\mathcal{F}}, B] = [B^{\mathcal{F}}, A] = 1$ and $A^{\mathcal{F}}$ and $B^{\mathcal{F}}$ are normal subgroups in $G = AB$.

Regarding mutual permutability, J. Bochtler in his PhD Thesis [19] ([8, Theorem 4.4.5]) proved that in a mutually permutable product of subgroups A and B, the nilpotent residual of each factor normalizes the other factor.

The next example shows that these properties fail for totally completely c-permutable products:

Example 3.4 ([2, 5]) Consider $V = \langle a, b\rangle \cong Z_5 \times Z_5$ and $Z_6 \cong C = \langle \alpha, \beta\rangle \leq \mathrm{Aut}(V)$ given by $a^\alpha = a^{-1}$, $b^\alpha = b^{-1}$, $a^\beta = b$, $b^\beta = a^{-1}b^{-1}$. Let $G = [V]C$ the corresponding semidirect product of V with C. Set $A = \langle\alpha\rangle$ and $B = V\langle\beta\rangle$. Then $G = AB$ is a totally completely c-permutable product, but $[A, B^{\mathcal{N}}] = [A, V] \neq 1$. On the other hand, $B^{\mathcal{N}} = V$ does not normalizes A.

Nevertheless, under this new permutability condition, it was proved in [4] that the nilpotent residuals of the factors are normal subgroups in the product:

Theorem 3.5 ([4]) *Let the group $G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Then $G_i^{\mathcal{N}}$ is a normal subgroup of G, for all $i \in \{1, \ldots, r\}$.*

Hence, we deduce that in such products if $\mathcal{F}$ is a formation with $\mathcal{N} \subseteq \mathcal{F}$, then the $\mathcal{F}$-residual of each factor is a subnormal subgroup in the product. Moreover, if $\mathcal{F}$ is a saturated formation containing $\mathcal{U}$, the $\mathcal{F}$-residual of each factor is a normal subgroup in the product, as we shall see in Corollary 4.2.

Another important structural property which holds in a totally permutable product $G = AB$ is that the commutator subgroup $[A, B]$ of the factors is a nilpotent normal subgroup in G (see [15]). This result follows as a direct consequence of the centralizing property of the nilpotent residuals of the factors. In spite of the failure of this result for totally completely c-permutable products as Example 3.4 shows, it is relevant that the property of the commutator remains true in this more general framework. This fact appears now as a deep result which involves the classification of finite simple groups.

Theorem 3.6 ([4]) *Let the group $G = AB$ be the product of totally completely c-permutable subgroups A and B. Then $[A, B] \leq F(G)$.*

For a totally permutable product $G = AB$, P. Hauck et al. in [27] proved that $[A, B] \leq Z_{\mathcal{U}}(G)$, where $Z_{\mathcal{U}}(G)$ denotes the $\mathcal{U}$-hypercentre of G (the largest normal subgroup of G such that every chief factor X/Y of G with $Y < X \leq Z_{\mathcal{U}}(G)$ is cyclic of prime order). Example 3.4 shows that this property is missed when permutability is weakened to complete conditional permutability, because in that example $Z_{\mathcal{U}}(G) = 1$.

Theorem 3.6 allows us however to derive the following centralizing property for totally completely c-permutable products:

Corollary 3.7 ([4]) *Let the group $G = AB$ be the product of totally completely c-permutable subgroups A and B. Then:*

(*i*) *If A is a normal subgroup of G, then B acts u-hypercentrally on A by conjugation* (see [24, IV. 6.2]). *In particular, $B^{\mathcal{U}}$ centralizes A.*

(*ii*) $[A^{\mathcal{U}}, B^{\mathcal{U}}] = 1$.

Another significant consequence of Theorem 3.6 has been the description of the structure of a monolithic primitive group which is a product of totally completely c-permutable subgroups. Corollary 3.8 and Lemma 3.9 below play an important role in the study of this kind of products and saturated formations presented in Section 4.

Corollary 3.8 ([4]) *Let the group $G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$, and $G_i \neq 1$ for all $i = 1, \ldots, r$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Let N be a minimal normal subgroup of G. Then:*

1. *If N is non-abelian, then there exists a unique $i \in \{1, \ldots, r\}$ such that $N \leq G_i$. Moreover, G_j centralizes N and $N \cap G_j = 1$ for all $j \in \{1, \ldots, r\}$, $j \neq i$.*
2. *If G is a monolithic primitive group, then the unique minimal normal subgroup N is abelian.*

Lemma 3.9 ([4]) *Let the group $1 \neq G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Assume in addition that G is a primitive group of type* 1 (see [24, A. 15.2]). *Let N be the unique minimal normal subgroup of G and p be a prime divisor of $|N|$. Then either G is supersoluble or the following conditions are satisfied:*

(*i*) *w.l.o.g. $N \leq G_1$;*

(*ii*) *$G_2 \cdots G_r$ is a cyclic group whose order divides $p - 1$;*

(*iii*) *there exists a maximal subgroup M of G with $\mathrm{Core}_G(M) = 1$ such that $M = (M \cap G_1)(G_2 \cdots G_r)$ and $M \cap G_1$ centralizes $G_2 \cdots G_r$.*

4 Products of groups and formations

Motivated by the previous research on products of totally permutable subgroups and formations (see [10, 11, 13, 14]) it is natural to ask whether analogous results can be achieved by weakening permutability to complete c-permutability. A first approach to this study for products of totally completely c-permutable subgroups and saturated formations of soluble groups containing $\mathcal{U}$ was carried out in [5]. The better knowledge of the structure of such products has allowed to extend the results in that paper to the non-soluble universe. Nevertheless, we give examples showing that the hypothesis of saturation for the formations involved can not be removed.

Theorem 4.1 ([4]) *Let $\mathcal{F}$ be a saturated formation containing $\mathcal{U}$. Let the group $G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Then:*

1. *If $G_i \in \mathcal{F}$ for all $i = 1, \ldots, r$, then $G \in \mathcal{F}$.*
2. *If $G \in \mathcal{F}$, then $G_i \in \mathcal{F}$ for all $i = 1, \ldots, r$.*

Part (1) was first proved for totally permutable products of subgroups and saturated formations $\mathcal{F}$ such that $\mathcal{U} \subseteq \mathcal{F}$ in [33] (for $r = 2$) and in [22] (for r arbitrary). Later on this result was also extended to non-saturated formations which contain $\mathcal{U}$ (see [11, 14]). In [10, 11, 14] the authors showed that the converse holds whenever $\mathcal{F}$ is a formation containing $\mathcal{U}$ such that either $\mathcal{F}$ is saturated or $\mathcal{F} \subseteq \mathcal{S}$, where $\mathcal{S}$ is the class of all soluble groups.

A stronger version of the previous theorem by means of $\mathcal{F}$-residuals can be stated as follows:

Corollary 4.2 ([4]) *Let $\mathcal{F}$ be a saturated formation containing $\mathcal{U}$. Let the group $G = G_1 \cdots G_r$ be the product of pairwise permutable subgroups $G_1, \ldots, G_r$, for $r \geq 2$. Assume that G_i and G_j are totally completely c-permutable subgroups for all $i, j \in \{1, \ldots, r\}$, $i \neq j$. Then:*

1. *$G_i^{\mathcal{F}} \trianglelefteq G$ for all $i = 1, \ldots, r$.*
2. *$G^{\mathcal{F}} = G_1^{\mathcal{F}} \cdots G_r^{\mathcal{F}}$.*

For totally permutable products Corollary 4.2(2) is also true if $\mathcal{F}$ is any formation of soluble groups containing $\mathcal{U}$, and part (1) is verified for $\mathcal{F}$ a formation such that $\mathcal{U} \subseteq \mathcal{F}$ (see [11]).

The following examples show that none of the statements in Theorem 4.1 remains true for arbitrary non-saturated formations containing $\mathcal{U}$, even in the universe of soluble groups:

Example 4.3 ([4]) We consider the set of all prime numbers $\mathbb{P}$ and define a mapping $f : \mathbb{P} \longrightarrow \{\text{classes of groups}\}$ by setting $f(5) = (1, Z_2, Z_4, Z_3)$ and $f(p)$ to be the class of abelian groups of exponent dividing $p - 1$ for all $p \neq 5$. Let $\mathcal{F}$ be the class of all soluble groups G such that $\mathrm{Aut}_G(S) \in f(p)$ for all p-chief factors S of G and for all primes p dividing the order of G. By [24, IV. 1.3] it follows that $\mathcal{F}$ is a formation of soluble groups, and clearly also $\mathcal{U} \subseteq \mathcal{F}$.

Now we consider again Example 3.4: Let $V = \langle a, b \rangle \cong Z_5 \times Z_5$ and $Z_6 \cong C = \langle \alpha, \beta \rangle \leq \mathrm{Aut}(V)$ given by

$$a^\alpha = a^{-1},\ b^\alpha = b^{-1}; \quad a^\beta = b,\ b^\beta = a^{-1}b^{-1}.$$

Let $G = [V]C$ be the corresponding semidirect product of V with C. Set $A = \langle \alpha \rangle$ and $B = V\langle \beta \rangle$. Then $G = AB$ is the product of totally completely c-permutable subgroups A and B. Observe that A and B are $\mathcal{F}$-groups. But $G \notin \mathcal{F}$, because $G/C_G(V) \cong Z_3 \times Z_2 \notin f(5)$. This shows that Theorem 4.1(1) is not valid if the formation under consideration is not assumed to be saturated.

We modify the construction of the formation $\mathcal{F}$ by considering $f(5) = (1, Z_2, Z_4, Z_6)$. It holds now that $G, A \in \mathcal{F}$ but $B \notin \mathcal{F}$ because $B/C_B(V) \cong Z_3 \notin f(5)$, which shows the necessity for the formation to be saturated in order to prove Theorem 4.1(2).

The behavior of $\mathcal{F}$-projectors, when $\mathcal{F}$ is a saturated formation of soluble groups containing $\mathcal{U}$, in products of totally completely c-permutable subgroups, studied in [5] can be also extended to the non-soluble universe.

Corollary 4.4 ([4]) *Let $\mathcal{F}$ be a saturated formation containing $\mathcal{U}$. Let the group $G = AB$ be the product of totally completely c-permutable subgroups A and B. Then there exist $\mathcal{F}$-projectors X of A and Y of B such that X is permutable with Y. In this case XY is an $\mathcal{F}$-projector of G.*

As stated in [4] it is an open question whether the above result can be extended to an arbitrary finite number of pairwise permutable factors. We mention that a positive answer holds for totally permutable products (see [11]).

On the other hand, in the spirit of Corollary 4.2, for mutually permutable products with nilpotent derived subgroup, the original result of M. Asaad and A. Shalaan [6], mentioned in the Introduction, can be generalized in the following sense: for a saturated formation $\mathcal{F}$ containing $\mathcal{U}$ and a group $G = G_1 \cdots G_r$ which is a pairwise mutually permutable product of the factors and with nilpotent derived subgroup, it holds that $G^{\mathcal{F}} = G_1^{\mathcal{F}} \cdots G_r^{\mathcal{F}}$ ([8, Theorem 5.2.23]).

In [19] (see also [8, Theorem 4.5.8]) it was proved that if $G = AB$ is the mutually permutable product of A and B, $\mathrm{Core}_G(A \cap B) = 1$ and $\mathcal{F}$ is a saturated formation containing $\mathcal{U}$, then $G^{\mathcal{F}} = A^{\mathcal{F}} B^{\mathcal{F}}$. This result appears at once as an extension of a previous result in [1] which states that a mutually permutable product of supersoluble subgroups A and B such that $\mathrm{Core}_G(A \cap B) = 1$ is supersoluble.

Nevertheless the situation is different in relation with $\mathcal{F}$-projectors, for saturated formations $\mathcal{F}$ (containing $\mathcal{U}$), and mutually permutable products. An example in [9] shows a mutually permutable product $G = AB$ with G' nilpotent and such that A and B possess corresponding $\mathcal{U}$-projectors whose product is a subgroup but not a $\mathcal{U}$-projector of G.

Acknowledgements This research has been supported by Proyecto MTM2010-19938-C03-02, Ministerio de Economia and Competitividad, Spain.

References

[1] M. J. Alejandre, A. Ballester-Bolinches and J. Cossey, Permutable products of supersoluble groups, *J. Algebra* **276** (2004), 453–461.

[2] M. Arroyo-Jordá, P. Arroyo-Jordá, A. Martínez-Pastor and M. D. Pérez-Ramos, On finite products of groups and supersolubility, *J. Algebra* **323** (2010), 2922–2934.

[3] M. Arroyo-Jordá, P. Arroyo-Jordá, A. Martínez-Pastor and M. D. Pérez-Ramos, in *Proceedings of the Meeting on Group Theory and its Applications*, Biblioteca de la Revista Matemática Iberoamerícana, (Madrid 2012), 1–11.

[4] M. Arroyo-Jordá, P. Arroyo-Jordá, A. Martínez-Pastor and M. D. Pérez-Ramos, On conditional permutability and factorized groups, *Annali di Matematica Pura ed Applicata*, 2013, DOI 10.1007/S10231-012-0319-1.

[5] M. Arroyo-Jordá, P. Arroyo-Jordá and M. D. Pérez-Ramos, On conditional permutability and saturated formations, *Proc. Edinburgh Math. Soc.* **54** (2011), 309–319.

[6] M. Asaad and A. Shaalan, On the supersolvability of finite groups, *Arch. Math.* **53** (1989), 318–326.

[7] A. Ballester-Bolinches, J. Cossey and M. C. Pedraza-Aguilera, On mutually permutable products of finite groups, *J. Algebra* **294** (2005), 127–135.

[8] A. Ballester-Bolinches, R. Esteban-Romero and M. Asaad, *Products of finite groups*, Expositions in Mathematics. **53** (De Gruyter, Berlin-New York 2010).

[9] A. Ballester-Bolinches and M. C. Pedraza-Aguilera, Mutually permutable products of finite groups II, *J. Algebra* **218** (1999), 563–572.

[10] A. Ballester-Bolinches, M. C Pedraza-Aguilera and M. D. Pérez-Ramos, On finite products of totally permutable groups, *Bull. Austral. Math. Soc.* **53** (1996), 441–445.
[11] A. Ballester-Bolinches, M. C. Pedraza-Aguilera and M. D Pérez-Ramos, Finite groups which are products of pairwise totally permutable subgroups, *Proc. Edinburgh Math. Soc.* **41** (1998), 567–572.
[12] A. Ballester-Bolinches, M. C. Pedraza-Aguilera and M. D. Pérez-Ramos, Mutually permutable products of finite groups, *J. Algebra* **213** (1999), 369–377.
[13] A. Ballester-Bolinches, M. C Pedraza-Aguilera and M. D. Pérez-Ramos, Totally and mutually permutable products of finite groups, in *Groups St. Andrews 1997 in Bath Vol. 1* (C. M. Campbell et al., eds.), London Math. Soc. Lecture Note Ser. **260** (CUP, Cambridge 1999), 65–68.
[14] A. Ballester-Bolinches and M. D. Pérez-Ramos, A question of R. Maier concerning formations, *J. Algebra* **182** (1996), 738–747.
[15] J. Beidleman and H. Heineken, Totally permutable torsion subgroups, *J. Group Theory* **2** (1999), 377–392.
[16] J. Beidleman and H. Heineken, Mutually permutable subgroups and group classes, *Arch. Math.* **85** (2005), 18–30.
[17] J. Beidleman and H. Heineken, Group classes and mutually permutable products, *J. Algebra* **297** (2006), 409–416.
[18] J. Bochtler, Radicals in mutually permutable products of finite groups, *J. Algebra* **321** (2009), 353–360.
[19] J. Bochtler, *Wechselseitig vertauschbare Produkte endlicher Gruppen*, Ph. D. thesis, Fakultät für Mathematik und Physik, Eberhand-Karls-Universität Tübingen, Tübingen, 2010.
[20] J. Bochtler and P. Hauck, Mutually permutable subgroups and Fitting classes, *Arch. Math.* **88** (2007), 385–388.
[21] A. Carocca, p-supersolubility of factorized finite groups, *Hokkaido Math. J.* **21** (1992), 395–403.
[22] A. Carocca, A note on the product of $\mathcal{F}$-subgroups in a finite group, *Proc. Edinburgh Math. Soc.* **39** (1996), 37–42.
[23] A. Carocca and R. Maier, Theorems of Kegel-Wielandt type, in *Groups St. Andrews 1997 in Bath Vol. 1* (C. M. Campbell et al., eds.), London Math. Soc. Lecture Note Ser. **260** (CUP, Cambridge 1999), 195–201.
[24] K. Doerk and T. Hawkes, *Finite Soluble Groups*, Expositions in Mathematics. **4** (Walter de Gruyter, Berlin-New York, 1992).
[25] W. Guo, K. P. Shum and A. N. Skiba, Conditionally permutable subgoups and supersolubility of finite groups, *Southeast Asian Bull. Math.* **29** (2005), 493–510.
[26] W. Guo, K. P. Shum and A. N. Skiba, Criterions of supersolubility for products of supersoluble groups, *Publ. Math. Debrecen* **68** (2006), 433–449.
[27] P. Hauck, A. Martínez-Pastor and M. D. Pérez-Ramos, Fitting classes and products of totally permutable groups, *J. Algebra* **252** (2002), 114–126.
[28] P. Hauck, A. Martínez-Pastor and M. D. Pérez-Ramos, Products of pairwise totally permutable groups, *Proc. Edinburgh. Math. Soc.* **46** (2003), 147–157.
[29] P. Hauck, A. Martínez-Pastor and M. D. Pérez-Ramos, Injectors and radicals in products of totally permutable groups, *Comm. Algebra* **31** (2003), 6135–6147.
[30] H. Li, and G. Qian, On PCM-subgroups of finite groups, *JP J. Algebra Number Theory Appl.* **12** (2008), 83–91.
[31] X. Liu, W. Guo and K. P Shum, Products of finite supersoluble groups, *Algebra Colloq.* **16** (2009), 333–340.
[32] X. Liu , B. Li and X. Yi, Some criteria for supersolubility in products of finite groups, *Front. Math. China* **3** (2008), 79–86.
[33] R. Maier, A completeness property of certain formations, *Bull. London Math. Soc.* **24**

(1992), 540–544.

[34] G. Qian and P. Zhu, Some sufficient conditions for supersolvability of groups, *J. Nanjing Normal Univ. (Nature Science)* **21** (1998), 15–21 (in Chinese).

A SURVEY ON THE NORMALIZER PROBLEM FOR INTEGRAL GROUP RINGS

ANDREAS BÄCHLE

Vrije Universiteit Brussel, DWIS, Pleinlaan 2, B-1050 Brussels, Belgium
Email: ABachle@vub.ac.be

Abstract

We give a survey of the normalizer problem for integral group rings. This question asks whether the normalizer of the group basis in the unit group of the group ring only contains the "obvious" units. It played an important role when M. Hertweck provided a counterexample for a long-standing conjecture, namely the isomorphism problem for integral group rings, which asks whether a finite group is determined by its integral group ring. We also give a quick account on the subgroup normalizer problem, which is a variation of the "classical" normalizer problem, where arbitrary subgroups of a group basis are considered.

1 Basic definitions and motivations

Let G be a (possibly infinite) group, R be a commutative ring with identity element and RG be the group ring of G with coefficients in R. We denote by $\mathrm{U}(RG)$ the group of units of the group ring RG. Clearly, when considering G as a subgroup of the group of units of RG, it is normalized by all elements of G and by central units, but it turned out that in many cases equality holds, i.e.,

$$\mathrm{N}_{\mathrm{U}(RG)}(G) = G \cdot \mathrm{Z}(\mathrm{U}(RG)). \tag{NP}$$

In this case we say that G has the *normalizer property* for the coefficient ring R, or (NP) holds for RG. Note that this property strongly depends on the coefficient ring R. Some authors say that a group has the normalizer property if (NP) is satisfied for all G-adapted coefficient rings R (cf. next section for a definition). The problem to decide whether a given group has the normalizer property is often referred to as *normalizer problem*. The question whether the normalizer property holds for all finite groups was raised by S. Jackowski and Z. Marciniak in [13, Question 3.7] and also made its way in the collection of important research questions in the book of S. Sehgal [24, Problem 43].

The normalizer problem gained popularity when M. Mazur established a connection to the long standing question of the isomorphism problem, namely whether a finite group is determined by the corresponding group ring, i.e., if for finite groups X and Y the implication

$$\mathbb{Z}X \simeq \mathbb{Z}Y \quad \Rightarrow \quad X \simeq Y \tag{IP}$$

holds. M. Mazur proved in [18] the following theorem.

Theorem 1.1 (Mazur) *Let G be a group and α_0, $\beta_0 \in \mathrm{Aut}(G)$. Consider the homomorphisms α, $\beta\colon C_\infty = \langle x\rangle \to \mathrm{Aut}(G)$ determined by $x \mapsto \alpha_0$ and $x \mapsto \beta_0$, respectively. Then*

1. *$G \rtimes_\alpha C_\infty \simeq G \rtimes_\beta C_\infty$ if and only if $\alpha_0\beta_0^{-1}$ is an inner automorphism of G.*
2. *$R[G \rtimes_\alpha C_\infty] \simeq R[G \rtimes_\beta C_\infty]$ if and only if $\alpha_0\beta_0^{-1}$ is an inner ring automorphism of RG.*

This connected the isomorphism problem (for infinite groups) to the question whether group automorphisms may become inner when considered as automorphism of the group ring (cf. next section). This was used by K. W. Roggenkamp and A. Zimmermann together with their discovery that outer group automorphisms become inner in a particular semi-local group ring [22] to find a counterexample to the isomorphism problem for (infinite) polycyclic groups over such rings [23]. Later, M. Hertweck constructed in his PhD thesis an example of a finite group not satisfying the normalizer property and cleverly adapted the just mentioned idea of M. Mazur to finally settle the isomorphism problem in the negative. He proved in [12, Theorem A, Theorem B], cf. also [8, Theorem A, Theorem B], the following.

Theorem 1.2 (Hertweck) *There is a metabelian group G of order $2^{25} \cdot 97^2$ such that* (NP) *does not hold for $\mathbb{Z}G$.*

Theorem 1.3 (Hertweck) *There are two non-isomorphic solvable groups X and Y of order $2^{21} \cdot 97^{28}$ with isomorphic integral group rings.*

For finite groups and the coefficient ring $\mathbb{Z}$ these are up to date the only known counterexamples to (NP) and (IP).

Our notation is mostly standard. For group elements x and y we put $x^y = y^{-1}xy$ and $[x, y] = x^{-1}y^{-1}xy$. By $[X, Y]$ we denote the group generated by all $[x, y]$ for all $x \in X,\ y \in Y$. For a subset $X \subseteq G$ the subgroup of G generated by X is denoted by $\langle X \rangle$. (If one of the sets in the previous situations is a singleton we will omit the set braces.) $\mathrm{Z}(G)$ denotes the center of the group G, $\mathrm{C}_G(X)$ and $\mathrm{N}_G(X)$ the centralizer and the normalizer of the subset X in the group G, respectively. By C_n we denote a cyclic group of order n, where $n \in \mathbb{N} \cup \{\infty\}$.

2 Some tools

The normalizer property can be restated in terms of automorphism groups. Denote by $\mathrm{Inn}(G)$ the group of inner automorphisms of G, i.e., the automorphisms induced by conjugation of elements of G and by $\mathrm{Aut}_{RG}(G)$ the group of automorphisms of G induced by units of RG normalizing G.

Proposition 2.1 *Let G be a group G and R a commutative ring with* 1*. Then the following are equivalent:*

1. (NP) *holds for RG.*
2. $\mathrm{Aut}_{RG}(G) = \mathrm{Inn}(G)$.
3. *For every $u \in \mathrm{N}_{\mathrm{U}(RG)}(G)$ there exists $g \in G$ such that $[gu, G] = 1$.*

The group $\mathrm{Aut}_{RG}(G)$ can be "bounded" by other groups which are defined in purely group-theoretical terms. For an element $x \in G$ denote by x^G its conjugacy class in G. Put $\mathrm{Aut}_c(G) = \{\varphi \in \mathrm{Aut}(G) \mid \forall\, x \in G : x\varphi \in x^G\}$, the group of *class-preserving automorphisms*. If all conjugacy classes are finite it can be easily seen

that $\mathrm{Aut}_{RG}(G) \leq \mathrm{Aut}_c(G)$ using class sums. The statement is also true for arbitrary groups, cf. [10, Theorem 17.3].

H. N. Ward [25] and independently D. Coleman [4] proved that if P is a p-subgroup of G for some rational prime p, then $\mathrm{N}_{\mathrm{U}(RG)}(P) = \mathrm{N}_G(P) \cdot \mathrm{C}_{\mathrm{U}(RG)}(P)$, provided the rational prime p is not invertible in R (this result is nowadays known as *Coleman lemma*). Let G be a group and R an integral domain of characteristic 0, then R is called *G-adapted* if, whenever there exists an element of order a rational prime p in G, then p is not invertible in R. Call an automorphism of a finite group G a *Coleman automorphism* if its restriction to any Sylow subgroup coincides with the restriction of an inner automorphism of G, and let $\mathrm{Aut}_{\mathrm{Col}}(G)$ be the group of all such automorphisms. Then the Coleman lemma shows that, for a finite group G and a G-adapted ring R, $\mathrm{Aut}_{RG}(G) \leq \mathrm{Aut}_{\mathrm{Col}}(G)$.

The group of class-preserving automorphisms has been a subject of study for more than a century. W. Burnside already asked in his textbook [2, Note B] "Does there exist any finite group G such that G has a non-inner class preserving automorphism?" and provided in 1913 several examples of such groups [3], namely the groups of lower unitriangular 3×3-matrices over fields with p^2 elements for primes $p \equiv \pm 3 \bmod 8$. Since then there were more examples examined having non-inner class-preserving automorphisms, cf. for example the survey article of M. Yadav [26]. On the other hand there were results ensuring that in certain cases the group $\mathrm{Aut}_c(G)$ is as small as possible, i.e., coincides with $\mathrm{Inn}(G)$. For example W. Feit and G. M. Seitz [5, Theorem C] showed, using the classification of finite simple groups, that $\mathrm{Aut}_c(G) = \mathrm{Inn}(G)$ for all finite simple groups, settling the normalizer problem at once for these groups. Also M. Hertweck showed that all elements of $\mathrm{Aut}_c(G)$ are inner automorphism if G is a finite group having an abelian normal subgroup such that the corresponding quotient is cyclic [10, Proposition 14.4].

A result by J. Krempa [13, Theorem 3.2] guarantees that the quotient group $\mathrm{Aut}_{\mathbb{Z}G}(G)/\mathrm{Inn}(G)$ is an elementary abelian 2-group. His proof makes use of the anti-involution * of the group ring RG induced by the anti-automorphism $g \mapsto g^{-1}$ of the group G, having for the coefficient ring $\mathbb{Z}$ the striking property that for a unit $u \in \mathbb{Z}G$, $uu^* = 1 \Leftrightarrow u \in \pm G$. This was generalized by M. Mazur in [19, Corollary 14] to the case where the coefficient ring is the ring of algebraic integers in a number field K such that the complex conjugation is central in the Galois group of the normal closure of K over $\mathbb{Q}$. Thus in this case it is enough to show that the group $(\mathrm{Aut}_{\mathrm{Col}}(G) \cap \mathrm{Aut}_c(G))/\mathrm{Inn}(G)$ is a $2'$-group to verify the normalizer property. It was observed by S. Jackowski and Z. Marciniak that the only rational primes that could occur as divisors of $|\mathrm{Aut}_c(G)|$ are the primes dividing $|G|$. M. Hertweck and W. Kimmerle showed that the corresponding statement holds true for the prime divisors of $|\mathrm{Aut}_{\mathrm{Col}}(G)|$. For details cf. [11, Proposition 1]. Using GAP [6] one could verify (NP) for groups G of order at most 161 and G-adapted rings R by calculating their groups of class-preserving and Coleman automorphisms.

3 Finite groups having the normalizer property

In the following theorems we list important classes of finite groups for which the normalizer property is known to be true.

Theorem 3.1 *Let G be a finite group. Then* (NP) *holds for $\mathbb{Z}G$ provided one of the following is true:*

1. *G has a normal Sylow 2-subgroup (Jackowski, Marciniak [13, Theorem 3.6]).*
2. *G is metabelian with abelian Sylow 2-subgroup (Marciniak, Roggenkamp [17, Proposition 12.3]).*
3. *G is metabelian, $A \trianglelefteq G$ is abelian, G/A is abelian and has a cyclic Sylow 2-subgroup (Li [15, Theorem 2.17]).*
4. *G is a Blackburn group, i.e., the intersection of all non-normal subgroups of G is non-trivial (Li, Parmenter, Sehgal [16, Theorem 1]).*
5. *G is a Frobenius group (Petit Lobão, Polcino Milies [21, Theorem 3.1]).*

Theorem 3.2 *Let G be a finite group and R a G-adapted ring. Then* (NP) *holds for RG provided one of the following is true:*

1. *G is quasi-nilpotent (Hertweck, Kimmerle [11, Corollary 16]).*
2. *G is solvable and no chief factor of $G/\mathrm{O}_2(G)$ is of order 2 (here $\mathrm{O}_2(G)$ denotes the largest normal 2-subgroup of G). (Hertweck, Kimmerle [11, Corollary 20]).*

4 Infinite groups having the normalizer property

When the investigation of the normalizer problem for infinite groups began, it turned out that most of the relevant information is already encoded in the so-called support subgroup of a unit in question, which is finite in many important cases; we give the relevant definitions. For an element $u = \sum_{g \in G} u_g g \in RG$ the set $\mathrm{supp}(u) = \{g \in G \mid u_g \neq 0\}$ of the elements of G where u has non-zero coefficients is called the *support of u*. The support subgroup of u is the subgroup $\langle \mathrm{supp}(u) \rangle$ of G generated by the elements of the support of u. For a group X let

$$\Delta(X) = \{x \in X \mid [X : \mathrm{C}_X(x)] < \infty\}$$

be the set of elements having finite conjugacy classes. This set is in fact a characteristic subgroup of X, the *FC-center*. (For details see [20, Chapter 4].) In [19, Corollary 1] it is proved by M. Mazur that $\langle \mathrm{supp}(u) \rangle$ is a normal subgroup of G contained in $\Delta(G)$ for $u \in \mathrm{N}_{\mathrm{U}(RG)}(G)$ provided $1 \in \mathrm{supp}(u)$ (the latter can always be arranged by multiplication by a group element). In [10, Theorem 18.5] M. Hertweck proved that in this situation $\langle \mathrm{supp}(u) \rangle$ is finite, if the ring R is $\Delta(G)$-adapted.

We will now list important classes of (infinite) groups for which the normalizer property is known.

Theorem 4.1 *Let G be a group. Then* (NP) *holds for $\mathbb{Z}G$ provided one of the following is true:*

1. *G has an abelian subgroup of index 2 (Li, Parmenter, Sehgal [16, Theorem 2]).*
2. *$\Delta(G)$ has no (non-trivial) 2-torsion elements (Jespers, Juriaans, de Miranda, Rogerio [14, Theorem 2.1]).*
3. *G is a torsion group and the 2-elements of G form a normal subgroup (Jespers, Juriaans, de Miranda, Rogerio [14, Theorem 2.2]).*

4. *G is locally nilpotent (Jespers, Juriaans, de Miranda, Rogerio [14, Theorem 2.4]).*
5. *All finite normal subgroups of G have a normal Sylow 2-subgroup (Hertweck [10, Corollary 19.11]).*
6. *G is a Blackburn group, i.e., it contains finite non-normal subgroups and the intersection of all such subgroups is non-trivial (Hertweck, Jespers, [9, Theorem 3.3]).*

For more general coefficient rings the following results were obtained:

Theorem 4.2 *Let G be a group. Then* (NP) *holds for RG provided one of the following is true:*

1. *G is nilpotent and R is $\Delta(G)$-adapted (Hertweck [10, Corollary 19.13]).*
2. *All finite normal subgroups of G have a normal p-subgroup containing its own centralizer in G, p is not invertible in R (Hertweck [10, Corollary 19.15]).*
3. *G is a locally finite Frobenius group and R is $\Delta(G)$-adapted (Hertweck [10, Corollary 19.17]).*

5 The normalizer of subgroups

When considering a subgroup $H \leq G$ of a group basis G one sees that it is normalized by all elements of $\mathrm{N}_G(H)$ and all units of RG centralizing H. Also in this situation one might ask if products of those "obvious" units are the only normalizing units, i.e., if

$$\mathrm{N}_{\mathrm{U}(RG)}(H) = \mathrm{N}_G(H) \cdot \mathrm{C}_{\mathrm{U}(RG)}(H) \qquad (\mathrm{NP},\ H \leq G)$$

holds.

The Coleman lemma is the verification of (NP, $H \leq G$) in the case that H is a p-group. As that lemma turned out to be quite useful in proofs of theorems dealing with units of integral group rings, it seems to be reasonable to investigate for which other groups results of this kind could be obtained. Note that the above question can be seen from two different standpoints. Either fix an isomorphism type of a group H. Does (NP, $H \leq G$) hold for all groups G in which H embeds? Or it can be seen as a question on G: Fix a group G and consider (NP, $H \leq G$) for all subgroups H of G. If (NP, $H \leq G$) is true for all $H \leq G$ we will say that G has the *subgroup normalizer property*, (SNP) for short. Clearly, (SNP) for G implies (NP) for G. We will report on results of [1].

Using a similar calculation as in the simplified proof of I. B. S. Passi for [10, Theorem 17.3] or in the proof of [21, Theorem 3.1] the following result was obtained [1, Proposition 3.23].

Proposition 5.1 *Let $H \leq G$. If H is cyclic, then* (NP, $H \leq G$) *holds for arbitrary rings R.*

Proof For an element $u = \sum_{g \in G} u_g g \in RG$ ($u_g \in R$) and a group element $x \in G$ denote by $\varepsilon_x(u) = \sum_{g \in x^G} u_g$ the partial augmentation of u at the conjugacy class of x. The map ε_x is R-linear and satisfies $\varepsilon_x(uv) = \varepsilon_x(vu)$ for all $u, v \in RG$. Denote by

$[RG, RG]_L$ the R-submodule of RG generated by all additive commutators $[x, y]_L = xy - yx$ for $x, y \in RG$. Note that $u \in [RG, RG]_L \Leftrightarrow \forall\, x \in G : \varepsilon_x(u) = 0$. Now let $H = \langle h \rangle$ and $u \in \mathrm{N}_{\mathrm{U}(RG)}(H)$. Then

$$u^{-1}hu - h = [u^{-1}h, u]_L \in [RG, RG]_L$$

and consequently $\varepsilon_x(u^{-1}hu) = \varepsilon_x(h)$ for all $x \in G$. As $u^{-1}hu \in H \leq G$ this implies $u^{-1}hu \in h^G$. Hence $u = gz$ for some $g \in \mathrm{N}_G(H)$ and $z \in \mathrm{C}_{\mathrm{U}(RG)}(H)$. □

In [1, Lemma 3.18] the following extension of [10, Lemma 19.4] for subgroups H of the group basis G was obtained:

Lemma 5.2 (Coleman lemma, relative version) *Let $H \leq G$ and R be a commutative ring with identity element. Let p be a rational prime which is not invertible in R. Let $u \in \mathrm{N}_{\mathrm{U}(RG)}(H)$. Then there exists $P \leq H$ with $|H : P| < \infty$, $p \nmid |H : P|$ and $x \in \mathrm{supp}(u) \cap \mathrm{N}_G(P)$ such that $x^{-1}u \in \mathrm{C}_{\mathrm{U}(RG)}(P)$.*

Using this lemma and a reduction to direct factors one can prove that the normalizer of subgroups in the unit group is as small as possible for a class of groups containing finite nilpotent groups:

Theorem 5.3 (Bächle, [1, Theorem 3.26, Theorem 3.42]) *Let G be a group. Then* (SNP) *holds for RG provided*

1. *G is a locally nilpotent torsion group and R is a G-adapted ring.*
2. *G admits a short exact sequence of the form $1 \to C_m \to G \to C_n \to 1$ with $m, n \in \mathbb{N} \cup \{\infty\}$, where m and n are coprime natural numbers or one of the two is a rational prime, and R is any commutative ring with* 1.

This implies that (SNP) holds for all dihedral groups or finite groups of square-free order and all coefficient rings. Also all finite subgroups of $\mathrm{O}(3, \mathbb{R})$, the three-dimensional orthogonal group over the reals, and all groups of order at most 47 satisfy (SNP), cf. [1, Corollary 3.48, Proposition 3.54].

When dealing with the subgroup version of the normalizer problem one is lead to the group $\mathrm{Aut}_{RG}(H)$ of automorphisms of H induced by units $u \in \mathrm{N}_{\mathrm{U}(RG)}(H)$. One can check that (NP, $H \leq G$) holds if and only if $\mathrm{Aut}_{RG}(H) = \mathrm{Inn}(H)$. But in contrast to the "classical" normalizer problem, where the group $\mathrm{Aut}_{RG}(G)$ is contained in $\mathrm{Aut}_c(G)$ and $\mathrm{Aut}_{\mathrm{Col}}(G)$, the groups $\mathrm{Aut}_c(H)$ and $\mathrm{Aut}_{\mathrm{Col}}(H)$ do not always contain $\mathrm{Aut}_{RG}(H)$. Here, requiring certain properties just for one prime seems to be fruitful as was done for the normalizer property in [11]. Let p be a rational prime. An automorphism φ of a finite group G is called *p-central* if there exists a Sylow p-subgroup P of G such that the restriction of φ to P is the identity. Employing the work of [7] and [11] on p-central automorphisms one obtains as application the following two results. If $H \trianglelefteq G$ and H is a finite simple group, then (NP, $H \leq G$) holds for RG if R is H-adapted [1, Corollary 3.38]. If H is a p-constrained group with $\mathrm{O}_{p'}(H) = 1$ for some prime p, and $H \trianglelefteq G$ or $H = \mathrm{N}_G(P)$ for a Sylow p-subgroup P of G, then (NP, $H \leq G$) holds for RG if p is not invertible in R (here $\mathrm{O}_{p'}(H)$ denotes the largest normal subgroup of H whose order is coprime to p) [1, Corollary 3.39]. Also for infinite simple groups H (or more general groups where all finite quotients

are p-groups for a fixed prime p) the property (NP, $H \leq G$) holds true for RG if p is not invertible in R [1, Proposition 3.19].

References

[1] A. Bächle, *On torsion subgroups and their normalizers in integral group rings*, PhD thesis, Universität Stuttgart, 2012, http://elib.uni-stuttgart.de/opus/volltexte/2013/7887/ (Last visited: June 30, 2013).

[2] W. Burnside, *Theory of groups of finite order*, 2nd Ed., Dover Publications, Inc., 1955. Reprint of the 2nd edition (Cambridge, 1911).

[3] W. Burnside, On the outer automorphisms of a group, *Proc. London Math. Soc. (2)* **11** (1913), 40–42.

[4] D. B. Coleman, On the modular group ring of a p-group, *Proc. Amer. Math. Soc.* **15** (1964), 511–514.

[5] W. Feit and G. M. Seitz, On finite rational groups and related topics, *Illinois J. Math.* **33** (1989), 103–131.

[6] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.6.4*; 2013, http://www.gap-system.org.

[7] F. Gross, Automorphisms which centralize a Sylow p-subgroup, *J. Algebra* **77** (1982), 202–233.

[8] M. Hertweck, A counterexample to the isomorphism problem for integral group rings, *Ann. of Math.* **154** (2001), 115–138.

[9] M. Hertweck and E. Jespers, Class-preserving automorphisms and the normalizer property for Blackburn groups, *J. Group Theory* **12** (2009), 157–169.

[10] M. Hertweck, *Contributions to the integral representation theory of groups*, Habilitationsschrift, Universität Stuttgart, 2004, http://elib.uni-stuttgart.de/opus/volltexte/2004/1638/ (Last visited: June 26, 2013).

[11] M. Hertweck and W. Kimmerle, Coleman automorphisms of finite groups, *Mat. Z.* **242** (2002), 203–215.

[12] M. Hertweck, *Eine Lösung des Isomorphieproblems für ganzzahlige Gruppenringen von endlichen Gruppen*, PhD thesis, Universität Stuttgart, 1998.

[13] S. Jackowski and Z. Marciniak, Group automorphisms inducing the identity map on cohomology, *J. Pure Appl. Algebra* **44** (1987), 241–250.

[14] E. Jespers, S. O. Juriaans, J. M. de Miranda, and J. R. Rogerio, On the Normalizer Problem, *J. Algebra* **247** (2002), 24–36.

[15] Y. Li, The normalizer of a metabelian group in its integral group ring, *J. Algebra* **256** (2002), 343–351.

[16] Y. Li, M. M. Parmenter, and S. Sehgal, On the Normalizer Property for Integral Group Rings, *Comm. Algebra* **27** (1999), 4217–4223.

[17] Z. Marciniak and K. W. Roggenkamp, The normalizer of a finite group in its integral group ring and Cech cohomology, *Algebra – Representation Theory: Proceedings of the NATO Advanced Study Institute* (Klaus W. Roggenkamp and Mirela Ştefănescu, eds.), NATO Science Series II: Mathematics, Physics and Chemistry, 2001.

[18] M. Mazur, On the isomorphism problem for integral group rings of infinite groups, *Expo. Math.* **13** (1995), 433–445.

[19] M. Mazur, The normalizer of a group in the unit group of its group ring, *J. Algebra* **212** (1999), 175–189.

[20] D. S. Passman, *The algebraic structure of group rings*, Wiley-Interscience, 1977.

[21] T. Petit Lobão and C. Polcino Milies, The normalizer property for integral group rings of Frobenius groups, *J. Algebra* **256** (2002), 1–6.

[22] K. W. Roggenkamp and A. Zimmermann, Outer group automorphisms may become inner in the integral group ring, *J. Pure Appl. Algebra* **103** (1995), 91–99.

[23] K. W. Roggenkamp and A. Zimmermann, A counterexample for the isomorphism-problem

of polycyclic groups, *J. Pure Appl. Algebra* **103** (1995), 101–103.

[24] S. Sehgal, *Units in integral group rings*, Pitman Monographs and Surveys in Pure and Applied Mathematics, Wiley, 1993.

[25] H. N. Ward, Some results on the group algebra of a group over a prime field, *Seminar in Group Theory, Harvard University*, Mimeographed Notes, (1960–1961), 13–19.

[26] M. K. Yadav, Class preserving automorphisms of finite p-groups: A survey, *Groups St Andrews 2009 in Bath, Vol. 2*, 569–579, London Math. Soc. Lecture Note Ser. 388, CUP, 2011.

A SURVEY ON CLIFFORD-FISCHER THEORY

AYOUB B. M. BASHEER* and JAMSHID MOORI†

*School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal (Pietermaritzburg), P Bag X01, Scottsville 3209, South Africa
Email: ayoubbasheer@gmail.com

†School of Mathematical Sciences, North-West University (Mafikeng), P Bag X2046, Mmabatho 2735, South Africa
Email: jamshid.moori@nwu.ac.za

Abstract

Bernd Fischer presented a powerful and interesting technique, known as *Clifford-Fischer theory*, for calculating the character tables of group extensions. This technique derives its fundamentals from the Clifford theory. The present article surveys the developments of Clifford-Fischer theory applied to group extensions (split and non-split) and in particular we focus on the contributions of the second author and his research groups including students.

1 Introduction

The character table of a finite group is a very powerful tool to study the group structure and to prove many results. Any finite group is either simple or has a non-trivial normal subgroup and hence will be of extension type (non-trivial). The classification of finite simple groups, more recent work in group theory, has been completed in 1985 and since then the researchers concentrated on the generation, subgroup structures of the finite simple groups and their automorphism groups. Few also studied the interplay between finite simple groups and combinatorial structures. A knowledge of the character table of a finite group G provides considerable information about G and hence it is of importance in the Physical Sciences as well as in Pure Mathematics. Character tables of finite groups can be constructed using various theoretical and computational techniques. The character tables of all the maximal subgroups of the sporadic simple groups are known, except for some maximal subgroups of the Monster $\mathbb{M}$ and the Baby Monster $\mathbb{B}$. There are several well-developed methods for calculating the character tables of group extensions and in particular when the kernel of the extension is an elementary abelian group. For example, the Schreier-Sims algorithm, the Todd-Coxeter coset enumeration method, the Burnside-Dixon algorithm and various other techniques. Bernd Fischer [18, 19, 20] presented a powerful and interesting technique for calculating the character tables of group extensions. This technique, which is known as *Clifford-Fischer matrices*, derives its fundamentals from the Clifford theory [17]. Let $\overline{G} = N{\cdot}G$, where $N \lhd \overline{G}$ and $\overline{G}/N \cong G$, be a finite group extension. If we know generators or a presentation of $\overline{G}$ and we are only interested in the calculation of the character table, then it could be computed by using Magma [16] or GAP [21] provided $\overline{G}$ is of a reasonable size. But Clifford-Fischer theory provides many other interesting information on the group and on the character table, in partic-

ular a character table produced by Clifford-Fischer theory is in a special format that could not be achieved by direct computations using GAP or Magma. Also applying Clifford-Fischer theory to both split and non-split extensions is making sense, since each group requires individual approach. The readers (particulary young researchers) will highly benefit from the theoretical background required for these computations. GAP and Magma are computational tools and would not replace good powerful and theoretical arguments. The following systematic steps show how to find the conjugacy classes and the character table of a group extension via the *coset analysis* technique and Clifford-Fischer theory.

The first step in constructing the character table of any finite group is to find its conjugacy classes. The basic idea of the coset analysis technique is to consider for each conjugacy class $[g_i]_G$, one coset $N\overline{g_i}$, where $\overline{g_i}$ is a pre-image of g_i in $\overline{G}$. Then we act N (by conjugation) on the coset $N\overline{g_i}$, followed by the action of $\overline{G}$ on the resulting orbits of the action of N on $N\overline{g_i}$. Corresponding to each class $[g_i]_G$, we construct a number of conjugacy classes of $\overline{G}$ (we denote the number of $\overline{G}$-classes correspond to $[g_i]_G$ by $c(g_i)$). That is each conjugacy class of $\overline{G}$ corresponds uniquely to a conjugacy class of G and hence the coset analysis organizes the $\overline{G}$-classes into cosets corresponding to the G-classes representatives.

The group $\overline{G}$ has dual action on the conjugacy classes of N and on $\mathrm{Irr}(N)$ and Brauer Theorem (see Theorem 5.1.5 of Mpono [26] for example) asserts that the number of orbits on the two actions is the same. Please note that orbits lengths of the two actions may be different. Indeed if N is non-abelian, then

$$\sum_{k=1}^{t} |{\theta_k}^{\overline{G}}| = |\mathrm{Irr}(N)| \neq |N| = \sum_{k=1}^{t} |[n_k]_N^{\overline{G}}|,$$

where

- t is the number of orbits on the action of $\overline{G}$ on the conjugacy classes of N or on $\mathrm{Irr}(N)$,
- θ_k and n_k are respective class representatives of characters and conjugacy classes of N,
- ${\theta_k}^{\overline{G}}$ and $[n_k]_N^{\overline{G}}$ are orbits of N containing θ_k and n_k, respectively, on the action of $\overline{G}$ on $\mathrm{Irr}(N)$ and on the conjugacy classes of N respectively.

For a representative character $\theta_k,\ k \in \{1, 2, \cdots, t\}$, we refer to the stabilizer of θ_k in $\overline{G}$ by the *inertia* group, denoted by $\overline{H}_k$. Note that $\overline{H}_k \leq \overline{G},\ \forall k$. Each inertia group contains N normally and the quotient group is referred to as the *inertia factor* group, denoted by H_k. Note that $H_k \leq G,\ \forall k$. Now for the characters $\theta_k,\ k \in \{1, 2, \cdots, t\}$, two cases are distinguished. Either all θ_k are extendable to ordinary characters of their respective inertia groups $\overline{H}_k,\ 1 \leq k \leq t$, or some are non-extendable. In the case of extendability of every character θ_k, the set of irreducible characters of $\overline{G}$ is given by

$$\mathrm{Irr}(\overline{G}) = \bigcup_{k=1}^{t} \left\{ (\psi_k \inf(\zeta))\uparrow_{\overline{H}_k}^{\overline{G}} |\ \zeta \in \mathrm{Irr}(\overline{H}_k/N) \right\}. \tag{1}$$

If for some $k \in \{1, 2, \cdots, t\}$, the character θ_k, is not extendable to an ordinary character of $\overline{H}_k$, then it is extendable to a projective character $\widetilde{\psi}_k$ of $\overline{H}_k$ with some

factor set $\overline{\alpha}_k^{-1}$ of the Schur multiplier of $\overline{H}_k$. Thus a more proper formula for Eq. (1) is given by

$$\mathrm{Irr}(\overline{G}) = \bigcup_{k=1}^{t} \left\{ (\widetilde{\psi}_k \inf(\zeta))\uparrow_{\overline{H}_k}^{\overline{G}} \mid \widetilde{\psi}_k \in \mathrm{IrrProj}(\overline{H}_k, \overline{\alpha}_k^{-1}),\ \zeta \in \mathrm{IrrProj}(\overline{H}_k/N, \alpha_k^{-1}) \right\}, \tag{2}$$

where the factor set α_k is obtained from $\overline{\alpha}_k$ as described in Corollary 7.3.3 of Whitely [40]. In fact this is method Clifford theory for obtaining the $\mathrm{Irr}(\overline{G})$. From Eq. (2) we can see that the character table of $\overline{G}$ is partitioned into t blocks $\mathcal{K}_1, \mathcal{K}_2, \cdots, \mathcal{K}_t$, where each block $\mathcal{K}_k$ of characters (ordinary or projective) is produced from the inertia group $\overline{H}_k$.

Note that to calculate the character table of $\overline{G}$ through Clifford theory (using Eq. (2)), we need to deal with the character tables (ordinary or projective) of the inertia groups. In practise we do not attempt to compute the character table of $\overline{H}_k$, simply because the character tables of these inertia groups are usually much larger and more complicated to compute than the character table of $\overline{G}$ itself. Bernd Fischer suggested to use the character tables of the inertia factor groups H_k together with some matrices, called by him *Clifford matrices* (throughout this paper we refer to them as *Fischer matrices*), to construct the character table of $\overline{G}$. Calculating the character table of $\overline{G}$ in this way is known as the Clifford-Fischer theory. Thus to apply Clifford-Fischer Theory we firstly need to determine the structures and the character tables (ordinary or projective) of all the inertia factors H_k together with the Fischer matrices. One of the biggest challenges in Clifford-Fischer theory is the determination of the type of the character table of H_k (projective or ordinary), which is to be used in the construction of the character table of $\overline{G}$. We may firstly assume that all the irreducible characters of N are extendible to their respective inertia groups and consequently all the character tables of the H_k that we need to use are the ordinary ones. However, in general, there is no reason guaranteeing that one can work with the ordinary characters of $H_k,\ 2 \leq k \leq t$ (see Section 5.1 of Basheer [8] for some partial results on extendability of characters). Thus in practice making the right choice of the appropriate projective character table of H_k, with factor set α_k, might be difficult unless the Schur multipliers of all the H_k are trivial. Otherwise there will be many combinations (for each H_k, there are many projective character tables associated with different factor sets of the Schur multiplier of H_k) and one has to test all the possible choices and eliminate the choices that lead to contradictions. Sometimes Eq. (5.7) of Basheer [8] might also be useful to prove that we need to use projective characters of some of the inertia factors of $\overline{G}$.

Once we have determined for each inertia factor group H_k, the appropriate projective character table with factor set α_k, the next step in our construction will be the determination of the fusions of the α_k-regular classes of H_k into classes of G.

The next step in our construction is the computations of the Fischer matrices. These are non-singular square unique (up to the permutations of rows and columns) associated to the conjugacy classes $[g_i]_G$, and denoted by $\mathcal{F}_i$. These matrices satisfy several interesting properties and certain orthogonality relations. There is an interesting interplay between the coset analysis and Clifford-Fischer Theory. Indeed the size of each Fischer matrix is $c(g_i)$, the number of $\overline{G}$-classes corresponding to $[g_i]_G$

obtained via the coset analysis technique. That is computations of the conjugacy classes of $\overline{G}$ using the coset analysis technique will determine the sizes of all Fischer matrices. Note that in some cases we have used information on the sizes of some of the Fischer matrices (which in turn were determined through the coset analysis) to help in determining the inertia factor groups (see the proof of Proposition 10.3.2 of [8]). Also we attach many information (obtained through the coset analysis) to the top and bottom of each Fischer matrix and these information are useful in the computations of the entries of Fischer matrices.

In this survey article we go briefly over the technique of the coset analysis and Clifford-Fischer theory applied to both split and non-split group extensions. The second author has a significant contribution to this domain. Indeed he developed the coset analysis technique in his PhD thesis [24] and in [25]. Then together with his MSc and PhD students (including the first author), they enriched this area of research by applying the above mentioned techniques to many various split and non-split group extensions in a considerable number of publications. For example, but not limited to, one can refer to [6], [9], [10], [11], [12], [13], [14], [15], [28], [29], [32], [33], [34], [35], [36], [38] or [40]. Barraclough produced an interesting PhD thesis [7], which contained a chapter on the method of Clifford-Fischer theory. He used this method to find the character table of any group of the form $2^{2\cdot}G{:}2$ for any finite group G. Also in 2007, H. Pahlings [37] calculated the Fischer matrices and the character table of the non-split extension $2_+^{1+22}\cdot Co_2$, which is the second largest maximal subgroup of the Baby Monster group $\mathbb{B}$. Then in 2010, H. Pahlings together with his student K. Lux published an interesting book [22] containing a full chapter on Clifford-Fischer theory that includes several examples on the application of the method.

2 Conjugacy Classes of Group Extensions

In the following we give a shortened description on how the coset analysis can be used to determine the conjugacy classes of any group extension.

For each $g \in G$ let $\overline{g} \in \overline{G}$ map to g under the natural epimorphism $\pi : \overline{G} \longrightarrow G$ and let $g_1 = N\overline{g}_1, g_2 = N\overline{g}_2, \cdots, g_r = N\overline{g}_r$ be representatives for the conjugacy classes of $G \cong \overline{G}/N$. Therefore $\overline{g}_i \in \overline{G}$, $\forall i$, and by convention we take $\overline{g}_1 = 1_{\overline{G}}$. The method of the coset analysis constructs for each conjugacy class $[g_i]_G$, $1 \leq i \leq r$, a number of conjugacy classes of $\overline{G}$. That is each conjugacy class of $\overline{G}$ corresponds uniquely to a conjugacy class of G. This method can be described briefly in the following steps:

- For fixed $i \in \{1, 2, \cdots, r\}$, act N (by conjugation) on the coset $N\overline{g}_i$ and let the resulting orbits be $Q_{i1}, Q_{i2}, \cdots, Q_{ik_i}$. If N is abelian (regardless to whether the extension is split or not), then $|Q_{i1}| = |Q_{i2}| = \cdots = |Q_{ik_i}| = |N|/k_i$.
- Act $\overline{G}$ on $Q_{i1}, Q_{i2}, \cdots, Q_{ik_i}$ and suppose f_{ij} orbits fuse together to form a new orbit Δ_{ij} and let the total number of the new resulting orbits in this action be $c(g_i)$ (that is $1 \leq j \leq c(g_i)$). Then $\overline{G}$ has a conjugacy class $[g_{ij}]_{\overline{G}}$ that contains Δ_{ij} and $|[g_{ij}]_{\overline{G}}| = |[g_i]_G| \times |\Delta_{ij}|$.
- Repeat the above two steps, for all $i \in \{1, 2, \cdots, r\}$.

Lemma 2.1 *For each* $i \in \{1, 2 \cdots, r\}$, *write* $g_i = N\overline{g}_i = \bigcup_{j=1}^{c(g_i)} \left(N\overline{g}_i \cap [g_{ij}]_{\overline{G}}\right) =$

$\bigcup_{j=1}^{c(g_i)} \Delta_{ij}$. *Then* $\{\overline{g}_{i1}, \overline{g}_{i2}, \cdots, \overline{g}_{ic(g_i)}\}$ *is a complete set of representatives for the conjugacy classes of* $\overline{G}$ *that correspond (under the natural epimorphism) to* $[g_i]_G$.

Proof One can refer to Barraclough [7] with slight difference in notations. □

Thus each $[g_i]_G$ affords $c(g_i)$ conjugacy classes in $\overline{G}$.

Remark 2.2 For fixed $i \in \{1, 2, \cdots, r\}$, the conjugacy class $[g_{ij}]_{\overline{G}}$ is partitioned into $|[g_i]_G|$ equal size subsets $\Delta_{ij1}, \Delta_{ij2}, \cdots, \Delta_{ij|[g_i]_G|}$, where $|\Delta_{iju}| = |\Delta_{ij}|$, for each $1 \leq u \leq |[g_i]_G|$ (we can take $\Delta_{ij1} = \Delta_{ij}$). Moreover, for fixed i and $s \in \{1, 2, \cdots, |[g_i]_G|\}$, the relation $\sum_{j=1}^{c(g_i)} |\Delta_{ijs}| = |N|$ holds. If the extension splits, then Δ_{i1s} is the intersection of $[g_{ij}]_{\overline{G}}$ with an element of $[g_i]_G$, for all $1 \leq s \leq |[g_i]_G|$.

Therefore information about every conjugacy class of $\overline{G}$ can be obtained by examining one coset $N\overline{g}_i = g_i \in G$ for each conjugacy class of G. The following two propositions relate the orders of the elements of $\overline{G}$ with those of G.

Proposition 2.3 *Let* $\overline{G} = N{:}G$, *where* N *is an abelian group. Also let* $\overline{G} \ni \overline{g} = ng$, *for some* $n \in N$ *and* $g \in G$. *Then* $o(g) \mid o(\overline{g})$.

Proof Let $o(\overline{g})$ and $o(g)$ be k and m respectively. We have $1_{\overline{G}} = \overline{g}^k = (ng)^k = nn^g n^{g^2} n^{g^3} \cdots n^{g^{k-1}} g^k$. Since G acts on N, we have $n, n^g, n^{g^2}, n^{g^3}, \cdots n^{g^{k-1}} \in N$ and therefore $nn^g n^{g^2} n^{g^3} \cdots n^{g^{k-1}} \in N$. Now since $N \cap G = \{1_{\overline{G}}\}$ and $nn^g n^{g^2} n^{g^3} \cdots n^{g^{k-1}}$ $g^k = 1_{\overline{G}}$, we must have $nn^g n^{g^2} n^{g^3} \cdots n^{g^{k-1}}$ and g^k equal to 1_N and 1_G respectively. Hence $m \mid k$. □

Proposition 2.4 *With the settings of* Proposition 2.3 *and its proof, assume further that* N *is an elementary abelian* p*-group. Then* $k \in \{m, pm\}$.

Proof See Mpono [26, 28]. □

Further results on the conjugacy classes of $\overline{G} = N{\cdot}G$, when N is abelian or the extension splits, can be found in many sources such as Ali [1, 6], Barraclough [7], Moori [24, 25], Mpono [26, 28], Rodrigues [38] or Whitely [40].

3 The Theory of Clifford-Fischer Matrices

We give a brief description on Clifford-Fischer theory for constructing the character table of a group extension $\overline{G}$.

Let $\overline{H} \trianglelefteq \overline{G}$ and let $\phi \in \mathrm{Irr}(\overline{H})$. For $\overline{g} \in \overline{G}$, define $\phi^{\overline{g}}$ by $\phi^{\overline{g}}(h) = \phi(\overline{g}h\overline{g}^{-1})$, for $h \in \overline{H}$. It follows that $\overline{G}$ acts on $\mathrm{Irr}(\overline{H})$ by conjugation and we define the *inertia group* of ϕ in $\overline{G}$ by $\overline{H}_\phi = \{\, \overline{g} \in \overline{G} \mid \phi^{\overline{g}} = \phi \,\}$. Also for a finite group $\mathcal{K}$, we let $\mathrm{IrrProj}(\mathcal{K}, \alpha^{-1})$ denotes the set of irreducible projective characters of $\mathcal{K}$ with factor set α^{-1}.

Theorem 3.1 (Clifford Theorem) *Let* $\chi \in \mathrm{Irr}(\overline{G})$ *and let* $\theta_1, \theta_2, \cdots, \theta_t$ *be representatives of orbits of* $\overline{G}$ *on* $\mathrm{Irr}(N)$. *For* $k \in \{1, 2, \cdots, t\}$, *let*

$$\theta_k^{\overline{G}} = \{\theta_k = \theta_{k1}, \theta_{k2}, \cdots, \theta_{ks_k}\}$$

and let $\overline{H}_k$ be the inertia group in $\overline{G}$ of θ_k. Then

$$\chi{\downarrow}_N^{\overline{G}} = \sum_{k=1}^{t} e_k \sum_{u=1}^{s_k} \theta_{ku}, \quad \textit{where} \quad e_k = \langle \chi{\downarrow}_N^{\overline{G}}, \theta_k \rangle.$$

Moreover, for fixed k

$$\begin{aligned} \mathrm{Irr}(\overline{H}_k, \theta_k) &:= \left\{ \psi_k \in \mathrm{Irr}(\overline{H}_k) \mid \langle \psi_k {\downarrow}_N^{\overline{H}_k}, \theta_k \rangle \neq 0 \right\} \\ &\leftrightarrow \left\{ \chi \in \mathrm{Irr}(\overline{G}) \mid \langle \chi{\downarrow}_N^{\overline{G}}, \theta_k \rangle \neq 0 \right\} := \mathrm{Irr}(\overline{G}, \theta_k) \end{aligned}$$

under the map $\psi_k \longmapsto \psi_k{\uparrow}_{\overline{H}_k}^{\overline{G}}$.

Proof See Theorems 4.1.5 and 4.1.7 of Ali [1] with the difference in notations. □

Theorem 3.2 *Further to the settings of Theorem 3.1, assume that for $k \in \{1, 2, \cdots, t\}$, there exists $\psi_k \in \mathrm{Irr}(\overline{H}_k, \theta_k)$. Then the irreducible characters of $\overline{G}$ are given by* Eq. (1).

Proof See Ali [1] or Whitley [40]. □

Remark 3.3 As mentioned in Section 1, it is by no means necessarily the case that there exists an extension ψ_k of θ_k to the inertia group (that is the case $\mathrm{Irr}(\overline{H}_k, \theta_k) = \emptyset$, the empty set, is feasible). However, there is always a *projective* extension $\widetilde{\psi_k} \in \mathrm{IrrProj}(\overline{H}_k, \overline{\alpha}_k^{-1})$ for some factor set $\overline{\alpha}_k$ of the Schur multiplier of $\overline{H}_k$ and Eq. (2) becomes the more proper formula for Equation (1) (see Remark 4.2.7 of Ali [1])

Note 3.4 Observe that if $\alpha_k \sim [1]$ in Equation (2), then we get Equation (1). That is $\mathrm{IrrProj}(\overline{H}_k, \overline{1}) = \mathrm{Irr}(\overline{H}_k)$ and $\mathrm{IrrProj}(H_k, 1) = \mathrm{Irr}(H_k)$. By convention we take $\theta_1 = \mathbf{1}_N$, the trivial character of N. Thus $\overline{H}_{\theta_1} = \overline{H}_1 = \overline{G}$ and thus $\overline{H}_1/N \cong G$. Since $\{\mathbf{1}_{\overline{G}}\} \subseteq \mathrm{Irr}(\overline{G}, \mathbf{1}_N)$ and such that $\mathbf{1}_{\overline{G}}{\downarrow}_N^{\overline{G}} = \mathbf{1}_N$, the block $\mathcal{K}_1$ will consists only of the ordinary irreducible characters of G.

We now fix some notations for the conjugacy classes.

- With π being the natural epimorphism from $\overline{G}$ onto G, we use the notation $U = \pi(\overline{U})$ for any subset $\overline{U} \subseteq \overline{G}$. We have seen from Section 2 that $\pi^{-1}([g_i]_G) = \bigcup_{j=1}^{c(g_i)} [g_{ij}]_{\overline{G}}$ for any $1 \leq i \leq r$. Let us assume that $\pi(g_{ij}) = g_i$ and by convention we may take $g_{11} = 1_{\overline{G}}$. Note that $c(g_1)$ is the number of $\overline{G}$-conjugacy classes obtained from N.
- $[g_{ij}]_{\overline{G}} \cap \overline{H}_k = \bigcup_{n=1}^{c(g_{ijk})} [g_{ijkn}]_{\overline{H}_k}$, where $g_{ijkn} \in \overline{H}_k$ and by $c(g_{ijk})$ we mean the number of $\overline{H}_k$-conjugacy classes that form a partition for $[g_{ij}]_{\overline{G}}$. Since $g_{11} = 1_{\overline{G}}$, we have $g_{11k1} = 1_{\overline{G}}$ and thus $c(g_{11k1}) = 1$ for all $1 \leq k \leq t$.
- $[g_i]_G \cap H_k = \bigcup_{m=1}^{c(g_{ik})} [g_{ikm}]_{H_k}$, where $g_{ikm} \in H_k$ and by $c(g_{ik})$ we mean the number of H_k-conjugacy classes that form a partition for $[g_i]_G$. Since $g_1 = 1_G$, we have $g_{1k1} = 1_G$ and thus $c(g_{1k1}) = 1$ for all $1 \leq k \leq t$. Also $\pi(g_{ijkn}) = g_{ikm}$ for some $m = f(j, n)$.

Proposition 3.5 *With the notations of Theorem 3.2 and the above settings, we have*

$$(\widetilde{\psi}_k \inf(\zeta))\uparrow_{\overline{H}_k}^{\overline{G}}(g_{ij}) = \sum_{m=1}^{c(g_{ik})} \zeta(g_{ikm}) \sum_{n=1}^{c(g_{ijk})} \frac{|C_{\overline{G}}(g_{ij})|}{|C_{\overline{H}_k}(g_{ijkn})|} \widetilde{\psi}_k(g_{ijkn}).$$

Proof See Ali [1] or Barraclough [7]. □

We proceed to define the Fischer matrix $\mathcal{F}_i$ corresponds to the conjugacy class $[g_i]_G$. We label the columns of $\mathcal{F}_i$ by the representatives of $[g_{ij}]_{\overline{G}}$, $1 \leq j \leq c(g_i)$ obtained by the coset analysis and below each g_{ij} we put $|C_{\overline{G}}(g_{ij})|$. Thus there are $c(g_i)$ columns. To label the rows of $\mathcal{F}_i$ we define the set $\overline{J}_i$ to be (this equivalent to the notation $R(g)$ used by Ali [1] (page 49), where g is a representative for a conjugacy class of G)

$$\overline{J}_i = \{(k, g_{ikm}) |\ 1 \leq k \leq t,\ 1 \leq m \leq c(g_{ik}),\ g_{ikm} \text{ is } \alpha_k^{-1}\text{-regular class}\},$$

or for more brevity we let

$$J_i = \{(k, m) |\ 1 \leq k \leq t,\ 1 \leq m \leq c(g_{ik}),\ g_{ikm} \text{ is } \alpha_k^{-1}\text{-regular class}\}. \tag{3}$$

Then each row of $\mathcal{F}_i$ is indexed by a pair $(k, g_{ikm}) \in \overline{J}_i$ or $(k, m) \in J_i$. For fixed $1 \leq k \leq t$, we let $\mathcal{F}_{ik}$ be a sub-matrix of $\mathcal{F}_i$ with rows correspond to the pairs $(k, g_{ik1}), (k, g_{ik2}), \cdots, (k, g_{ikr_{ik}})$ or for brevity $(k, 1), (k, 2), \cdots, (k, r_k)$. Now let

$$a_{ij}^{(k,m)} := \sum_{n=1}^{c(g_{ijk})} \frac{|C_{\overline{G}}(g_{ij})|}{|C_{\overline{H}_k}(g_{ijkn})|} \widetilde{\psi}_k(g_{ijkn}) \tag{4}$$

(for which $\pi(g_{ijkn}) = g_{ikm}$). For each i, corresponding to the conjugacy class $[g_i]_G$, we define the Fischer matrix $\mathcal{F}_i = \left(a_{ij}^{(k,m)}\right)$, where $1 \leq k \leq t,\ 1 \leq m \leq c(g_{ik}),\ 1 \leq j \leq c(g_i)$. The Fischer matrix $\mathcal{F}_i$,

$$\mathcal{F}_i = \left(a_{ij}^{(k,m)}\right) = \left(\begin{array}{c} \mathcal{F}_{i1} \\ \hline \mathcal{F}_{i2} \\ \hline \vdots \\ \hline \mathcal{F}_{it} \end{array}\right)$$

together with additional information required for their definition are presented in Table 1 below. In this table the last entries give the weights m_{ij} are defined by

$$m_{ij} = [N_{\overline{G}}(N\overline{g}_i) : C_{\overline{G}}(g_{ij})] = |N| \frac{|C_G(g_i)|}{|C_{\overline{G}}(g_{ij})|}. \tag{5}$$

These weights are required for computing the entries of $\mathcal{F}_i$ (see Proposition 3.6).

Fischer matrices satisfy some interesting properties, which help in computations of their entries. We gather these properties in the following Proposition.

Proposition 3.6 (i) $\sum_{k=1}^{t} c(g_{ik}) = c(g_i)$,

(ii) $\mathcal{F}_i$ *is non-singular for each* i,

Table 1.

$\mathcal{F}_i$

g_i		g_{i1}	g_{i2}	$\cdots$	$g_{ic(g_i)}$
$\lvert C_{\overline{G}}(g_{ij})\rvert$		$\lvert C_{\overline{G}}(g_{i1})\rvert$	$\lvert C_{\overline{G}}(g_{i2})\rvert$	$\cdots$	$\lvert C_{\overline{G}}(g_{ic(g_i)})\rvert$
(k,m)	$\lvert C_{H_k}(g_{ikm})\rvert$				
$(1,1)$	$\lvert C_G(g_i)\rvert$	$a_{i1}^{(1,1)}$	$a_{i2}^{(1,1)}$	$\cdots$	$a_{ic(g_i)}^{(1,1)}$
$(2,1)$	$\lvert C_{H_2}(g_{i21})\rvert$	$a_{i1}^{(2,1)}$	$a_{i2}^{(2,1)}$	$\cdots$	$a_{ic(g_i)}^{(2,1)}$
$(2,2)$	$\lvert C_{H_2}(g_{i22})\rvert$	$a_{i1}^{(2,2)}$	$a_{i2}^{(2,2)}$	$\cdots$	$a_{ic(g_i)}^{(2,2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(2,r_2)$	$\lvert C_{H_2}(g_{i2r_{i2}})\rvert$	$a_{i1}^{(2,r_2)}$	$a_{i2}^{(2,r_2)}$	$\cdots$	$a_{ic(g_i)}^{(2,r_2)}$
$(u,1)$	$\lvert C_{H_u}(g_{iu1})\rvert$	$a_{i1}^{(u,1)}$	$a_{i2}^{(u,1)}$	$\cdots$	$a_{ic(g_i)}^{(u,1)}$
$(u,2)$	$\lvert C_{H_u}(g_{iu2})\rvert$	$a_{i1}^{(u,2)}$	$a_{i2}^{(u,2)}$	$\cdots$	$a_{ic(g_i)}^{(u,2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
(u,r_u)	$\lvert C_{H_u}(g_{iur_{iu}})\rvert$	$a_{i1}^{(u,r_u)}$	$a_{i2}^{(u,r_u)}$	$\cdots$	$a_{ic(g_i)}^{(u,r_u)}$
$(t,1)$	$\lvert C_{H_t}(g_{it1})\rvert$	$a_{i1}^{(t,1)}$	$a_{i2}^{(t,1)}$	$\cdots$	$a_{ic(g_i)}^{(t,1)}$
$(t,2)$	$\lvert C_{H_t}(g_{it2})\rvert$	$a_{i1}^{(t,2)}$	$a_{i2}^{(t,2)}$	$\cdots$	$a_{ic(g_i)}^{(t,2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
(t,r_t)	$\lvert C_{H_t}(g_{itr_{it}})\rvert$	$a_{i1}^{(t,r_t)}$	$a_{i2}^{(t,r_t)}$	$\cdots$	$a_{ic(g_i)}^{(t,r_t)}$
m_{ij}		m_{i1}	m_{i2}	$\cdots$	$m_{ic(g_i)}$

(iii) $a_{ij}^{(1,1)} = 1,\ \forall\ 1 \le j \le c(g_i)$,

(iv) *If $N\overline{g}_i$ is a split coset, then $a_{i1}^{(k,m)} = |C_G(g_i)|/|C_{H_k}(g_{ikm})|$, for $i \in \{1, 2, \cdots, r\}$. In particular for the identity coset we have $a_{11}^{(k,m)} = [G : H_k]\theta_k(1_N)$, for $(k,m) \in J_1$,*

(v) *If $N\overline{g}_i$ is a split coset, then $|a_{ij}^{(k,m)}| \le |a_{i1}^{(k,m)}|$ for all $1 \le j \le c(g_i)$. Moreover if $|N| = p^\alpha$, for some prime p, then $a_{ij}^{(k,m)} \equiv a_{i1}^{(k,m)} (\mathrm{mod}\ p)$,*

(vi) *For each $1 \le i \le r$, the weights m_{ij} satisfy the relation $\sum_{j=1}^{c(g_i)} m_{ij} = |N|$,*

(vii) ***Column Orthogonality Relation:***

$$\sum_{(k,m)\in J_i} |C_{H_k}(g_{ikm})| a_{ij}^{(k,m)} \overline{a_{ij'}^{(k,m)}} = \delta_{jj'} |C_{\overline{G}}(g_{ij})|,$$

(viii) ***Row Orthogonality Relation:***

$$\sum_{j=1}^{c(g_i)} m_{ij} a_{ij}^{(k,m)} \overline{a_{ij}^{(k',m')}} = \delta_{(k,m)(k',m')} a_{i1}^{(k,m)} |N|.$$

Proof Proofs for many assertions of Proposition 3.6 can be founded in Moori's students theses, for example see Ali [1] or Mpono [26] and some other assertions

are provided in Schiffer [39] as well as in Basheer and Moori [9, 10] and Lux and Pahlings [22]. □

4 Character Tables of Group Extensions

Let $\overline{G} = N{\cdot}G$, where $N \lhd \overline{G}$ and $\overline{G}/N \cong G$, be a group extension. To construct the character table of $\overline{G}$ in the format of Clifford-Fischer theory we need to have

- the conjugacy classes of $\overline{G}$ obtained though the coset analysis method,
- the character tables (ordinary or projective) of the inertia factor groups,
- the fusions of classes of the inertia factors into classes of G,
- the Fischer matrices of $\overline{G} = N{\cdot}G$.

For fixed $1 \leq k \leq t$ and $1 \leq i \leq r$, let $\mathcal{K}_{ik}$ be the fragment of the projective character table of H_k, with factor set α_k^{-1}, consisting of columns correspond to the conjugacy classes $g_{ik1}, g_{ik2}, \cdots, g_{ikr_{ik}}$ of H_k (those are the α_k^{-1}-regular classes of H_k that fuse to $[g_i]_G$ and thus $r_{ik} = c(g_{ik})$). Then the characters of $\overline{G}$ on the classes $[g_{ij}]_{\overline{G}}$, $1 \leq j \leq c(g_i)$, is given by the matrix $\mathcal{K}_{ik}\mathcal{F}_{ik}$, where $\mathcal{F}_{ik}$ is the sub-matrix of $\mathcal{F}_i$ defined previously with rows correspond to the pairs $(k, g_{ik1}), (k, g_{ik2}), \cdots, (k, g_{ikr_{ik}})$. Note that the size of $\mathcal{K}_{ik}$ is $|\mathrm{IrrProj}(H_k, \alpha_k^{-1})| \times r_{ik}$ and the size of $\mathcal{F}_{ik}$ is $r_{ik} \times c(g_i)$. Therefore the character table of $\overline{G}$ will have the form

	g_1			g_2			$\cdots$	g_r		
	g_{11}	g_{12}	$\cdots$ $g_{1c(g_1)}$	g_{21}	g_{22}	$\cdots$ $g_{2c(g_2)}$	$\cdots$	g_{r1}	g_{r2}	$\cdots$ $g_{rc(g_r)}$
$\mathcal{K}_1$	$\mathcal{K}_{11}\mathcal{F}_{11}$			$\mathcal{K}_{12}\mathcal{F}_{12}$			$\cdots$	$\mathcal{K}_{1r}\mathcal{F}_{1r}$		
$\mathcal{K}_2$	$\mathcal{K}_{21}\mathcal{F}_{21}$			$\mathcal{K}_{22}\mathcal{F}_{22}$			$\cdots$	$\mathcal{K}_{2r}\mathcal{F}_{2r}$		
$\vdots$	$\vdots$			$\vdots$			$\ddots$	$\vdots$		
$\mathcal{K}_t$	$\mathcal{K}_{t1}\mathcal{F}_{t1}$			$\mathcal{K}_{t2}\mathcal{F}_{t2}$			$\cdots$	$\mathcal{K}_{tr}\mathcal{F}_{tr}$		

Note 4.1 From Note 3.4 we know that characters of $\overline{G}$ consisted in $\mathcal{K}_1$ are just $\mathrm{Irr}(G)$ and therefore the size of $\mathcal{K}_{1i}\mathcal{F}_{1i}$, for each $1 \leq i \leq r$, is $|\mathrm{Irr}(G)| \times c(g_i)$. In particular, columns of $\mathcal{K}_{11}\mathcal{F}_{11}$ are the degrees of irreducible characters of G repeated themselves $c(g_1)$ times, where we know that $c(g_1)$ is number of $\overline{G}$-conjugacy classes obtained from the normal subgroup N.

5 Two Examples

Here we give two examples on the applications of Clifford-Fischer Theory to split and non-split group extensions. These examples are fully discussed in the PhD thesis of the first author.

5.1 On the Split Extension Group 2_-^{1+6}:$((3^{1+2}$:8):2)

In [15] we calculated the inertia factors, Fischer matrices and the ordinary character table of the split extension 2^{10}:$(U_5(2)$:2) by means of Clifford-Fischer Theory. The second inertia factor group of 2^{10}:$(U_5(2)$:2) is a split extension group of the form 2_-^{1+6}:$((3^{1+2}$:8):2) := $\overline{G}$. The group $\overline{G}$ is a maximal subgroup, of index 3, in

$2_-^{1+6}{:}3_-^{1+2}{:}2S_4$, which in turn is the second largest maximal subgroup of the automorphism group of the unitary group $U_5(2)$. In [14] we used the coset analysis to compute the conjugacy classes of $\overline{G}$. Corresponding to the 14 classes of $G = (3^{1+2}{:}8){:}2$, we obtain 41 conjugacy classes for $\overline{G}$. For example the group G has two classes of involutions represented by 2_1 and 2_2 with respective centralizer sizes 48 and 12. Corresponding to $[2_2]_G$ we get five conjugacy classes in $\overline{G}$ with information listed in Table 2.

Table 2. Few conjugacy classes of $\overline{G}$ obtained via the coset analysis method

$[g_i]_G$	k_i	m_{ij}	$[g_{ij}]_{\overline{G}}$	$o(g_{ij})$	$\lvert[g_{ij}]_{\overline{G}}\rvert$	$\lvert C_{\overline{G}}(g_{ij})\rvert$
$g_3 = 2_2$	$k_3 = 9$	$m_{31} = 8$	g_{31}	8	288	192
		$m_{32} = 8$	g_{32}	8	288	192
		$m_{33} = 24$	g_{33}	2	576	96
		$m_{34} = 48$	g_{34}	8	1728	32
		$m_{35} = 48$	g_{35}	4	1728	32

Following [14] the action of $\overline{G}$ on N produced four orbits of lengths 1, 1, 54 and 72 and it follows that the action of $\overline{G}$ on $\mathrm{Irr}(N)$ will also produce four orbits of characters. Through various theoretical and computational aspects we were able to determine the structures of the inertia factor groups. These are the groups $H_1 = H_2 = (3^{1+2}{:}8){:}2$, $H_3 = QD_{16}$ and $H_4 = D_{12}$, where QD_{16} and D_{12} are the quasihedral and dihedral groups of orders 16 and 12 respectively. The determination of these inertia factors included the computations of the Schur multipliers of some of these groups, some computations with GAP on the structures of the maximal subgroups of G and some results on extendability of characters such as Theorem 5.1.18 of Mpono [26].

For the Fischer matrices of $\overline{G}$ we have used the arithmetical properties of the Fischer matrices, given by Proposition 3.6, to calculate some of the entries of these matrices. In addition to these properties, we established in [14] further properties that the Fischer matrices of $\overline{G}$ satisfy. These additional properties are given by Lemmas 10.3.3, 10.3.4 and Note 10.3.1 of [8]. These properties helped in reducing the number of unknowns in every Fischer matrix of size $c(g_i)$ to $c(g_i)^2 - 4c(g_i) + 4$.

Using the row and column orthogonality relations given by Proposition 3.6 we have built an algebraic system of equations. With the help of the symbolic mathematical package Maxima [23], we were able to solve these systems of equations and hence we have computed all the Fischer matrices of $\overline{G}$ which listed in Section 4 of [14]. As an example below we give, in Table 3, the Fischer matrix $\mathcal{F}_3$ corresponding to class $g_3 = 2_2$.

Based on Table 2 and the Fischer matrix $\mathcal{F}_3$ below, an example on how to construct the partial character table of $\overline{G}$ on the conjugacy classes listed in Table 2 is given in [14]. The full character table of $\overline{G}$ in the format of Clifford-Fischer Theory is available in [8].

5.2 On the Non-Split Extension $\overline{G}_n = 2^{2n\cdot}Sp(2n, 2)$

In [11] we established some general results on the non-split extension group $\overline{G}_n = 2^{2n\cdot}Sp(2n, 2)$, where $Sp(2n, 2)$ acts faithfully on 2^{2n}. Firstly the group $\overline{G}_n$ can be

Table 3.
$\mathcal{F}_3$

g_3		g_{31}	g_{32}	g_{33}	g_{34}	g_{35}
$o(g_{3j})$		8	8	2	4	8
$\lvert C_{\overline{G}}(g_{3j})\rvert$		192	192	96	32	32
(k,m)	$\lvert C_{H_k}(g_{3km})\rvert$					
$(1,1)$	12	1	1	1	1	1
$(2,1)$	12	$2\sqrt{2}i$	$-2\sqrt{2}i$	0	0	0
$(3,1)$	4	3	3	3	-1	-1
$(4,1)$	12	1	1	-1	-1	1
$(4,2)$	4	3	3	-3	1	-1
m_{3j}		8	8	16	48	48

constructed in terms of permutations of a set of cardinality at least 2^{2n+1}, i.e., $\overline{G}_n \leq S_{2^{2n+1}}$, but $\overline{G}_n \not\leq S_{2^{2n+1}-1}$ (in fact $\overline{G}_n \leq A_{2^{2n+1}}$). Moreover the group $\overline{G}_n$ acts transitively on a $2^{2n+1}-2$ points, that it fixes 2 points of the set $\{1,2,\cdots,2^{2n+1}\}$. Hence the resulting permutation character of this action is of degree $2^{2n+1}-2$. The group $\overline{G}_n/2^{2n} \cong Sp(2n,2)$ acts faithfully on 2^{2n}, it yields two orbits of lengths 1 and $2^{2n}-1$ and it is necessarily that the lengths of the orbits of $\overline{G}_n$ on $\mathrm{Irr}(2^{2n})$ are 1 and $2^{2n}-1$ also. The respective inertia factor groups are $H_1 = G_n = Sp(2n,2)$ and $H_2 = 2^{2n-1}{:}Sp(2n-2,2)$, the affine symplectic group. From Section 3, it follows that the irreducible characters of $\overline{G}_n$ are distributed into two blocks of characters $\mathcal{K}_1$ and $\mathcal{K}_2$ corresponding to the ordinary characters of $H_1 = Sp(2n,2)$ and a projective character table of $H_2 = 2^{2n-1}{:}Sp(2n-2,2)$ respectively. Thus the number of irreducible characters of $\overline{G}_n$ is given by the following formula:

$$|\mathrm{Irr}(2^{2n\cdot}Sp(2n,2))| = |\mathrm{Irr}(Sp(2n,2))| + |\mathrm{IrrProj}(2^{2n-1}{:}Sp(2n-2,2),\alpha^{-1})|, \qquad (6)$$

for some factor set α of the Schur multiplier of H_2. Another result about the group $\overline{G}_n$ is that for any $n \geq 2$, the identity Fischer matrix of $\overline{G}_n$ will have the form:

$\mathcal{F}_1$

$g_1 = 1_{Sp(2n,2)}$		g_{11}	g_{12}
$o(g_{1j})$		1	2
$\lvert C_{\overline{G}}(g_{1j})\rvert$		$\lvert\overline{G}_n\rvert$	$\lvert\overline{G}_n\rvert/2^{2n}-1$
(k,m)	$\lvert C_{H_k}(g_{1km})\rvert$		
$(1,1)$	$\lvert G_n\rvert$	1	1
$(2,1)$	$\lvert G_n\rvert/2^{2n}-1$	$2^{2n}-1$	-1
m_{1j}		1	$2^{2n}-1$

Thus the degree of any irreducible character of $\overline{G}_n$ contained in the block $\mathcal{K}_2$ is a multiple of $2^{2n}-1$. We also proved that for $n \in \{2,3,4,5,6\}$, we only need the ordinary character table of H_2 for the construction of the character table of $\overline{G}_n$. The following table lists the number of ordinary irreducible characters of $\overline{G}_n = 2^{2n\cdot}Sp(2n,2)$ for small values of n.

Table 4. The number of ordinary irreducible characters of $\overline{G}_n$, $n \in \{2,3,4,5,6\}$

n	$\lvert\mathrm{Irr}(Sp(2n,2))\rvert$	$\lvert\mathrm{Irr}(2^{2n-1}{:}Sp(2n-2,2))\rvert$	$\lvert\mathrm{Irr}(2^{2n}{\cdot}Sp(2n,2))\rvert$
2	11	10	21
3	30	37	67
4	81	114	195
5	198	322	520
6	477	839	1316

We conjectured that for any $n \in \mathbb{N}^{\geq 2}$, we only need to use the ordinary character table of H_2 to construct the character table of $\overline{G}_n$. In [10] and [11] we applied the coset analysis technique, found the inertia factor groups, computed the Fischer matrices and the character tables (via Clifford-Fischer matrices) of the groups $\overline{G}_3$ and $\overline{G}_4$ respectively.

Acknowledgments

The first author would like to thank his supervisor (second author) for his advice and support. The financial support from the National Research Foundation (NRF) of South Africa and the University of KwaZulu-Natal are also acknowledged.

References

[1] F. Ali, *Fischer-Clifford Theory For Split and Non-Split Group Extensions* (PhD Thesis, University of Natal, Pietermaritzburg 2001).

[2] F. Ali and J. Moori, The Fischer-Clifford matrices and character table of a maximal subgroup of Fi_{24}, *Algebra Colloq.* **17** (2010), 389–414.

[3] F. Ali and J. Moori, Fischer-Clifford matrices of the non-split group extension $2^{6}{\cdot}U_4(2)$, *Quaest. Math.* **31** (2008), 27–36.

[4] F. Ali and J. Moori, The Fischer-Clifford matrices and character table of the group $2^8{:}Sp_6(2)$, *Int. J. Math. Game Theory Algebra* **14** (2004), 123–135.

[5] F. Ali and J. Moori, Fischer-Clifford matrices and character table of the group $2^7{:}Sp_6(2)$, *Algebra Colloq.* **14** (2004), 101–121.

[6] F. Ali and J. Moori, The Fischer Clifford matrices of a maximal subgroup of $Fi_{24}^{'}$, *Represent. Theory* **7** (2003), 300–321.

[7] R. W. Barraclough, *Some Calculations Related To The Monster Group* (PhD Thesis, University of Birmingham, Birmingham 2005).

[8] A. B. M. Basheer, *Clifford-Fischer Theory Applied to Certain Groups Associated with Symplectic, Unitary and Thompson Groups* (PhD Thesis, University of KwaZulu-Natal, Pietermaitzburg 2012). Accessed online through http://researchspace.ukzn.ac.za/xmlui/handle/10413/6674?show=full

[9] A. B. M. Basheer and J. Moori, Fischer matrices of Dempwolff group $2^{5}{\cdot}GL(5,2)$, *Int. J. Group Theory* **1** (2012), 43–63.

[10] A. B. M. Basheer and J. Moori, On the non-split extension group $2^{6}{\cdot}Sp(6,2)$, *Bull. Iranian Math. Soc.* **39** (2013), 1189–1212.

[11] A. B. M. Basheer and J. Moori, On the non-split extension $2^{2n}{\cdot}Sp(2n,2)$ and the character table of $2^{8}{\cdot}Sp(8,2)$, *Bull. Iranian Math. Soc.*, to appear.

[12] A. B. M. Basheer and J. Moori, Fischer matrices of the group $2_{+}^{1+8}{\cdot}A_9$, submitted.

[13] A. B. M. Basheer and J. Moori, On a group of the form $3^7{:}Sp(6,2)$, submitted.

[14] A. B. M. Basheer and J. Moori, Clifford-Fischer theory applied to a group of the form 2^{1+6}_{-}:((3^{1+2}:8):2), submitted.
[15] A. B. M. Basheer and J. Moori, On a group of the form 2^{10}:($U_5(2)$:2), to be submitted.
[16] W. Bosma and J. J. Cannon, *Handbook of Magma Functions* (Department of Mathematics, University of Sydeny, November 1994).
[17] A. H. Clifford, Representations induced in an invariant subgroup, *Ann. of Math.* **38** (1937), 533–550.
[18] B. Fischer, *Clifford matrizen, manuscript* (1982).
[19] B. Fischer, *Unpublished manuscript* (1985).
[20] B. Fischer, Clifford matrices, *Representation theory of finite groups and finite-dimensional Lie algebras* (eds G. O. Michler and C. M. Ringel; Birkhäuser, Basel, 1991), 1–16.
[21] The GAP Group, *GAP – Groups, Algorithms, and Programming,* Version 4.4.10; 2007, http://www.gap-system.org
[22] K. Lux and H. Pahlings, *Representations of Groups: A Computational Approach* (Cambridge University Press, Cambridge 2010).
[23] Maxima, *A Computer Algebra System.* Version 5.18.1; 2009, http://maxima.sourceforge.net
[24] J. Moori, *On the Groups G^+ and $\overline{G}$ of the form 2^{10}:M_{22} and 2^{10}:$\overline{M}_{22}$* (PhD Thesis, University of Birmingham 1975).
[25] J. Moori, On certain groups associated with the smallest Fischer group, *J. London Math. Soc.* **2** (1981), 61–67.
[26] Z. E. Mpono, *Fischer Clifford Theory and Character Tables of Group Extensions* (PhD Thesis, University of Natal, Pietermaritzburg 1998).
[27] J. Moori and Z. Mpono, Fischer-Clifford matrices and the character table of a maximal subgroup of $\overline{Fi}_{22}$, *Int. J. Math. Game Theory Algebra* **10** (2000), 1–12.
[28] J. Moori and Z. Mpono, The Fischer-Clifford matrices of the group 2^6:$SP_6(2)$, *Quaest. Math.* **22** (1999), 257–298.
[29] J. Moori, Z. E. Mpono and T. T. Seretlo, A group 2^7:S_8 in $\overline{Fi}_{22}$, *Southeast Asian Bull. Math.*, to appear.
[30] J. Moori and K. Zimba, Fischer-Clifford matrices of $B(2, n)$, *Quaest. Math.* **29** (2006), 9–37.
[31] J. Moori and B. G. Rodrigues, On Frattini extensions, *Sci. Math. Jpn.* **55** (2002), 215–221.
[32] J. Moori and T. T. Seretlo, On the Fischer-Clifford matrices of a maximal subgroup of the Lyons Group Ly, *Bull. Iranian Math. Soc.*, to appear.
[33] J. Moori and T. T. Seretlo, A Group of the form 2^6:A_8 as an inertia factor group of 2^8:$O^+_8(2)$, submitted.
[34] J. Moori and T. T. Seretlo, On two non-split extension groups associated with HS and HS:2, submitted.
[35] J. Moori and T. T. Seretlo, On an inertia factor group of $O^+_{10}(2)$, submitted.
[36] J. Moori and K. Zimba, Fischer-Clifford Matrices of $B(2, n)$, *Quaest. Math.* **29** (2005), 9–37.
[37] H. Pahlings, The character table of $2^{1+22}_{+}\cdot Co_2$, *J. Algebra* **315** (2007), 301–325.
[38] B. G. Rodrigues, *On The Theory and Examples of Group Extensions* (MSc Thesis, University of Natal, Pietermaritzburg 1999).
[39] U. Schiffer, *Cliffordmatrizen* (Diplomarbeit, Lehrstul D Fur Matematik, RWTH, Aachen 1995).
[40] N. S. Whitely, *Fischer Matrices and Character Tables of Group Extensions* (MSc Thesis, University of Natal, Pietermaritzburg 1993).

A GENERALISATION ON THE SOLVABILITY OF FINITE GROUPS WITH THREE CLASS SIZES FOR NORMAL SUBGROUPS

ANTONIO BELTRÁN* and MARÍA JOSÉ FELIPE†

*Departamento de Matemáticas, Universidad Jaume I, 12071 Castellón, Spain
E-mail: abeltran@mat.uji.es

†Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, 46022 Valencia, Spain
E-mail: mfelipe@mat.upv.es

Abstract

Let N be a normal subgroup of a finite group G. In the recent past years some results have appeared concerning the influence of the G-class sizes of N, that is, with the sizes of the conjugacy classes in G contained in N, on the structure of N. In this survey, we present the main results and techniques used for proving that any normal subgroup of G which has exactly three G-conjugacy class sizes is solvable. Thus, we obtain a generalisation for normal subgroups of the classical N. Itô's theorem which asserts that those finite groups having three class sizes are solvable, and in particular, a new proof of it is provided.

1 Introduction

The solvability of a finite group G with three conjugacy class sizes is a complex problem solved by N. Itô in [22]. He proved that such groups are solvable by appealing to Feit-Thompson's theorem and some deep classification theorems by M. Suzuki. This result was simplified by J. Rebmann in [25] when G is an F-group (that is, G has no pair of non-central elements such that the centraliser of one element properly contains the other centraliser). Then he determined the structure of F-groups by using results of R. Baer ([8] and [9]) and M. Suzuki ([27]) about groups with a non-trivial normal partition. Afterwards, A.R. Camina proved in [14], by using the description of finite groups with dihedral Sylow 2-subgroups given by D. Gorenstein and J.H. Walter, that if G is not an F-group and has three class sizes, then G is a direct product of an abelian subgroup and a subgroup whose order involves no more than two primes. Forty years later, the structure of these groups has been completely determined (up to nilpotent groups, which in this context are p-groups) by S. Dolfi and E. Jabara in [15], who based their proof on the solvability of this type of groups.

Let G be a finite group and N be a normal subgroup of G. Recent research works have put forward that the set of sizes of those conjugacy classes of G contained in N, also called G-class sizes of N and denoted by $\mathrm{cs}_G(N)$, exerts a strong influence on the structure of N. If $\mathrm{cs}(N)$ denotes the set of class sizes of N, we stress that $|\mathrm{cs}(N)| \leq |\mathrm{cs}_G(N)|$ does not hold for every normal subgroup N of G. For instance, the smallest example which is not a p-group is a group of order 72 defined as follows: Let $G = \mathrm{S}_3 \wr \mathbb{Z}_2$ and let $N = \mathrm{S}_3 \times \mathrm{S}_3$, which is normal in G. Then $\mathrm{cs}(N) = \{1, 2, 3, 4, 6, 9\}$ while $\mathrm{cs}_G(N) = \{1, 4, 6, 9, 12\}$.

Nevertheless, it is surprising that the G-class sizes of normal subgroups still seem to keep control on their structure. This has emerged as a new useful tool to obtain information regarding normal subgroups, and moreover, this approach has the advantage of enabling to argue by induction on the order of N. As a first step, the nilpotency of normal subgroups having two G-class sizes is proved in [6] and thereby we obtain a generalisation for normal subgroups of the celebrated Itô's Theorem on groups having two class sizes. It is worth mentioning that while the proof of Itô's result is quite elementary, the proof of the above extension requires the Classification of the Finite Simple Groups (CFSG). In this survey, we present the main results and different methods used in order to prove the following theorem.

Theorem A *If N is a normal subgroup of a finite group G and $|\mathrm{cs}_G(N)| = 3$, then N is solvable.*

We remark that the proof of Theorem A needs CFSG too. As mentioned above, several previous works and the use of different techniques have been necessary in order to complete it. For example, the classification of the simple CP-groups (those groups having only elements of prime power order) appears in a natural way. The classification of non-abelian simple groups whose prime graph is a forest is also used, as well as certain properties relating the G-class sizes to the Fitting subgroup and the centre of a normal subgroup, some properties of the Schur multiplier of simple groups, and determining the structure of normal sections of G involving certain hypothesis on the G-class sizes.

The proof of Theorem A is divided into two cases, which call for different treatments. In section 2, we analyse the case in which $\mathrm{cs}_G(N) = \{1, m, n\}$ and m does not divide n, and in section 3, we study the remaining case. This latter case is more difficult and we will present several previous research works in different subsections. We also pose an open problem in the last section.

All groups will be finite. If x is any element of a group G, we denote by x^G the conjugacy class of x in G and by $|x^G|$ the G-conjugacy class size of x. This is also called the index of x in G. For the rest of the notation, we will follow [20].

2 The case in which m does not divide n

In this section, we present the main tools employed in order to prove the solvability of a normal subgroup N of G with three G-class sizes just when $\mathrm{cs}_G(N) = \{1, m, n\}$ and m does not divide n and, consequently, the structure of N is determined (these results are obtained in [2]). The proof of the solvability sets up a generalisation and a subsequent classification of the concept of F-group for normal subgroups. As mentioned in the introduction, the definition and the classification of F-groups was originally given by J. Rebmann in [25]. Our extension is the following.

Definition 2.1 A non-central normal subgroup N of a finite group G is said to be an F-normal subgroup if for every $x, y \in N \setminus \mathbf{Z}(G)$, such that $\mathbf{C}_G(x) \subseteq \mathbf{C}_G(y)$, then $\mathbf{C}_G(x) = \mathbf{C}_G(y)$.

It is not difficult to prove the following property. As for the concept of partition and its properties, we refer the reader to the survey [28].

Theorem 2.2 *If N is an* F*-normal subgroup of G, then $N/(N \cap \mathbf{Z}(G))$ has a non-trivial normal abelian partition.*

The authors employed Baer and Suzuki's results on groups having a non-trivial normal partition so as to classify F-normal subgroups, which is given in the following (Theorem 7 of [2]).

Theorem 2.3 *Let G be a group and N be an* F*-normal subgroup of G. Then N satisfies one of the following conditions:*

(1) *$N/\mathbf{Z}(N)$ is a Frobenius group, with Frobenius kernel $L/\mathbf{Z}(N)$ and Frobenius complement $K/\mathbf{Z}(N)$, where K and L are abelian, and N is an* F*-group.*

(2) *$N/\mathbf{Z}(N)$ is a Frobenius group, with Frobenius kernel $L/\mathbf{Z}(N)$ and Frobenius complement $K/\mathbf{Z}(N)$, where K is abelian, and $L/\mathbf{Z}(N)$ is of prime-power order, and L is an* F*-normal subgroup.*

(3) *$N/\mathbf{Z}(N) \cong \mathrm{S}_4$ and V is non-abelian, for $V/\mathbf{Z}(N)$, the Klein four-group of $N/\mathbf{Z}(N)$. In particular, N is an* F*-group.*

(4) *N has abelian Fitting subgroup of index p, where p divides $|\mathbf{F}(N)/\mathbf{Z}(N)|$, and in particular, N is an* F*-group.*

(5) *$N = P \times \mathbf{Z}(N)_{p'}$, where $P \in \mathrm{Syl}_p(N)$.*

(6) *$N/\mathbf{Z}(N) \cong \mathrm{PSL}(2, p^h)$ or $\mathrm{PGL}(2, p^h)$, where p is a prime, and $p^h \geq 4$.*

In all cases but case (5)*, we have $\mathbf{Z}(N) = N \cap \mathbf{Z}(G)$.*

As a consequence of this classification, the solvability of F-normal subgroups with three G-class sizes is obtained. The main part consists in showing that case (6) is not possible. In fact, when N is not solvable, the proof reduces to the case in which N is quasi-simple and then, by means of some properties of the Schur multiplier, a contradiction is reached.

Theorem 2.4 *Let N be an* F*-normal subgroup of G, such that $|\mathrm{cs}_G(N)| = 3$. Then N is solvable.*

Notice that if $\mathrm{cs}_G(N) = \{1, m, n\}$ where m does not divide n, then N clearly is an F-normal subgroup of G and then, by Theorem 2.4, N is solvable. Finally, the normal structure is fully determined (Theorem A of [2]), basically by showing that cases (3) and (4) of Theorem 2.3 cannot happen.

Theorem 2.5 *Let N be a normal subgroup of a finite group G such that $\mathrm{cs}_G(N) = \{1, m, n\}$, where $m < n$ and m does not divide n. Then one of the following conditions is satisfied:*

(i) *$N = P \times A$, where $P \in \mathrm{Syl}_p(N)$, for some prime p, and $A \subseteq \mathbf{Z}(G)$.*

(ii) *N is a quasi-Frobenius group. More precisely, $N/\mathbf{Z}(N)$ is a Frobenius group, with Frobenius kernel $L/\mathbf{Z}(N)$ and Frobenius complement $K/\mathbf{Z}(N)$, and either K and L are abelian, and $\mathrm{cs}(N) = \{1, |L/\mathbf{Z}(N)|, |K/\mathbf{Z}(N)|\}$; or K is abelian, and $L/\mathbf{Z}(N)$ is of prime-power order, and $\mathrm{cs}(N) = \{1, |L/\mathbf{Z}(N)|, |K/\mathbf{Z}(N)||x^L| \ : x \in L \setminus \mathbf{Z}(N)\}$.*

As regards the above theorem, we point out that the structure of groups having three conjugacy class sizes is completely determined in [15] and the authors used the solvability due to Itô to attain its proof. Thus, the structure given in Theorem 2.5 for N when m does not divide n is a partial generalisation for normal subgroups of this classification, which is summarized in the following.

Theorem 2.6 *A finite group G has three class sizes if and only if, up to an abelian factor, either*

(1) *G is a p-group for some prime p or*

(2) *$G = KL$ with $K \trianglelefteq G$, $(|K|, |L|) = 1$ and one of the following occurs*

   (a) *both K and L are abelian, $\mathbf{Z}(G) < L$ and G is a quasi-Frobenius group,*

   (b) *K is abelian, L is a non-abelian p-group, for some prime p and $\mathbf{O}_p(G)$ is an abelian subgroup of index p in L and $G/\mathbf{O}_p(G)$ is a Frobenius group or*

   (c) *K is a p-group with two class sizes for some prime p, L is abelian, $\mathbf{Z}(K) = \mathbf{Z}(G) \cap K$ and G is quasi-Frobenius.*

Some easy examples show that each of the two types of normal structure described in Theorem 2.5 really occurs. For instance, let G be the group of the library of the small groups of GAP with number Id(324,8). Using GAP, it is easy to check that G has an abelian normal subgroup $N \cong \mathbb{Z}_3 \times \mathbb{Z}_3$, with $\mathrm{cs}_G(N) = \{1, 2, 3\}$. This provides an example of groups described in Theorem 2.5 (i).

On the other hand, the group $G = \mathrm{GL}(2, 3)$ with normal subgroup $N = \mathrm{SL}(2, 3)$ is an example of a group of type (ii), with $\mathrm{cs}(N) = \{1, 4, 6\}$ and $\mathrm{cs}_G(N) = \{1, 6, 8\}$, and where the inverse images of the Frobenius kernel and Frobenius complement of the Frobenius group $N/\mathbf{Z}(N)$ are abelian. Now, we consider $L = \langle x, y\rangle$ an extraspecial group of order p^3 and exponent for some odd prime p. Let $K = \langle a\rangle \times \langle b\rangle$ with a of order 8 and b of order 2. Assume that $\langle a^2 b\rangle$ acts trivially on L and $K/\langle a^2 b\rangle = \langle \alpha\rangle$ acts on L by $y^\alpha = x^{-1}$ and $x^\alpha = y$, where α is an element of order 4. Assume that $N = LK$. Note that $\mathrm{cs}(N) = \{1, 4p, p^2\}$. Let $G = N \times A$, where A is any finite group. Then $\mathrm{cs}_G(N) = \mathrm{cs}(N)$, which provides another example of a group of type (ii) in Theorem 2.5.

An easy consequence is to determine the structure of the normal subgroups having three G-class sizes when these satisfy that the non-trivial sizes are coprime numbers.

Corollary 2.7 *Let G be a finite group and N be a normal subgroup of G such that $\mathrm{cs}_G(N) = \{1, m, n\}$ with $(m, n) = 1$. Then either N is quasi-Frobenius and the inverse image of the kernel and complement are abelian, or $N = P \times A$ with $A \subseteq \mathbf{Z}(G)$ and $P \in \mathrm{Syl}_p(N)$, for some prime p.*

The group $N = \mathrm{A}_4$, considered as a normal subgroup of $G = \mathrm{S}_4$, is an example of a Frobenius group with $\mathrm{cs}(N) = \{1, 3, 4\}$ and $\mathrm{cs}_G(N) = \{1, 3, 8\}$. Thus, this corollary extends the ordinary case of groups with three class sizes such that the non-trivial sizes are coprime. This appeared first in [23] and was reformulated later in [11] within the context of graphs and we state it next.

Theorem 2.8 *Let $\mathrm{cs}(G) = \{1, m, n\}$ be the set of conjugacy class sizes of a finite group G. Then $(m, n) = 1$ if and only if $G/\mathbf{Z}(G)$ is a Frobenius group where the inverse image of the kernel and complement are abelian.*

3 The case in which m divides n

The approach to the solvability of a normal subgroup N with $\mathrm{cs}_G(N) = \{1, m, n\}$ and with m being a divisor of n needs different methods from those employed when m does not divide n. We remark that these techniques also differ from those employed by Camina in [14] in the ordinary case. As a result, we obtain a new proof of this case when we take $N = G$.

The general outline to obtain the solvability of N in this case consists in proving first that every element in N whose class size is m lies in $\mathbf{F}(N)$, the Fitting subgroup of N. In order to obtain this result, we employ the structure of normal subgroups with two G-class sizes of p-regular elements, which are analysed in section 3.1. The important fact is that since m divides n, then m actually divides all nontrivial G-class sizes of all elements of N. As we will see in section 3.2, this allows to obtain relevant properties of $\mathbf{F}(N)$. On the other hand, once this property is proved, we have that every element of N lying outside $\mathbf{F}(N)$ has G-class size n. We will explore this condition in section 3.3. to show interesting results on the structure of the normal section $N/\mathbf{F}(N)$.

3.1 Normal subgroups with two G-class sizes of p-regular elements

Some recent results indicate that the structure of the p-complements of a finite group for some prime p is closely related to the set of sizes of its p-regular conjugacy classes, that is, the conjugacy classes of p'-elements of G (see for instance [7] or [3]). However, studying this relation may be a complex problem, even when G is assumed to be p-solvable. It is easy to find examples of groups having p-complements, for some prime p, such that the class sizes of the elements of these p-complements do not divide the respective class sizes in the whole group. In fact, if H is a p-complement of G, then the cardinality of the set of conjugacy class sizes of H is not necessarily bounded by the cardinality of the set of p-regular class sizes of G. For instance, using GAP we see that the small group Id(600,57) has four class sizes of 3-regular elements whereas it has a 3-complement with five class sizes.

We consider again the solvability of a normal subgroup N when $\mathrm{cs}_G(N) = \{1, m, n\}$ and m divides n. We observe that if we take a p-element z of index m, for some prime p, and if y is a p-regular element of $\mathbf{C}_G(z)$, then

$$|y^{\mathbf{C}_G(z)}| = |\mathbf{C}_G(z) : \mathbf{C}_G(z) \cap \mathbf{C}_G(y)| = |\mathbf{C}_G(z) : \mathbf{C}_G(yz)|.$$

Thus, the class size of y in $\mathbf{C}_G(z)$ is equal to 1 or n/m, and consequently, $\mathbf{C}_N(z)$ is a normal subgroup of $\mathbf{C}_G(z)$ having at most two p-regular $\mathbf{C}_G(z)$-class sizes. Determining the structure of $\mathbf{C}_N(z)$ has been necessary in several steps in the proof of Theorem A. This situation led us to study the structure of normal subgroups having two p-regular G-class sizes. We obtained an extension (Theorem A of [1]), under the p-solvability hypothesis, of the main theorem of [6] which establishes the nilpotency of the p-complements of normal subgroups having two G-class sizes.

Theorem 3.1 *Let N be a normal subgroup of a finite group p-solvable G. Suppose that the G-conjugacy class of every p-regular element of N has size 1 or m for some fixed integer m. Then N has abelian p-complements or $N = RP \times A$, where R and P*

are a Sylow r-subgroup for some prime r and a Sylow p-subgroup of N respectively, and A is a central group of G.

Eliminating the p-solvability condition in the above theorem is a more complicated problem and the following result has been fundamental to solve it.

Theorem 3.2 *Let N be a normal subgroup of a group G such that the G-conjugacy class size of every p-regular element of N is 1 or m, for some integer m. Then either N has abelian p-complements or all p-regular elements of $N/(N \cap \mathbf{Z}(G))$ have prime power order.*

In order to eliminate the p-solvability hypothesis in Theorem 3.1, we take a chief factor N/K of G such that $N \cap \mathbf{Z}(G) \subseteq K$ and then, by using the above theorem, we get that the prime graph of N/K is a forest. The problem can be reduced to the case in which N/K is simple and then, we use the classification of the non-abelian simple groups whose prime graph, $\Gamma(G)$, is a forest (see [24]). This has been decisive for our purposes.

Theorem 3.3 *Let G be a finite non-abelian simple group. If $\Gamma(G)$ is a forest then G is one of the following simple groups:* A_5, A_6, A_7, A_8, M_{11}, M_{22}, $\mathrm{PSL}_4(3)$, $\mathrm{B}_2(3)$, $\mathrm{G}_2(3)$, $\mathrm{U}_4(3)$, $\mathrm{U}_5(2)$,${}^2\mathrm{F}_4(2)'$*, or belongs to one of the families:* $\mathrm{PSL}_2(q)$, $\mathrm{PSL4}_3(q)$, $\mathrm{PSU}_3(q)$, $\mathrm{Sz}(q^2)$ *with $q^2 = 2^f$ or $q = 2^{f^2}$ with f an odd prime, and* $\mathrm{Ree}(3^f)$*, with f an odd prime.*

Finally, the application of certain properties of the Schur multiplier of these simple groups has allowed to conclude the desired result which appears as Theorem 8 of [4].

Theorem 3.4 *Let N be a normal subgroup of a finite group G. Let p a prime number and suppose that N has two G-conjugacy class sizes of every p-regular elements. Then N is solvable.*

3.2 Properties of the Fitting subgroup of N

Let us go back to the proof of the fact that every element in N of G-class size m lies in $\mathbf{F}(N)$ when m divides n. In order to prove it, we use several properties which have been developed in [10] and have interest on their own. The first one (Theorem 5 of [10]) is an extension for normal subgroups of a result by Dolfi and Jabara (Theorem 3.2 of [15]) which asserts that when G is a solvable group and m divides s for every $s \in \mathrm{cs}(G)$, $s \neq 1$, we have that if $g \in G$ and $|g^G| = m$, then $g \in \mathbf{F}(G)$. This property suitably works for solvable normal subgroups and G-class sizes.

Theorem 3.5 *Suppose that N is a normal solvable subgroup of a group G and suppose that m divides s for every $s \in \mathrm{cs}_G(N)$, $s \neq 1$. If $g \in N$ and $|g^G| = m$, then $g \in \mathbf{F}(N)$.*

We do not know whether Theorem 3.5 could hold for non-solvable subgroups (and for Dolfi and Jabara's original result either) and therefore, this problem remains open. However, when N is non-solvable we can still attain some useful properties under the hypothesis that there is an integer dividing all non-trivial G-class sizes. We point out that the following result (Theorem 7 of [10]) relies on the CFSG.

Theorem 3.6 *Suppose that N is a normal nonsolvable subgroup of a group G and suppose that an integer m divides $|x^G|$ for every $x \in N \setminus \mathbf{Z}(N)$. Then m divides $|\mathbf{Z}(N)|$.*

The main application of our development is Theorem 3.8, which is not trivial at all. It relies on the Classification, precisely, on the following property concerning the Schur multiplier of a simple group.

Lemma 3.7 *A finite nonabelian simple group does not have a nontrivial conjugacy class whose size divides the order of its Schur multiplier.*

The following theorem is one of the turning points when dealing with obtaining the solvability of N.

Theorem 3.8 *If N is a nonabelian normal subgroup of a finite group G and if $|\mathrm{cs}_G(N)| = 3$, then $\mathbf{Z}(N)$ is properly contained in $\mathbf{F}(N)$.*

3.3 Normal sections

Once we have obtained that every element in N of G-class size m belongs to $\mathbf{F}(N)$, we deduce that every element in $N \setminus \mathbf{F}(N)$ has index n. Let us examine for a moment ordinary conjugacy classes. In [19], I. M. Isaacs considered finite groups G which contain a proper normal subgroup N such that all of the conjugacy classes of G which lie outside N have equal sizes. A nonabelian group with this property is said to satisfy condition (*). It turns out that either G/N is cyclic or else every nonidentity element of G/N has prime order. Moreover, Isaacs provides examples for these two situations: a Frobenius group G with kernel N and cyclic complement is an example of the first kind of groups, and for the second case, it is not difficult to produce a group G with a normal subgroup N such that G/N isomorphic to the symmetric group S_3 satisfying condition (*). Extending Isaacs' definition, we give in [5] the following definition.

Definition 3.9 A normal section N/K of a group G satisfies condition (*) over G when N is a nonabelian normal subgroup of G such that all the G-conjugacy classes in N lying outside of K have equal size.

The main tool employed by Isaacs to cope with this situation is the concept of partition relative to a normal subgroup. In order to provide a better understanding of the properties of normal sections under condition (*), we will state the original results on normal partitions relative to a section G/N, which, as we said, were introduced by Isaacs in [19].

Definition 3.10 Suppose that $N \trianglelefteq G$ and $G = N \cup (\cup_i H_i)$ where $H_i \subset G$ are subgroups satisfying $H_i \cap H_j \subseteq N$ when $i \neq j$. In this situation we say that G is partitioned relative to N.

Proposition 3.11 *Let G satisfy (*) and suppose that G/N contains an element of order p^2 for some prime p, then G has abelian Sylow p-subgroups.*

Proposition 3.12 *Let $N \trianglelefteq G$ and suppose that G is partitioned relative to N and that G/N is abelian. Let p be a prime divisor of $|G : N|$ and suppose that a Sylow p-subgroup of G is normal in G. Then G/N is an elementary abelian p-group.*

Now, we focus our attention on a normal section N/K of a group. Relative partitions also appear naturally when studying such a normal section satisfying condition (*). It is easy to see the following results.

Lemma 3.13 *If N/K satisfies (*) over G and $\mathbf{Z}(N) \subseteq K$, then N is partitioned relative to K.*

Lemma 3.14 *Suppose that N/K satisfies (*) over G. Let $x \in N \setminus K$.*

i) *If $xK \in N/K$ is not a p-element, then there exists a Sylow p-subgroup P of $\mathbf{C}_N(x)$ such that $P \subseteq \mathbf{Z}(\mathbf{C}_G(x))$.*

ii) *If the order of $xK \in N/K$ is divisible by two distinct primes, then $\mathbf{C}_N(x) \subseteq \mathbf{Z}(\mathbf{C}_G(x))$ and in particular, $\mathbf{C}_N(x)$ is abelian.*

These are the main properties which finally lead to obtain Theorem 3.17.

Proposition 3.15 *Let N/K be a normal section satisfying (*) over G and suppose that N/K has elements of order p^2 for a certain prime p. Let $P \in \mathrm{Syl}_p(N)$ and let $A = \langle x \in P | x^p \notin K \rangle$. Then*

i) *A is abelian and $P \cap K \subseteq A \trianglelefteq \mathbf{N}_G(P)$.*

ii) *If $\overline{A} = A/(P \cap K)$ and $\overline{P} = P/(P \cap K)$, for every generator a of A we have $\mathbf{C}_P(a) = A$ and $\mathbf{C}_{\overline{P}}(\overline{a}) = \overline{A}$.*

By using Proposition 3.12 and Lemma 3.13 we obtain the following.

Lemma 3.16 *Let N/K satisfy (*) over G and $\mathbf{Z}(N) \subseteq K$. Let $xK \in N/K$ whose order is not a prime number and let p be a prime divisor of the order of xK. If P is a Sylow p-subgroup of K, then P is abelian, K has a normal p-complement and $xK \in \mathbf{C}_N(P)K/K$.*

Finally, we attain the main theorem about normal sections, which is Theorem B of [5].

Theorem 3.17 *Let N/K be a normal section satisfying (*) over G.*

i) *If $\mathbf{Z}(N) \not\subseteq K$, then N/K is a p-group for some prime p and N/K is either abelian or has exponent p.*

ii) *If $\mathbf{Z}(N) \subseteq K$, then either N/K is cyclic or is a* CP*-group. If N/K is not a* CP*-group, then N has abelian Hall π-subgroups and normal π-complement, where π is the set of prime divisors of $|N/K|$.*

3.4 CP-groups and final arguments

If we continue with our reasoning, we see that $N/\mathbf{F}(N)$ is a nontrivial normal section of G satisfying (*). If we want to prove that a normal subgroup N with $|\mathrm{cs}_G(N)| = 3$ is solvable, then by Theorem 3.17 ii), the only case to be studied is when $N/\mathbf{F}(N)$ is a

non-solvable CP-group. We appeal then to the structure theorem of nonsolvable CP-groups (Theorem 3.18), which are those groups having all elements of prime-power order. The structure of finite solvable CP-groups was given by G. Higman fifty years ago ([17]), and the structure of non-solvable CP-groups and the classification of the simple CP-groups have been recently obtained by H. Heineken in [16]. It can be summarised as follows.

Theorem 3.18 *If G is a finite, non-solvable* CP*-group, then there are normal subgroups B, C of G such that $1 \subseteq B \subseteq C \subseteq G$ and B is a 2-group, C/B is non-abelian and simple, and G/C is a p-group for some prime p and cyclic or generalised quaternion. In particular, if G is a finite non-abelian simple* CP*-group, then G is isomorphic to one of the following groups:* $\mathrm{L}_2(q)$*, for* $q = 5, 7, 8, 9, 17$*,* $\mathrm{L}_3(4)$*,* $\mathrm{Sz}(8)$ *or* $\mathrm{Sz}(32)$*.*

In our proof, we are able to reduce to the case in which N/B is simple, where B denotes the radical solvable subgroup of N. Taking into account the classification in Theorem 3.18, we carry out a case-by-case analysis on each of the eight simple groups appearing in such theorem. Without going into details in these cases, we only remark that the key facts in the analysis are the order of the elements of the simple groups, Proposition 3.15 and Lemma 3.16. With all of them we produce a contradiction in each case. The most laborious cases are $\mathrm{L}_2(5)$ and $\mathrm{L}_2(9)$ (that is, the alternating groups A_5 and A_6).

4 Open problem

We note that a complete classification of the structure of normal subgroups N with $\mathrm{cs}_G(N) = \{1, m, n\}$ and m dividing n is still open. As we said in the introduction, in the ordinary case, the structure of a finite group G satisfying $\mathrm{cs}(G) = \{1, m, n\}$ with m dividing n was already studied by Camina, although it was completely determined in [15]. In fact, such a group satisfies that either $G/\mathbf{Z}(G)$ is a p-group for some prime p, or $\mathbf{F}(G)$ is an abelian subgroup of index p. Nevertheless, when dealing with G-class sizes in normal subgroups, such structure does not hold as we show in the following example.

Let

$$L = \langle x, y | x^3 = y^3 = 1, [x, y]^3 = 1, [x, [x, y]] = [y, [x, y]] = 1 \rangle$$

be the extraspecial group of order 3^3 and exponent 3. If $z = [x, y]$, then $\mathbf{Z}(L) = \langle z \rangle$. Let $\langle a \rangle$ be the automorphism of L defined by $x^a = x^2$ and $y^a = y^2$. The set of fixed points of a on L is exactly $\mathbf{Z}(L)$. On the other hand, let us consider an automorphism α of order 3 acting non-trivially on the quaternion group Q of order 8. Observe that α exactly fixes the elements in $\mathbf{Z}(Q)$. We form the group $G := Q\langle\alpha\rangle \times L\langle a \rangle$ and take the normal subgroup $N = Q \times L$. It follows that $\mathrm{cs}_G(N) = \{1, 6, 36\}$ and, however, $N/\mathbf{Z}(N)$ is not a p-group and $\mathbf{F}(N)$ is not abelian either.

Acknowledgements This research is supported by the Valencian Government, Proyecto PROMETEO/2011/30, by the Spanish Government, Proyecto MTM2010-19938-C03-02 and the first author is also supported by grant Fundació Bancaixa P11B2010-47.

References

[1] Z. Akhlaghi, A. Beltrán, M. J. Felipe and M. Khatami, Normal subgroups and p-regular G-class sizes. J. Algebra **336** (2011), 236-241.

[2] Z. Akhlaghi, A. Beltrán, M. J. Felipe and M. Khatami, Structure of normal subgroups with three G-class sizes. Monatsh. Math. **167** (1) (2012), 1-12.

[3] Z. Akhlaghi, A. Beltrán, M. J. Felipe and M. Khatami, Finite p-solvable groups with three p-regular class sizes. Proc. Edinb. Math. Soc. **56** (2013), 371-386.

[4] Z. Akhlaghi, A. Beltrán A and M. J. Felipe, The influence of p-regular class sizes on normal subgroups. J. Group Theory **16** (2013), 585-593.

[5] Z. Akhlaghi, A. Beltrán and M. J. Felipe, Normal sections, class sizes and solvability of finite groups. J. of Algebra **399** (2014), 220-231.

[6] E. Alemany, A. Beltrán and M. J. Felipe, Nilpotency of normal subgroups having two G-class sizes. Proc. Amer. Math. Soc. **139** (2011), 2663-2669.

[7] E. Alemany, A. Beltrán and M. J. Felipe, Itô's theorem on groups with two class sizes revisited, Bull. Aust. Math. Soc. **85** (2012), 476-481.

[8] R. Baer, Einfache Partitionen endlicher Gruppen mit nicht-trivialer Fittingscher Untergruppe. (German) Arch. Math. (Basel) **12** (1961), 81-89.

[9] R. Baer, Partitionen endlicher Gruppen. (German) Math. Z. **75** (1961), 333-372.

[10] A. Beltrán and M. J. Felipe, Solvability of normal subgroups and G-class sizes, Publ. Math. Debrecen. In press.

[11] E. A. Bertram, M. Herzog and A. Mann, On a graph related to conjugacy classes of groups, Bull. London Math. Soc. **22** (1990), 569-575.

[12] R. Brandl, Finite groups all of whose elements are of prime power order. Boll. Unione Mat. Ital. A (5) **18** (1981), no 3, 491-493.

[13] A. R. Camina, Arithmetical conditions on the conjugacy class numbers of a finite group. J. Lond. Math. Soc. (2) **5** (1972), 127-132.

[14] A. R. Camina, Finite groups of conjugate rank 2, Nagoya Math. J. **53** (1974), 47-57.

[15] S. Dolfi and E. Jabara, The structure of finite groups of conjugate rank 2. Bull. Lond. Math. Soc. **41** (2009), 916-926.

[16] H. Heineken, On groups all of whose elements have prime power order. Math. Proc. R. Ir. Acad, **106A** (2) (2006), 191-198.

[17] G. Higman, Groups and rings having automorphisms without nontrivial fixed elements, J. Lond. Math. Soc. **32** (1957), 335-242.

[18] B. Huppert, *Character Theory of Finite Groups.* De Gruyter Expositions in Mathematics **25**. Berlin, New York 1998.

[19] I. M. Isaacs, Groups with many equal classes, Duke Math. J. **37** (1970), 501-506.

[20] I. M. Isaacs, *Finite Group Theory.* Graduate Studies in Mathematics **92**. American Mathematical Society, 2008.

[21] N. Itô, On finite groups with given conjugate type I. Nagoya Math. J. **6** (1953), 17-28.

[22] N. Itô, On finite groups with given conjugate types. II. Osaka J. Math. **7** (1970), 231-251.

[23] S. L. Kazarin, On groups with isolated conjugacy classes, Izv. Vyssh. Uchebn. Zaved. Mat. **7** (1981), 40-45.

[24] M. S. Lucido, Groups in which the prime graph is a tree, Bollettino U.M.I. **8** 5-B (2002), 131-148.

[25] J. Rebmann, F-Gruppen. (German) Arch. Math. (Basel) **22** (1971), 225-230.

[26] The GAP Group, GAP - Groups, Algorithms and Programming, Vers. 4.4.12 (2008). http://www.gap-system.org

[27] M. Suzuki, On a finite group with a partition. Arch. Math. (Basel) **12** (1961), 241-254.

[28] G. Zappa, Partitions and other coverings of finite groups. Special issue in honor of Reinhold Baer (1902-1979). Illinois J. Math. **47** 1-2 (2003), 571-580.

AUTOMORPHISM GROUPS OF COMPACT NON-ORIENTABLE RIEMANN SURFACES

E. BUJALANCE*, F.J. CIRRE*, J.J. ETAYO†, G. GROMADZKI§
and E. MARTÍNEZ‡

*Departamento de Matemáticas Fundamentales, Facultad de Ciencias, UNED, Paseo Senda del Rey, 9, Madrid 28040, Spain, supported by MTM2011-23092
Email: eb@mat.uned.es, jcirre@mat.uned.es

†Departamento de Álgebra, Facultad de Matemáticas, Universidad Complutense, Madrid 28040, Spain, supported by MTM2011-22435, UCM910444
Email: jetayo@mat.ucm.es

§Institute of Mathematics, Gdańsk University, Wita Stwosza 57, 80-952 Gdańsk, Poland, supported by NCN 2012/05/B/ST1/02171
Email: greggrom@mat.ug.edu.pl

‡Departamento de Matemáticas Fundamentales, Facultad de Ciencias, UNED, Paseo Senda del Rey, 9, Madrid 28040, Spain, supported by MTM2011-23092
Email: emartinez@mat.uned.es

Abstract

This paper is a survey on automorphism groups of compact unbordered non-orientable surfaces with dianalytic structures, called non-orientable Riemann surfaces. We deal with different aspects of groups acting on these surfaces, as the maximum order and minimum genus problems, the determination of the symmetric crosscap number of a finite group and the fixed point set of an automorphism. The guiding theme linking the sections of this survey comes from the Riemann uniformization theorem, which relates these surfaces to the combinatorial group theory of non-Euclidean crystallographic groups. Some other areas are also related to these surfaces, amongst which we mention real algebraic curves.

1 Introduction

It is easy to see that a classical analytic structure existing on orientable topological surfaces cannot exist on non-orientable ones. One can, however, relax a bit the notion of analyticity of a structure allowing the complex conjugation to be involved for transition functions between charts with overlapping domains. Such structures, called *dianalytic*, can exist on non-orientable closed surfaces. Sometimes they are also called *conformal* – motivated by the fact that the transition functions are angle-preserving but not necessarily orientation-preserving – while usually for classical structures of Riemann surfaces the term *analytic* is reserved. This way we obtain the notion of unbordered non-orientable Klein surface. These surfaces were called *non-orientable Riemann surfaces* by Singerman in [35], and we will keep here this convention. Similarly, a self-homeomorphism of such a surface is said to be an *automorphism* if its local forms are either analytic or the composite of complex conjugation with an analytic function.

The topic we present here is at the intersection of several important disciplines of mathematics: theory of functions of one complex variable (classical Riemann surfaces), algebraic geometry (algebraic curves), low-dimensional topology (mapping class groups, three-dimensional hyperbolic manifolds), Galois theory (automorphisms of algebraic function fields in one variable), differential geometry (moduli spaces, complex surfaces), combinatorics (regular maps on surfaces) and combinatorial group theory. This last is the main tool used to prove the results we present here, and that makes this survey suitable to this volume.

Such objects and their automorphisms are interesting by themselves and have a vast literature. Let us however, for the sake of completeness, explain more precisely their connection with algebraic geometry and Galois theory as mentioned above. Firstly, non-orientable Riemann surfaces can be seen as complex curves allowing defining equations over the reals $\mathbb{R}$ but having no $\mathbb{R}$-rational points. Such curves are called *purely imaginary real algebraic curves*. Next, in the theory of algebraic function fields in one variable over the reals these curves can be seen as not formally real fields in which -1 is not a square. (A field is formally real if -1 is not a sum of squares.) Moreover, their groups of automorphisms are the Galois groups of such fields of rational functions viewed as extensions over the reals.

The importance of such surfaces and their automorphism groups seems to have been perceived already by Felix Klein himself, but the foundations for their modern study have been given by Alling and Greenleaf in their monograph [1]. The principal tool comes from the Riemann uniformization theorem, which allows to relate the topology of compact surfaces with conformal structures and conformal actions on them, to the algebra of classical cocompact Fuchsian groups or, more generally, non-euclidean crystallographic groups. Combinatorial algebraic foundations for these groups had been established by Macbeath in the early sixties of the last century and developed later by Singerman, May and Bujalance, among others.

Throughout the paper, we will only deal with surfaces of topological genus bigger than 2, in order to assure finiteness of their automorphism groups.

2 Preliminaries

This is a survey of known results on the groups of automorphisms of non-orientable Riemann surfaces (for a survey on bordered Klein surfaces, see [5]). Our purpose is to give a general overview of the state of the art of this topic, and not to analyze in detail the techniques used in each paper. The interested reader is referred to the appropriate reference in the Bibliography. So in general proofs are omitted. Nevertheless we consider it appropriate to outline the general approach to be used in this area. The most fruitful technique turns out to be the combinatorial theory of *non-euclidean crystallographic groups* (NEC groups in short). The first presentations for NEC groups appeared in [38] and their structure was clarified by the introduction of signatures in [32].

An NEC group will mean here a discrete and co-compact subgroup of the group of all isometries of the hyperbolic plane $\mathcal{H}$ (including the orientation reversing ones). The *canonical Fuchsian subgroup* Γ^+ of an NEC group Γ comprises the orientation preserving isometries of Γ. We say that Γ is a *proper* NEC group if $\Gamma \neq \Gamma^+$. The

algebraic presentation of such a group Γ is well known and is concentrated in its so called *signature*. We shall not give here the general definition of the signature of an NEC group since it can be quite complicated. Just for some specific NEC groups, their signature and their presentation will be explicitly described.

A non-orientable Riemann surface S is canonically doubly covered by a Riemann surface R whose topological genus is said to be *the algebraic genus* of S. If S has algebraic genus $g \geq 2$ then it can be represented as the orbit space $\mathcal{H}/\Gamma$ of the hyperbolic plane under the action of an NEC group Γ with no non-trivial elements of finite order. In this situation we say that Γ is a *surface group* and that it uniformizes S. Such NEC group Γ has signature $(g; -; [-]; \{-\})$ where g is the topological genus of $\mathcal{H}/\Gamma$, and presentation

$$\langle\, d_1, \ldots, d_g \mid d_1^2 \ldots d_g^2 \,\rangle.$$

The orbit space $\mathcal{H}/\Gamma$ has algebraic genus $g-1$.

Elementary properties of covering spaces and of the mentioned Riemann double cover R allow to show that a finite group G acts as a group of dianalytic automorphisms of the non-orientable surface $\mathcal{H}/\Gamma$ if and only if there exist a proper NEC group Λ and an epimorphism $\theta : \Lambda \to G$ such that $\ker\theta = \Gamma$ and $\theta(\Lambda^+) = G$. Such epimorphisms with a surface group as a kernel will be called *smooth epimorphisms,* for short. Observe however that there is another way to see such surface and its group of automorphisms. Namely there is a fixed-point-free orientation-reversing involution σ of R so that the mentioned double covering $R \to S$ induces a conformal isomorphism between S and the orbit space R/σ. Consequently we obtain

$$\mathrm{Aut}(S) = \mathrm{C}_{\mathrm{Aut}^+(R)}(\sigma),$$

the centralizer of σ in the group $\mathrm{Aut}^+(R)$ of orientation-preserving automorphisms of R.

At the beginning of the eighties, the first author of this paper developed a combinatorial method to study the relation between the signatures of Λ and Γ in function of G. The method is described in Chapter 2 in [7] and it is an essential ingredient in most of the proofs of the results presented here.

The area $\mu(\Delta)$ of a fundamental region of an NEC group Δ depends only on the algebraic structure of the group itself and we have the following, crucial for our considerations, *Hurwitz-Riemann formula*:

$$[\Delta_1 : \Delta_2] = \frac{\mu(\Delta_2)}{\mu(\Delta_1)},$$

where Δ_2 is a subgroup of finite index in Δ_1.

3 Prescribed families of groups acting on non-orientable surfaces

In the beginning of the seventies, Singerman in [35] considered the problem of finding the *largest possible groups of automorphisms* of non-orientable surfaces. He showed that a necessary and sufficient condition for a finite group G to be a group of automorphisms of a non-orientable Riemann surface is the existence of a proper NEC group Λ and a smooth epimorphism $\theta : \Lambda \to G$ such that $\theta(\Lambda^+) = G$, where Λ^+ is

the canonical Fuchsian group of Λ. As a corollary, he showed that if G is a group of automorphisms of the non-orientable Riemann surface S, then G is (isomorphic to) a group of conformal automorphisms of the two-sheeted orientable covering surface of S. Using this, together with known results on Hurwitz groups, he showed that the largest order of a group of automorphisms of a non-orientable Riemann surface of topological genus g is $84(g-2)$. A group G attaining this bound is called H*-group, and it is characterized by the existence of three generators c_0, c_1, c_2 which satisfy the relations

$$(c_0)^2 = (c_1)^2 = (c_2)^2 = (c_0c_1)^2 = (c_1c_2)^3 = (c_0c_2)^7 = 1,$$

and such that c_0c_1 and c_1c_2 generate G. In particular, every H*-group is a Hurwitz group. The converse is not true, as Singerman showed that the linear fractional group PSL(2,7) (which is known to be Hurwitz) is not an H*-group. Other examples of Hurwitz groups which are not H*-groups are the alternating groups A_{21}, A_{22} and A_{29}, as Etayo and Martínez proved in [20]. Singerman also showed that there are infinitely many values of g for which the bound $84(g-2)$ is attained, and this led him to find new infinite families of Hurwitz groups.

Conditions under which $\mathrm{PSL}(2,q)$ is an H*-group were found later on by W. Hall in her unpublished thesis [30] as follows:

The group $\mathrm{PSL}(2,q)$ *is an* H**-group if and only if one of the following conditions hold:*

i) $q = p$ *is prime and* $p \equiv 1$ *or* $13 \pmod{28}$.

ii) $q = p$ *is prime,* $p \equiv -1$ *or* $-13 \pmod{28}$, *and* $(3 - \tau_i^2)$ *is a square for two values of* i *where* τ_1, τ_2, τ_3 *are the roots of* $\xi^3 + \xi^2 - 2\xi - 1 = 0$ *in* $\mathrm{GF}(q)$.

iii) $q = p^3$ *for* $p = 2$ *or* p *prime,* $p \equiv 5,\ 9,\ -3$ *or* $-11 \pmod{28}$.

The other result on simple H*-groups was obtained by Conder in [13] who proved that the alternating group A_n is an H*-group for $n > 167$, as well as for a given list of values of $n < 167$.

Using computational methods, Conder [14] found the list of the largest orders of a group of automorphisms of a non-orientable surface of given topological genus g, for g between 3 and 302. It turns out that the $84(g-2)$ bound is attained for very few genera in that range, namely for $g = 8$, 15 and 147 only.

Singerman's bound may be considered as a particular case of the general problem of finding the minimum genus of surfaces for which a given finite group G is a group of automorphisms. Bujalance in [2] carried out the preliminary step to this by studying the case of cyclic groups. He analyzed the possible signatures that a proper NEC group Γ may have in order for it to admit a smooth epimorphism onto a cyclic group. This allowed him to obtain a precise lower bound for the genus of a non-orientable surface with an automorphism of order n, in terms of the prime decomposition of n. As a corollary, he reproved Hall's bound in [30] for the largest order of an automorphism acting on a non-orientable surface of topological genus g, namely, $2g$ if g is odd, and $2g-2$ if g is even.

An interesting property of these surfaces (with a cyclic automorphism group of the largest possible order) is that they admit more automorphisms than those in the cyclic group, as proved in [6]. This problem is related to the *question of extendability* of

group actions, which can be stated as follows: given a group G acting on a surface S, does S admit more automorphisms than those in G or, on the contrary, G is the full group $\mathrm{Aut}(S)$ of all automorphisms of S? In the case of cyclic actions, this problem was considered in [4], where it was shown that if such a cyclic action is realised by means of a non-maximal NEC signature, then the action always extends. The special case where the full automorphism group is cyclic of the largest possible order (for given genus) was also considered. In addition, the smallest algebraic genus of a non-orientable surface on which a given cyclic group acts as the full automorphism group was determined. This is called the *full cross-cap genus* of the cyclic group, following the definition of the symmetric cross-cap number, see Section 6.

As said above, the largest cyclic group acting on a non-orientable surface of topological genus g has order $2g$ if g is odd and $2g-2$ if g is even. Bujalance, Gromadzki and Turbek in [9] showed that this order almost characterizes the group itself. In fact, if G is a group of order $2g$ or $2g-2$ then either G is cyclic (and this occurs if and only if the surface is hyperelliptic) or G is an extension of a cyclic group by C_2. It is worth mentioning that defining equations for all but one family of such surfaces were also obtained.

At the end of the eighties, the order of the largest *supersoluble* group of automorphisms acting on surfaces was known for orientable or bordered surfaces. Gromadzki in [27] completed the panorama by showing that if G is a supersoluble group acting on a non-orientable surface of topological genus g then $|G| \leq 12(g-2)$. The bound is sharp since he also showed that a necessary and sufficient condition for the existence of a such a surface having a supersoluble group of automorphisms of order $12(g-2)$ is that $g = 3^n + 2$ for some $n \geq 0$.

Analogous results for *soluble* groups were found by the same author in [28]. In this case the bound is $24(g-2)$, and an infinite series of values of g for which this bound is attained was also given. It is worth mentioning the relation found by Gromadzki between soluble groups attaining this bound and M*-groups, that is, maximal groups of automorphisms of bordered surfaces. More precisely, he found that a soluble group of order $24(g-2)$ acts on a non-orientable surface of topological genus g if and only if it can be viewed as an M*-group acting on a non-orientable bordered surface with maximal symmetry of algebraic genus $2g-3$ with $g-2$ boundary components. Other relations between H*-groups and M*-groups can be found in [5, Section 5].

It is well known that for each $g \geq 2$ there exists a compact Riemann surface of genus g with $8g+8$ automorphisms. Thus, if $\nu(g)$ denotes the largest number of automorphisms of a compact Riemann surface of genus g then the above shows that $\nu(g) \geq 8g+8$. The search of similar bounds for non-orientable surfaces has been considered by Conder, Maclachlan, Todorovic Vasiljevic and Wilson in [17], see also [36]. The authors first defined smooth epimorphisms onto the dihedral group of order $4g$ for each g odd, and onto a group of order $8(g-2)$ for each g even. So $\nu(g) \geq 4g$ if g is odd, while $\nu(g) \geq 8(g-2)$ if g is even. Next, they achieved improvements of the bounds for all $g \not\equiv 3 \pmod{12}$, namely, $\nu(g) \geq 8(g+2)$ for $g \equiv 1, 4, 7, 10 \pmod{12}$, while $\nu(g) \geq 8(g-2)$ for $g \equiv 0, 2, 5, 6, 8, 11 \pmod{12}$, and $\nu(g) \geq 6(g+1)$ for $g \equiv 9 \pmod{12}$. Finally, the authors also showed that these bounds are sharp for infinitely many values of g in each congruence class modulo 12.

4 Groups acting on prescribed families of non-orientable surfaces

Another direction in the study of automorphism groups of non-orientable Riemann surfaces is to fix a family of such surfaces and describe which groups act on them. As usual, one of the first examples to be considered is the family of hyperelliptic surfaces. A non-orientable Riemann surface S is hyperelliptic if it admits a dianalytic involution $\varphi : S \to S$ such that the orbit space $S/\langle\varphi\rangle$ has algebraic genus 0 (so in this case $S/\langle\varphi\rangle$ is either the projective plane or the closed disc). In [3], using non-euclidean crystallographic groups and Teichmüller spaces, the authors gave the complete list of full automorphism groups of hyperelliptic non-orientable surfaces which are double covers of the closed disc.

The case of double covers of the real projective plane was solved in [12] using a different approach. In this paper, the author directly worked with polynomial equations defining such double covers, and this allowed him to obtain explicit formulae of generators of the full automorphism groups.

A natural generalization of the notion of hyperellipticity is that of q-hyperellipticity: a non-orientable surface S is said to be q-hyperelliptic if it admits a dianalytic involution φ such that the orbit space $S/\langle\varphi\rangle$ has algebraic genus q. Hence, hyperelliptic means 0-hyperelliptic. J. A. Bujalance in [10] characterized q-hyperellipticity by means of NEC groups, and this allowed him to show that the q-hyperelliptic involution φ is unique and central in $\mathrm{Aut}(S)$ provided that $g > 4q + 2$, where g is the topological genus of S.

Under this assumption, the same author together with Estrada determined in [11] bounds for the order of the automorphism group of a non-orientable q-hyperelliptic surface S such that the orbit space $S/\langle\varphi\rangle$ has no boundary. They also proved that the bounds are attained.

5 On fixed points of automorphisms

An interesting feature of the fixed point set of an automorphism of a non-orientable surface is that it may contain closed curves. In the classical case of Riemann surfaces, Macbeath gave a formula to count the number of (isolated) points fixed by each non-identity element of a cyclic group of automorphisms. It was given in terms of the cyclic group and its universal covering transformation Fuchsian group. Izquierdo and Singerman in [31] showed that Macbeath's formula generalizes to non-orientable surfaces except when the element is an involution. In this case, in addition to isolated fixed points, the automorphism may fix curves (called ovals). The authors calculated the number of ovals and isolated fixed points of an involution in terms of the universal covering transformation NEC group. In addition, they also determined whether the ovals are twisted or not. (An oval is twisted if it has a Möbius band neighbourhood, and untwisted if it has an annular neighbourhood.)

The formulae given by Izquierdo and Singerman for cyclic groups were generalized by Gromadzki in [29] for an arbitrary finite group G of automorphisms of a non-orientable surface S. He gave a formula for the number of isolated points of an automorphism $f \in G$ in terms of the normalizer of f in G and the branched indices of the covering map $S \to S/G$. The same data is used to give another formula for the number of ovals of an involution.

6 The symmetric crosscap number of a group

Every finite group G may act as an automorphism group of non-orientable Riemann surfaces, [2, Theorem 2.5]. The minimum genus of these surfaces is called the symmetric crosscap number of G and it is denoted by $\tilde{\sigma}(G)$. The systematic study of the symmetric crosscap number was begun by C. L. May in [33], although previous results from other authors are also to be noted, see for instance [2], [19] and [26]. A very early result was obtained by W. Hall in [30]. As said above, she determined which groups $\mathrm{PSL}(2,q)$ attain the maximal bound $84(g-2)$.

Four types of problems arise naturally when dealing with the symmetric crosscap number $\tilde{\sigma}(G)$.

1) First of all, to obtain $\tilde{\sigma}(G)$ for any given group G, and for the groups belonging to an infinite family.
2) Second, to obtain $\tilde{\sigma}(G)$ for all groups G with $o(G) < n$ for a given (small) value of n.
3) Third, for a given value of g, to obtain all groups G such that $g = \tilde{\sigma}(G)$. Evidently this question is feasible only for low values of g.
4) Finally, to determine which values of g are in fact $\tilde{\sigma}(G) = g$ for a group G. The set of such values is called the symmetric crosscap spectrum and there exists a conjecture according to which $g = 3$ is the unique positive integer not belonging to the spectrum.

These four problems are intertwined and we will describe the present status of all of them.

Let $\mathcal{H}/\Gamma$ be a non-orientable Riemann surface on which G acts as an automorphism group. Then there exists another NEC group Λ such that $G = \Lambda/\Gamma$. From the Hurwitz-Riemann relation we have $g - 2 = o(G)|\Lambda|$, where $o(G)$ denotes the order of G and $|\Lambda| = \mu(\Lambda)/2\pi$ is the reduced area of Λ. Then

$$\tilde{\sigma}(G) \leq g = 2 + o(G)|\Lambda|,$$

and so to obtain the symmetric crosscap number is equivalent to find a group Λ and an epimorphism $\theta : \Lambda \to G$, such that $\Gamma = \ker(\theta)$ is a surface NEC group with $G = \theta(\Lambda^+)$ and minimal $|\Lambda|$.

The symmetric crosscap number of groups belonging to an infinite family. The symmetric crosscap number of groups belonging to several infinite families has been obtained. For Abelian groups of odd order it was calculated in [19] and this result was extended to all Abelian groups in [26] in the following terms.

Let G be an Abelian group different from C_n, $\mathrm{C}_2 \oplus \mathrm{C}_n$ (n even) and $\mathrm{C}_2 \oplus \mathrm{C}_2 \oplus \mathrm{C}_2$ (these groups have symmetric crosscap number 1 or 2, see below). We distinguish two cases. In the first one, G has non-cyclic 2-Sylow subgroup. The result is as follows:

Let G be an Abelian group of order N having non-cyclic 2-Sylow subgroup and suppose that $G = \mathrm{C}_2^s \oplus H$, where s is as big as possible and $H = \mathrm{C}_{m_1} \oplus \cdots \oplus \mathrm{C}_{m_k}$ is the canonical decomposition, where $m_1, \ldots, m_l$ are odd and $m_{l+1}, \ldots, m_k$ are even.

Then $(\tilde{\sigma}(G)-2)/N$ *is equal to*

$$\begin{cases} k-1-\displaystyle\sum_{i=1}^{k-s}\frac{1}{m_i} & \text{if } s-(k-l)\leq 0 \\ k-1 & \text{if } s-(k-l)=2l \\ k-1+\dfrac{s-k-l+1}{4} & \text{if } s-(k-l)>2l \\ k-1-\displaystyle\sum_{i=1}^{(k+l-s)/2}\frac{1}{m_i} & \text{if } 0<s-(k-l)<2l,\ s-(k-l) \text{ even} \\ k-1-\dfrac{1}{2m_{(k+l-s+1)/2}}-\displaystyle\sum_{i=1}^{(k+l-s-1)/2}\frac{1}{m_i} & \text{if } 0<s-(k-l)<2l,\ s-(k-l) \text{ odd.} \end{cases}$$

The second case covers the remaining possibilities:

Let G *be an Abelian group of order* N *having cyclic* 2*-Sylow subgroup, or* N *odd. Let* $G=\mathrm{C}_{m_1}\oplus\cdots\oplus\mathrm{C}_{m_k}$ *be its canonical decomposition. Then*

$$\tilde{\sigma}(G)=N\left(-1+\sum_{i=1}^{r}\left(1-\frac{1}{m_i}\right)\right)+2.$$

May obtained in [33] the symmetric crosscap number of dicyclic groups and Hamiltonian groups without odd order part. The result corresponding to the first family of groups is the following:

Let DC_n *be the dicyclic group of order* $4n$*. If* $n\neq 3$*, then* $\tilde{\sigma}(\mathrm{DC}_n)=2n+2$*. Besides,* $\tilde{\sigma}(\mathrm{DC}_3)=7$*.*

In [21], [22] and [23], the symmetric crosscap number of the groups $\mathrm{C}_m\oplus\mathrm{D}_n$, $\mathrm{D}_m\oplus\mathrm{D}_n$, $\mathrm{DC}_3\oplus\mathrm{C}_n$ and $A_4\oplus\mathrm{C}_n$ have been calculated. These families of groups whose orders are in arithmetic progressions are a useful tool for the study of the symmetric crosscap spectrum, see below.

The symmetric crosscap number of groups of small order. In [24] the symmetric crosscap number of all groups G of order less than 32 is obtained. For each group G the corresponding NEC group Λ and the epimorphism $\theta:\Lambda\to G$ are given.

M. Conder announced in a Conference in Castro Urdiales (2010) to have obtained these values for the groups G with $o(G)<128$ in terms of the `SmallGroupLibrary` identification. The result is still not published, but the list can be consulted in Conder's web page [15].

Groups with small symmetric crosscap number. The groups having symmetric crosscap numbers 1 and 2 have been classified by Tucker in [37]. The groups of genus 1 are C_n, D_n, A_4, S_4 and A_5. Of genus 2 we have $\mathrm{C}_2\oplus\mathrm{C}_n$ and $\mathrm{C}_2\oplus\mathrm{D}_n$, in both cases with n even. May proved that there is no group with symmetric crosscap number 3 in [33].

The groups with $\tilde{\sigma}(G) = 4$ or 5 have been obtained in [8]. There are two groups with symmetric crosscap number 4 which are $\mathrm{C}_2 \oplus A_4$ and $\mathrm{C}_2 \oplus S_4$. The eight groups G with $\tilde{\sigma}(G) = 5$ are $\mathrm{C}_3 \oplus \mathrm{C}_3$, $((3,3,3;2))$, $\mathrm{C}_3 \oplus \mathrm{D}_3$, $\langle 5,4,2\rangle$, $\mathrm{D}_3 \oplus \mathrm{D}_3$, $(4,4 \mid 2,3)$, $(2,4,6;2)$ and S_5.

In the above unpublished result by M. Conder, the groups G with $\tilde{\sigma}(G) \leq 65$ were also expressed by the `SmallGroupLibrary` identification, see [16].

The symmetric crosscap spectrum. A natural question arises: which numbers are the symmetric crosscap number of a group? The aforementioned results on families of groups are a useful tool in order to cover most of the numbers. In particular, the groups $\mathrm{C}_m \oplus \mathrm{D}_n$ cover all numbers of the forms $4k$, $4k+1$ and $4k+2$, [21, Prop. 3.2]. So only the numbers $4k+3$ with $k \geq 1$ need to be studied. For each n congruent to 11 modulo 12, there exists an Abelian group G with $\tilde{\sigma}(G) = n$, see [26, Prop. 6.2] and [21, Prop. 3.3].

Finally, concerning the numbers congruent to 3 or 7 modulo 12, the groups $\mathrm{DC}_3 \oplus \mathrm{C}_n$ and $A_4 \oplus \mathrm{C}_n$ cover nine classes modulo 144 among the numbers congruent to 7 modulo 12, as well as the numbers congruent to 39, 87 and 135 modulo 144, which are of the form $12k+3$, [23]. The whole set $12k+7$ has been dealt by Conder in an unpublished work, by means of a subgroup of a semidirect product $\mathrm{C}_{3n} \rtimes S_4$. So that, the present state of the question is summarized in the following:

Let $S = \{3, 15, 27, 51, 63, 75, 99, 111, 123\}$. If g is a number non-congruent to x modulo 144, for $x \in S$, there exists a group G such that $\tilde{\sigma}(G) = g$.

It remains to prove the conjecture that 3 is actually the unique gap. In order to enforce this conjecture we know the following facts. First, Etayo and Martínez exhibit in [23] a group $\mathrm{PSL}(2,p)$ whose symmetric crosscap number is congruent with x modulo 144 for each $x \in S$. Besides, infinitely many numbers in six of those nine classes are proved to be the symmetric crosscap number of groups $(3,3 \mid 3,k)$ in [24]. Finally, in [25] the same authors have found groups for all remaining numbers g up to 206. However, the procedure does not use a family of groups which may fill all the gaps. All in all, this information supports the following

Conjecture: For all $n \neq 3$ there exists a group G such that $\tilde{\sigma}(G) = n$.

Related results were recently obtained for the corresponding parameters on orientable surfaces. The strong symmetric genus $\sigma^0(G)$ is the least genus of an orientable Riemann surface on which the group G acts conformally. May and Zimmerman proved in [34] that for every n there is a group G which $\sigma^0(G) = n$. If one allows anticonformal automorphisms, the parameter is called the symmetric genus. Its spectrum has been studied by Conder and Tucker in [18]. They prove that the possible gaps of the spectrum are numbers congruent with 8 or 14 (mod 18), and also conjecture that there is no gap at all.

Both conjectures could be related, but a complete proof of them seems to be difficult.

The authors wish to thank the anonymous referee for his/her helpful suggestions which have contributed to improve the final version.

References

[1] N. L. Alling and N. Greenleaf, Foundations of the theory of Klein surfaces, Lecture Notes in Math. **219**, Springer-Verlag (1971).

[2] E. Bujalance Cyclic groups of automorphisms of compact non-orientable Klein surfaces without boundary, *Pacific J. Math.* **109** (1983), no. 2, 279–289.

[3] E. Bujalance, J. A. Bujalance, G. Gromadzki and E. Martínez, The groups of automorphisms of non-orientable hyperelliptic Klein surfaces without boundary, in *Groups Korea 1988 (Pusan, 1988)*, Lecture Notes in Math. **1398**, (Springer, Berlin, 1989), 43–51.

[4] E. Bujalance, F. J. Cirre and M. D. E. Conder, Extensions of finite cyclic group actions on non-orientable surfaces, *Trans. Amer. Math. Soc.* **365** (2013), 4209–4227.

[5] E. Bujalance, F. J. Cirre, J. J. Etayo, G. Gromadzki and E. Martínez, A survey on the minimum genus and maximum order problems for bordered Klein surfaces, in *Groups St Andrews 2009 in Bath, Vol. 1*, London Math. Soc. Lecture Note Ser. **387**, (CUP, Cambridge, 2011), 161–182.

[6] E. Bujalance, F. J. Cirre and P. Turbek, Riemann surfaces with real forms which have maximal cyclic symmetry, *J. Algebra* **283** (2005), no. 2, 447–456.

[7] E. Bujalance, J. J. Etayo, J. M. Gamboa and G. Gromadzki, *Automorphisms Groups of Compact Bordered Klein Surfaces,* Lecture Notes in Math. **1439**, (Springer-Verlag, Berlin, 1990).

[8] E. Bujalance, J. J. Etayo and E. Martínez, The full group of automorphisms of non-orientable unbordered Klein surfaces of topological genus 3, 4 and 5, *Rev. Mat. Compl.* (to appear).

[9] E. Bujalance, G. Gromadzki and P. Turbek, On non-orientable Riemann surfaces with $2p$ or $2p+2$ automorphisms, *Pacific J. Math.* **201** (2001), no. 2, 267–288.

[10] J. A. Bujalance, Hyperelliptic compact non-orientable Klein surfaces without boundary, *Kodai Math. J.* **12** (1989), no. 1, 1–8.

[11] J. A. Bujalance and B. Estrada, q-hyperelliptic compact non-orientable Klein surfaces without boundary, *Int. J. Math. Math. Sci.* **31** (2002), no. 4, 215–227.

[12] F. J. Cirre, Automorphism groups of real algebraic curves which are double covers of the real projective plane, *Manuscripta Math.* **101** (2000), no. 4, 495–512.

[13] M. D. E. Conder, Generators for alternating and symmetric groups, *J. London Math. Soc. (2)* **22** (1980), 75–86.

[14] M. D. E. Conder, http://www.math.auckland.ac.nz/∼conder/MaximumGroupOrdersByGenus-nonorientable.txt

[15] M. D. E. Conder, http://www.math.auckland.ac.nz/∼conder/CrosscapNumberSmallGroups127.txt

[16] M. D. E. Conder, http://www.math.auckland.ac.nz/∼conder/GroupsWithCrosscapNumber3to65.txt

[17] M. D. E. Conder, C. Maclachlan, S. Todorovic Vasiljevic and S. Wilson, Bounds for the number of automorphisms of a compact non-orientable surface, *J. London Math. Soc. (2)* **68** (2003), no. 1, 65–82.

[18] M. D. E. Conder and T W Tucker, The symmetric genus spectrum of finite groups, *Ars Math. Contemp.* **4** (2011), 271–289.

[19] J. J. Etayo, Sobre grupos de automorfismos de superficies de Klein, Doctoral Thesis, Universidad Complutense, 1983.

[20] J. J. Etayo and E. Martínez, Alternating groups, Hurwitz groups and H*-groups, *J. Algebra* **283** (2005), no. 1, 327–349

[21] J. J. Etayo and E. Martínez, The symmetric crosscap number of the groups $C_m \times D_n$, *Proc. Royal Soc. Edinburgh* **138 A** (2008), 1197–1213.

[22] J. J. Etayo and E. Martínez, The action of the groups $D_m \times D_n$ on unbordered Klein surfaces, *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM* **105** (2011), 97–108.

[23] J. J. Etayo and E. Martínez, The symmetric crosscap number of the groups $DC_3 \times C_n$ and $A_4 \times C_n$, *Houston J. Math.* **38** (2012), 345–358.
[24] J. J. Etayo and E. Martínez, The symmetric crosscap number of the groups of small order, *J. Algebra Appl.* **12** (2013), no. 2, 1250164, 16 pp.
[25] J. J. Etayo and E. Martínez, Filling the gaps of the symmetric crosscap spectrum, (preprint).
[26] G. Gromadzki, Abelian groups of automorphisms of compact non-orientable Klein surfaces without boundary, *Comment. Math. Prace Mat.* **28** (1989), no. 2, 197–217.
[27] G. Gromadzki, Supersoluble groups of automorphisms of non-orientable Riemann surfaces, *Bull. London Math. Soc.* **22**, (1990), no. 6, 561–568.
[28] G. Gromadzki, On soluble groups of automorphisms of non-orientable Klein surfaces, *Fund. Math.* **141** (1992), no. 3, 215–227.
[29] G. Gromadzki, On fixed points of automorphisms of non-orientable unbordered Klein surfaces, *Publ. Mat.* **53** (2009), no. 1, 73–82.
[30] W. Hall, Automorphisms and coverings of Klein surfaces, PhD Thesis, University of Southampton, 1977. http://eprints.soton.ac.uk/259986/
[31] M. Izquierdo and D. Singerman, On the fixed-point set of automorphisms of non-orientable surfaces without boundary, in *The Epstein birthday schrift,* Geom. Topol. Monogr. **1**, Geom. Topol. Publ., Coventry, (1998), 295–301.
[32] A. M. Macbeath, The classification of non-euclidean plane crystallographic groups, *Canad. J. Math.* **19** (1967), 1192–1205.
[33] C. L. May, The symmetric crosscap number of a group, *Glasgow Math. J.* **41** (2001), 399–410.
[34] C. L. May and J. Zimmerman, There is a group of every strong symmetric genus, *Bull. London Math. Soc.* **35** (2003), 433–439.
[35] D. Singerman, Automorphisms of compact non-orientable Riemann surfaces, *Glasgow Math. J.* **12** (1971), 50–59.
[36] S. Todorovic Vasiljevic, Bounds on the number of automorphisms of a compact non-orientable surface, PhD Thesis, University of Auckland, 2001.
[37] T. W. Tucker, Symmetric embeddings of Cayley graphs in non-orientable surfaces, in *Graph Theory, Combinatorics and Applications*, eds I. Alavy et al., 1991, 1105–1120.
[38] H. C. Wilkie, On non-Euclidean crystallographic groups, *Math. Z.* **91** (1966), 87–102.

WHAT ARE THE $\mathcal{C}_2$-GROUPS?

INNA CAPDEBOSCQ* and CHRISTOPHER PARKER†

*Mathematics Institute, Zeeman Building, University of Warwick, Coventry, CV4 7AL
Email: I.Capdeboscq@warwick.ac.uk

†School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT
Email: c.w.parker@bham.ac.uk

The Classification of Finite Simple Groups (CFSG) is a theorem that states that if G is a finite simple group, then it is either:

(i) A cyclic group of prime order,

(ii) A an alternating group $\mathrm{Alt}(n)$ for some $n \geq 5$,

(iii) A simple group of Lie type, or

(iv) A sporadic simple group.

The original proof of the CFSG was a patchwork of interrelated individual theorems with spectacular contributions by many mathematicians. It was finally completed in 2004 with the monumental work of M. Aschbacher and S. Smith, which classified the so-called quasithin groups [AS].

The project of D. Gorenstein, R. Lyons and R. Solomon (the GLS-project) aims to produce what is known as the Generation-2 proof of the Classification Theorem. The outcome of this impressive work will be a new, coherent proof of the CFSG and it will be (and is being) published in the special series of monographs by the AMS [GLS1].

The basic strategy of the proof of the CFSG is to consider a simple group X which is a minimal counterexample to the theorem. This is a finite group all of whose proper simple subsections are listed in the statement of the CFSG (in general, a group satisfying this condition is called a $\mathcal{K}$-*proper* group). Thanks to the celebrated result of W. Feit and J.G. Thompson [FT], one pays special attention to the centralisers of involutions (elements of order 2) of X. According to the structure of those centralisers, the group X is always either of *even type* or of *odd type* (this is by definition [GLS1, p.58]). Thus the proof of CFSG is reduced to the classification of groups of even type and classification of groups of odd type. While the classification of groups of odd type is currently nearing its completion in the work of Lyons and Solomon [GLS5], [GLS6], [GLS7], the classification of groups of even type remains to be researched. Let us therefore discuss this important notion ([GLS1, 21.3]). We abide by the definitions in [GLS3], in particular, the description of the automorphisms of groups of Lie type (which is given in [GLS3, 2.5.13]) is especially important to us.

First of all it is useful to consider the structure of a finite group G in general. Recall that a group is H called *quasisimple* provided H is *perfect*, that is $H = [H, H]$, and $H/Z(H)$ is simple. For any finite group G, a *component* of G is a subnormal quasisimple subgroup. It is a fact that follows from the three subgroups lemma and induction, that any two distinct components commute. The *layer* of G is

$$E(G) = \langle H \mid H \text{ is a component of } G \rangle.$$

Thus, if $L_1, L_2, \ldots, L_n$, are the distinct components of G, then

$$E(G) = L_1 \circ L_2 \circ \cdots \circ L_n$$

where $\circ$ denotes a commuting product.

The *Fitting subgroup* of G, $F(G)$, is better known. It is the largest nilpotent normal subgroup of G. Recall that the largest odd order normal subgroup of a group H is denoted by $O(H)$ and $O_2(H)$ is the largest normal 2-subgroup of H. Thus $F(G) = O_2(G) \times O(F(G))$.

The *generalized Fitting subgroup* of G is the product

$$F^*(G) = F(G)E(G) = F(G) \circ L_1 \circ L_2 \circ \cdots \circ L_n.$$

The most fundamental property of $F^*(G)$ is that $C_G(F^*(G)) = Z(F^*(G))$. It is the structure of $F^*(C_G(x))$, for $x \in G$ an involution, that dictates whether or not G has odd or even type.

Definition 1.1 A $\mathcal{K}$-proper simple group G is said to be of *even type*[1] if and only if the following conditions hold:

(i) G contains an elementary abelian subgroup of order at least 8.

(ii) if $x \in G$ is an involution, then $C_G(x)$ has no non-trivial odd order normal subgroups and

$$F^*(C_G(x)) = O_2(C_G(x))E(C_G(x)) = O_2(C_G(x)) \circ L_1 \circ \ldots \circ L_n$$

where each L_i is a component of $C_G(x)$ and $L_i \in \mathcal{C}_2$ for $1 \leq i \leq n$.

If G does not contain an elementary abelian subgroup of order 8, then the structure of a Sylow 2-subgroup of G is very limited. The determination of the simple groups without an elementary abelian subgroup of order 8 was presented in [ABG] as a culmination of many significant contributions especially including work of Walter (on abelian and dihedral Sylow 2-subgroups) and Lyons for the characterisation of $\mathrm{PSU}_3(4)$. This theorem is proved in volume six of the GLS-project and is subsumed into Theorem $\mathcal{C}_2^*$ [GLS6]. We distill it here as follows.

Theorem 1.2 *If G is a finite simple group with no elementary abelian subgroup of order 8, then one of the following holds:*

(i) G has dihedral Sylow 2-subgroups and $G \cong \mathrm{PSL}_2(q)$, q odd, or $\mathrm{Alt}(7)$;

(ii) G has semidihedral (quasidihedral) Sylow 2-subgroups and $G \cong \mathrm{PSL}_3(q)$, $q \equiv -1 \pmod 4$, or $\mathrm{PSU}_3(q)$, $q \equiv 1 \pmod 4$ or M_{11};

(iii) G has wreathed Sylow 2-subgroups and $G \cong \mathrm{PSL}_3(q)$, $q \equiv 1 \pmod 4$, or $\mathrm{PSU}_3(q)$, $q \equiv -1 \pmod 4$; or

(iv) $G \cong \mathrm{PSU}_3(4)$.

In definition of groups of even type condition 1.1(ii) involves the mysterious set $\mathcal{C}_2$ which is a part of the title of this paper. We now populate this set. Recall that Lie(2) consists of the groups of Lie type defined in characteristic 2.

[1] This is equivalent to the Definition 21.3 given in [GLS1]

Definition 1.3 [2] A quasisimple group $\mathcal{K}$-group is a *$\mathcal{C}_2$-group* if and only if K has no odd order normal subgroup and one of the following holds:

(i) $K \in \mathrm{Lie}(2)$ or $K \cong 2^{\cdot 2}\mathrm{B}_2(8)$, $2^{2\cdot 2}\mathrm{B}_2(8)$, $2^{\cdot}\mathrm{Sp}_6(2)$, $2^{\cdot}\mathrm{PSU}_6(2)$, $2^{2\cdot}\mathrm{PSU}_6(2)$, $2^{\cdot}\mathrm{PSL}_3(4)$, $2^{2\cdot}\mathrm{PSL}_3(4)$, $2^{\cdot}\mathrm{PSU}_4(3)$, $2^{\cdot}\Omega_8^+(2)$, $2^{2\cdot}\Omega_8^+(2)$, $2^{\cdot}\mathrm{G}_2(4)$, $2^{\cdot}\mathrm{F}_4(2)$, $2^{\cdot 2}\mathrm{E}_6(2)$, $2^{2\cdot 2}\mathrm{E}_6(2)$;

(ii) $K \cong \mathrm{PSL}_2(q)$ where q is a Fermat or Mersenne prime;

(iii) $K \cong \mathrm{PSL}_3(3)$, $\mathrm{PSL}_4(3)$, $\mathrm{PSU}_4(3)$, $\mathrm{G}_2(3)$, $2^{\cdot}\mathrm{PSU}_4(3)$; or

(iv) $K \cong \mathrm{M}_{11}$, M_{12}, $2^{\cdot}\mathrm{M}_{12}$, M_{22}, $2^{\cdot}\mathrm{M}_{22}$, $4^{\cdot}\mathrm{M}_{22}$, M_{23}, M_{24}, J_2, $2^{\cdot}\mathrm{J}_2$, J_3, J_4, HS, $2^{\cdot}$HS, Suz, $2^{\cdot}$Suz, Ru, $2^{\cdot}$Ru, Co_1, $2^{\cdot}\mathrm{Co}_1$, Co_2, Fi_{22}, $2^{\cdot}\mathrm{Fi}_{22}$, Fi_{23}, Fi'_{24}, Th, B, $2^{\cdot}$B or M.

The definition of $\mathcal{C}_2$ at first sight looks technical and complicated. Let us explain the motivation for this definition and so reveal its inner coherence.

Suppose for a moment that $G \in \mathrm{Lie}(2)$ and α is its automorphism of order 2. If α is inner, then a straightforward consequence of the Borel-Tits Theorem [GLS3, Theorem 3.1.3] is that $F^*(C_G(\alpha)) = O_2(C_G(\alpha))$ and if α is an outer automorphism of G, then [GLS3, Propositions 4.9.1 and 4.9.2] imply that $F^*(C_G(\alpha)) \in \mathrm{Lie}(2)$. Let us try to distinguish such a group from say $K = \mathrm{G}_2(3)$. In this case there is a unique conjugacy class of inner involutions α. For such involutions we have $C_K(\alpha) \cong (\mathrm{SL}_2(3) \circ \mathrm{SL}_2(3)) : 2$ and so $F^*(C_K(\alpha)) = O_2(C_K(\alpha))$. As for the outer automorphisms of order 2, we have $F^*(C_K(\alpha)) \cong \mathrm{PSL}_2(8) \in \mathrm{Lie}(2)$. Likewise, for the other $\mathcal{C}_2$-groups defined in characteristic 3, the fact that $F^*(\mathrm{SL}_2(3))$ is a 2-group , $\mathrm{PSL}_2(9) \cong \mathrm{PSp}_4(2)'$, $\mathrm{PSp}_4(3) \cong \mathrm{PSU}_4(2)$ and $\mathrm{PSU}_3(3) \cong \mathrm{G}_2(2)'$ reveals a similar picture but in these cases without the distinction between inner and outer involutions. Inductively, allowing components in $\mathcal{C}_2$, we have revealed what is almost a common feature of the groups in $\mathcal{C}_2$, the exceptions being the groups $\mathrm{PSL}_2(q)$ (for q a Fermat or Mersenne prime or 9) and $\mathrm{PSL}_3(4)$ with α acting as a graph-field automorphism. Such exceptions have the centralizer of an outer automorphism of order 2 containing an odd order non-trivial normal subgroup. We record this "internal closure" observation (with the exceptions) as follows.

Proposition 1.4 *Let K be a $\mathcal{C}_2$-group and α be an automorphism of order 2. Then exactly one of the following holds.*

(i) $F^*(C_K(\alpha)) = O_2(C_K(\alpha))$ or $F^*(C_K(\alpha)) = O_2(C_K(\alpha)) \circ L$ with $L \in \mathcal{C}_2$.

(ii) $K \cong \mathrm{PSL}_2(q)$, q a Fermat or Mersenne prime or 9, α is an outer automorphism of K which is inner-diagonal and $O(C_K(\alpha))$ is a non-trivial cyclic group.

(iii) $K/Z(K) \cong \mathrm{PSL}_3(4)$, α is a graph-field automorphism of K and $O(C_K(\alpha))$ is elementary abelian of order 9.

A further common feature of $\mathcal{C}_2$-groups is the following statement.

Lemma 1.5 *Suppose that $K \in \mathcal{C}_2$ and $R \in \mathrm{Syl}_2(K)$. Then $C_{\mathrm{Aut}(K)}(R)$ is a 2-group.*

[2] This is equivalent to Definition 12.1 of [GLS1]

Proof Suppose first that $K/O_2(K)$ is in Lie(2) (including $K \cong \mathrm{PSL}_2(9)$). Assume that $\beta \in \mathrm{Aut}(K)$ has odd order and centralizes R. Set $X = \mathrm{Inn}(K)\langle\beta\rangle$. Then

$$F^*(N_X(R/Z(K))) = O_2(N_X(R/Z(K))) \leq R/Z(K)$$

is a 2-group by [GLS3, Corollary 3.1.4]. Since β centralizes R, we now have $\beta \in F^*(N_X(R/Z(K)))$ which is nonsense. Thus the result is true for $K/Z(K) \in$ Lie(2).

The remaining members K of $\mathcal{C}_2$ have $\mathrm{Out}(K)$ a 2-group. So we only need to consider $C_K(R)$. Let $t \in R$ be such that $tZ(K) \in Z(R/Z(K))$ is an involution. Clearly $C_K(R) \leq C_K(t)$.

Suppose that $K \cong \mathrm{PSL}_2(q)$ where q is a Fermat or Mersenne prime. Since $C_K(t)$ is a dihedral 2-group by [GLS3, Theorem 4.5.1], the result follows.

If $K/Z(K) \cong \mathrm{PSL}_3(3)$, $\mathrm{PSL}_4(3)$, $\mathrm{PSU}_4(3)$ or $\mathrm{G}_2(3)$, then [GLS3, Tables 4.5.1 and 4.5.2] yield $F^*(C_K(t)) = O_2(C_K(t))$ and so the result holds in this case.

Finally, suppose that $K/Z(K)$ is a sporadic simple group. Since $C_K(R) \leq C_K(t)$, the result now follows by examination of of [GLS3, Tables 5.3] and their "notes". □

Recall that for a prime number r and group H, $m_r(H)$ denotes the minimal number of generators of a maximal abelian r-subgroup of H. We call $m_r(H)$ the *r-rank* of H. A *2-local subgroup* of G is, by definition, the normalizer of a non-trivial 2-subgroup of G. The measure of largeness used in the CFSG (called the *rank* of G) $e(G)$, is the maximum of $m_r(H)$ as H runs through the 2-local subgroups of G and r runs through the odd primes dividing $|G|$. Rank 1 and rank 2 groups are known as quasithin groups and were the subject of the aforementioned monograph of Aschbacher and Smith. Even type groups of rank at least 4 are the subject of current study of Lyons and Solomon. This leaves groups of even type and rank 3 as an outstanding problem. In such groups, the only components which can appear in the centralizer of an involution are the $\mathcal{C}_2$-group that have r-rank at most 3 for all odd primes r. We we will call this subset of $\mathcal{C}_2$-groups, ${}^3\mathcal{C}_2$-groups.

The remainder of this paper is devoted to cataloguing various properties of ${}^3\mathcal{C}_2$-groups which are exploited in our work on rank 3 groups.

Lemma 1.6 *Let K be an element of ${}^3\mathcal{C}_2$. Then the isomorphism type of K and the r-rank of K for various primes r are listed in Table 1. Moreover, for every $K \in {}^3\mathcal{C}_2$ and an odd prime r, one has:*

$$m_r(\mathrm{Aut}(K)) \leq m_r(K) + 1.$$

The notation $\mathcal{FM}$ stands for the set of Fermat and Mersenne primes, and for $a \in \mathbb{N}$ with $a \geq 1$, $a_+ := \frac{3+(-1)^a}{2}$ and $a_- := \frac{3-(-1)^a}{2}$.

Proof To obtain the complete list of elements of ${}^3\mathcal{C}_2$ and to complete Table 1, we need to determine the r-ranks of the members of $\mathcal{C}_2$.

Suppose first that $K \cong \mathrm{PSL}_2(q)$ for $q \in \mathcal{FM}$, $q = 9$ or $q = 2^a$. Then [G, Lemma 15.1.1] yields a precise statement about $m_2(K)$, $m_3(K)$ and $m_r(K)$. These are listed in Table 1 and we have $K \in {}^3\mathcal{C}_2$. We continue to examine all the other elements of $\mathcal{C}_2$.

Suppose that $K \cong \mathrm{PSL}_3(3)$. Then $m_3(K) = 2$ by [GLS3, Theorem 3.1.3]. Since $|K| = 2^4 \cdot 3^3 \cdot 13$, $K \in {}^3\mathcal{C}_2$. Now Theorem 1.2 gives us that $m_2(K) = 2$. If $K \cong$

K	$m_2(K)$	$m_3(K)$	$m_r(K)$, $r \geq 5$
$\mathrm{PSL}_2(q), q \in \mathcal{FM}$	2	1	≤ 1
$\mathrm{PSL}_2(9) \cong \mathrm{PSp}_4(2)'$	2	2	≤ 1
$\mathrm{PSL}_3(3)$	2	2	≤ 1
$\mathrm{PSU}_3(3) \cong G_2(2)'$	2	2	≤ 1
$\mathrm{PSL}_2(2^{a+2})$	$a+2$	1	≤ 1
$\mathrm{PSL}_3(2^a)$	$2a$	a_+	≤ 2
$\mathrm{PSU}_3(2^a)$	a	a_-	≤ 2
$\mathrm{PSL}_4(2^a)$	$4a$	$1+a_+$	≤ 3
$\mathrm{PSU}_4(2^a)$	$4a$	$1+a_-$	≤ 3
$\mathrm{PSL}_5(2)$	6	2	≤ 1
$\mathrm{PSU}_5(4)$	8	2	≤ 3
$\mathrm{PSL}_6(2)$	9	3	≤ 2
$\mathrm{PSL}_7(2)$	12	3	≤ 2
$\mathrm{PSp}_4(2^{a+1})$	$3(a+1)$	2	≤ 2
$\mathrm{PSp}_6(2^a)$	$6a$	3	≤ 3
$\Omega_8^-(2^a)$	$6a$	3	≤ 3
${}^2\mathrm{B}_2(2^{2a+1})$	a	0	≤ 1
$\mathrm{G}_2(2^{a+1})$	$3(a+1)$	2	≤ 2
${}^2\mathrm{F}_4(2)'$	5	2	≤ 2
${}^2\mathrm{F}_4(2^{2a+1})$	$5a$	2	≤ 2
${}^3\mathrm{D}_4(2^a)$	$5a$	2	≤ 2
$2\cdot\mathrm{PSL}_3(4)$	5	2	≤ 1
$2^2\cdot\mathrm{PSL}_3(4)$	6	2	≤ 1
$2\cdot\mathrm{Sp}_6(2)$	4	3	≤ 2
$2\cdot{}^2\mathrm{B}_2(8)$	4	0	≤ 1
$2^2\cdot{}^2\mathrm{B}_2(8)$	5	0	≤ 1
$2\cdot\mathrm{G}_2(4)$	5	2	≤ 2
M_{11}	2	2	≤ 1
$\mathrm{M}_{12}/2\cdot\mathrm{M}_{12}$	3/4	2	≤ 1
$\mathrm{M}_{22}/2\cdot\mathrm{M}_{22}$	4/5	2	≤ 1
$4\cdot\mathrm{M}_{22}$	4	2	≤ 1
M_{23}	4	2	≤ 1
M_{24}	6	2	≤ 1
$\mathrm{J}_2/2\cdot\mathrm{J}_2$	4/3	2	≤ 2
J_3	4	3	≤ 1
J_4	11	2	≤ 2
$\mathrm{HS}/2\cdot\mathrm{HS}$	4/5	2	≤ 2
$\mathrm{Ru}/2\cdot\mathrm{Ru}$	6/7	2	≤ 2

Table 1. The Ranks of K

$\mathrm{PSU}_3(3)$, then $m_3(K) = 2$ by [GLS3, Theorem 3.1.3]. Since $|K| = 2^5 \cdot 3^3 \cdot 7$, $K \in {}^3\mathcal{C}_2$ and by Theorem 1.2, $m_2(K) = 2$. If $K/Z(K) \cong \mathrm{PSL}_4(3), \mathrm{PSU}_4(3)$ or $\mathrm{G}_2(3)$, $m_3(K) \geq 4$ by [GLS3, Theorem 3.1.3]. Hence, $K \notin {}^3\mathcal{C}_2$.

Suppose now that K is a simple group of Lie type in characteristic 2. We will use the results of Section 4.10 of [GLS3] to establish the r-ranks of K for odd primes r.

Recall that $\Phi_m(x)$ is the mth cyclotomoic polynomial. Decomposing the order of K into a product of such polynomials is necessary when applying the results from [GLS3].

We begin by considering $K \cong \Omega_8^+(q)$. Then

$$|K| = q^{12}\Phi_1^4(q)\Phi_2^4(q)\Phi_3(q)\Phi_4^2(q)\Phi_6(q)$$

and so applying of [GLS3, Theorem 4.10.3], we obtain that $m_r(K) = 4$ for an odd prime r which divides $q+1$ or $q-1$. Thus K and any group containing a subgroup isomorphic to K is not contained in ${}^3\mathcal{C}_2$. In particular, the groups $\mathrm{PSp}_{2n}(q)$, $\Omega_{2n}^{\pm}(q)$ for $n \geq 5$ are not contained in ${}^3\mathcal{C}_2$.

Since B_4 is a subdiagram of extended Dynkin diagram F_4 and $\mathrm{B}_4(2^a) = \mathrm{C}_4(2^a)$, $\mathrm{F}_4(q)$ contains $\mathrm{PSp}_8(q)$ and so this group is also not in ${}^3\mathcal{C}_2$. We may exclude $\mathrm{E}_6(q)$ and ${}^2\mathrm{E}_6(q)$ from ${}^3\mathcal{C}_2$ as, by [GLS3, Proposition 4.9.2], $\mathrm{F}_4(q)$ is a subgroup of the fixed points of a graph automorphism of those groups. Because $\mathrm{E}_6(q)$ is contained in $\mathrm{E}_7(q)$ and $\mathrm{E}_8(q)$ these groups are not in ${}^3\mathcal{C}_2$.

Therefore out of the families of exceptional, symplectic and orthogonal groups the only possible surviving families are ${}^2\mathrm{B}_2(2^a)$, $\mathrm{G}_2(2^a)$, $\mathrm{PSp}_4(2^a)$, $\mathrm{PSp}_6(2^a)$, $\Omega_8^-(2^a)$, ${}^3\mathrm{D}_4(2^a)$ and ${}^2\mathrm{F}_4(2^a)'$ and we shall return to these groups at a later stage in the proof where we determine the various r-ranks.

Let us now study the linear groups. Suppose first that $K \cong \mathrm{PSL}_5(q)$ with $q > 2$. Then

$$|K| = q^{10}\Phi_1^4(q)\Phi_2^2(q)\Phi_3(q)\Phi_4(q)\Phi_5(q)/(q-1,5).$$

Consider the equation $\Phi_1(q) = 2^a - 1 = 5^b$ for $b \in \mathbb{N}$. Since $a \geq 2$, taking the identity modulo 4, we get a contradiction. Thus there exists prime $r \neq 5$ that divides $q-1$. Now [GLS3, Theorem 4.10.3] yields $m_r(K) = 4$ for such a prime. Thus $K \notin {}^3\mathcal{C}_2$ and so neither are the groups containing K. In particular, if $n \geq 5$ and $q > 2$, then $\mathrm{PSL}_n(q)$ is not a member of ${}^3\mathcal{C}_2$.

Suppose that $q = 2$. Since $\mathrm{PSL}_8(2) \geq \mathrm{PSp}_8(2)$, for $n \geq 8$, $\mathrm{PSL}_n(2)$ is not a member of ${}^3\mathcal{C}_2$. Thus the only possible linear groups in ${}^3\mathcal{C}_2$ are $\mathrm{PSL}_n(2^a)$ for $n \leq 4$ and $\mathrm{PSL}_n(2)$ for $n = 5, 6, 7$.

Consider $K \cong \mathrm{PSU}_6(q)$. Then $|K|$ is divisible by $\Phi_2^4(q)$ and [GLS3, Theorem 4.10.3] implies that $m_r(K) \geq 4$ for some odd prime r with r dividing $q+1$. Thus K and all groups containing K are not in ${}^3\mathcal{C}_2$. Suppose that $K \cong \mathrm{PSU}_5(q)$ with $q = 2^a \neq 4$. Then

$$|K| = q^{10}\Phi_1^2(q)\Phi_2^4(q)\Phi_4(q)\Phi_6(q)\Phi_{10}(q)/(q+1,5).$$

We investigate the equation $\Phi_2(q) = 2^a + 1 = 5^b$ for $b \in \mathbb{N}$. If b is odd, reducing modulo 8 gives us that $1 \equiv 5 \pmod 8$ if $2^a \geq 8$ and $3 \equiv 5 \pmod 8$ if $2^a = 2$, a contradiction which shows that b is even. Hence, $b = 2B$ for some $B \in \mathbb{N}$. It follows that $2^a = 5^{2B} - 1 = (5^B - 1)(5^B + 1)$, which is obviously incorrect as the right side is divisible by 3, a contradiction. Thus there is an odd prime different to 5 which

divides $\Phi_2(q)$ so long as $q \neq 4$. Therefore the unitary groups are not members of ${}^3\mathcal{C}_2$ unless possibly $K \cong \mathrm{PSU}_5(4)$ or $\mathrm{PSU}_n(2^a)$ for $n \leq 4$.

We have now narrowed down the list of possibilities for groups of Lie type in characteristic 2 to be elements of ${}^3\mathcal{C}_2$. Let us now see that the r-ranks of the remaining groups indeed do not exceed 3. We show this below and record our findings in Table 1.

Let $K \cong \mathrm{PSL}_3(2^a)$ and set $q = 2^a$. Then

$$|K| = q^3\Phi_1^2(q)\Phi_2(q)\Phi_3(q)/(q-1,3)$$

By [GLS3, Theorem 4.10.3], $m_r(K) \leq 2$ for all odd primes r. If a is odd, 3 divides $q+1$ and so $m_3(K) = 1 = \frac{3+(-1)^a}{2} = a_+$. If a is even, 3 divides $\Phi_1(q)$. Since K contains a subgroup isomorphic to $\mathrm{PSL}_3(4)$, $m_3(K) \geq m_3(\mathrm{PSL}_3(4)) = 2$, and so $m_3(K) = 2$. Moreover, using [GLS3, Theorem 4.10.3] again we obtain that $m_3(\mathrm{Inndiag}(K)) = 2$. Since the only non-inner-diagonal outer automorphisms of K of odd order are the field automorphisms and those form a cyclic subgroup of $\mathrm{Out}(K)$, it follows that $m_r(\mathrm{Aut}(K)) \leq m_r(K) + 1$ for all odd primes r.

Let $K \cong \mathrm{PSU}_3(2^a)$. Then

$$|K| = q^3\Phi_1(q)\Phi_2^2(q)\Phi_6(q)/(q+1,3)$$

with $q = 2^a$. Again [GLS3, Theorem 4.10.3], $m_r(K) \leq 2$ for all odd primes r. If a is even, 3 divides $q-1$ and so $m_3(K) = 1 = \frac{3-(-1)^a}{2} = a_-$. If a is odd, 3 divides $q+1$. Since K contains a subgroup isomorphic to $\mathrm{PSU}_3(2)$, $2 \geq m_3(K) \geq m_3(\mathrm{PSU}_3(2)) = 2$, and so $m_3(K) = 2$. Using [GLS3, Theorem 4.10.3] again we obtain that $m_3(\mathrm{Inndiag}(K)) = m_3(\mathrm{SU}_3(2^a)) = 2$. Since the only non-inner-diagonal outer automorphisms of K of odd order are the field automorphisms and those form a cyclic subgroup of $\mathrm{Out}(K)$, it follows that $m_r(\mathrm{Aut}(K)) \leq m_r(K) + 1$ for all odd primes r.

Suppose now that $K \cong \mathrm{PSL}_4(2^a)$ or $\mathrm{PSU}_4(2^a)$. Since K has no outer inner-diagonal automorphisms, we have $K \cong \mathrm{SL}_4(2^a)$ or $\mathrm{SU}_4(2^a)$ respectively. Thus application of [GLS3, Theorem 4.10.3] is very straightforward and gives the desired answers.

If $K \cong \mathrm{PSL}_5(2) (\cong \mathrm{SL}_5(2))$, $|K| = 2^{10} \cdot 3^2 \cdot 5 \cdot 7 \cdot 31$. Since K naturally contains the subgroup $\mathrm{SL}_2(2) \times \mathrm{SL}_2(2)$, the results of Table 1 follow immediately.

Suppose that $K \cong \mathrm{PSU}_5(4)$. Then $|K| = 2^{20} \cdot 3^2 \cdot 5^4 \cdot 13 \cdot 17 \cdot 41$. Clearly $m_3(K) = 2$ and $m_r(K) \leq 1$ for $r > 5$. So suppose that $r = 5$. Then, as $\mathrm{GU}_1(4) \wr \mathrm{Sym}(5) = 5 \wr \mathrm{Sym}(5)$ is a subgroup of $\mathrm{GU}_5(4)$, we have $m_5(K) = 3$ and $m_5(\mathrm{Inndiag}(K)) = 4$. As K does not admit non-inner-diagonal outer automorphisms of odd order, this finishes the proof for $\mathrm{PSU}_5(4)$.

If $K \cong \mathrm{PSL}_6(2)$, $|K| = 2^{15} \cdot 3^4 \cdot 5 \cdot 7^2 \cdot 31$. Since the Sylow 3-subgroup of K is isomorphic to $3 \wr 3$ contained in $\mathrm{SL}_2(2) \wr 3$ and K does not admit outer automorphisms of odd order, the tabulated results follow immediately.

If $K \cong \mathrm{PSL}_7(2) (\cong \mathrm{SL}_7(2))$, $|K| = 2^{21} \cdot 3^4 \cdot 5 \cdot 7^2 \cdot 31 \cdot 127$. Then we obtain the results in Table 1 as the Sylow 3-subgroup of K is again isomorphic to $3 \wr 3$ and K does not admit outer automorphisms of odd order.

We now consider the remaining families of groups. We have

$$\begin{aligned}
|{}^2\mathrm{B}_2(q)| &= q^2\Phi_1(q)\Phi_4(q)\\
|\mathrm{PSp}_4(q)| &= q^4\Phi_1^2(q)\Phi_2^2(q)\Phi_4(q)\\
|\mathrm{G}_2(q)| &= q^6\Phi_1^2(q)\Phi_2^2(q)\Phi_3(q)\Phi_6(q)\\
|\mathrm{PSp}_6(q)| &= q^4\Phi_1^3(q)\Phi_2^3(q)\Phi_3(q)\Phi_4(q)\Phi_6(q)\\
|\Omega_8^-(q)| &= q^{12}\Phi_1^3(q)\Phi_2^3(q)\Phi_3(q)\Phi_4(q)\Phi_6(q)\Phi_8(q)\\
|{}^2\mathrm{F}_4(q)| &= q^{12}\Phi_1^2(q)\Phi_2^2(q)\Phi_4^2(q)\Phi_6(q)\Phi_{12}(q)\\
|{}^3\mathrm{D}_4(q)| &= q^{12}\Phi_1^2(q)\Phi_2^2(q)\Phi_3^2(q)\Phi_6^2(q)\Phi_{12}(q).
\end{aligned}$$

In all of these cases the centre of the universal group is trivial. Hence [GLS3, Theorem 4.10.3] gives us the desired upper bounds on the odd ranks of K and as 3 divides $\Phi_1(q)$ or $\Phi_2(q)$, we always have $m_3(K) = 2$. Since the Tits group ${}^2\mathrm{F}_4(2)'$ has index 2 in ${}^2\mathrm{F}_4(2)$, Table 1 is also true for this group. Finally, as $\mathrm{Out}(K)$ is cyclic for all the groups listed, we have $m_r(\mathrm{Aut}(K)) \leq m_r(K) + 1$.

The 2-ranks of the groups of Lie type in characteristic 2, are given in [GLS3, Theorem 3.3.3].

Suppose now that K is a quasisimple group of Lie type of characteristic 2 with $Z(K) \neq 1$. Since $Z(K)$ is a 2-group, for an odd prime r, the structure of Sylow r-subgroup of K is the same as in $K/Z(K)$, the results are as in the table. As for the 2-ranks, if $K/Z(K) \cong \mathrm{PSL}_3(4)$ or $Sp_6(2)$, then $m_2(K)$ is given by [GLS3, Proposition 6.4.4].

Suppose that $K/Z(K) \cong {}^2\mathrm{B}_2(8)$, $|Z(K)| = 2$ and $T \in \mathrm{Syl}_2(K)$. Since T is not quaternion or cyclic, T contains a fours subgroup. Ass all the involutions in $Z(T/Z(K))$ are central and $N_K(T)$ acts transitively on $Z(T/Z(K))$ we deduce that $m_2(T) = 4$. Now suppose that $|Z(K)| = 4$. Then as $\mathrm{Aut}(K)$ acts transitively on $Z(K)$ we obtain $m_2(K) = 5$.

Let $K/Z(K) \cong \mathrm{G}_2(4)$. From the character table of $\mathrm{G}_2(4)$ and $2\dot{}\mathrm{G}_2(4)$, we see that the 2-central involutions of $K/Z(K)$ lift to involutions and the other class of involutions of $K/Z(K)$ lift to elements of order 4. Thus, if A is an elementary abelian subgroup of K of maximal 2-rank, then $(A/Z(K))^\#$ consists of 2-central involutions of $K/Z(K)$. In particular, $|A| \geq 2^5$. Let χ be a character of degree 300 for $K/Z(K)$ and let $B = A/Z(K)$. Then the inner product of characters $(\chi|_B, 1_B)$ is non-negative. Thus, for $z \in B^\#$),

$$(|B| - 1)\chi(z) + \chi(1) = -(|B| - 1)20 + 300 \geq 0.$$

Hence $|B| \leq 2^4$. This shows that $|A| = 2^5$ and $m_2(K) = 5$ as claimed.

Finally, for sporadic groups, the members of ${}^3\mathcal{C}_2$ and their ranks can be read from [GLS3, Table 5.6.1]. □

While we know the general shape of the centralisers of involutions of elements of ${}^3\mathcal{C}_2$, let us state explicitly the structure of $F^*(C_K(\alpha))$ (and whenever obvious of $C_K(\alpha)$) for an involutory automorphism α of $K \in {}^3\mathcal{C}_2$. In the statement below E_{2^a} denotes an elementary abelian group of order 2^a, $2_\pm^{1+2a}$ an appropriate (+ or −)

extraspecial group of order 2^{1+2a}, $\mathrm{Dih}(n)$ dentes a dihedral group of order n and $\mathrm{Frob}(20)$ a Frobenius group of order 20.

Proposition 1.7 *Let K be a simple group contained in ${}^3\mathcal{C}_2$. Assume that $\alpha \in \mathrm{Aut}(K)$ is an involution. Then one of the following holds.*

(i) $K \cong \mathrm{PSL}_2(p)$ *where* $p = 2^a + 1$ *is a Fermat prime, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{Dih}(2^a)$*, or*
- *(b)* α *is outer inner-diagonal and* $C_K(\alpha) \cong \mathrm{Dih}(2^a + 2)$.

(ii) $K \cong \mathrm{PSL}_2(p)$ *where* $p = 2^a - 1$ *is a Mersenne prime, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{Dih}(2^a)$*, or*
- *(b)* α *is outer inner-diagonal and* $C_K(\alpha) \cong \mathrm{Dih}(2^a - 2)$.

(iii) $K \cong \mathrm{PSL}_2(9) (\cong \mathrm{PSp}_4(2)')$*, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{Dih}(8)$*, or*
- *(b)* α *is outer and one of the following holds:*
  - *i.* α *is a field automorphism and* $C_K(\alpha) \cong \mathrm{PGL}_2(3) \cong \mathrm{Sym}(4)$.
  - *ii.* α *is an outer inner-diagonal automorphism and* $C_K(\alpha) \cong \mathrm{Dih}(10)$.

(iv) $K \cong \mathrm{PSL}_3(3)$*, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{GL}_2(3)$*, or*
- *(b)* α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PGL}_2(3) \cong \mathrm{Sym}(4)$.

(v) $K \cong \mathrm{PSU}_3(3) (\cong \mathrm{G}_2(2)')$*, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{GU}_2(3) \cong 4 \circ \mathrm{GL}_2(3)$*, or*
- *(b)* α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PGL}_2(3) \cong \mathrm{Sym}(4)$.

(vi) $K \cong \mathrm{PSL}_2(2^a)$ *with* $a \geq 2$*, and either*
- *(a)* α *is inner and* $C_K(\alpha) \cong \mathrm{E}_{2^a}$*, or*
- *(b)* a *is even,* α *is a field automorphism and* $C_K(\alpha) \cong \mathrm{PSL}_2(2^{a/2})$.

(vii) $K \cong \mathrm{PSL}_3(2^a)$ *with* $a \geq 2$ *and either*
- *(a)* α *is inner,* $|C_K(\alpha)| = \frac{(q-1)}{(3,q-1)} 2^{3a}$ *and* $F^*(C_K(\alpha))$ *is a Sylow 2-subgroup of* K*, or*
- *(b)* α *is outer and one of the following holds:*
  - *i.* α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_2(2^a) \cong \mathrm{PSL}_2(2^a)$.
  - *ii.* a *is even,* α *is a field automorphism and* $C_K(\alpha) \cong \mathrm{PSL}_3(2^{a/2})$.
  - *iii.* a *is even,* α *is a graph-field automorphism and* $C_K(\alpha) \cong \mathrm{PSU}_3(2^{a/2})$.

(viii) $K \cong \mathrm{PSU}_3(2^a)$ *with* $a \geq 2$*, and either*
- *(a)* α *is inner,* $|C_K(\alpha)| \cong \frac{(q+1)}{(3,q+1)} 2^{3a}$ *and* $F^*(C_K(\alpha))$ *is a Sylow 2-subgroup of* K*, or*
- *(b)* α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_2(2^a) \cong \mathrm{PSL}_2(2^a)$.

(ix) $K \cong \mathrm{PSL}_4(2^a)$ *with* $a \geq 1$*, and either*
- *(a)* α *is inner and* $F^*(C_K(\alpha))$ *is a 2-group, or*
- *(b)* α *is outer and one of the following holds:*
  - *i.* α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_4(2^a)$.

ii. α *is a graph automorphism and* $F^*(C_K(\alpha))$ *is a* 2*-group.*

iii. a *is even,* α *is a field automorphism and* $C_K(\alpha) \cong \mathrm{PSL}_4(2^{a/2})$.

iv. a *is even,* α *is a graph-field automorphism and* $C_K(\alpha) \cong \mathrm{PSU}_4(2^{a/2})$.

(x) $K \cong \mathrm{PSU}_4(2^a)$ *with* $a \geq 1$*, and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is outer and one of the following holds:*

i. α *is a graph automorphism with* $C_K(\alpha) \cong \mathrm{PSp}_4(2^a)$.

ii. α *is a graph automorphism and* $F^*(C_K(\alpha))$ *is a* 2*-group.*

(xi) $K \cong \mathrm{PSL}_5(2)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_4(2)$.

(xii) $K \cong \mathrm{PSU}_5(4)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_4(4)$.

(xiii) $K \cong \mathrm{PSL}_6(2)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is outer and one of the following holds:*

i. α *is a graph automorphism with* $C_K(\alpha) \cong \mathrm{PSp}_6(2)$.

ii. α *is a graph automorphism and* $F^*(C_K(\alpha))$ *is a* 2*-group.*

(xiv) $K \cong \mathrm{PSL}_7(2)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is a graph automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_6(2)$.

(xv) $K \cong \mathrm{PSp}_4(2^a)$ *with* $a \geq 2$*, and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is outer and one of the following holds:*

i. a *is odd, alpha is a graph-field automorphism and* $C_K(\alpha) \cong {}^2\mathrm{B}_2(2^a)$.

ii. a *is even, alpha is a field automorphism and* $C_K(\alpha) \cong \mathrm{PSp}_4(2^{a/2})$.

(xvi) $K \cong \mathrm{PSp}_6(2^a)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) a *is even and* α *is a field automorphism with* $C_K(\alpha) \cong \mathrm{PSp}_6(2^{a/2})$.

(xvii) $K \cong \Omega_8^-(2^a)$ *and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) α *is outer and one of the following holds.*

i. α *is a graph automorphism with* $C_K(\alpha) \cong \mathrm{PSp}_6(2^a)$.

ii. α *is a graph automorphism and* $F^*(C_K(\alpha))$ *is a* 2*-group.*

(xviii) $K \cong {}^2\mathrm{B}_2(2^a)$*,* α *is inner and* $C_K(\alpha)$ *is a Sylow* 2*-subgroup of* K.

(xix) $K \cong \mathrm{G}_2(2^a)$*,* $a \geq 2$*, and either*

(a) α *is inner and* $F^*(C_K(\alpha))$ *is a* 2*-group, or*

(b) *a is even, α is a field automorphism and $C_K(\alpha) \cong \mathrm{G}_2(2^{a/2})$.*

(xx) $K \cong {}^2\mathrm{F}_4(2^a)'$, *$\alpha$ is inner and $F^*(C_K(\alpha))$ is a 2-group.*

(xxi) $K \cong {}^3\mathrm{D}_4(2^a)$, *and either*

(a) *α is inner and $F^*(C_K(\alpha))$ is a 2-group, or*

(b) *a is even, α is a field automorphism and $C_K(\alpha) \cong {}^3\mathrm{D}_4(2^{a/2})$.*

(xxii) $K \cong \mathrm{M}_{11}$, *α is inner in class $2A$ and $C_K(\alpha) \cong \mathrm{GL}_2(3)$.*

(xxiii) $K \cong \mathrm{M}_{12}$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 2 \times \mathrm{Sym}(5)$.*

(b) *α is inner in class $2B$ and $C_K(\alpha) \cong 2^{1+4}_+.\mathrm{Sym}(3)$ with $F^*(C_K(\alpha))$ a 2-group.*

(c) *α is outer in class $2C$ and $C_K(\alpha) \cong 2 \times \mathrm{Alt}(5)$.*

(xxiv) $K \cong \mathrm{M}_{22}$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong \mathrm{E}_{2^4}.\mathrm{Sym}(4)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is outer in class $2B$ and $C_K(\alpha) \cong \mathrm{E}_{2^3}.\mathrm{SL}_3(2)$ with $F^*(C_K(\alpha))$ a 2-group.*

(c) *α is outer in class $2C$ and $C_K(\alpha) \cong \mathrm{E}_{2^4}.\mathrm{Frob}(20)$ with $F^*(C_K(\alpha))$ a 2-group.*

(xxv) $K \cong \mathrm{M}_{23}$, *α is inner in class $2A$ and $C_K(\alpha) \cong \mathrm{E}_{2^4}.\mathrm{SL}_3(2)$ with $F^*(C_K(\alpha))$ a 2-group.*

(xxvi) $K \cong \mathrm{M}_{24}$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 2^{1+6}_+.\mathrm{SL}_3(2)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is inner in class $2B$ and $C_K(\alpha) \cong \mathrm{E}_{2^6}.\mathrm{Sym}(5)$ with $F^*(C_K(\alpha))$ a 2-group.*

(xxvii) $K \cong \mathrm{J}_2$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 2^{1+4}_-.\mathrm{Alt}(5)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is inner in class $2B$ and $C_K(\alpha) \cong \mathrm{E}_{2^2} \times \mathrm{Alt}(5)$.*

(c) *α is outer in class $2C$ and $C_K(\alpha) \cong \mathrm{PGL}_2(7)$.*

(xxviii) $K \cong \mathrm{J}_3$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 2^{1+4}_-.\mathrm{Alt}(5)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is outer in class $2B$ and $C_K(\alpha) \cong \mathrm{PSL}_2(17)$.*

(xxix) $K \cong \mathrm{J}_4$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 2^{1+12}_+.(3\dot{}\mathrm{M}_{22}.2)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is inner in class $2B$ and $C_K(\alpha) \cong \mathrm{E}_{2^{11}}(\mathrm{M}_{22}.2)$ with $F^*(C_K(\alpha))$ a 2-group.*

(xxx) $K \cong \mathrm{HS}$ *and*

(a) *α is inner in class $2A$ and $C_K(\alpha) \cong 4 \circ 2^{1+4}_+.\mathrm{Sym}(5)$ with $F^*(C_K(\alpha))$ a 2-group.*

(b) *α is inner in class $2B$ and $C_K(\alpha) \cong 2 \times \mathrm{Aut}(\mathrm{Alt}(6))$.*

(c) α *is outer in class* $2C$ *and* $C_K(\alpha) \cong \mathrm{E}_{2^4}.\mathrm{O}_4^-(2)$ *with* $F^*(C_K(\alpha))$ *a 2-group.*

(d) α *is outer in class* $2D$ *and* $C_K(\alpha) \cong \mathrm{Sym}(8)$.

(xxxi) $K \cong \mathrm{Ru}$ *and*

(a) α *is inner in class* $2A$ *and* $C_K(\alpha) \cong \underline{2^{11}}.\mathrm{Sym}(5)$ [3] *with* $F^*(C_K(\alpha))$ *a 2-group.*

(b) α *is inner in class* $2B$ *and* $C_K(\alpha) \cong \mathrm{E}_{2^2} \times {}^2\mathrm{B}_2(8)$.

Proof For the first five cases we refer to [GLS3, Table 4.5.1 and Proposition 4.9.1].

The groups parts (vi) to (xxi) are simple groups of Lie type in characteristic 2. Thus, for these groups, the Borel-Tits Theorem [GLS3, Theorem 3.1.3] implies that every inner involutory automorphism α of K has $F^*(C_K(\alpha))$ is a 2-group. Suppose now that α is an outer automorphism of K. Then α is either a graph, field or graph-field automorphism and, according to the conventions in [GLS3, Definition 2.5.13], the latter two can only occur when a is even. In particular, K is neither ${}^2\mathrm{B}_2(2^a)$ nor ${}^2\mathrm{F}_4(2^a)'$ (be aware the outer automorphism of ${}^2\mathrm{F}_4(2)$ is covered by the Borel-Tits theorem and in fact there is no involution in ${}^2\mathrm{F}_4(2) \setminus {}^2\mathrm{F}_4(2)'$). Suppose that $K \cong \mathrm{PSL}_n(2^a)$ or $\mathrm{PSU}_n(2^a)$ and α has order 2. If α is a graph automorphism, then the structure of $C_K(\alpha)$ is presented in [GLS3, Proposition 4.9.2]. Thus, if n is odd, $C_K(\alpha) \cong \mathrm{Sp}_{n-1}(2^a)$, while, if . n is even, either $C_K(\alpha) \cong \mathrm{Sp}_n(2^a)$ or $F^*(C_K(\alpha))$ is a 2-group and, in this latter case, both possibilities occur.

Now, $\mathrm{PSU}_n(2^a)$ does not admit any other outer automorphisms of order 2 ([GLS3, Definition 2.5.13]). Neither does $\mathrm{PSL}_n(2^a)$ when a is odd. Suppose though that a is even. Then [GLS3, Proposition 4.9.1] says that for α a field automorphism and $C_K(\alpha)$ contains $\mathrm{PSL}_n(2^{a/2})$ and is contained in $\mathrm{PGL}_n(2^{a/2})$ and that for α a graph-field automorphism $C_K(\alpha)$ contains $\mathrm{PSU}_n(2^{a/2})$ and is contained in $\mathrm{PGU}_n(2^{a/2})$. For the values of n and a being considered, this means that $C_K(\alpha) \cong \mathrm{PSL}_n(2^{a/2})$ is α is a field automorphism or $C_K(\alpha) \cong \mathrm{PSU}_n(2^{a/2})$ if α is a graph-field automorphism unless perhaps $n = 3$ and a is even. Since $\mathrm{PSL}_3(2^a)$ does not contain $\mathrm{PGL}_3(2^{a/2})$ or $\mathrm{PGU}_3(2^{a/2})$ by [GLS3, Theorem 6.5.3], the statements involving field and graph-field automorphisms in (vii)–(xiv) are true. This completes the discussion of linear and unitary groups.

Suppose that $K \cong \mathrm{Sp}_4(2^a)$ with $a \geq 2$ or $\mathrm{Sp}_6(2^a)$, $\mathrm{G}_2(2^a)$, ${}^3\mathrm{D}_4(2^a)$ with $a \geq 1$. If a is even, it follows from [GLS3, Definition 2.5.13] that the only outer automorphisms of K of order 2 are the field ones. Thus by [GLS3, Proposition 4.9.], $C_K(\alpha) \cong \mathrm{Sp}_4(2^{a/2})$, $\mathrm{Sp}_6(2^{a/2})$, $\mathrm{G}_2(2^{a/2})$ and ${}^3\mathrm{D}_4(2^{a/2})$ respectively. If however a is odd, then [GLS3, Definiton 2.5.13 together with Proposition 4.9.1] imply that for K to admit an outer involutory automorphisms, $K \cong \mathrm{Sp}_4(2^a)$. In this case α is a graph-field automorphism and $C_K(\alpha) \cong {}^2\mathrm{B}_2(a^a)$. Thus (xv), (xvi), (xix) and (xxi) hold.

If $K \cong \Omega_8^-(2^a)$, then [GLS3, Definition 2.5.13 and Theorem 4.9.2(3)] yields that α is a graph automorphism and either $C_K(\alpha) \cong \mathrm{Sp}_6(2^a)$ or $F^*(C_K(\alpha))$ is a 2-group (and both cases occur). This proves (xvii). This completes the consideration of all possibilities with K a Lie type group in characteristic 2.

The remaining results about the sporadic simple groups follow from inspection of [GLS3, Tables 5.3]. □

[3] The notation $\underline{2^{11}}$ means a 2 group of order 2^{11}.

Here is an important property of ${}^3\mathcal{C}_2$-groups that can be obtained as an immediate consequence of the previous statement.

Corollary 1.8 *Let K be a simple group contained in ${}^3\mathcal{C}_2$. If z is a 2-central involution of K (i.e., z is an involution contained in the centre of a Sylow 2-subgroup of K), then $F^*(C_K(z))$ is a 2-group.*

While we now know "everything" about the structure of the involutory automorphisms of $\mathcal{C}_2$-groups, it is often handy to look at the 2-subgroups of $\mathrm{Aut}(K)$ and to know the structure of centralisers of those.

Lemma 1.9 *Let K be a ${}^3\mathcal{C}_2$-group and W is a non-trivial 2-subgroup of $\mathrm{Aut}(K)$. If J is a component of $C_K(W)$, then $J \in \mathcal{C}_2$.*

Proof This follows by repeated use of Proposition 1.7. □

In fact, the components from the above statement disappear very quickly. The next result is true in general for groups in Lie(2) but not for arbitrary members of $\mathcal{C}_2$.

Lemma 1.10 *Let K be a simple group contained in ${}^3\mathcal{C}_2$. Assume that $W \leq \mathrm{Aut}(K)$ is an elementary abelian group of order 8. Then $F^*(C_K(W)) = O_2(C_G(W))$.*

Proof Assume that $K \in {}^3\mathcal{C}_2$ is a counterexample to the statement. In particular, $K \ncong \mathrm{PSL}_2(p)$ with K a Mersenne or Fermat prime, then there is no eights subgroup in $\mathrm{Aut}(K) \cong \mathrm{PGL}_2(p)$.

If for any involution w in W, $F^*(C_K(w))$ is a 2-group, then as $C_K(W) \leq C_K(w)$, [GLS2, Corollary 5.12] implies that $F^*(C_K(W))$ is a 2-group. Thus all the involutions in W have centralizers with components in K.

Using of [GLS3, Theorem 2.5.12] for groups of Lie type and Tables 5.3 of [GLS3], we observe that $\mathrm{Out}(K)$ does not contain an eights subgroup. Thus W must contain an inner automorphism of K. It follows from Proposition 1.7 that K is one of the following sporadic simple groups: M_{12}, J_2, HS or Ru.

Suppose that $K \cong \mathrm{M}_{12}$. Then W contains $2A$ involutions and possibly $2C$ involutions. If α is a $2A$ involution of K, then $C_{\mathrm{Aut}(K)}(\alpha) \cong D_8 Y \mathrm{Sym}(5)$, and so for every eights group $W \leq C_{\mathrm{Aut}(K)}(\alpha)$, $F^*(C_K(W)) \leq F^*(C_{C_K(\alpha)}(W))$ is a 2-group. So $K \ncong \mathrm{M}_{12}$.

Suppose that $K \cong \mathrm{J}_2$. Then, if $W \leq K$, all the involutions in W must be in class $2B$. Since the centralizer of such an element is isomorphic to $\mathrm{E}_{2^2} \times \mathrm{Alt}(5)$, we obtain a contradiction. So $W \not\leq K$. Hence W contains an element w in class $2C$. So $C_K(w) \cong \mathrm{PGL}_2(7)$ and we observe a contradiction.

Suppose that $K \cong \mathrm{HS}$. If $W \leq K$, W is contained in a subgroup isomorphic to $2 \times \mathrm{Aut}(\mathrm{Alt}(6))$ and we have a contradiction. So suppose that $W \not\leq K$. Then there exists $\alpha \in W$ in class 2D. In particular, $W \leq C_K(\alpha) \cong 2 \times \mathrm{Sym}(8)$. It follows that $W \cap E(C_K(\alpha)) \neq 1$ and so, as $E(C_K(\alpha)) \cong \mathrm{PSL}_4(2)$, $F^*(C_K(W))$ is a 2-group by Proposition 1.7 (ix).

Finally, if $K \cong \mathrm{Ru}$, then W is contained in a subgroup isomorphic to $\mathrm{E}_{2^2} \times {}^2\mathrm{B}_2(8)$ and again we have a contradiction. □

In fact, just like for involutory automorphisms, $O(C_K(W))$ is is often trivial for the 2-subgroups $W \leq \mathrm{Aut}(K)$.

Lemma 1.11 *Suppose that K is a simple group contained in ${}^3\mathcal{C}_2$. Assume that $W \leq \mathrm{Aut}(K)$, $W = \Omega_1(W)$ and $O(C_K(W)) \neq 1$. Then either*

(i) $K \cong \mathrm{PSL}_2(p)$ *with p a Fermat or Mersenne prime, $|W| = 2$ and $O(C_K(W))$ is cyclic;*

(ii) $K \cong \mathrm{PSL}_2(9)$, $|O(C_K(W))| = 5$ *and* $|W| = 2$*;*

(iii) $K \cong \mathrm{PSL}_3(4)$, $|O(C_K(W))| = 9$ *and* $|W| = 2$, *or* $|O(C_K(W))| = 3$ *and* $|W| = 4$*;*

(iv) $K \cong \mathrm{M}_{12}$, $|O(C_K(W))| = 3$, $|W| = 4$ *or* $W \cong \mathrm{Dih}(8)$*;*

(v) $K \cong \mathrm{J}_2$, $|O(C_K(W))| = 3$, $|W| = 4$ *or* $W \cong \mathrm{Dih}(8)$*; or*

(vi) $K \cong \mathrm{HS}$, $|O(C_K(W))| = 5$, $|W| = 4$ *or* $W \cong \mathrm{Dih}(8)$.

Proof This follows from a recursive use of Proposition 1.7. □

Going back to the centralisers of involutory automorphisms of ${}^3\mathcal{C}_2$-groups, we may look at those in the "opposite direction".

Lemma 1.12 *Let K be an element of ${}^3\mathcal{C}_2$. For each such K Table 2 indicates the groups $K^* \in {}^3\mathcal{C}_2$ and an involution $z \in \mathrm{Aut}(K^*)$ such that $E(C_{K^*}(z)) \cong K$. We note that such K^* is usually called a pumpup of K and that the absence of any sporadic simple groups K in this table is a consequences of no such K^* existing.*

Proof The proof follows from Proposition 1.7. □

References

[ABG] J.L. Alperin, R. Brauer, D. Gorenstein. Finite simple groups of 2-rank two. Collection of articles dedicated to the memory of Abraham Adrian Albert. *Scripta Math.* **3–4:29** (1973) 191–214.

[AS] M. Aschbacher, G. Seitz. Involutions in Chevalley Groups Over Fields of Even Order. *Nagoya Math. J.* **63** (1976), 1–91.

[G] D. Gorenstein. Finite Groups. Second Edition, Chelsea, New York, 1980.

[GLS1] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 1. Amer. Math. Soc. Surveys and Monographs **40**, #**1** (1995).

[GLS2] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 2. Amer. Math. Soc. Surveys and Monographs **40**, #**2** (1997).

[GLS3] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 3. Amer. Math. Soc. Surveys and Monographs **40**, #**3** (1998).

[GLS4] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 4. Amer. Math. Soc. Surveys and Monographs **40**, #**4** (1999).

[GLS5] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 5. Amer. Math. Soc. Surveys and Monographs **40**, #**5** (2002).

[GLS6] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 6. Amer. Math. Soc. Surveys and Monographs **40**, #**6** (2005).

[GLS7] D. Gorenstein, R. Lyons, R. Solomon. The Classification of the Finite Simple Groups, Number 7. Amer. Math. Soc. Surveys and Monographs **40**, #**7**. In preparation.

[FT] W. Feit, J.G. Thompson. Solvability of Groups of Odd Order. *Pacific J. Math.* **13** (1963) 775–1029.

K	$\langle K^*, z\rangle$	type of z
$\mathrm{PSL}_2(q), q \in \mathcal{FM} \setminus \{5,7,17\}$	—	—
$\mathrm{PSL}_2(17)$	$\mathrm{Aut}(\mathrm{J}_3)$	$2B$
$\mathrm{PSL}_2(7) \cong \mathrm{PSL}_3(2)$	$\mathrm{PSL}_3(4){:}2$	field
	$\mathrm{Aut}(\mathrm{J}_2)$	$2C$
$\mathrm{PSL}_2(4) \cong \mathrm{PSL}_2(5) \cong \mathrm{Alt}(5)$	$\mathrm{PSL}_2(2^4){:}2$	field
	$\mathrm{PSL}_3(4){:}2$	graph
	$\mathrm{PSU}_3(4){:}2$	graph
	M_{12}	$2A$
	$\mathrm{Aut}(\mathrm{M}_{12})$	$2C$
	J_2	$2B$
$\mathrm{PSL}_2(9) \cong \mathrm{PSp}_4(2)' \cong \mathrm{Alt}(6)$	$\mathrm{PSp}_4(4){:}2$	field
	$\mathrm{PSL}_4(2){:}2$	graph
	$\mathrm{PSU}_4(2){:}2$	graph
	$\mathrm{PSL}_5(2){:}2$	graph
	HS	$2B$
$\mathrm{PSL}_3(3)$	—	—
$\mathrm{PSU}_3(3)$	$\mathrm{G}_2(4){:}2$	field
$\mathrm{PSL}_2(2^a)$, $a \geq 3$	$\mathrm{PSL}_2(2^{2a}){:}2$	field
	$\mathrm{PSL}_3(2^a){:}2$	graph
	$\mathrm{PSU}_3(2^a){:}2$	graph
$\mathrm{PSL}_3(2^a)$	$\mathrm{PSL}_3(2^{2a}){:}2$	field
$\mathrm{PSU}_3(2^a)$, $a \geq 2$	$\mathrm{PSL}_3(2^{2a}){:}2$	graph-field
$\mathrm{PSL}_4(2) \cong \mathrm{Alt}(8)$	$\mathrm{Aut}(\mathrm{HS})$	$2D$
$\mathrm{PSL}_4(2^a)$	$\mathrm{PSL}_4(2^{2a}){:}2$	field
$\mathrm{PSU}_4(2^a)$	$\mathrm{PSL}_4(2^{2a}){:}2$	graph-field
$\mathrm{PSL}_5(2)$	—	—
$\mathrm{PSU}_5(4)$	—	—
$\mathrm{PSL}_6(2)$	—	—
$\mathrm{PSL}_7(2)$	—	—
$\mathrm{PSp}_4(2^a)$, $a \geq 2$	$\mathrm{PSp}_4(2^{2a}){:}2$	field
	$\mathrm{PSL}_4(2^a){:}2$	graph
	$\mathrm{PSU}_4(2^a){:}2$	graph
$\mathrm{PSp}_4(4)$	$\mathrm{PSU}_5(4){:}2$	graph
$\mathrm{PSp}_6(2^a)$	$\mathrm{PSp}_6(2^{2a}){:}2$	field
	$\Omega_8^-(2^a){:}2$	graph
$\mathrm{PSp}_6(2)$	$\mathrm{PSL}_6(2){:}2$	graph
	$\mathrm{PSL}_7(2){:}2$	graph
$\Omega_8^-(2^a)$	—	—
${}^2\mathrm{B}_2(2^a)$, $a \geq 3$	$\mathrm{PSp}_4(2^a){:}2$	graph-field
${}^2\mathrm{B}_2(8)$	Ru	$2B$
$\mathrm{G}_2(2^a)$, $a \geq 2$	$\mathrm{G}_2(2^{2a}){:}2$	field
${}^2\mathrm{F}_4(2^a)$	—	—
${}^3\mathrm{D}_4(2^a)$	${}^3\mathrm{D}_4(2^{2a}){:}2$	field

Table 2. Pumpups in ${}^3\mathcal{C}_2$

RESURRECTING WELLS' EXACT SEQUENCE AND BUCKLEY'S GROUP ACTION

JILL DIETZ

St. Olaf College, Northfield, Minnesota 55410, USA
Email: dietz@stolaf.edu

Abstract

We give a historical perspective on the Wells exact sequence and Buckley's interpretation of it, and include a survey of applications and extensions of their work from the 1970's to the present.

1 Introduction

In 1971, Charles Wells [26] constructed an exact sequence for the automorphism group of a group extension. The paper received a bit of attention at the time, but seems to have been largely ignored until the last decade. In particular, a series of papers in the *Journal of Algebra* over the last few years (see [12], [13], [16]) have brought attention to the sequence and its applications.

This paper presents a historical view of the Wells exact sequence, emphasizing the game-changing nature of Joseph Buckley's interpretation of Wells' result in the context of group actions. Indeed, one goal of this paper is to give Buckley credit–at least equal to Wells–for describing an invaluable method for investigating automorphisms of group extensions. A survey of applications of Wells' and Buckley's work is provided.

The paper is organized as follows: Section 2 begins with background information on group extensions and their automorphisms; Section 3 focuses on the 1970's and gives a history of Wells' work and subsequent papers by Buckley and others; Section 4 focuses on the "resurrection" of Wells' and Buckley's work in the last decade, but does not contain detailed statements of theorems; and Section 5 concludes with a survey of applications and extensions of the Wells exact sequence, including details missing in Section 4.

2 Background Information

In this section we give background information on group extensions and their automorphisms. This information is well known and available in many other places, but we will use this opportunity to establish some notation. For group extensions, we will follow the notation in [20] as closely as possible since Derek Robinson's book is widely available, and because both the book and an old paper of his [18] are commonly used resources by those in the field. In particular, we use the following conventions:

- if f is a function, then either xf or x^f will denote the image of x under f;
- if $f : A \to B$ and $g : B \to C$ then $fg : A \xrightarrow{f} B \xrightarrow{g} C$ will denote the composition of f and g;

- conjugation will be denoted by $x^y = y^{-1}xy$ and the corresponding conjugation homomorphism is denoted by $\iota(y)$ so that $x^{\iota(y)} = y^{-1}xy$;
- ZN is the center of the group N.

2.1 Group extensions

A *group extension* is a short exact sequence of groups of the form

$$\mathbf{e} : 1 \to N \xrightarrow{\mu} G \xrightarrow{\varepsilon} Q \to 1.$$

Thus μ is injective, ε is surjective, and $\operatorname{Im}\mu = \operatorname{Ker}\varepsilon$. Although some authors present a different viewpoint, we will call both $\mathbf{e}$ and the group G itself an *extension of N by Q*. For convenience, we will write $\mathbf{e}$ as

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q,$$

and assume that μ is inclusion, so that N is considered a normal subgroup of G (*n.b.* Robinson does not assume μ is inclusion in his book, but he does in a recent paper [22]).

A *transversal* is a function $t : Q \to G$ satisfying $t\varepsilon = 1_Q$. Note that every element $g \in G$ can be uniquely represented as $g = q^t n$ for some $q \in Q$ and $n \in N$. A transversal need not be a homomorphism; indeed, a *factor set* $\alpha : Q \times Q \to N$ measures how far away t is from being a homomorphism. That is, if $q_1, q_2 \in Q$, then α is defined by

$$q_1^t q_2^t = (q_1 q_2)^t (q_1, q_2)^\alpha.$$

The group extension *splits* if there is a transversal that is a homomorphism. A splitting exists if and only if $G \cong N \rtimes Q$ is a semi-direct product.

We see that Q essentially acts on N by conjugation as follows. Define a function $\lambda : Q \to \operatorname{Aut} N$ by $q^\lambda = \iota(q^t)$; that is,

$$n^{q^\lambda} = (q^t)^{-1} n (q^t)$$

for $q \in Q$ and $n \in N$. Clearly λ depends on the choice of transversal t. If $t' : Q \to G$ is another transversal, then $t\varepsilon = 1_Q = t'\varepsilon$ means that q^t and $q^{t'}$ differ by an element of $\operatorname{Ker}\varepsilon = \operatorname{Im}\mu = N$. Hence, conjugation by q^t is equal to conjugation by $q^{t'}$ modulo an inner automorphism of N. Thus we get a homomorphism

$$\chi : Q \to \operatorname{Out} N$$

defined by $q^\chi = q^\lambda \operatorname{Inn} N$. The homomorphism χ is called the *coupling* or *twisting* of the extension $\mathbf{e}$, and is independent of the choice of transversal. We see that χ is uniquely defined by the extension $\mathbf{e}$, and it tracks the way Q twists the normal subgroup N inside of G.

2.2 Equivalence classes of group extensions

Let

$$\mathbf{e}' : N \xrightarrow{\mu'} G' \xrightarrow{\varepsilon'} Q$$

be another extension of N by Q with coupling $\chi' : Q \to \operatorname{Out} N$.

We say that $\mathbf{e}'$ is *equivalent* to $\mathbf{e}$ if there is an homomorphism $\gamma : G \to G'$ that *restricts* to the identity on N and *induces* the identity on Q; that is, if $n^\gamma = n$ for all $n \in N$, and $(g^\gamma)^{\varepsilon'} = g^\varepsilon$ for all $g \in G$. That is, the diagram below commutes.

$$\begin{array}{ccccc} N & \rightarrowtail & G & \overset{\varepsilon}{\twoheadrightarrow} & Q \\ \| & & \gamma\downarrow & & \| \\ N & \rightarrowtail & G' & \overset{\varepsilon'}{\twoheadrightarrow} & Q \end{array}$$

It can be shown that equivalent extensions have the same coupling.

Let $\mathcal{E}(Q, N)$ be the set of equivalence classes of extensions of N by Q, and let $\mathcal{E}_\chi(Q, N)$ be the subset of equivalence classes of extensions of N by Q with coupling χ. In either case, let $[\mathbf{e}]$ denote the equivalence class of the extension $\mathbf{e}$.

2.3 Automorphisms of group extensions

An *automorphism* of the group extension $\mathbf{e}$ is an isomorphism $\gamma : G \to G$ that is invariant on N; that is, $N^\gamma = N$. We will denote the group of all such automorphisms by $\operatorname{Aut} \mathbf{e}$, but other common notations are $\operatorname{Aut}(G; N)$ and $\operatorname{Aut}_N G$.

Any element $\gamma \in \operatorname{Aut} \mathbf{e}$ restricts to an automorphism $\gamma|_N$ of N, and hence induces an automorphism of Q that we will denote $\gamma|_Q$. More specifically $n^{\gamma|_N} = n^\gamma$ for all $n \in N$, and $q^{\gamma|_Q} = (q^t)(\gamma\varepsilon)$ for all $q \in Q$, where t is a transversal. If t' is another transversal, then q^t and $q^{t'}$ differ by an element of N. Since $\gamma\varepsilon$ is the identity on N by the exactness of $\mathbf{e}$, we see that $\gamma|_Q$ is well-defined.

Define a homomorphism

$$\rho : \operatorname{Aut} \mathbf{e} \to \operatorname{Aut} N \times \operatorname{Aut} Q$$

by $\rho(\gamma) = (\gamma|_N, \gamma|_Q)$. An issue of fundamental importance is to identify the image of ρ, denoted $\operatorname{Im} \rho$. This is the subgroup of *inducible pairs* $(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$ for which there is some $\gamma \in \operatorname{Aut} \mathbf{e}$ satisfying $\gamma|_N = \theta$ and $\gamma|_Q = \sigma$.

3 History

In this section we focus on the 1970's, describing Wells' original exact sequence and Buckley's insightful observation that $\operatorname{Im} \rho$ is a stabilizer group.

3.1 Charles Wells, 1971

The main result in Wells' paper [26] is the construction of an exact sequence built around $\operatorname{Aut} \mathbf{e}$, but Wells also gives a precise (and difficult) description of the elements in $\operatorname{Aut} \mathbf{e}$. We begin with the description of $\operatorname{Aut} \mathbf{e}$ as a way of motivating Wells' exact sequence.

The pair of functions (λ, α) for an extension $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ with coupling $\chi : Q \to \operatorname{Aut} N$ and transversal $t : Q \to G$ as described in Section 2 is called an *associated pair.* Recall that $\chi = \lambda \operatorname{Inn} N$ and $\alpha : Q \times Q \to N$ is a factor set for t. Wells proves that there is a bijection between $\operatorname{Aut} \mathbf{e}$ and a set of triples in $\operatorname{Aut} N \times \operatorname{Aut} Q \times N^Q$, where N^Q denotes the set of all functions $\phi : Q \to N$, satisfying

certain conditions. We will use the equations from section 4 of [18], which seem a bit more usable than Wells' equations, and adopt Robinson's and Wells' convention of writing N additively.

Theorem 3.1 (Wells [26], Robinson [18]) *Using the notation above, there is a bijection between* $\operatorname{Aut}\mathbf{e}$ *and triples* $(\theta, \sigma, \phi) \in \operatorname{Aut} N \times \operatorname{Aut} Q \times N^Q$ *satisfying the following equations for all* $n \in N$ *and* $q_1, q_2 \in Q$*:*

$$q_1^\lambda \cdot \theta = \theta \cdot q_1^{\sigma\lambda} \cdot \iota(q_1^{\sigma\phi}) \tag{1}$$

$$((q_1, q_2)^\alpha)\theta = -(q_1^\sigma q_2^\sigma)\phi + (q_1^\sigma, q_2^\sigma)\alpha + (q_1^{\sigma\phi})(q_2^{\sigma\lambda}) + (q_2^\sigma)\phi \tag{2}$$

Under this bijection, $\gamma \in \operatorname{Aut}\mathbf{e}$ *is associated with the triple* $(\gamma|_N, \gamma|_Q, \phi)$*, where* ϕ *is defined by equation (1) above. Conversely, a triple* (θ, σ, ϕ) *is associated with the automorphism* γ *defined by*

$$(q^t n)^\gamma = (q^\sigma)^t \cdot (q^\sigma)\phi \cdot (n^\theta),$$

for $q \in Q$ *and* $n \in N$*.*

In particular, we note that a pair $(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$ is inducible if and only if there is a function $\phi : Q \to N$ satisfying the equations above.

The complexity of the three equations in Theorem 3.1 may have led Wells to the more tractable notion of "compatibility." A pair $(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$, is a *compatible pair* for χ if it satisfies

$$\bar{\theta}^{-1} q^\chi \bar{\theta} = (q^\sigma)^\chi, \tag{3}$$

where $q \in Q$ and $\bar{\theta} = \theta \operatorname{Inn} N \in \operatorname{Out} N$. Wells denotes the subgroup of $\operatorname{Aut} N \times \operatorname{Aut} Q$ all all compatible pairs simply by C, but these pairs depend on χ so a more clear notation (used by several authors) is $\operatorname{Comp}(\chi)$. One can check that inducible pairs are compatible so we have

$$\operatorname{Im}\rho \subseteq \operatorname{Comp}(\chi).$$

Next, Wells constructs a set map $\omega : \operatorname{Comp}(\chi) \to H^2(Q, ZN)$ by building a 2-cocycle from a compatible pair (θ, σ). By compatibility, we know that $\theta^{-1} q^\lambda \theta = (q^\sigma)^\lambda$ modulo an inner automorphism of N. The function $\phi : Q \to N$ associates $q \in Q$ with the element n that induces the inner automorphism. Then Wells defines a 2-cocycle $k : Q \times Q \to ZN$ that essentially measures how ϕ deviates from satisfying equation (2). He notes that $k = 1 \in H^2(Q, ZN)$ if and only if (θ, σ) is inducible, thus $\operatorname{Ker}\omega = \operatorname{Im}\rho$. (Another interpretation of ω via group actions will be given in Section 5, equation (4).)

The last piece needed for Wells' exact sequence is $\operatorname{Ker}\rho$. First note that ZN has the structure of a Q-module via $\lambda : Q \to \operatorname{Aut} N$, where

$$m^q := m^{q\lambda} = (q^t)^{-1} m (q^t)$$

for $m \in ZN$ and $q \in Q$.

Wells denotes $\operatorname{Ker}\rho$ by $Z^1_\chi(Q, ZN)$, but it is more commonly denoted $\operatorname{Der}(Q, ZN)$, where $\operatorname{Der}(Q, ZN)$ is the group of *derivations* $\delta : Q \to ZN$ satisfying

$$(q_1 q_2)^\delta = (q_1^\delta)^{q_2} + (q_2^\delta), \ \forall q_i \in Q.$$

Define $\psi : \mathrm{Der}(Q, ZN) \to \mathrm{Aut}\,\mathbf{e}$ by

$$g(\delta^\psi) = g(g^\varepsilon)^\delta, \text{ where } g \in G.$$

It is easy to see that δ^ψ restricts to the identity on N and induces the identity on Q, so is in $\mathrm{Ker}\,\rho$.

Now we can describe Wells' exact sequence.

Theorem 3.2 (Wells [26]) *Given a group extension of the form*

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q,$$

there is an exact sequence

$$0 \to \mathrm{Der}(Q, ZN) \xrightarrow{\psi} \mathrm{Aut}\,\mathbf{e} \xrightarrow{\rho} \mathrm{Comp}(\chi) \xrightarrow{\omega} H^2(Q, ZN),$$

where ω is a set map, and $\mathrm{Comp}(\chi)$ is the subgroup of $\mathrm{Aut}\,N \times \mathrm{Aut}\,Q$ consisting of compatible pairs.

A few notes concerning the theorem:

- By Wells' time it was well known that when N is abelian there is a bijection between $H^2(Q, N)$ and $\mathcal{E}_\chi(Q, N)$ (see Urs Stammbach's 1973 book [25], for example), so cohomology classes were known to contain information about equivalent and isomorphic group extensions, but much of the prior work was focused on computing cohomology groups via spectral sequences and other means. Wells may not have been the first to study $\mathrm{Aut}\,\mathbf{e}$, but his exact sequence is the most enduring tool.
- It turns out that ω depends on $[\mathbf{e}]$ so we will sometimes adopt the notation in [13] and write $\omega(\mathbf{e})$ if the dependency is important.
- A pair $(\theta, \sigma) \in \mathrm{Aut}\,N \times \mathrm{Aut}\,Q$ is inducible if and only if $(\theta, \sigma)^\omega$ is the identity in $H^2(Q, ZN)$.

3.2 Joseph Buckley, 1974

Buckley [1] was not particularly interested in the Wells exact sequence–he barely mentions the set map ω–but he provided a new interpretation of $\mathrm{Comp}(\chi)$ as a stabilizer group that has proved to be invaluable. Indeed, we will generally refer to "Buckley's group action" in conjunction with the Wells exact sequence in order to give Buckley the credit he deserves.

There is a right group action of $\mathrm{Aut}\,N \times \mathrm{Aut}\,Q$ on $\mathcal{E}(Q, N)$ defined by $[\mathbf{e}] \cdot c = [\mathbf{e}^\mathbf{c}]$, where $c = (\theta, \sigma) \in \mathrm{Aut}\,N \times \mathrm{Aut}\,Q$ and $\mathbf{e}^\mathbf{c}$ is the extension

$$N \overset{\theta^{-1}\mu}{\rightarrowtail} G \overset{\varepsilon\sigma}{\twoheadrightarrow} Q.$$

The coupling associated with $\mathbf{e}^\mathbf{c}$ is $\sigma^{-1}\chi\iota(\bar{\theta})$. Note that instead of $\mathbf{e}^\mathbf{c}$, Buckley writes $\mathbf{e}(\theta, \sigma)$.

We see that $\mathbf{e}^\mathbf{c}$ has coupling χ if and only if $\sigma^{-1}\chi\iota(\bar{\theta}) = \chi$. This is equivalent to (3) so we see that the subgroup $\mathrm{Comp}(\chi)$ acts on the right of the set $\mathcal{E}_\chi(Q, N)$.

Theorem 3.3 (Buckley [1]) *Given a group extension of the form*

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q,$$

there is an exact sequence

$$0 \to \mathrm{Der}(Q, ZN) \xrightarrow{\psi} \mathrm{Aut}\,\mathbf{e} \xrightarrow{\rho} \mathrm{Comp}(\chi) \xrightarrow{\omega} H^2(Q, ZN),$$

where ω is a set map, and the image of ρ is the stabilizer of $[\mathbf{e}]$ under the action of $\mathrm{Comp}(\chi)$ *on $\mathcal{E}_\chi(Q, N)$.*

The rest of Buckley's paper concerns extensions with trivial coupling, and uses the theorem above applied to p-groups to help prove a result about the size of $\mathrm{Aut}\, G$.

3.3 Others

In the mid-1960s, prior to Wells' paper, Wolfgang Gaschütz used cohomological techniques to first prove the existence of outer automorphisms of finite p-groups [5], and then prove that non-abelian finite p-groups have outer p-automorphisms [6]. A series of papers extended Gaschütz's celebrated theorems, including one by Peter Schmid [23]. Also using cohomological techniques, including the Wells exact sequence, Schmid found conditions under which $\mathrm{Out}\, P$ has a non-trivial normal p-subgroup. (Ping Jin [11] takes this theorem a step further, as we will see in Section 4.)

It is worth noting that Kung-Wei Yang, a colleague of Buckley's at Western Michigan University, wrote a paper [27] appearing in 1974 that might have been a precursor to Buckley's work. Yang mentions the Wells exact sequence, but his study of equivalent extensions does not yet include the language of group actions that makes Buckley's work so enduring. Still, something good was brewing at Western Michigan.

Derek Robinson [18] wrapped up the 1970's with his thorough presentation of applications of cohomology to group theory in 1982, including a complete discussion of Wells' work. Robinson's study of Wells' exact sequence in special cases resulted in new proofs of the theorems of Gaschütz', Schmid and others. This remains a "go-to" reference, because of its scope and attention to detail.

4 Rebirth

After Robinson's paper in 1982 and a related follow-up in 1984 [19], there are no references (at least none that this author could find in a thorough search of the *MathSciNet*© database) to either Wells' work or Buckley's until the early 2000's. As noted above, in 2002 Jin [11] used the Wells exact sequence and other cohomological tools to extend Schmid's work to a larger class of groups (finite p-nilpotent groups). At this time, Jin makes no mention of group actions on $\mathcal{E}(Q, N)$ in general, nor of Buckley's work in particular.

Also in 2002, Wim Malfait [14] examines $\mathrm{Im}\,\rho$ in detail and essentially extends the Wells exact sequence to a 27 term, cubic commutative diagram with exact rows and columns. Malfait does not mention Buckley's work directly, but was clearly influenced by it–probably via Robinson's work in [18] and [19]–since he describes a group action of $\mathrm{Aut}\, N \times \mathrm{Aut}\, Q$ on $H^2(Q, ZN)$ and identifies $\mathrm{Im}\,\rho$ as a stabilizer.

In 2003, John Martino and Stewart Priddy [15] seemingly rediscover Buckley's influence and tweak his work by considering the action of $\mathrm{Comp}(\chi)$ on $\mathcal{E}_\chi(Q,N)$ as two separate actions, and then computing what they call the "intersection orbit group." The intersection orbit group gives Martino and Priddy a useful tool for computing $|\mathrm{Im}\,\rho|$. One might surmise that Martino and Priddy became aware of Buckley's work on $\mathrm{Aut}\,\mathbf{e}$ because Martino and Buckley were colleagues at Western Michigan University.

In 2007, this author [3] uses the Wells exact sequence and Buckley's group action to prove that if the extension $\mathbf{e}$ splits and N is abelian, then the extension

$$\mathbf{e_W} : 1 \to \mathrm{Der}(Q,N) \to \mathrm{Aut}\,\mathbf{e} \to \mathrm{Im}\,\rho \to 1$$

splits. Based on a query in Wells' original paper, we obtain a further characterization of the conditions under which $\mathbf{e_W}$ splits in [4]. We can say with absolute certainty that Dietz learned about the work of Wells and Buckley via Martino and Priddy.

In the mid-2000's, Marek Golasiński and Daciberg Lima Gonçalves computed automorphisms of groups, especially automorphisms of semi-direct products, as a means of counting homotopy types of spherical space forms (see [7] for one example). They did not specifically reference Wells or Buckley until later in the decade, perhaps after learning about the Wells exact sequence via [3] and [24], and through private communication with Dietz. In [9] Golasiński and Gonçalves interpret the Wells exact sequence in the case $G = N \rtimes Q$ and independently find that $\mathbf{e_W}$ splits when N is abelian. They further compute $\mathrm{Im}\,\rho$ when G is a split metacyclic group. Using the splitting of $\mathbf{e_W}$ in the case that A is a finitely generated abelian group with $A \cong T(A) \oplus F(A)$, where $T(A)$ is the torsion part of A and $F(A)$ is free abelian of finite rank, Golasiński and Gonçalves [8] write $\mathrm{Aut}\,A$ as a semi-direct product that is essentially determined by an action based on Buckley's group action.

Most of the results listed above *use* the work of Wells and Buckley, but do not *advance* their ideas in substantial ways. This begins to change toward the end of the decade, with three interesting papers published in the *Journal of Algebra.*

In 2007, Jin [12] restricts his attention to automorphisms of $\mathbf{e}$ that restrict to the identity on Q. More specifically, $(\theta, 1_Q)$ is a compatible pair if and only if $\theta \in C_{\mathrm{Aut}\,N}(Q^\chi)$. Then Jin defines an action of $C_{\mathrm{Aut}\,N}(Q^\chi)$ on $H^2(Q, ZN)$ in such a way that Wells' set map ω becomes a derivation when restricted to $C_{\mathrm{Aut}\,N}(Q^\chi)$. This is the first time we see Wells' function ω described as anything but a set map.

There are two notions related to inducibility that we remark on here. First recall that a pair $(\theta, \sigma) \in \mathrm{Aut}\,N \times \mathrm{Aut}\,Q$ is inducible if there is some $\gamma \in \mathrm{Aut}\,\mathbf{e}$ such that $\gamma|_N = \theta$ and $\gamma|_Q = \sigma$; that is the group of inducible pairs is equal to $\mathrm{Im}\,\rho$.

We will say that an element $\theta \in \mathrm{Aut}\,N$ *extends* to an automorphism of G if there exists $\gamma \in \mathrm{Aut}\,G$ such that $\gamma|_N = \theta$. Necessarily, $\gamma \in \mathrm{Aut}\,\mathbf{e}$. Of particular interest are the extendable automorphisms of N that induce the identity on Q, so that $(\theta, 1_Q) \in \mathrm{Im}\,\rho$.

On the other hand, we will say that an element $\sigma \in \mathrm{Aut}\,Q$ *lifts* to an automorphism of G if there exists $\gamma \in \mathrm{Aut}\,G$ such that $\gamma|_Q = \sigma$. Necessarily, $\gamma \in \mathrm{Aut}\,\mathbf{e}$. Of particular interest are the liftable automorphisms of Q that induce the identity on N, so that $(1_N, \sigma) \in \mathrm{Im}\,\rho$.

Jin [12] describes necessary and sufficient conditions for $(\theta, 1_Q)$ to be an inducible

pair. In particular, he reduces this particular extension problem to an extension problem involving Sylow subgroups.

In 2010, I.B.S. Passi, Mahender Singh, and Manoj Yadav [16] continue in this same vein. Not only do they consider extensions of automorphisms of N as well as lifts of automorphisms of Q, but they reduce the questions to analogous ones involving just some Sylow subgroups rather than all. Along the way, they establish two variations of the Wells exact sequence.

In 2010, Jin is joined by Heguo Liu [13], and together the pair proves in full generality that $\omega(\mathbf{e})$ is a derivation:

$$\omega(\mathbf{e}) \in \mathrm{Der}(\mathrm{Comp}(\chi), H^2(Q, ZN)).$$

They further show that the Wells map depends on the extension $\mathbf{e}$ and that the class of $\omega(\mathbf{e})$ in $H^1(\mathrm{Comp}(\chi), H^2(Q, ZN))$ is "the obstruction to every compatible pair being inducible in some extension."

These exciting developments have several consequences that we will say more about in Section 5.

Finally we reach 2013, when Robinson [22] uses properties of restriction and corestriction maps in cohomology (rather than referring directly to the Wells exact sequence) to further extend the results in [12] and [16] on the inducibility of general pairs of the form $(\theta, \sigma) \in \mathrm{Aut}\, N \times \mathrm{Aut}\, Q$ and the role of Sylow subgroups. Much of this work, including detailed proofs, appears in 2012 [21], but the 2013 paper is easier to locate.

The last 10 years have seen the resurrection of the Wells exact sequence in the service of understanding automorphisms of group extensions, and recognition of the value of Buckley's group action point-of-view.

5 Survey of Results

In this section we provide some details on applications and extensions of the Wells exact sequence and Buckley's group action, including all of the theorems described in Section 4. We no longer adhere to a chronological order in this section, but try to group together similar ideas.

5.1 The Wells map is a derivation

We begin with the details needed to understand two important theorems of Jin and Liu in [13].

Clearly inspired by Buckley, Jin and Liu investigate the interplay of three group actions: one is attributed to Buckley, and the other two were certainly known by Robinson in [18], who gives credit to a comprehensive series of lecture notes by Karl Gruenberg [10].

First, there is the action of $\mathrm{Comp}(\chi)$ on $\mathcal{E}_\chi(Q, N)$ described by Buckley that we recall here: If $c = (\theta, \sigma) \in \mathrm{Comp}(\chi)$ and $[\mathbf{e}] \in \mathcal{E}_\chi(Q, N)$, then $[\mathbf{e}] \cdot c = [\mathbf{e}^c]$, where $\mathbf{e}^\mathbf{c}$ is the extension

$$\mathbf{e}^\mathbf{c} : N \overset{\theta^{-1}\mu}{\rightarrowtail} G \overset{\varepsilon\sigma}{\twoheadrightarrow} Q.$$

Second, there is an action of $H^2(Q, ZN)$ on $\mathcal{E}_\chi(Q, N)$ described as follows. Let (λ, α) be an associated pair for an extension $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ with coupling χ and transversal t. We can define a new extension

$$\mathbf{e}(\lambda, \alpha) : N \rightarrowtail G(\lambda, \alpha) \twoheadrightarrow Q,$$

where the group $G(\lambda, \alpha)$ is the set of pairs $(q, n) \in Q \times N$ with the binary operation

$$(q_1, n_1)(q_2, n_2) = (q_1 q_2, (q_1, q_2)^\alpha + (n_1)q_2^\lambda + n_2), \text{ where } q_i \in Q, n_i \in N.$$

It can be shown that $\mathbf{e}$ and $\mathbf{e}(\lambda, \alpha)$ are equivalent extensions. Now let $[\beta] \in H^2(Q, ZN)$ where $\beta \in Z^2(Q, ZN)$, and define $\alpha\beta : Q \times Q \to N$ by

$$(q_1, q_2)^{(\alpha\beta)} = (q_1, q_2)^\alpha + (q_1, q_2)^\beta,$$

then $(\lambda, \alpha\beta)$ is an associated pair for an extension of N by Q with coupling χ. Thus, we can define a right action of $H^2(Q, ZN)$ on $\mathcal{E}_\chi(Q, N)$ by

$$[\mathbf{e}] \cdot [\beta] = [\mathbf{e}(\lambda, \alpha)] \cdot [\beta] = [\mathbf{e}(\lambda, \alpha\beta)].$$

This turns out to be a regular action so that given $c \in \mathrm{Comp}(\chi)$, there is a unique element $h \in H^2(Q, ZN)$ satisfying $[\mathbf{e^c}] \cdot h = [\mathbf{e}]$. It can be shown that the map

$$\omega(\mathbf{e}) : \mathrm{Comp}(\chi) \to H^2(Q, ZN)$$

defined by the equation

$$[\mathbf{e^c}] \cdot c^{\omega(\mathbf{e})} = [\mathbf{e}] \tag{4}$$

matches Wells' original definition of ω.

Third, there is an action of $\mathrm{Comp}(\chi)$ on $H^2(Q, ZN)$. Let $c = (\theta, \sigma) \in \mathrm{Comp}(\chi)$ and $[\zeta] \in H^2(Q, ZN)$ where $\zeta \in Z^2(Q, ZN)$. Define $[\zeta] \cdot c = [\zeta]^c = [\zeta^c]$, where $\zeta^c \in Z^2(Q, ZN)$ is defined by

$$(q_1, q_2)^{\zeta^c} = (q_1^{\sigma^{-1}}, q_2^{\sigma^{-1}})\zeta\theta.$$

This last action allows one to form the semi-direct product

$$\Gamma = \mathrm{Comp}(\chi) \ltimes H^2(Q, ZN).$$

Combining the first two actions, Jin and Liu prove their Theorem A and Corollary B, stated next.

Theorem 5.1 (Jin and Liu [13]) *Given a group extension of the form*

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$$

with coupling χ, there is a group action of $\Gamma = \mathrm{Comp}(\chi) \ltimes H^2(Q, ZN)$ on $\mathcal{E}_\chi(Q, N)$ defined by

$$[\mathbf{e}] \cdot (ch) = [\mathbf{e}]^c \cdot h,$$

where $[\mathbf{e}] \in \mathcal{E}_\chi(Q, N)$, $c \in \mathrm{Comp}(\chi)$, and $h \in H^2(Q, ZN)$. Furthermore, the stabilizer of $[\mathbf{e}]$ in Γ is a complement of $H^2(Q, ZN)$ in Γ, and the set of all such stabilizers is a single conjugacy class of subgroups of Γ.

Theorem 5.2 (Jin and Liu [13]) *For any group extension of the form*

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$$

with coupling χ, the Wells map $\omega(\mathbf{e})$ is a derivation from $\mathrm{Comp}(\chi)$ *into* $H^2(Q, ZN)$ *under the natural action of* $\mathrm{Comp}(\chi)$ *on* $H^2(Q, ZN)$*. Further, if* $\mathbf{e}'$ *is another extension of N by Q with coupling χ, then $\omega(\mathbf{e})$ and $\omega(\mathbf{e}')$ differ by an inner derivation.*

5.2 Variations on the Wells exact sequence

Several authors describe variations on the Wells exact sequence, each using his or her own notation. Here we use the notation from [16] as the common language. We will use superscripts on $\mathrm{Aut}\, G$ to denote automorphisms that *centralize* a subgroup by fixing its elements, and subscripts to denote automorphisms that *normalize* a subgroup by acting invariantly on it. More specifically we have

- $\mathrm{Aut}\, \mathbf{e} = \mathrm{Aut}_N(G)$ is the group of automorphisms γ for which $\gamma(N) = N$.
- $\mathrm{Aut}^N(G)$ is the group of automorphisms γ for which $\gamma(n) = n$ for all $n \in N$.
- $\mathrm{Aut}_N^Q(G)$ is the group of automorphisms in $\mathrm{Aut}_N(G)$ that induce the identity on Q.
- $\mathrm{Aut}^{N,Q}(G)$ is the group of automorphisms in $\mathrm{Aut}^N(G)$ that induce the identity on Q.

We will also need to consider subgroups of $\mathrm{Comp}(\chi)$. Let

$$C_1 = \{\theta \in \mathrm{Aut}\, N \mid (\theta, 1_Q) \in \mathrm{Comp}(\chi)\}$$

and

$$C_2 = \{\sigma \in \mathrm{Aut}\, Q \mid (1_N, \sigma) \in \mathrm{Comp}(\chi)\}.$$

The first variation on Wells' exact sequence that we will mention comes from Robinson in [18] (Theorem 4.4), where he considers extensions that have injective coupling. We can deduce that $\chi : Q \to \mathrm{Out}\, N$ is injective exactly when $C_G(N) = ZN$. Let $\rho_N : \mathrm{Aut}\, \mathbf{e} \to \mathrm{Aut}\, N$ be the restriction of an automorphism to N. It turns out that $\mathrm{Ker}\, \rho_N = \mathrm{Ker}\, \rho = \mathrm{Der}(Q, ZN)$ when χ is injective. If $\theta \in \mathrm{Im}\, \rho_N$ then there exists $\sigma \in \mathrm{Aut}\, Q$ such that (θ, σ) is inducible. The compatibility condition in equation (3) implies that $\bar{\theta}$ normalizes Q^χ so ρ_N maps into $N_{\mathrm{Aut}\, N}(Q^\chi)$.

Theorem 5.3 (Robinson [18]) *Let $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ be an extension with injective coupling χ, then there is an exact sequence*

$$0 \to \mathrm{Der}(Q, ZN) \to \mathrm{Aut}\, \mathbf{e} \to N_{\mathrm{Aut}\, N}(Q^\chi) \to H^2(Q, ZN).$$

Robinson eventually uses this theorem and others to construct outer automorphisms of free abelianized extensions.

Next we see a special case of the Wells exact sequence applied to central products in [3] (Theorem 5.17). Let $G = A \circ B$ be a central product so that $[A, B] = 1$, and let $C = A \cap B$. From G we can build the extension

$$B \rightarrowtail G \twoheadrightarrow A/C$$

that has trivial coupling χ. In this case $\mathrm{Comp}(\chi) = \mathrm{Aut}(A/C) \times \mathrm{Aut}\, B$.

Theorem 5.4 (Dietz [3]) *Let $G = A \circ B$ be as above, then there is an exact sequence*

$$1 \to \mathrm{Hom}(A, ZB) \to \mathrm{Aut}_B^C(G) \xrightarrow{\nu} \mathrm{Aut}(A/C) \times \mathrm{Aut}^C B \xrightarrow{\omega} H^2(A/C, ZB)$$

where ν is the restriction of ρ to $\mathrm{Aut}_B^C(G)$ and ω is the Wells map restricted to $\mathrm{Aut}(A/C) \times \mathrm{Aut}^C B \leq \mathrm{Aut}(A/C) \times \mathrm{Aut}\, B$.

In 2007, Jin [12] concentrates on the subgroup C_1 of $\mathrm{Comp}(\chi)$ and we see the Wells map as a derivation for the first time in Theorem A. Note that an element $(\theta, 1_Q)$ is compatible if and only if

$$\bar{\theta}^{-1} q^\chi \bar{\theta} = q^\chi$$

by equation (3); that is, $(\theta, 1_Q) \in C_1$ if and only if $\theta \in C_{\mathrm{Aut}\, N}(Q^\chi)$.

Theorem 5.5 (Jin [12]) *Given a group extension of the form*

$$\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q,$$

there is an exact sequence

$$0 \to \mathrm{Der}(Q, ZN) \to C_{\mathrm{Aut}\, G}(Q) \to C_{\mathrm{Aut}\, G}(Q^\chi) \xrightarrow{\omega} H^2(Q, ZN),$$

where the derivation ω is the restriction of Wells' map ω to compatible pairs of the form $(\theta, 1_Q)$.

Continuing in the same vein, but imposing the restriction that N is abelian, we get three exact sequences in [16] (Theorems 1 and 2).

Theorem 5.6 (Passi, Singh, and Yadav [16]) *Let $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ be an extension with N abelian, then there exist the following two exact sequences:*

$$1 \to \mathrm{Aut}^{N,Q}(G) \to \mathrm{Aut}_N^Q(G) \xrightarrow{\rho_1} C_1 \xrightarrow{\omega_1} H^2(Q, N)$$

and

$$1 \to \mathrm{Aut}^{N,Q}(G) \to \mathrm{Aut}^N(G) \xrightarrow{\rho_2} C_2 \xrightarrow{\omega_2} H^2(Q, N),$$

where ρ_i is the restriction of ρ to the indicated subgroup and ω_i is the restriction of ω to C_i.[1] *Note that ω_i is not necessarily a homomorphism.*

Theorem 5.7 (Passi, Singh, and Yadav [16]) *Let $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ be a central extension (i.e., $N \leq Z(G)$), then there is an exact sequence*

$$1 \to \mathrm{Aut}^{N,Q}(G) \to \mathrm{Aut}_N(G) \xrightarrow{\rho} \mathrm{Aut}\, N \times \mathrm{Aut}\, Q \xrightarrow{\omega} H^2(Q, N).$$

Robinson alludes to the three sequences above in his discussion of extensions with abelian kernel, but he does not directly mention a variation on the Wells exact sequence except in the case that the coupling is injective (see Theorem 4.4 in [18]).

Finally, we will mention again that Malfait [14] produces a 27-term, cubic commutative diagram related to the Wells exact sequence. We will not reproduce his Theorem 4.10 here.

[1] This statement is true modulo a coboundary. That is, the functions ω_i (denoted λ_i in the original) may not be exactly equal to the restriction of Wells' ω, but they will be cohomologous.

5.3 Splitting the Wells exact sequence

In Wells' original paper he states that "it would be interesting to know precisely under what conditions" $\operatorname{Aut}\mathbf{e}$ splits over $Z^1(Q, ZN)$. In such a case we would know

$$\operatorname{Aut}\mathbf{e} \cong \operatorname{Der}(Q, ZN) \rtimes \operatorname{Im}\rho,$$

and together with information on inducible pairs (see Subsection 5.5), one would know exactly what $\operatorname{Aut}\mathbf{e}$ is. Wells' question is not precisely answered, but there are a few things known about when the sequence

$$\mathbf{e_W} : 1 \to \operatorname{Der}(Q, N) \to \operatorname{Aut}\mathbf{e} \to \operatorname{Im}\rho \to 1$$

and its variations split.

Though Robinson and others surely knew this early on, we see the following in print in 2007 (Theorem 4.6).

Theorem 5.8 (Dietz [3]) *Let $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ be a split extension with N abelian, then $\mathbf{e_W}$ splits.*

Note that Golasiński and Gonçalves prove the same result in [9], and John Curran proves it directly in [2] without using cohomological methods.

Using the notation in [16], let $C_i^* = \{\gamma \in C_i \mid \gamma^{\omega_i} = 1\}$. Then the exact sequences in Theorem 5.6 yield the following exact sequences when N is abelian:

$$1 \to \operatorname{Aut}^{N,Q}(G) \to \operatorname{Aut}_N^Q(G) \xrightarrow{\rho_1} C_1^* \to 1 \tag{5}$$

and

$$1 \to \operatorname{Aut}^{N,Q}(G) \to \operatorname{Aut}^N(G) \xrightarrow{\rho_2} C_2^* \to 1. \tag{6}$$

When the extension is central, Theorem 5.7 yields the exact sequence

$$1 \to \operatorname{Aut}^{N,Q}(G) \to \operatorname{Aut}_N(G) \xrightarrow{\rho} C^* \to 1, \tag{7}$$

where $C^* = \{(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q \mid (\theta, \sigma)^\omega = 1\}$. Theorem 8 in [16] states:

Theorem 5.9 (Passi, Singh, and Yadav [16]) *Let G be a finite group and N an abelian normal subgroup of G such that $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ splits. Then sequences* (5) *and* (6) *split. Furthermore, if the extension is central then the sequence* (7) *splits.*

Finally, in 2012 we get some information on the splitting homomorphism for $\mathbf{e_W}$ (Theorem 2.2 of [4]). Assume $G = N \rtimes Q$ and for $(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$, define $\theta * \sigma : G \to G$ by

$$(nq)^{\theta * \sigma} = n^\theta q^\sigma.$$

If $\theta * \sigma$ is a homomorphism, it will be in $\operatorname{Aut}\mathbf{e}$ and (θ, σ) will be an inducible pair.

Theorem 5.10 (Dietz [4]) *Let $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ be a split extension and consider $G = N \rtimes Q$. As long as $\operatorname{Der}(Q, ZN)$ is non-trivial, the extension*

$$\mathbf{e_W} : 1 \to \operatorname{Der}(Q, N) \to \operatorname{Aut}\mathbf{e} \to \operatorname{Im}\rho \to 1$$

splits if and only if $\theta * \sigma \in \operatorname{Aut}\mathbf{e}$ for all $(\theta, \sigma) \in \operatorname{Im}\rho$. In this case, the assignment $(\theta, \sigma) \mapsto \theta * \sigma$ is a splitting homomorphism.

5.4 The role of Sylow subgroups

Questions of inducibility can be reduced to questions about lifting and extending automorphisms that come from Sylow subgroups.

Below we state Theorem D in [12] that gives a necessary and sufficient condition for a pair $(\theta, 1_Q) \in \operatorname{Aut} N \times \operatorname{Aut} Q$ to be inducible, where we identify Q with G/N.

Theorem 5.11 (Jin [12]) *Let N be a normal subgroup of G. Then $\theta \in \operatorname{Aut} N$ extends to an automorphism of G inducing the identity on G/N if and only if for each Sylow subgroup P/N of G/N, θ extends to an automorphism of P inducing the identity on P/N.*

Next are Theorem 7 and part of Theorem 4 from [16], rewritten in the language established in this paper, and with Q identified with G/N. The first theorem below extends Jin's theorem above, but under the restriction that N is an abelian group. The second theorem below concerns lifting automorphisms of G/N.

Theorem 5.12 (Passi, Singh, and Yadav [16]) *Let N be an abelian normal subgroup of a finite group G. Then $\theta \in \operatorname{Aut} N$ extends to an automorphism of G inducing the identity on G/N if and only if for some Sylow p-subgroup P/N of G/N, for each prime number p dividing $|G/N|$, θ extends to an automorphism of P inducing the identity on P/N.*

Theorem 5.13 (Passi, Singh, and Yadav [16]) *Let N be an abelian normal subgroup of a finite group G. Then $\sigma \in \operatorname{Aut}(G/N)$ lifts to an automorphism of G that is the identity on N provided the restriction of σ to some Sylow p-subgroup P/N of G/N, for each prime number p dividing $|G/N|$, lifts to an automorphism of P that is the identity on N.*

Thus we see that the theorems above give conditions under which pairs of the form $(\theta, 1_Q)$ and $(1_N, \sigma)$ in $\operatorname{Aut} N \times \operatorname{Aut} Q$ are inducible.

Robinson extends all three theorems from above by considering the inducibility of general pairs of the form (θ, σ). Before we state Theorem 2 in [22], we need some notation.

Assume Q is finite and let $\pi(Q) = \{p_1, p_2, \ldots, p_k\}$ be the complete set of distinct primes dividing $|Q|$. Let $P_i = R_i/N$ be a Sylow p_i-subgroup of Q, where $R_i \leq G$. Then we have *subextensions* of the form

$$\mathbf{e_i} : N \rightarrowtail R_i \twoheadrightarrow P_i$$

with couplings $\chi_i = \chi|_{P_i}$. Let $(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$, then P_i^σ is a Sylow p_i-subgroup of Q, hence there exists $q_i \in Q$ such that $P_i^\sigma = q_i P_i q_i^{-1}$. Thus $P_i^{\sigma\iota(q_i)} = P_i$ and we see that $\sigma\iota(q_i) \in \operatorname{Aut} P_i$. Let $q_i^t = g_i$, then $\theta\iota(g_i) \in \operatorname{Aut} N$.

Theorem 5.14 (Robinson [22]) *With the notation above, the pair*

$$(\theta, \sigma) \in \operatorname{Aut} N \times \operatorname{Aut} Q$$

is inducible to $\operatorname{Aut} \mathbf{e}$ *if and only if* $(\theta\iota(g_i), \sigma\iota(q_i))$ *is inducible to* $\operatorname{Aut} \mathbf{e_i}$ *for all* $i = 1, 2, \ldots, k$.

Robinson further investigates the case when Q is not necessarily finite, but is *locally finite*. The set-up to describe Theorem 3 in [22] will take us too far astray, so we refer the reader to Robinson's paper.

5.5 Inducible pairs

Certainly all five theorems in Subsection 5.2 concern inducible pairs, but there are other results.

In 1977, Robinson [17] uses the exact sequence described below to investigate groups with finitely many automorphisms, but in [18] he rephrases it in terms of Wells' work. Essentially he proves in Theorem 4.3 of [18] that if N is abelian then

$$\operatorname{Im}\rho = C_{\operatorname{Comp}(\chi)}([\mathbf{e}]).$$

Theorem 5.15 (Robinson [18]) *Let* $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ *be an extension with* N *abelian and coupling* χ*, then there is an exact sequence*

$$1 \to \operatorname{Der}(Q, N) \to \operatorname{Aut}\mathbf{e} \to C_{\operatorname{Comp}(\chi)}([\mathbf{e}]) \to 1.$$

Furthermore, if the extension is central so that χ is trivial, then it is easy to see that $\operatorname{Comp}(\chi) = \operatorname{Aut} N \times \operatorname{Aut} Q$ and the sequence above becomes

$$1 \to \operatorname{Hom}(Q_{ab}, N) \to \operatorname{Aut}\mathbf{e} \to C_{\operatorname{Aut} N \times \operatorname{Aut} Q}([\mathbf{e}]) \to 1.$$

In 2004 we see conditions under which a pair (θ, σ) is inducible if and only if both $(\theta, 1_Q)$ and $(1_N, \sigma)$ are inducible. Suppose $\mathbf{e} : N \rightarrowtail G \twoheadrightarrow Q$ is split so that we can identify Q with a subgroup of G and write $G = N \rtimes Q$. Decompose $\rho : \operatorname{Aut}\mathbf{e} \to \operatorname{Aut} N \times \operatorname{Aut} Q$ as $\rho = \rho_N \times \rho_Q$ where $\rho_N(\gamma) = \gamma|_N$ and $\rho_Q(\gamma) = \gamma|_Q$ for $\gamma \in \operatorname{Aut}\mathbf{e}$. Putting the work of [3] into established language and notation, and combining Theorems 4.5 and 4.7, we get the following theorem.

Theorem 5.16 (Dietz [3]) *Let* $G = N \rtimes Q$*. If either* $C_1 = \operatorname{Aut} N \times \{1_Q\}$ *or* $C_2 = \{1_N\} \times \operatorname{Aut} Q$*, then* $\operatorname{Im}\rho = \operatorname{Im}\rho_N \times \operatorname{Im}\rho_Q$*. If we further know that* N *is abelian, then the previous hypothesis implies* $\operatorname{Im}\rho \cong C_1 \times C_2$*.*

The first part of the theorem above says that if either ρ_N or ρ_Q is surjective, then a pair (θ, σ) is inducible if and only if both $(\theta, 1_Q)$ and $(1_N, \sigma)$ are inducible.

Jin and Liu call the class of $\omega(\mathbf{e})$ in $H^1(\operatorname{Comp}(\chi), H^2(Q, ZN))$ the *associated cohomology element* with the coupling χ and denote it by $[\chi]$. The vanishing of $[\chi]$ has consequences on the inducibility of a pair in $\operatorname{Aut} N \times \operatorname{Aut} Q$. We site Corollary C and Theorem D of [13] below.

Theorem 5.17 (Jin and Liu [13]) *The following statements are equivalent:*

1. $[\chi]$ *vanishes in* $H^1(\operatorname{Comp}(\chi), H^2(Q, ZN))$*.*
2. *For some extension* $\mathbf{e}$ *with coupling* χ*, the stabilizer of* $[\mathbf{e}]$ *under the action of* Γ *on* $\mathcal{E}_\chi(Q, N)$ *is equal to the stabilizer of* $[\mathbf{e}]$ *under Buckley's action of* C *on* $\mathcal{E}_\chi(Q, N)$*.*
3. *There exists an extension* $\mathbf{e}$ *with coupling* χ *such that* $\rho : \operatorname{Aut}\mathbf{e} \to \operatorname{Comp}(\chi)$ *is surjective.*

An element $c \in \mathrm{Comp}(\chi)$ is *absolutely inducible* for χ if c is inducible for each extension with coupling χ.

Theorem 5.18 (Jin and Liu [13]) *If $[\chi]$ vanishes, then an element $c \in \mathrm{Comp}(\chi)$ is absolutely inducible if and only if c acts trivially on $H^2(Q, ZN)$.*

5.6 An algorithm for determining automorphisms of p-groups

Martino and Priddy [15] use their notion of an intersection orbit group and the Wells exact sequence in the case of trivial coupling to describe an inductive procedure for computing $\mathrm{Aut}\, P$ when P is a p-group.

Let $\{\Gamma^n(P)\}$ be the *mod p lower central series*

$$P = \Gamma^0(P) \geq \Gamma^1(P) \geq \Gamma^2(P) \geq \cdots \geq \Gamma^n(P) = \{1\}$$

with

$$\Gamma^n(P) = \langle [g_1, g_2, \ldots, g_s]^{p^k} \mid sp^k > n \rangle,\ n \geq 1$$

where

$$[g_1, g_2, \ldots, g_s] = [g_1, [g_2, [\ldots [g_{s-1}, g_s] \ldots]]]$$

is the s-fold iterated commutator. We see that $\Gamma^1(P) = \Phi(P)$ is the Frattini subgroup of P, and $P/\Gamma^1(P)$ is an elementary abelian p-group. Since P is finite, $\Gamma^n(P) = 1$ for some n.

Since $\Gamma^i(P)$ is characteristic in P, the natural homomorphism

$$\rho_V : \mathrm{Aut}\, P \to \mathrm{Aut}(P/\Gamma^1(P))$$

factors as

$$\mathrm{Aut}\, P \to \cdots \to \mathrm{Aut}(P/\Gamma^{i+1}(P)) \to \mathrm{Aut}(P/\Gamma^i(P)) \to \cdots \to \mathrm{Aut}(P/\Gamma^1(P)). \quad (8)$$

Applying Wells' and Buckley's ideas to various extensions associated with the mod p lower central series, Martino and Priddy show how to lift an element $f \in \mathrm{Im}\, \rho_V$ up the "ladder" in (8) to $\hat{f} \in \mathrm{Aut}\, P$. Note that $\mathrm{Aut}(P/\Gamma^1(P)) \cong \mathrm{GL}_r(\mathbb{F}_p)$ for some r, so this group is well understood; nonetheless, the inductive procedure is difficult, and actually applying it requires knowledge of the second cohomology group of many different groups. (There are several interesting examples in [15] where the inductive process ends quickly.)

The theorems given above are certainly not an exhaustive list of applications of the Wells exact sequence and Buckley's group action. In particular, we have not included results on outer automorphism groups as seen in [23], [11], [18], and others. However, this survey should give one a sense of the history of the Wells exact sequence over the last 40 years, and some interesting new views on it in recent years.

References

[1] Joseph Buckley, Automorphism groups of isoclinic p-groups, *J. London Math. Soc.* **12** (1975), 37-44.

[2] M. John Curran, Automorphisms of semidirect products, *Math. Proc. R. Ir. Acad.* **108** (2008), 205–210.

[3] Jill Dietz, On automorphisms of products of groups, *Groups St Andrews 2005, Vol. 1*, London Math. Soc. Lecture Note Ser. **339** (CUP, Cambridge, 2007), 288–305.
[4] Jill Dietz, Automorphism groups of semi-direct products, *Comm. Algebra* **40** (2012), 3308–3316.
[5] Wolfgang Gaschütz, Kohomologische Trivialitäten und äußere Automorphismen von p-Gruppen. *Math. Z.* **88** (1965), 432–433.
[6] Wolfgang Gaschütz, Nichtabelsche p-Gruppen besitzen äussere p-automorphismen. *J. Algebra* **4** (1966), 1–2.
[7] Marek Golasiński and Daciberg Lima Gonçalves, Spherical space forms–Homotopy types and self-equivalences for the groups $\mathbb{Z}/a \rtimes \mathbb{Z}/b$ and $\mathbb{Z}/a \rtimes (\mathbb{Z}/b \times \mathbb{Q}_{2^i})$, *Topology and its Applications* **146/147** (2005), 451–470.
[8] Marek Golasiński and Daciberg Lima Gonçalves, On automorphisms of finite abelian p-groups, *Math. Slovaca* **58** (2008), 405–412.
[9] Marek Golasiński and Daciberg Lima Gonçalves, On automorphisms of split metacyclic groups, *Manuscripta Math.* **128** (2009), 251–273.
[10] Karl Gruenberg, *Cohomological topics in group theory*, Lecture Notes Math. **143**, Springer-Verlag, Berlin-New York, 1970.
[11] Ping Jin, Outer automorphism groups of finite p-nilpotent groups, *Comm. Algebra* **30** (2002), 4369–4375.
[12] Ping Jin, Automorphisms of groups, *J. Algebra* **312** (2007), 562–569.
[13] Ping Jin and Heguo Liu, The Wells exact sequence for the automorphism group of a group extension, *J. Algebra* **324** (2010), 1219–1228.
[14] Wim Malfait, The (outer) automorphism group of a group extension, *Bull. Belg. Math. Soc.* **9** (2002), 361–372.
[15] John Martino and Stewart Priddy, Group extensions and automorphism group rings, *Homology Homotopy Appl.* **5** (2003), 53–70.
[16] I. B. S. Passi, Mahender Singh, and Manoj Yadav, Automorphisms of abelian group extensions, *J. Algebra* **324** (2010), 820–830.
[17] Derek J. S. Robinson, A contribution to the theory of groups with finitely many automorphisms, *Proc. London Math. Soc.* **35** (1977), 34–54.
[18] Derek J. S. Robinson, Applications of cohomology to the theory of groups, *Groups St Andrews 1981*, London Math. Soc. Lecture Note Ser. **71** (CUP, Cambridge-New York 1982), 46–80.
[19] Derek J. S. Robinson, Automorphisms of group extensions, *Algebra and its applications*, Lecture Notes in Pure and Appl. Math. **91** (Dekker, New York 1984), 163–167.
[20] Derek J. S. Robinson, *A Course in the Theory of Groups, Second Edition*, Graduate Texts Math. **80**, Springer-Verlag, New York, 1996.
[21] Derek J. S. Robinson, Inducibility of automorphism pairs in group extensions, in *Encuentro en Teoria de Grupos y sus Aplicaciones (Zaragoza 2011)*, Revista Matematica Iberoamericana, (Madrid 2012), 233–241.
[22] Derek J. S. Robinson, Automorphisms of group extensions, *Note di Matematica* **33** (2013), 121–129.
[23] Peter Schmid, Normal p-subgroups in the group of automorphisms of a finite p-group, *Math. Z.* **147** (1976), 271–277.
[24] Mark Schulte, Automorphisms of metacyclic p-groups with cyclic maximal subgroups, *Rose-Hulman Undergraduate Research Journal* **2** (2001).
[25] Urs Stammbach, *Homology in group theory*, Lecture Notes Math. **359**, Springer-Verlag, Berlin-New York, 1973.
[26] Charles Wells, Automorphisms of group extensions, *Trans. Amer. Math. Soc.* **155** (1971), 189–194.
[27] Kung-Wei Yang, Isomorphisms of group extensions, *Pac. J. Math* **50** (1974), 299–304.

RECENT WORK ON BEAUVILLE SURFACES, STRUCTURES AND GROUPS

BEN FAIRBAIRN

Department of Economics, Mathematics and Statistics, Birkbeck, University of London, Malet Street, London, WC1E 7HX
Email: b.fairbairn@bbk.ac.uk

Abstract

Beauville surfaces are a class of complex surfaces defined by letting a finite group G act on product of Riemann surfaces. These surfaces possess many attractive geometric properties several of which are dictated by properties of the group G. In this survey we discuss the groups that may be used in this way. *En route* we discuss several open problems, questions and conjectures.

1 Introduction

Roughly speaking (precise definitions will be given in the next section), a Beauville surface is a complex surface $\mathcal{S}$ defined by taking a pair of complex curves, i.e., Riemann surfaces, $\mathcal{C}_1$ and $\mathcal{C}_2$ and letting a finite group G act freely on their product to define $\mathcal{S}$ as a quotient $(\mathcal{C}_1 \times \mathcal{C}_2)/G$. These surfaces have a wide variety of attractive geometric properties: they are surfaces of general type; their automorphism groups [50] and fundamental groups [20] are relatively easy to compute (being closely related to G — see Section 7.2 and 7.3); these surfaces are rigid surfaces in the sense of admitting no nontrivial deformations [10] and thus correspond to isolated points in the moduli space of surfaces of general type [37].

Much of this good behaviour stems from the fact that the surface $(\mathcal{C}_1 \times \mathcal{C}_2)/G$ is uniquely determined by a particular pair of generating sets of G known as a 'Beauville structure'. This converts the study of Beauville surfaces to the study of groups with Beauville structures, i.e., Beauville groups.

Beauville surfaces were first defined by Catanese in [20] as a generalisation of an earlier example of Beauville [14, Exercise X.13(4)] (native English speakers may find the English translation [15] somewhat easier to read and get hold of) in which $\mathcal{C} = \mathcal{C}'$ and the curves are both the Fermat curve defined by the equation $X^5 + Y^5 + Z^5 = 0$ being acted on by the group $(\mathbb{Z}/5\mathbb{Z})\times(\mathbb{Z}/5\mathbb{Z})$ (this choice of group may seem somewhat odd at first, but the reason will become clear later). Bauer, Catanese and Grunewald went on to use these surfaces to construct examples of smooth regular surfaces with vanishing geometric genus [11]. Early motivation came from the consideration of the 'Friedman-Morgan speculation' — a technical conjecture concerning when two algebraic surfaces are diffeomorphic which Beauville surfaces provide counterexamples to. More recently, they have been used to construct interesting orbits of the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ (connections with Gothendeick's theory of dessins d'enfant make it possible for this group to act on the set of all Beauville surfaces). We will discuss this in slightly more detail in Section 7.6. Furthermore, Beauville's original

example has also recently been used by Galkin and Shinder in [34] to construct examples of exceptional collections of line bundles.

Like any survey article, the topics discussed here reflect the research interests of the author. Slightly older surveys discussing related geometric and topological matters are given by Bauer, Catanese and Pignatelli in [12, 13]. Other notable works in the area include [7, 51, 58, 62].

We remark that throughout we shall use the standard 'Atlas' natation for finite groups and related concepts as described in [24], excepting that we will occasionally deviate to minimise confusion with similar notation for geometric concepts.

In Section 2 we will introduce the preliminary definitions before proceeding in Section 3 to discuss the case of the finite simple groups. We then go on in Section 4 to discuss the abelian and nilpotent groups. Next, we focus our attention on special types of Beauville structures when we discuss strongly real Beauville structures in Section 5 and mixed Beauville structures in Section 6. Finally, we discuss a miscellany of related but less well studied topics in Section 7.

2 Preliminaries

Definition 2.1 A surface $\mathcal{S}$ is a *Beauville surface of unmixed type* if

- the surface $\mathcal{S}$ is isogenous to a higher product, that is, $\mathcal{S} \cong (\mathcal{C}_1 \times \mathcal{C}_2)/G$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are algebraic curves of genus at least 2 and G is a finite group acting faithfully on $\mathcal{C}_1$ and $\mathcal{C}_2$ by holomorphic transformations in such a way that it acts freely on the product $\mathcal{C}_1 \times \mathcal{C}_2$, and
- each $\mathcal{C}_i/G$ is isomorphic to the projective line $\mathbb{P}_1(\mathbb{C})$ and the covering map $\mathcal{C}_i \to \mathcal{C}_i/G$ is ramified over three points.

There also exists a concept of Beauville surfaces of mixed type but we shall postpone our discussion of these until Section 6. In the first of the above conditions the genus of the curves in question needs to be at least 2. It was later proved by Fuertes, González-Diez and Jaikin-Zapirain in [32] that in fact we can take the genus as being at least 6. The second of the above conditions implies that each $\mathcal{C}_i$ carries a regular dessin in the sense of Grothendieck's theory of *dessins d'enfants* (children's drawings) [45]. Furthermore, by Belyĭ's Theorem [16] this ensures that $\mathcal{S}$ is defined over an algebraic number field in the sense that when we view each Riemann surface as being the zeros of some polynomial we find that the coeffcients of that polynomial belong to some number field. Equivalently they admit an orientably regular hypermap [52], with G acting as the orientation-preserving automorphism group. A modern account of dessins d'enfants and proofs of Belyĭ's theorem may be found in the recent book of Girondo and González-Diez [38].

This can also be described instead in terms of uniformisation and the language of Fuchsian groups [40, 60].

What makes this class of surfaces so good to work with is the fact that all of the above definition can be 'internalised' into the group. It turns out that a group G can be used to define a Beauville surface if and only if it has a certain pair of generating sets known as a Beauville structure.

Definition 2.2 Let G be a finite group. Let $x, y \in G$ and let

$$\Sigma(x,y) := \bigcup_{i=1}^{|G|} \bigcup_{g \in G} \{(x^i)^g, (y^i)^g, ((xy)^i)^g\}.$$

An *unmixed Beauville structure* for the group G is a set of pairs of elements $\{\{x_1, y_1\}, \{x_2, y_2\}\} \subset G \times G$ with the property that $\langle x_1, y_1 \rangle = \langle x_2, y_2 \rangle = G$ such that

$$\Sigma(x_1, y_1) \cap \Sigma(x_2, y_2) = \{e\}.$$

If G has a Beauville structure we say that G is a *Beauville group*. Furthermore we say that the structure has *type*

$$((o(x_1), o(y_1), o(x_1y_1)), (o(x_2), o(y_2), o(x_2y_2))).$$

Traditionally, authors have defined the above structure in terms of so-called 'spherical systems of generators of length 3', meaning $\{x, y, z\} \subset G$ with $xyz = e$, but we omit $z = (xy)^{-1}$ from our notation in this survey. (The reader is warned that this terminology is a little misleading since the underlying geometry of Beauville surfaces is hyperbolic thanks to the below constraint on the orders of the elements.) Furthermore, many earlier papers on Beauville structures add the condition that for $i = 1, 2$ we have that

$$\frac{1}{o(x_i)} + \frac{1}{o(y_i)} + \frac{1}{o(x_iy_i)} < 1,$$

but this condition was subsequently found to be unnecessary following Bauer, Catanese and Grunewald's investigation of the wall-paper groups in [9]. A triple of elements and their orders satisfying this condition are said to be hyperbolic. Geometrically, the type gives us considerable amounts of geometric information about the surface: the Riemann-Hurwitz formula

$$g(\mathcal{C}_i) = 1 + \frac{|G|}{2}\left(1 - \frac{1}{o(x_i)} - \frac{1}{o(y_i)} - \frac{1}{o(x_iy_i)}\right)$$

tells us the genus of each of the curves used to define the surface $\mathcal{S}$ and by a theorem of Zeuthen-Segre this in turn gives us the Euler number of the surface $\mathcal{S}$ since

$$e(\mathcal{S}) = 4\frac{(g(\mathcal{C}_1) - 1)(g(\mathcal{C}_2) - 1)}{|G|}$$

which in turn gives us the holomorphic Euler-Poincaré characteristic of $\mathcal{S}$, namely $4\chi(\mathcal{S}) = e(\mathcal{S})$ (see [20, Theorem 3.4]).

Furthermore, if a group can be generated by a pair of elements of orders a and b whose product has order c then G is a homomorphic image of the triangle group

$$T_{a,b,c} = \langle x, y, z | x^a = y^b = z^c = xyz = 1 \rangle.$$

Homomorphic images of the triangle group $T_{2,3,7}$ are known as Hurwitz groups. In several places in the literature, knowing that a particular group is a Hurwitz group has proved useful for deciding its status as a Beauville group. For a discussion of known results on Hurwitz groups see the excellent surveys of Conder [22, 23].

3 Finite Simple Groups

A necessary condition for a group to be a Beauville group is that it is 2-generated. In [1, 59] it is proved that all non-abelian finite simple groups are 2-generated. For a long time it was conjectured that every non-abelian finite simple group, aside from the alternating group A_5, is a Beauville group [10, Conjecture 7.17], providing a rich source of examples. Various authors proved special cases of this [10, 31, 33]. The full result comes from the following Theorem which is proved by the author, Magaard and Parker in [27, 28].

Theorem 3.1 *With the exceptions of $SL_2(5)$ and $PSL_2(5)(\cong A_5 \cong SL_2(4))$, every finite quasisimple group is a Beauville group.*

Similar results were proved at around the same time by Garion, Larsen and Lubotzky in [36] (using probabilistic results concerning triangle groups from the PhD thesis of Marion [55]) and by Guralnick and Malle in [46] using the theory of linear algebraic groups. Since the overriding ideas behind the proofs given in [27, 36, 46] are in many ways quite general we sketch these ideas in the hope that they may be useful in proving other conjectures that appear later in this survey.

First note that the alternating groups can be dealt with using classical permutation group theory. Furthermore, the low rank groups of Lie type may be dealt with using explicit matrix calculations (see for instance the work of Fuertes and Jones in [33] concerning the groups $PSL_2(q)$, ${}^2B_2(2^{2n+1})$ and ${}^2G_2(3^{2n+1})$.) The sporadic simple groups are easily dealt with on a case by case basis with structure constant calculations being useful for the larger groups. The real difficulty lies with the groups of Lie type of unbounded rank.

Let G be a finite simple group of Lie type of characteristic p. To ensure that we can choose elements of the group G whose product behaves as we require we use a theorem of Gow [44] (a slight generalisation of this result to quasisimple groups is given in [27, Theorem 2.6]). An element of G is said to be 'semisimple' if its order is coprime to p and is said to be 'regular semisimple' if its centralizer in G has order coprime to p.

Theorem 3.2 *Let G be a finite simple group of Lie type of characteristic p and let $s \in G$ be a semisimple element. Let $R_1, R_2 \subset G$ be conjugacy classes of regular semisimple elements of G. Then there exist elements $x \in R_1$ and $y \in R_2$ such that $s = xy$.*

To ensure that the conjugacy part of the definition of a Beauville structure is satisfied we aim to choose $x_1, x_2, y_1, y_2 \in G$ such that $o(x_1)o(y_1)o(x_1y_1)$ is coprime to $o(x_2)o(y_2)o(x_2y_2)$. This is made possible by a classical theorem of Zsigmondy [63] (or rather Bang [2] in the case $p = 2$.) Whilst [2] and [63] are over a century old and therefore difficult to read and get hold of, a more recent account of a proof is given by Lünburg in [54].

Theorem 3.3 *For any positive integers a and n there exists a prime that divides $a^n - 1$ but not $a^k - 1$ for any $k < n$ with the following exceptions:*

- *$a = 2$ and $n = 6$; and*

- $a+1$ *is a power of* 2 *and* $n=2$.

The real significance of the above results stems from the fact that most groups of Lie type have an order that is a product of numbers of the form p^k-1 and so the above result guarantees the existence of a rich supply of distinct primes that can be taken as being the orders of the elements of our Beauville structure.

It remains to decide if a given triple will generate the group. Since our elements have orders given by Theorem 3.3 we can use a theorem of Guralnick, Pentilla, Praeger and Saxl [47] concerning subgroups of the general linear group $GL_n(p^a)$ containing elements of these orders and closely related results of Niemeyer and Praeger [56] for the other classical groups to show that no proper subgroups contain our elements. It follows that our chosen elements will generate the group.

4 Abelian and Nilpotent Groups

The abelian Beauville groups were essentially classified by Catanese in [20, page 24] and the full argument is given explicitly in [9, Theorem 3.4] where the following is proved.

Theorem 4.1 *Let G be an abelian group. Then G is a Beauville group if, and only if, $G=(\mathbb{Z}/n\mathbb{Z})\times(\mathbb{Z}/n\mathbb{Z})$ where $n>1$ is coprime to 6.*

This explains why Beauville's original example used the group $(\mathbb{Z}/5\mathbb{Z})\times(\mathbb{Z}/5\mathbb{Z})$: it is the smallest abelian Beauville group.

Theorem 4.1 has been put to great use by González-Diez, Jones and Torres-Teigell in [42] where several structural results concerning the surfaces defined by abelian Beauville groups are proved. For each abelian Beauville group they describe all the surfaces arising from that group, enumerate them up to isomorphism and impose constraints on their automorphism groups. As a consequence they show that all such surfaces are defined over $\mathbb{Q}$.

After the abelian groups, the next most natural class of finite groups to consider are the nilpotent groups. In [3, Lemma 1.3] Barker, Boston and the author note the following easy Lemma.

Lemma 4.2 *Let G and G' be Beauville groups and let $\{\{x_1,y_1,\},\{x_2,y_2\}\}$ and $\{\{x_1',y_1',\},\{x_2',y_2'\}\}$ be their respective Beauville structures. Suppose that*

$$gcd(o(x_i),o(x_i'))=gcd(o(y_i),o(y_i'))=1$$

for $i=1,2$. Then $\{\{(x_1,x_1',),(y_1,y_1')\},\{(x_2,x_2'),(y_2,y_2')\}\}$ is a Beauville structure for the group $G\times G'$.

Recall that a finite group is nilpotent if, and only if, it isomorphic to the direct product of its Sylow subgroups. It thus follows that this lemma, and its easy to prove converse, reduces the study of nilpotent Beauville groups to that of Beauville p-groups. Note that Theorem 4.1 gives us infinitely many examples of Beauville p-groups for every prime $p>3$ — simply let n be any power of p. Early examples of Beauville 2-groups and 3-groups were constructed by Fuertes, González-Diez and Jaikin-Zapirain in [32] where a Beauville group of order 2^{12} and another of order 3^{12}

were constructed. Even earlier than this, two Beauville 2-groups of order 2^8 arose as part of a classification due to Bauer, Catanese and Grunewald in [11] of certain classes of surfaces of general type.

More recently, in [3], Barker, Boston and the author classified the Beauville p-groups of order at most p^4 and made substantial progress on the cases of groups of order p^5 and p^6. In particular, the number of Beauville p-groups of order p^4 is two for every $p > 3$ and zero otherwise, but for p^5 we have the following.

Conjecture 4.3 *For all $p \geq 5$, the number of Beauville p-groups of order p^5 is given by $p + 10$.*

In [3, Theorem 1.4] we prove that there are at least $p + 8$ Beauville groups of order p^5. Furthermore, the above conjecture has been verified computationally for all primes p such that $5 \leq p \leq 19$. Perhaps more interestingly, other results proved in [3] verify that the proportion of 2-generated p-groups of order p^5 that are Beauville tends to 1 as p tends to infinity, however this fails to to be true for p-groups of order p^6.

Question 4.4 If $n > 6$ what is the behaviour, as p tends to infinity, of the proportion of 2-generated p-groups that are Beauville?

Another consequence of this work was determining the smallest Beauville p-group for all primes. In the below presentations, if no relationship between two generators is specified by a relation or relator then it should be assumed that the two generators commute.

Theorem 4.5 *The smallest Beauville p-groups are as follows.*

- *For $p = 2$ the group*

$$\langle x, y \,|\, x^4, y^4, [x, y^2]^2 2, [x, y^3]^2, [x^2, y^3] \rangle$$

  of order 2^7.

- *For $p = 3$ the group*

$$\langle x, y, z, w, t \,|\, x^3, y^3, z^3, w^3, t^3, y^x = yz, z^x = zw, z^y = zt \rangle$$

  of order 3^5.

- *For $p \geq 5$ the group*

$$\langle x, y, z \,|\, x^5, y^5, z^5, [x, y] = z \rangle$$

  of order p^3.

Further examples are given by the following unpublished constructions due to Jones and Wolfart.

Theorem 4.6 *Let G be a finite group of exponent $n = p^e > 1$ for some prime $p \geq 5$, such that the abelianisation G/G' of G is isomorphic to $\mathbb{Z}_n \times \mathbb{Z}_n$. Then G has a Beauville structure.*

Corollary 4.7 *Let G be a 2-generated finite group of exponent p for some prime $p \geq 5$. Then G has a Beauville structure.*

As noted earlier Beauville p-groups for $p > 3$ are in bountiful supply. Several examples of Beauville 2-groups and 3-groups are constructed by Barker, Boston, Peyerimhoff and Vdovina in [5, 6] by considering sections of groups defined using projective planes. More recently, in [4], Barker, Boston, Peyerimhoff and Vdovina using similar ideas constructed the first infinite family of Beauville 2-groups. At the time of writing, as far as the author is aware, only finitely many Beauville 3-groups are known leading to the following natural problem.

Problem 4.8 Construct infinitely many Beauville 3-groups.

We conclude this section with the following remarks. Nigel Boston has recently undertaken some substantial and as yet unpublished computations regarding the relationship between p-groups' status as Beauville groups and their position on the so-called 'O'Brien Trees' [57]. Whilst little global pattern appears to exist in general, there does appear to be some mysterious relationship with an invariant known as the 'nuclear rank' of the group — see [17]. Since defining this concept is somewhat involved we shall say no more about this here.

5 Strongly Real Beauville Groups

Given any complex surface $\mathcal{S}$ it is natural to consider the complex conjugate surface $\overline{\mathcal{S}}$. In particular it is natural to ask if the surfaces are biholomorphic.

Definition 5.1 Let $\mathcal{S}$ be a complex surface. We say that $\mathcal{S}$ is *real* if there exists a biholomorphism $\sigma : \mathcal{S} \to \overline{\mathcal{S}}$ such that σ^2 is the identity map.

As noted earlier this geometric condition can be translated into algebraic terms.

Definition 5.2 Let G be a Beauville group and let $X = \{\{x_1, y_1\}, \{x_2, y_2\}\}$ be a Beauville structure for G. We say that G and X are *strongly real* if there exists an automorphism $\phi \in \mathrm{Aut}(G)$ and elements $g_i \in G$ for $i = 1, 2$ such that

$$g_1\phi(x_i)g_1^{-1} = x_i^{-1} \text{ and } g_2\phi(y_i)g_2^{-1} = y_i^{-1}$$

for $i = 1, 2$.

It is often, but not always, convenient to take $g_1 = g_2$.

Our first examples come immediately from Theorem 4.1 since for any abelian group the function $x \mapsto -x$ is an automorphism.

Corollary 5.3 *Every Beauville structure of an abelian Beauville group is strongly real.*

A little more generally, when it comes to strongly real Beauville p-groups the examples given by Theorem 4.1 are, as far as the author is aware, the only known examples. Furthermore, the Beauville 2-groups constructed by Barker, Boston, Peyerimhoff and Vdovina in [4] are explicitly shown to not be strongly real. However, a combination of Corollary 5.3 and the fact that p-groups in general tend to have large automorphism groups [18, 19] it seems likely that most Beauville p-groups are in fact strongly real Beauville groups. This makes the following problem particularly pressing.

Problem 5.4 Construct examples of strongly real Beauville p-groups.

In [25] the following conjecture, a refinement of an earlier conjecture of Bauer, Catanese and Grunewald [9, Section 5.4], is made.

Conjecture 5.5 *All non-abelian finite simple groups apart from A_5, M_{11} and M_{23} are strongly real Beauville groups.*

Only a few cases of this conjecture are known.

- In [31] Fuertes and González-Diez showed that the alternating groups A_n ($n \geq 7$) and the symmetric groups S_n ($n \geq 5$) are strongly real Beauville groups by explicitly writing down permutations for their generators and the automorphisms used and applying some of the classical theory of permutation groups to show that their elements had the properties they claimed. It was subsequently found that the group A_6 is also strongly real.
- In [33] Fuertes and Jones proved that the simple groups $PSL_2(q)$ for prime powers $q > 5$ and the quasisimple groups $SL_2(q)$ for prime powers $q > 5$ are strongly real Beauville groups. As with the alternating and symmetric groups, these results are proved by writing down explicit generators, this time combined with a celebrated theorem usually (but historically inaccurately) attributed to Dickson for the maximal subgroups of $PSL_2(q)$. (For a full statement of this result and related theorems as well a detailed historical account of the maximal subgroups of low dimensional classical groups see the excellent survey of King in [53].) General lemmas for lifting structures from a group to its covering groups are also used.
- In [26] the author determined which of the sporadic simple groups are strongly real Beauville groups, including the '27^{th} sporadic simple group', the Tits group ${}^2F_4(2)'$. Only the Mathieu groups M_{11} and M_{23} are not strongly real For all of the other sporadic groups smaller than the Baby Monster group $\mathbb{B}$ explicit words in the 'standard generators' [61] for a strongly real Beauville structure were given. For the Baby Monster group $\mathbb{B}$ and Monster group $\mathbb{M}$ character theoretic methods were used.
- In [25] the author also verified this conjecture for the Suzuki groups ${}^2B_2(2^{2n+1})$. Again, this was achieved by writing down explicit elements of the group which using the list of maximal subgroups of the Suzuki group are shown to generate.
- In [25] the author extended earlier computations of Bauer, Catanese and Grunewald, verifying this conjecture for all non-abelian finite simple groups of order at most 100 000 000.

We remark that several of the groups mentioned in the above bullet points are not simple. More generally we ask the following.

Question 5.6 Which groups are strongly real Beauville groups?

Finally, we remark that in [25] the author constructs many further examples of strongly real Beauville groups. This includes the characteristically simple groups A_n^k for moderate values of k and sufficiently large values of n, the groups $S_n \times S_n$ for

$n \geq 5$ and the almost simple sporadic groups. This last calculation combined with the earlier remarks on the symmetric group lead to the following conjecture.

Corollary 5.7 *A split extension of a simple group is a Beauville group if, and only if, it is a strongly real Beauville group.*

6 The Mixed Case

When we defined Beauville surfaces and groups we considered the action of a group G on the product of two curves $\mathcal{C}_1 \times \mathcal{C}_2$. In an unmixed structure this action comes solely from the action of G on each curve individually, however there is nothing to stop us considering an action on the product that interchanges the two curves and it is precisely this situation that we discuss in this section. Recall from Definition 2.2 that given $x, y \in G$ we write

$$\Sigma(x,y) := \bigcup_{i=1}^{|G|} \bigcup_{g \in G} \{(x^i)^g, (y^i)^g, ((xy)^i)^g\}.$$

Definition 6.1 Let G be a finite group. A *mixed Beauville structure* for G is a quadruple (G^0, g, h, k) where G^0 is an index 2 subgroup and $g, h, k \in G$ are such that

- $\langle g, h \rangle = G^0$;
- $k \notin G^0$;
- for every $\gamma \in G^0$ we have that $(k\gamma)^2 \notin \Sigma(g,h)$ and
- $\Sigma(g,h) \cap \Sigma(g^k, h^k) = \{e\}$

A Beauville surface defined by a mixed Beauville structure is called a *mixed Beauville surface* and group possessing a mixed Beauville structure is called a *mixed Beauville group*.

In terms of the curves defining the surface, the group G^0 is the stabiliser of the curves with the elements of $G \setminus G^0$ interchanging the two terms of $\mathcal{C}_1 \times \mathcal{C}_2$. Moreover it is only possible for a Beauville surface $(\mathcal{C}_1 \times \mathcal{C}_2)/G$ to come from a mixed Beauville structure if $\mathcal{C}_1 \cong \mathcal{C}_2$. The above conditions also ensure that $\{\{g,h\},\{g^k,h^k\}\} \subset G^0 \times G^0$ is a Beauville structure for G^0.

In general, mixed Beauville structures are much harder to construct than their unmixed counterparts. The following lemma of Fuertes and González-Diez imposes a strong condition on a group with a mixed Beauville structure [31, Lemma 5].

Lemma 6.2 *Let $(\mathcal{C}_1 \times \mathcal{C}_2)/G$ be a mixed Beauville surface and let G^0 be the subgroup of G consisting of the elements which do not interchange the two curves. Then the order of any element in $G \setminus G^0$ is divisible by 4.*

Clearly no simple group can have a mixed Beauville structure since it is necessary to have a subgroup of index 2 and the cyclic group of order 2 is not a Beauville group, however that does not preclude the possibility of almost simple groups having mixed Beauville structures. The above lemma was originally used to show that no

symmetric group has a mixed Beauville structure. In [26] the author used the above to show that no almost simple sporadic group has a mixed Beauville structure (though the almost simple Tits group ${}^2F_4(2)$ is not excluded by the above lemma) and in general most almost simple groups are ruled out by it (though as the groups $P\Sigma L_2(p^2)$ show there are infinitely many exceptions to this). A further restriction comes from [9, Theorem 4.3] where Bauer, Catanese and Grunewald prove that G^0 must be non-abelian. Various geometric constraints are proved by Torres-Teigell in his PhD thesis [60]. Most notably the genus of a mixed Beauville surface is odd and at least 17. Furthermore, this bound is sharp. This naturally leads to the following problem.

Problem 6.3 Find mixed Beauville structures.

The earliest examples of groups that do possess mixed Beauville structures were given by Bauer, Catanese and Grunewald in [9]. Their general construction is of the form $(H \times H) : (\mathbb{Z}/4\mathbb{Z})$, the generator of the group $\mathbb{Z}/4\mathbb{Z}$ acting on the direct product by interchanging its two factors and $G^0 = H \times H \times \mathbb{Z}/2\mathbb{Z}$.

Lemma 6.4 *Let H be a finite group and let $x_1, y_1, x_2, y_2 \in H$. Suppose that*

(1) $o(x_1)$ and $o(y_1)$ are even;

(2) $\langle x_1^2, y_1^2, x_1y_1\rangle = H$;

(3) $\langle x_2, y_2\rangle = H$ and

(4) $o(x_1)o(y_1)o(x_1y_1)$ is coprime to $o(x_2)o(y_2)o(x_2y_2)$.

If the above conditions are satisfied then (G^0, x, y, g) is a mixed Beauville structure for some $g \in (H \times H) : (\mathbb{Z}/4\mathbb{Z})$ where $x = (x_1, x_2, 2), y = (y_1, y_2, 2) \in H \times H \times \mathbb{Z}/2\mathbb{Z}$ (note that $2 \in \mathbb{Z}/4\mathbb{Z}$ generates the subgroup isomorphic to $\mathbb{Z}/2\mathbb{Z}$). Furthermore, if H is a perfect group then we cane replace condition (2) with the condition

(2') $\langle x_1, y_1\rangle = H$.

Note that in [9] this last hypothesis was incorrectly stated in terms of the perfectness of G rather than H. Bauer, Catanese and Grunewald go on to use the above lemma to construct examples in the cases with the property that if H is taken to be a sufficiently large alternating group or a special linear groups $SL_2(p)$ with $p \neq 2, 3, 5, 17$ (though their argument also does not apply in the case $p = 7$), then $(H \times H) : (\mathbb{Z}/4\mathbb{Z})$ has a mixed Beauville structure. Given the extent to which mixed Beauville groups are in short supply it would be interesting to see if the above construction can be used in other cases.

Problem 6.5 Find other groups H that the above lemma can be applied to.

In [29] the author and Pierro prove a slight generalisation of Lemma 6.4 that replaces the cyclic group of order 4 with the dicyclic group of order $4k$ defined by the presentation

$$\langle x, y \,|\, x^{2k} = y^4 = 1, x^y = x^{-1}, x^k = y^2\rangle$$

for some positive integer k. In particular, when finding examples of groups that satisfy the hypotheses of this generalisation (which is sufficient to show that such groups satisfy the hypotheses of Lemma 6.4) we obtain new examples of mixed Beuville groups from the groups H and $H \times H$ where H is any of the alternating groups A_n

($n \geq 6$), the linear groups $PSL_2(q)$ ($q \geq 7$ odd), the unitary groups $PSU_3(q)$ ($q \geq 3$), the Suzuki groups ${}^2B_2(2^{2n+1})$ ($n \geq 1$), the small Ree groups ${}^2G_2(3^{2n+1})$ ($n \geq 1$), the large Ree groups ${}^2F_4(q)$ ($q \geq 8$), the Steinberg triality groups ${}^3D_4(q)$ ($q \geq 2$) and the sporadic simple groups (including the Tits group ${}^2\mathrm{F}_4(2)'$) as well as the groups $PSL_2(2^n) \times PSL_2(2^n)$ ($n \geq 3$).

What about p-groups? If p is odd then again, the absence of index 2 subgroups ensures that there exist no mixed Beauville p-groups. In the construction described above the technical constraints on H ensure that it cannot be a 2-group, stopping this providing a source of examples. Early examples of mixed Beauville 2-groups were given by Bauer, Catanese and Grunewald constructed in [11] where they constructed two mixed Beauville groups of order 2^8. Even so, the lack of known Beauville 2-groups makes the following a natural problem.

Problem 6.6 Construct infinitely many mixed Beauville 2-groups.

7 Miscellanea

7.1 $PSL_2(q)$ and $PGL_2(q)$

In [10, Question 7.7] Bauer, Catanese and Grunewald asked the following question,

> Existence and classification of Beauville surfaces, i.e.,
>
> a) which finite groups G can occur?
>
> b) classify all possible Beauville surfaces for a given finite group G.

In [35] Garion answered the above in the case of the groups $PSL_2(q)$ and $PGL_2(q)$. For $PSL_2(q)$ we have the following.

Theorem 7.1 *Let $G = PSL_2(q)$ where $5 < q = p^e$ for some prime number p and some positive integer e. Let $\tau_1 = (r_1, s_1, t_1)$, $\tau_2 = (r_2, s_2, t_2)$ be two hyperbolic triples of integers. Then G admits an unmixed Beauville structure of type (τ_1, τ_2) if, and only if, the following hold:*

(i) the group G is a quotient of the triangle groups T_{r_1,s_1,t_1} and T_{r_2,s_2,t_2} with torsion-free kernel;

(ii) if $p = 2$ or e is odd or $q = 9$, then $r_1s_1t_1$ is coprime to $r_2s_2t_2$. If p is odd, e is even and $q > 9$, then $g = gcd(r_1s_1t_1, r_2s_2t_2) \in \{1, p, p^2\}$. Moreover, if p divides g and τ_1 (respectively τ_2) is up to a permutation (p, p, n) then $n \neq p$ and n is a 'good G-order'.

Here by 'good G-order' we mean the following. Let q be an odd prime power and let $n > 1$ be an integer. Then n is a good G-order if either

- n divides $(q-1)/2$ and a primitive root of unity a of order $2n$ in $\mathbb{F}_q$ has the property that $-a = c^2$ for some $c \in \mathbb{F}_q$ or
- n divides $(q+1)/2$ and a primitive root of unity a of order $2n$ in $\mathbb{F}_q^2$ has the property that $-a = c^2$ for some $c \in \mathbb{F}_{q^2}$ such that $c^{q+1} = 1$.

A similar theorem is given for the groups $PGL_2(q)$.

Given that generic lists of maximal subgroups of other low rank groups of Lie type are well known in numerous other cases, it seems likely that analogous results for these groups can also be obtained. We thus reiterate Bauer, Catanese and Grunewald's earlier question in this case.

Problem 7.2 Obtain results analogous to the above for other classes of finite simple groups.

7.2 Fundamental Groups of Beauville Surfaces

We mentioned in the introduction that Beauville surfaces have fundamental groups that are easy to work with. To make this vague remark a little more specific we note the following. Suppose that if G is a Beauville group with a Beauville structure of type $((a_1, b_1, c_1), (a_2, b_2, c_2))$, then for $i = 1, 2$ there exist surjective homomorphisms $\rho_i : T_{a_i,b_i,c_i} \to G$. The direct product $\ker(\rho_1)\times\ker(\rho_2)$ is the fundamental group of the product $\mathcal{C}_1\times\mathcal{C}_2$. The fundamental group of the surface $(\mathcal{C}_1\times\mathcal{C}_2)/G$ is now an extension of a normal subgroup $\ker(\rho_1)\times\ker(\rho_2)$ by G, or more precisely the inverse image in $T_{a_1,b_1,c_1} \times T_{a_2,b_2,c_2}$ of the diagonal subgroup of $G \times G$ under the epimorphism $\rho_1 \times \rho_2$. It turns out that this simple description of the fundamental group is responsible for the rigidity of Beauville surfaces and this in turn ensures that the topological and geometric features of the surfaces are closely intertwined - see [51, Section 9] for details.

Unsurprisingly, since a Beauville group dictates so many features of its corresponding Beauville surface which in turn determines its fundamental group we also have the reverse relationship whereby the fundamental group determines the original Beauville group. The following is proved by González-Diez and Torres-Teigell [41, 60]. (It also worth noting related results given by Bauer, Catanese and Grunewald in [10] and by Catanese in [20]).

Theorem 7.3 *Two Beauville surfaces are isometric if and only if their fundamental groups are isomorphic.*

The fundamental group is one of the most basic tools in algebraic topology. It is, however, somewhat limited in its usefulness and topologists have found several important higher dimensional analogues of the fundamental group and so it is natural to pose the following question.

Question 7.4 Do the higher homotopy/homology/cohomology groups of a Beauville surface have similar descriptions in terms of triangle groups and the corresponding Beauville group and to what extent do they uniquely determine the surface?

By way of partial progress on this question in [8] Bauer, Catanese and Frapporti recently showed that for any Beauville surface $\mathcal{S}$ the homology group $H_1(\mathcal{S}, \mathbb{Z})$ is finite. They also give a much more detailed discussion of geometric aspects of the study of fundamental groups of Beauville surfaces and related objects as well as computer calculations of these objects in some cases.

7.3 Automorphism Groups of Beauville Surfaces

In [50] Jones investigated the automorphism groups of unmixed Beauville surfaces. Some of these results were obtained independently by Fuertes and González-Diez in [30] and were later extended to mixed Beauville surfaces by González-Diez and Torres-Teigell in [40, Section 5.3].

Theorem 7.5 *The automorphism group Aut($\mathcal{S}$) of a Beauville surface $\mathcal{S} = (\mathcal{C}_1 \times \mathcal{C}_2)/G$ has a normal subgroup Inn($\mathcal{S}$) $\triangleleft$ $Z(G)$ with Aut($\mathcal{S}$)/Inn($\mathcal{S}$) isomorphic to a subgroup of the wreath product $S_3 \wr S_2$. In particular Aut($\mathcal{S}$) is a finite soluble group of order dividing $72|Z(G)|$ and of derived length at most 4.*

Here the subgroup $\mathrm{Inn}(\mathcal{S})$ consists of automorphisms preserving the two curves (or more precisely, induced by automorphisms of $\mathcal{C} \times \mathcal{C}'$ preserving them) though it does not necessarily contain all of them: they form a subgroup of index at most 2 in $\mathrm{Aut}(\mathcal{S})$, whereas $\mathrm{Inn}(\mathcal{S})$ can have index up to 72. The results in the mixed case are similar.

7.4 Beauville Genus Spectra

In [10, Question 7.7(b)] Bauer, Catanese and Grunewald ask us to classify all possible Beauville surfaces for a given finite group G.

As a partial answer to this, in [32, Section 4] Fuertes, González-Diez and Jaikin-Zapirain introduce the concept of Beauville genus spectrum which we define as follows.

Definition 7.6 Let G be a finite group. The *Beauville genus spectrum* of G is the set $Spec(G)$ of pairs of integers (g_1, g_2) such that $g_1 \leq g_2$ and there are curves $\mathcal{C}_1$ and $\mathcal{C}_2$ of genera g_1 and g_2 with an action of G on $\mathcal{C}_1 \times \mathcal{C}_2$ such that $(\mathcal{C}_1 \times \mathcal{C}_2)/G$ is a Beauville surface.

By the Riemann-Hurwitz formula each g_i is bounded above by $1 + |G|/2$ and so this set is always finite. Fuertes, González-Diez and Jaikin-Zapirain determine the Beauville spectra of several small groups.

Proposition 7.7
1. $Spec(S_5) = \{(19, 21)\}$
2. $Spec(PSL_2(7)) = \{(8, 49), (15, 49), (17, 22), (22, 33), (22, 49)\}$
3. $Spec(S_6) = \{(49, 91), (91, 121), (91, 169), (121, 169), (151, 169)\}$
4. *If* $gcd(n, 6) = 1$ *and* $n > 1$ *then*

$$Spec((\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})) = \left\{ \left(\frac{(n-1)(n-2)}{2}, \frac{(n-1)(n-2)}{2} \right) \right\}.$$

Unpublished calculations of the author's PhD student, Emilio Pierro, has added a few more finite simple and almost simple groups to the above list, the largest being the Mathieu group M_{23}. Furthermore, the Beauville genus structures of $PSL_2(q)$ and $PGL_2(q)$ may be deduced from the results discussed in Subsection 7.1. This naturally leads us to ask the following.

Problem 7.8 Determine the Beauville genus spectrum of more groups.

7.5 Characteristically Simple Groups

Characteristically simple groups are usually defined in terms of characteristic subgroups, but for finite groups this turns out to be equivalent to the following.

Definition 7.9 A finite group G is said to be *characteristically simple* if G is isomorphic to the direct product H^k where H is a finite simple group for some positive integer k.

If we fix H then for large values of k the group H^k will not be 2-generated and therefore will not be Beauville. For more modest values of k there is, however, still hope. These groups have recently been investigated by Jones in [48, 49] where the following conjecture is investigated.

Conjecture 7.10 *Let G be a finite characteristically simple group. Then G is Beauville if and only if it is 2-generated and not isomorphic to the alternating group A_5.*

Theorem 3.1 shows that this conjecture is true for the characteristically simple group H^k in the case $k = 1$ for every non-abelian finite simple group H. If G is abelian then this conjecture holds by Theorem 4.1 following the convention that a cyclic group is not considered to be 2-generated. In [39] the above conjecture is verified for the alternating groups and in [49] it is verified for the linear groups $PSL_2(q)$ and $PSL_3(q)$, the unitary groups $PSU_3(q)$, the Suzuki groups ${}^2B_2(2^{2n+1})$, the small Ree groups ${}^2G_2(3^{2n+1})$ and the sporadic simple groups. In addition to the above the author has performed computations that verify the above conjecture for all characteristically simple groups of order at most 10^{30}. As an amusing aside we note that this shows that whilst A_5 is not a Beauville group, the direct product of nineteen copies of A_5 is!

In [25] the author considers which of the characteristically simple groups are strongly real Beauville groups. The main conjecture is the following.

Conjecture 7.11 *If G is a finite simple group of order greater than 3, then $G \times G$ is a strongly real Beauville group.*

It is likely that many larger direct products are also strongly real, however the precise statement of a conjecture along these lines is likely to be much more complicated. For example, a straightforward computation verifies that neither of the groups $\mathrm{M}_{11}\times\mathrm{M}_{11}\times\mathrm{M}_{11}$ and $\mathrm{M}_{23}\times\mathrm{M}_{23}\times\mathrm{M}_{23}$ are strongly real despite the fact that both of the groups $\mathrm{M}_{11}\times\mathrm{M}_{11}\times\mathrm{M}_{11}\times\mathrm{M}_{11}$ and $\mathrm{M}_{23}\times\mathrm{M}_{23}\times\mathrm{M}_{23}\times\mathrm{M}_{23}$ are.

The above conjecture has been verified for the alternating groups (though slightly stronger results are true in this case), the sporadic simple groups, the linear groups $PSL_2(q)$ $(q > 5)$, the Suzuki groups ${}^2B_2(2^{2n+1})$, the sporadic simple groups (including the Mathieu groups M_{11} and M_{23}, despite the statement of Conjecture 5.5) and all of the finite simple groups of order at most 100 000 000.

7.6 Orbits of the Absolute Galois Group

The task of understanding the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is of central importance in algebraic number theory and is related to the Inverse Galois Problem (it is

equivalent to asking if every finite group is a quotient of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ under a topologically closed normal subgroup) and this is arguably the hardest open problem in algebra today. As things stand $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ remains very poorly understood. A natural approach to understanding any group is to study some action(s) of the group. An immeidate consequence of Belyĭ's Theorem is that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on the set of all Beauville surfaces. Recently there has been much interest in constructing orbits consisting of mutually non-homeomorphic pairs of Beauville surfaces. In [41, 43] González-Diez, Jones and Torres-Teigell have constructed arbitrarily large orbits of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ consisting of mutually non-homeomorphic pairs of Beauville surfaces defined by the Beauville groups $PSL_2(q)$ and $PGL_2(q)$.

Problem 7.12 Construct arbitrarily large orbits of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ consisting of mutually non-homeomorphic pairs of Beauville surfaces using other groups.

A slightly different motivation for addressing the above problem comes from the following. Knowing whether or not $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts faithfully on the set of Beauville surfaces is equivalent to the longstanding question of whether or not $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts faithfully on the set of regular dessins. This was recently resolved by González-Diez and Jaikin-Zapirain in [39] by showing that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts faithfully on the set of Beauville surfaces.

References

[1] M. Aschbacher and R. Guralnick, Some applications of the first cohomology group, *J. Algebra* **90** (1984), 446–460.

[2] A. S. Bang, Talteoretiske undersølgelser, Tidskrifft, Math. (5)4 (1886), 130–137.

[3] N. W. Barker, N. Boston and B. T. Fairbairn, A note on Beauville p-groups, *Exp. Math.* **21** (2012), 298–306; arXiv:0907.2028v3.

[4] N. W. Barker, N. Boston, N. Peyerimhoff and A. Vdovina, An infinite family of 2-groups with mixed Beauville structures, preprint 2013, to appear *Int. Math. Res. Not.*; arXiv:1304.4480v1.

[5] N. W. Barker, N. Boston, N. Peyerimhoff and A. Vdovina, Regular algebraic surfaces isogenous to a higher product constructed from group representations using projective planes, preprint 2011, arXiv:1109.6053.

[6] N. W. Barker, N. Boston, N. Peyerimhoff and A. Vdovina, New examples of Beauville surfaces, *Monatsh. Math.* **166** (2012), 319–327.

[7] I. Bauer, Product-Quotient Surfaces: Result and Problems, preprint 2012, arXiv: 1204.3409.

[8] I. Bauer, F. Catanese and D. Frapporti, The fundamental group and torsion group of Beauville surfaces, preprint 2014, arXiv:1402.2109

[9] I. Bauer, F. Catanese and F. Grunewald, Beauville surfaces without real structures, *Geometric Methods in Algebra and Number Theory*, 1–42, Progr. Math. 235, Birkhäuser, 2005.

[10] I. Bauer, F. Catanese and F. Grunewald, Chebycheff and Belyi polynomials, dessins d'enfants, Beauville surfaces and group theory, *Mediterr. J. Math.* **3** (2006), 121–146.

[11] I. Bauer, F. Catanese and F. Grunewald, The classification of surfaces with $p_g = q = 0$ isogenous to a product of curves, *Pure Appl. Math. Q.* **4** (2008), 547–586.

[12] I. Bauer, F. Catanese and R. Pignatelli, Surfaces of general type with geometric genus zero: a survey, *Complex and Differential Geometry*, 1–48, Springer Proc. Math. 8, Springer, 2011.

[13] I. C. Bauer, F. Catanese and R. Pignatelli, Complex surfaces of general type: some recent progress, *Global Aspects of Complex Geometry*, 1–58, Springer, 2006.

[14] A. Beauville, Surfaces algébriques complexes, *Astérisque* **54**, 1978.
[15] A. Beauville, *Complex Algebraic Surfaces*, London Math. Soc. Student Texts 34, Cambridge University Press, 1996.
[16] G. V. Belyĭ, On Galois extensions of a maximal cyclotomic field, *Math. USSR Izv.* **14** (1980), 247–256.
[17] N. Boston, A Survey of Beauville p-groups, *Proceedings of Conference on Beauville Surfaces and Group, Newcastle 2012* (eds I. Bauer, S. Garion and A. Vdovina), to appear.
[18] N. Boston, Embedding 2-groups in groups generated by involutions, *J. Algebra* **300** (2006), 73-76.
[19] N. Boston, M. R. Bush and F. Hajir, Heuristics for p-class towers of imaginary quadratic fields, preprint 2011, arXiv:1111.4679v1.
[20] F. Catanese, Fibered surfaces, varieties isogenous to a product and related moduli spaces, *Amer. J. Math.* **122** (2000), 1–44.
[21] F. Catanese, Moduli spaces of surfaces and real structures, *Ann. of Math. (2)* **158** (2003), 577–592.
[22] M. D. E. Conder, Hurwitz groups: a brief survey, *Bull. Amer. Math. Soc.* **23** (1990), 359–370.
[23] M. D. E. Conder, An update on Hurwitz groups, *Groups Complex. Cryptol.* **2** (2010), 35–49.
[24] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, *Atlas of Finite Groups*, Clarendon Press, 1985.
[25] B. T. Fairbairn, Strongly Real Beauville Groups, *Proceedings of Conference on Beauville Surfaces and Groups, Newcastle 2012* (eds I. Bauer, S. Garion and A. Vdovina), Springer, 2014, to appear.
[26] B. T. Fairbairn, Some Exceptional Beauville Structures, *J. Group Theory* **15** (2012), 631–639; arXiv:1007.5050.
[27] B. T. Fairbairn, K. Magaard and C. W. Parker, Generation of finite simple groups with an application to groups acting on Beauville surfaces, *Proc. London Math. Soc. (3)* **107** (2013), 744–798.
[28] B. T. Fairbairn, K. Magaard and C. W. Parker, Corrigendum: Generation of finite simple groups with an application to groups acting on Beauville surfaces, to appear.
[29] B. T. Fairbairn and E. Pierro, New examples of mixed Beauville groups, preprint 2014.
[30] Y. Fuertes and G. González-Diez, On the number of automorphisms of unmixed Beauville surfaces, preprint.
[31] Y. Fuertes and G. González-Diez, On Beauville structures on the groups S_n and A_n, *Math. Z.* **264** (2010), 959–968.
[32] Y. Fuertes, G. González-Diez and A. Jaikin-Zapirain, On Beauville surfaces, *Groups Geom. Dyn.* **5** (2011), 107–119.
[33] Y. Fuertes and G. Jones, Beauville surfaces and finite groups, *J. Algebra* **340** (2011), 13–27.
[34] S. Galkin and E. Shnider, Exceptional collections of line bundles on the Beuville surface, preprint 2012, arXiv:1210.3339.
[35] S. Garion, On Beauville Structures for $PSL(2, q)$, preprint 2010, arXiv:1003.2792v2.
[36] S. Garion, M. Larsen and A. Lubotzky, Beauville surfaces and finite simple groups, *J. Reine Angew. Math.* **666** (2012), 225–243.
[37] S. Garion and M. Penegini, New Beauville surfaces, moduli spaces and finite groups, *Comm. Algebra* **42** (2014), 2126–2155; arXiv:0910.5402.
[38] E. Girondo and G. González-Diez, *Introduction to Compact Riemann Surfaces and Dessins d'Enfants*, London Math. Soc. Student Texts 79, Cambridge University Press, 2011.
[39] G. González-Diez and A. Jaikin-Zapirain, The absolute Galois group acts faithfully on regular dessins and on Beauville surfaces, preprint 2012, http://www.uam.es/personal_

pdi/ciencias/gabino/Jule03.pdf.
[40] G. González-Diez and D. Torres-Teigell, An introduction to Beauville surfaces via uniformization, *Quasiconformal Mappings, Riemann Surfaces, and Teichmüller Spaces*, 123–151, Contemp. Math. 575, Amer. Math. Soc., 2012.
[41] G. González-Diez and D. Torres-Teigell, Non-homeomorphic Galois conjugate Beauville structures on $PSL(2,p)$, Adv. Math. 229 (2012), 3096–3122.
[42] G. González-Diez, G. A. Jones and D. Torres-Teigell, Beauville surfaces with abelian Beauville group, arXiv:1102.4552.
[43] G. González-Diez, G. A. Jones and D. Torres-Teigell, Arbitrarily large Galois orbits of non-homeomorphic surfaces, arXiv:1110.4930.
[44] R. Gow, Commutators in finite simple groups of Lie type, *Bull. London Math. Soc.* **32** (2000), 311–315.
[45] A. Grothendieck, Esquisse d'un Programme, *Geometric Galois Actions 1. Around Grothendieck's Esquisse d'un Programme*, 5–84, London Math. Soc. Lecture Note Ser. 242, Cambridge University Press, 1997.
[46] R. Guralnick and G. Malle, Simple groups admit Beauville structures, *J. London Math. Soc. (2)* **85** (2012), 694–721.
[47] R. Guralnick, T. Pentilla, C. Praeger and J. Saxl, Linear groups with orders having certain large prime divisors, *Proc. London Math. Soc. (3)* **78** (1999), 167–214.
[48] G. A. Jones, Characteristically simple Beauville groups, I: Cartesian powers of alternating groups, preprint 2013, arXiv:1304.5444v1.
[49] G. A. Jones, Characteristically simple Beauville groups, II: Low rank and sporadic groups, preprint 2013, arXiv:1304.5450v1.
[50] G. A. Jones, Automorphism groups of Beauville surfaces, *J. Group Theory* **16** (2013) 353–381.
[51] G. A. Jones, Beauville surfaces and groups: a survey, *Fields Inst. Comm.*, to appear.
[52] G. A. Jones and D. Singerman, Belyi functions, hypermaps and Galois groups, *Bull. London Math. Soc.* **28** (1996) 561–590.
[53] O. King, The subgroup stucture of finite classical groups in terms of geometric configurations, *Surveys in Combinatorics 2005*, 29–56, London Math. Soc. Lecture Note Ser. 327. Cambridge University Press, 2005.
[54] H. Lünburg, Ein einfacher Beweis für den Satz von Zsigmondy über primitive Primteiler von $a^N - 1$, *Geometries and Groups*, 219–222, Lecture Notes Math. 893, Springer, 1981.
[55] C. Marion, Triangle groups and finite simple groups, PhD thesis, Imperial College London, 2009
[56] A. C. Niemeyer and C. E. Praeger, A recognition algorithm for classical groups over finite fields, *Proc. London Math. Soc. (3)* **77** (1998), 117–169.
[57] E. A. O'Brien, The p-group generation algorithm, *J. Symbolic Comput.* **9** (1990), 677–698.
[58] J. Širáň, How symmetric can maps on surfaces be?, *Surveys in Combinatorics 2013*, 161–238, London Math. Soc. Lecture Note Ser. 409, Cambridge University Press, 2013.
[59] R. Steinberg, Generators for simple groups, *Canad. J. Math.* **14** (1962), 277–283.
[60] D. Torres-Teigell, Triangle groups, dessins d'enfants and Beauville surfaces, PhD thesis, Universidad Autonoma de Madrid, 2012.
[61] R. A. Wilson, Standard generators for sporadic simple groups, *J. Algebra* **184** (1996), 505–515.
[62] J. Wolfart, ABC for polynomials, dessins d'enfants and uniformization—a survey, Elementare und analytische Zahlentheorie, 313–345, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main, 20, Franz Steiner Verlag Stuttgart, Stuttgart, 2006, http://www.math.uni-frankfurt.de/~wolfart/.
[63] K. Zsigmondy, Zur Theorie der Potenzreste, *Monatsh. Math. Phys.* **3** (1892), 265–284.

SOMETHING FOR NOTHING: SOME CONSEQUENCES OF THE SOLUTION OF THE TARSKI PROBLEMS

BENJAMIN FINE*, ANTHONY GAGLIONE†, GERHARD ROSENBERGER§ and DENNIS SPELLMAN‡

*Department of Mathematics, Fairfield University, Fairfield, Connecticut 06430, United States

†Department of Mathematics, United States Naval Academy, Annapolis, Maryland 21402, United States

§Fachbereich Mathematik, University of Hamburg, Bundestrasse 55, 20146 Hamburg, Germany

‡Department of Statistics, Temple University, Philadelphia, Pennsylvania 19122, United States

1 Introduction

Alfred Tarski in 1940 made three well-known conjectures concerning nonabelian free groups (see Section 2). There had been various partial solutions until complete positive solutions were presented during the past 15 years by Kharlampovich and Myasnikov (see [51]–[59]) and independently by Z. Sela (see [78]–[83]). In the Kharlampovich-Myasnikov approach the proof arose from a detailed study of fully residually free groups (called limit groups in Sela's approach), the development of algebraic geometry over free groups, and an elimination process involving solutions of equations over free groups based on work of Makhanin and Razborov (see [51]–[59]). These steps were mirrored, with somewhat different terminology, by Sela, who called his approach diophantine geometry over free groups.

The positive solution of the Tarski conjectures provides a straightforward proof of Magnus's theorem in surface groups which we present. This result was proved directly by J. Howie [46] and independently by O. Bogopolski [10]. We will present this proof in Section 4. This type of proof leads to several different types of questions.

- Which additional nontrivial free group results are true in surface groups but difficult to obtain directly?
- What first-order properties of nonabelian free groups are true beyond the class of elementary free groups?

After showing a proof of Magnus's Theorem based on the solution of the Tarski problems we give several examples of other free group results holding in surface groups. Using this technique we give a proof of a theorem of D. Lee on C-test words. We then consider and prove certain other results that hold in elementary free groups, in particular surface groups, including the retract theorem of Turner [86] and the property of conjugacy separability.

After this we turn to the second type of question and survey a large number of recent results. In particular we first consider groups satisfying certain quadratic properties that we call *Lyndon properties* and show that the class of groups satisfying these properties are closed under many amalgam constructions. Elementary free groups satisfy these properties and these amalgam results extend the class of groups

satisfying Lyndon properties beyond the class of elementary free groups. We then introduce a class of groups that generalize a theorem of B. Baumslag [2] and then generalized by Gaglione and Spellman [42] and independently Remeslennikov [73]. All elementary free groups satisfy these theorems and we show that the classes of groups satisfying these results are fairly extensive. In the next section we provide some background material.

2 The Tarski Problems and Elementary Free Groups

The original *Tarski Problems* (or *Tarski Conjectures*) asked, among other things, whether all nonabelian free groups satisfy the same first-order or elementary theory.

Recall that a first-order sentence in group theory has logical symbols $\forall, \exists, \vee, \wedge, \sim$ but no quantification over sets. A first-order theorem in a free group is a theorem that says a first-order sentence is true in all nonabelian free groups. We make this a bit more precise:

We start with a first-order language appropriate for group theory. This language, which we denote by L_0, is the first-order language with equality containing a binary operation symbol $\cdot$, a unary operation symbol $^{-1}$ and a constant symbol 1. A *universal sentence* of L_0 is one of the form $\forall \overline{x}\{\phi(\overline{x})\}$ where $\overline{x}$ is a tuple of distinct variables, $\phi(\overline{x})$ is a formula of L_0 containing no quantifiers and containing at most the variables of $\overline{x}$. Similarly an *existential sentence* is one of the form $\exists \overline{x}\{\phi(\overline{x})\}$ where $\overline{x}$ and $\phi(\overline{x})$ are as above. A *universal-existential sentence* is one of the form $\forall \overline{x} \exists \overline{y}\{\phi(\overline{x}, \overline{y})\}$. Similarly defined is an *existential-universal sentence*. It is known that every sentence of L_0 is logically equivalent to one of the form $Q_{1x_1} \dots Q_{nx_n} \phi(\overline{x})$ where $\overline{x} = (x_1, \dots, x_n)$ is a tuple of distinct variables, each Q_i for $i = 1, \dots, n$ is a quantifier, either $\forall$ or $\exists$, and $\phi(\overline{x})$ is a formula of L_0 containing no quantifiers and containing free at most the variables $x_1, \dots, x_n$. Further vacuous quantifications are permitted. Finally a *positive sentence* is one logically equivalent to a sentence constructed using (at most) the connectives $\vee, \wedge, \forall, \exists$.

If G is a group then the *universal theory* of G consists of the set of all universal sentences of L_0 true in G. We denote the universal theory of a group G by $Th_\forall(G)$. Since any universal sentence is equivalent to the negation of an existential sentence it follows that two groups have the same universal theory if and only if they have the same *existential theory*. The set of all sentences of L_0 true in G is called the *first-order theory* or the *elementary theory* of G. We denote this by $Th(G)$. We note that being *first-order* or *elementary* means that in the intended interpretation of any formula or sentence all of the variables (free or bound) are assumed to take on as values only individual group elements — never, for example, subsets of nor functions, on the group in which they are interpreted.

We say that two groups G and H are *elementarily equivalent* (symbolically $G \equiv H$) if they have the same first-order theory, that is $Th(G) = Th(H)$.

Group monomorphisms which preserve the truth of first-order formulas are called elementary embeddings. Specifically, if H and G are groups and $f : H \to G$ is a monomorphism then f is an *elementary embedding* provided whenever $\phi(x_0, \dots, x_n)$ is a formula of L_0 containing free at most the distinct variables $x_0, \dots, x_n$ and $(h_0, \dots, h_n) \in H^{n+1}$ then $\phi(h_0, \dots, h_n)$ is true in H if and only if $\phi(f(h_0), \dots, f(h_n))$

is true in G. If H is a subgroup of G and the inclusion map $i : H \to G$ is an elementary embedding then we say that G is an *elementary extension* of H.

Two very important concepts in the elementary theory of groups are *completeness* and *decidability*. Given a nonempty class of groups $\mathcal{X}$ closed under isomorphism then we say its first-order theory is *complete* if given a sentence ϕ of L_0 then either ϕ is true in every group in $\mathcal{X}$ or ϕ is false in every group in $\mathcal{X}$. The first-order theory of $\mathcal{X}$ is *decidable* if there exists a recursive algorithm which, given a sentence ϕ of L_0 decides whether or not ϕ is true in every group in $\mathcal{X}$.

The positive solution to the Tarski Problems, given by Kharlampovich and Myasnikov (see [51]–[59] and independently by Sela (see [78]–[83]) is given in the next three theorems:

Theorem 2.1 (Tarski 1) *Any two nonabelian free groups are elementarily equivalent. That is any two nonabelian free groups satisfy exactly the same first-order theory.*

Theorem 2.2 (Tarski 2) *If the nonabelian free group H is a free factor in the free group G then the inclusion map $H \to G$ is an elementary embedding.*

In addition to the completeness of the theory of the nonabelian free groups the question of its *decidability* also arises. The *decidability* of the theory of nonabelian free groups means the question of whether there exists a recursive algorithm which, given a sentence ϕ of L_0, decides whether or not ϕ is true in every nonabelian free group. Kharlampovich and Myasnikov, in addition to proving the two above Tarski conjectures, also proved the following.

Theorem 2.3 (Tarski 3) *The elementary theory of the nonabelian free groups is decidable.*

Prior to the solution of the Tarski problems it was asked whether there exist non-free *elementary free groups*, that is whether there exists non-free groups that have exactly the same first-order theory as the class of nonabelian free groups. Elementary free groups are also known as *elementarily free groups*. The answer was yes, and both the Kharlampovich-Myasnikov solution and the Sela solution provide a complete characterization of the finitely generated elementary free groups. In the Kharlampovich-Myasnikov formulation these are given as a special class of what are termed NTQ groups (see [51]–[59]).

What is important for this paper is that the orientable surface groups of genus $g \geq 2$ are elementary free. Recall that a *surface group* is the fundamental group of a compact surface. If the surface is orientable it is an orientable surface group otherwise a nonorientable surface group.

If S_g denotes the orientable surface group of genus g then S_g has a one-relator presentation with a quadratic relator.

$$S_g = \langle a_1, b_1, \ldots, a_g, b_g; [a_1, b_1] \ldots [a_g, b_g] = 1 \rangle.$$

Groups with presentations similar to this play a major role in the structure theory of fully residually free groups and NTQ groups (see [51]–[59]).

Further if N_g denotes the nonorientable surface group of genus g then N_g has a one-relator presentation with a quadratic relator.

$$N_g = \langle a_1, \ldots, a_g; a_1^2 \cdots a_g^2 = 1 \rangle.$$

We note that the solution to the Tarski Problems implies that any first-order theorem holding in the class of nonabelian free groups must also hold in most surface groups. In many cases proving these results directly is very nontrivial.

Theorem 2.4 (see [51]–[59], [78]–[83]) *An orientable surface group of genus $g \geq 2$ is elementary free, that is has the same elementary theory as the class of nonabelian free groups. Further the nonorientable surface groups N_g for $g \geq 4$ are also elementary free.*

3 Surface Groups and Magnus's Theorem

Magnus proved the following theorem about the normal closures of elements in nonabelian free groups:

Theorem 3.1 (Magnus) *Let F be a finitely generated nonabelian free group and $R, S \in F$. Then if $N(R) = N(S)$, it follows that R is conjugate to either S or S^{-1}. Here $N(g)$ denotes the normal closure in F of the element g.*

J. Howie [46] and independently O. Bogopolski and Bogopolski–Sviridov [11] gave a proof of this for surface groups. Howie's proof was for orientable surface groups while Bogopolski and Sviridov also handled the nonorientable case. Their proofs were nontrivial and Howie's proof used the topological properties of surface groups. Howie further developed, as part of his proof of Magnus's theorem for surface groups, a theory of one-relator surface groups. These are surface groups modulo a single additional relator. Bogopolski and Bogopolski-Sviridov proved in addition that Magnus's Theorem holds in even a wider class of groups.

With some work it can be determined that Magnus's result is actually a first-order theorem on nonabelian free groups and hence from the theorems concerning the solution of the Tarski problems it holds automatically in all elementary free groups. In particular Magnus' theorem will hold in surface groups, both orientable and nonorientable of appropriate genus. If G is a group and $g \in G$ then $N(g)$, as in the statement of Magnus's Theorem above, will denote the normal closure in G of the element g.

Theorem 3.2 *Let G be an elementary free group and $R, S \in G$. Then if $N(R) = N(S)$, it follows that R is conjugate to either S or S^{-1}.*

Before exhibiting the proof of this result we mention the following two corollaries which extend Magnus's Theorem to surface groups and recover the results of Howie [46], Bogopolski [10] and Bogopolski-Sviridov [11].

Corollary 3.3 ([46], [10]) *Let S_g be an orientable surface group of genus $g \geq 2$. Then S_g satisfies Magnus's theorem, that is if $u, v \in S_g$ and $N(u) = N(v)$ it follows that u is conjugate to either v or v^{-1}.*

Corollary 3.4 ([11]) *Let N_g be a nonorientable surface group of genus $g \geq 4$. Then N_g satisfies Magnus's theorem, that is if $u, v \in S_g$ and $N(u) = N(v)$ it follows that u is conjugate to either v or v^{-1}. The genus $g \geq 4$ is essential here.*

We now present a proof of Theorem 3.2. From Theorem 3.2 the two corollaries describing this result in surface groups follow easily based on the solution to the Tarski problems coupled with the facts that orientable surface groups of genus $g \geq 2$ and nonorientable surface groups of genus $g \geq 4$ are elementary free.

Proof To prove the theorem we show that Magnus's theorem is actually a first-order result in nonabelian free groups. Therefore the result will hold in any elementary free group.

Magnus's theorem can be given by a sequence of elementary sentences of the form (see also [GLS]).

$$\{\forall R, S \in G, \forall g \in G, \exists g_1, \ldots, g_t, h_1, \ldots, h_k\}$$
$$(g^{-1}Rg = g_1^{-1}S^{\pm 1}g_1 \ldots g_t^{-1}S^{\pm 1}g_t) \wedge (g^{-1}Sg = h_1^{-1}R^{\pm 1}h_1 \ldots h_k^{-1}R^{\pm 1}h_k)\}$$
$$\Longrightarrow \{\exists x \in G(x^{-1}Rx = S \vee x^{-1}Rx = S^{-1})\}$$

Magnus's theorem is therefore a first-order result proving Theorem 3.2. □

As described prior to the proof it follows that any elementary free group and hence surface groups of the appropriate genus satisfy Magnus's theorem. This recovers the results in [46], [11], [10]. Actually more is true. An examination of the sentences capturing that Magnus's theorem (Theorem 3.1) is first-order shows that the sentences are universal-existential. Hence the theorem holds in the almost locally free groups of Gaglione and Spellman [43].

Before continuing we mention that Magnus's theorem is related to some interesting consequences for one-relator groups and their automorphisms.

Lemma 3.5 *If $G = \langle X; R\rangle$ a one-relator group and $\alpha \in Aut(F)$ then G is isomorphic to $\alpha(G) = \langle X; \alpha(R)\rangle$.*

The converse is not true. That is, there exist examples of two one-relator presentations $\langle X; R\rangle$ and $\langle X; S\rangle$ of a one-relator group G such that there is no $\alpha \in Aut(F)$ with $S = \alpha(R)$. An example can be found in the book by Collins and Zieschang [19]. However surprisingly the result has been shown to be *generically* true (Kapovich, Schupp, Shpilrain [49]). This means that a measure can be put on the set of one-relator presentations such that the asymptotic density of those one-relator groups that satisfy the above lemma is one.

4 Questions and Something for Nothing

The proof of Magnus's theorem for surface groups given in the last section is a type of something for nothing result. That is nontrivial proofs, such as those of Howie and Bogopolski, of results in certain classes of groups fall out directly from the solution to the Tarski problems. These types of proofs and results lead to several different types of questions:

1. Which additional nontrivial free group results are true in surface groups but difficult to obtain directly?
2. What first-order properties of nonabelian free groups are true beyond the class of elementary free groups?

In the next section we consider the first question and present a series of results true in all elementary free groups and in particular surface groups of the appropriate genus.

5 Results in Elementary Free Groups

As a simple example of the first type of question we consider the well known property concerning commutativity in free groups. It is well known (see [68]) that nonabelian free groups have cyclic centralizers of nontrivial elements. This is a consequence of the following somewhat weaker result.

Theorem 5.1 ([68]) *Let F be a nonabelian free group. If $x, y \in F$ and x, y commute then both x and y are powers of a single element $w \in F$.*

This result is given by the sentence

$$\forall\{x, y \in F\}([x, y] = 1) \implies \exists\{w \in F\}\exists\{m, n \in \mathbb{Z}\}(x = w^m \wedge y = w^n)$$

This is not first-order in the language of group theory since we must quantify over the integers which are not included in the language L_0. Hence this result is not necessarily true in elementary free groups. As an example, let D be a nonprincipal ultrafilter on $\mathbb{Z}$ (see [8]). Let $F = \langle a_1, a_2; \rangle$ the free group of rank 2 on a_1, a_2 and let ${}^*F = F^{\mathbb{Z}}/D$ be the corresponding ultrapower so that ${}^*F$ is elementary free (see [8]). Consider the elements

$$[(a_1)_{k\in\mathbb{Z}}]_D = [(\dots, a_1, a_1, \dots, a_1, \dots)]_D$$

and

$$[(a_1^k)_{k\in\mathbb{Z}}]_D = [(\dots, a_1^{-2}, a_1^{-1}, 1, a_1, a_1^2, \dots)]_D.$$

These commute but there is no fixed element B of which they are both powers.

However the following result can be proved directly:

Theorem 5.2 *Let G be a finitely generated elementary free group. Then G has cyclic centralizers of nontrivial elements. It follows that if $x, y \in G$ and x,y commute then both x and y are powers of a single element $w \in G$.*

Proof Let G be a finitely generated elementary free group. Then G is finitely generated and fully residually free. It follows from the fact that finitely generated fully residually free groups are commutative transitive that G has abelian centralizers. Applying Szmielew's criteria for elementary equivalence of abelian groups (see [85]) it follows that in any elementary free group the centralizer of any nontrivial element

is elementarily equivalent to the infinite cyclic group. In particular such centralizers must satisfy the sentences:

$$\forall x_1, x_2(x_1x_2 = x_2x_1);$$
$$\exists x(x \neq 1);$$

for each integer $n \geq 2$ the sentence

$$\forall x((x^n = 1) \to (x = 1));$$

and the sentence

$$\forall x_1, x_2, x_3 \exists y((x_1x_2^{-1} = y^2) \vee (x_1x_3^{-1} = y^2) \vee (x_2x_3^{-1} = y^2));$$

asserting that, modulo 2, there are at most 2 distinct elements.

A result of Gaglione, Lipschutz and Spellman (Lemma 3.6 in [41]) shows that up to isomorphism the only finitely generated group M which can satisfy these properties simultaneously is the infinite cyclic group. Here we will repeat the proof given there.

Suppose not and M is a finitely generated abelian group satisfying the above sentences. Then M contains a rank 2 free abelian direct factor A and suppose that $M = A \times B$.

Now let $(a_1, a_2, a_3) \in A^3$. Then there is $a \in A$, $b \in B$ such that

$$a_1a_2^{-1} = a^2b^2 \vee a_1a_3^{-1} = a^2b^2 \vee a_2a_3^{-1} = a^2b^2.$$

Since the product is direct $b^2 = 1$ is the only possibility. Then, writing $A(X^2)$ for the subgroup of A generated by the squares, $a_1 \equiv a_2 \bmod A(X^2)$ or $a_1 \equiv a_3 \bmod A(X^2)$ or $a_2 \equiv a_3 \bmod A(X^2)$. Since $(a_1, a_2, a_2) \in A^3$ was arbitrary, the index $[A : A(X^2)] \leq 2$. However if A has rank 2 it follows that $[A : A(X^2)] = 4$. This contradiction shows that M is cyclic. □

As a corollary we get that the result must be true in surface groups a fact that can also be obtained directly from the amalgam structure of such groups or from their faithful representations in $PSL(2, \mathbb{C})$.

Corollary 5.3 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$. If $x, y \in G$ and x, y commute then both x and y are powers of a single element $w \in F$.*

An example that is less trivial and is not obvious in a surface group is the following. The next theorem can be easily proved in free groups.

Theorem 5.4 *Let F be a free group and n, k nonzero integers. For all $x, y \in F$ if $[x^n, y] = [x, y^k]$ then either $n = k = 1$ or x, y commute and both are powers of a single element.*

The first part of the result that either $n = k$ or $[x, y] = 1$ is first-order given by a sequence of elementary sentences, one for each $(n, k) \in \mathbb{Z}^2 \setminus \{(1, 1)\}$ with neither n nor k zero;

$$\forall x, y \in F([x^n, y] = [x, y^k]) \implies [x, y] = 1$$

Therefore this part of the result must hold in any elementary free group. Further if the elementary free group is finitely generated the second part must also hold.

Corollary 5.5 *Let G be an elementary free group. If $x, y \in G$ and if $[x^n, y] = [x, y^k]$ then either $n = k = 1$ or x, y commute. If G is finitely generated then both x and y are powers of a single element $w \in G$.*

Since surface groups are finitely generated we have the following.

Corollary 5.6 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$. If $x, y \in G$ and if $[x^n, y] = [x, y^k]$ then either $n = k = 1$ or x, y commute and then both x and y are powers of a single element $w \in G$.*

Csorgo, Fine and Rosenberger [21] proved the following extension of this.

Theorem 5.7 ([21]) *Suppose F is a nonabelian free group and $x, y, u, v \in F$ with $[x, y] \neq 1$ and u, v in the subgroup generated by x, y. Then if $[x, y]$ is conjugate to a power of $[u, v]$ within $\langle x, y \rangle$, that is there exists a k with $[x, y] = g([u, v]^k)g^{-1}$ for some $g \in \langle x, y \rangle$, and $[x, y^m] = [u, v^n]$ it follows that $m = n$. Further if $m = n \geq 2$ then y is conjugate within $\langle x, y \rangle$ to v or v^{-1}.*

As with Magnus's theorem this can be shown to be given by a sequence of first-order sentences and is hence a first-order result. Therefore this holds in any elementary free group.

Theorem 5.8 *Let G be an elementary free group and $x, y, u, v \in G$ with $[x, y] \neq 1$ and u, v in the subgroup generated by x, y. Then if $[x, y]$ is conjugate to a power of $[u, v]$ within $\langle x, y \rangle$, that is there exists a k with $[x, y] = g([u, v]^k)g^{-1}$ for some $g \in \langle x, y \rangle$, and $[x, y^m] = [u, v^n]$ it follows that $m = n$. Further if $m = n \geq 2$ then y is conjugate within $\langle x, y \rangle$ to v or v^{-1}.*

In particular we get the extension to surface groups.

Corollary 5.9 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$ and suppose that $x, y, u, v \in G$ with $[x, y] \neq 1$ and u, v in the subgroup generated by x, y. Then if $[x, y]$ is conjugate to a power of $[u, v]$ within $\langle x, y \rangle$, that is there exists a k with $[x, y] = g([u, v]^k)g^{-1}$ for some $g \in \langle x, y \rangle$, and $[x, y^m] = [u, v^n]$ it follows that $m = n$. Further if $m = n \geq 2$ then y is conjugate within $\langle x, y \rangle$ to v or v^{-1}.*

5.1 The Retract Theorem and Turner Groups

An element g in a group G is a *test element* if whenever $f(g) = g$ for some endomorphism of G then f must be an automorphism. This concept dates back to Nielsen who showed that $[x, y]$ is a test element in the free group on $\{x, y\}$. Test elements in a free group are called *test words* (see [39]).

Turner [86] gave the following characterization of test words in finitely generated free groups. This is now referred to as either the *Retract Theorem* or *Turner's Theorem*.

Theorem 5.10 *Let F be a finitely generated nonabelian free group. Then an element $g \in F$ is a test word if and only if g lies in no proper retract.*

The question whether Turner's theorem is first-order or not was considered in [23] where it was shown that the theorem is not first-order. We call an element g in a group G *nonprojectible* if it lies in no proper retract of G. We then call a group G a *Turner group* if for $g \in G$ being nonprojectible in G implies that g is a test element. Equivalently G is a Turner group if and only if the Retract Theorem holds. Hence Turner's theorem says that nonabelian free groups are Turner groups.

A group G is *stably hyperbolic* if G is hyperbolic and for any endomorphism $\phi : G \to G$ for all n there is an $m \geq n$ such that $\phi^m(G)$ is hyperbolic. A result of O'Neill and Turner (see [71]) shows that stably hyperbolic groups are Turner groups. Using this result we can prove that finitely generated elementary free groups are Turner groups, that is they satisfy the Retract Theorem.

Theorem 5.11 *Let G be a finitely generated elementary free group. Then G is a Turner group, that is G satisfies the Retract Theorem and hence the test elements in G are precisely those elements that avoid any proper retract.*

Proof Let G be a finitely generated elementary free group. Then G is finitely generated and fully residually free. It follows from the classification in [24] that any two generator subgroup is either free of rank 2 or free abelian of rank 2. Further since it is finitely generated and elementary free it has cyclic centralizers. Since G is fully residually free and has cyclic centralizers it is hyperbolic (see [51]–[54]). Full residual freeness is preserved by subgroups as is cyclic centralizers and therefore any finitely generated subgroup of G is also hyperbolic. Let $\phi : G \to G$ be an endomorphism. Since G is finitely generated then $\phi^n(G)$ is also a finitely generated fully residually free group for any natural number n. Hence $\phi^n(G)$ is hyperbolic for any n and therefore G is stably hyperbolic. Therefore G satisfies the Retract Theorem from the result of O'Neill and Turner. □

As in the previous cases this then extends to surface groups of appropriate genus.

Corollary 5.12 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$. Then G is a Turner group.*

In [71] it was proved directly that there are test elements in surface groups. However this also follows directly from the previous corollary since not every element in either S_g or N_g falls in a proper retract.

Corollary 5.13 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$. Then G has test elements.*

In [23] the following results were proved showing that Turner's Theorem is not first-order and not the model class of any set of sentences of L_0.

Theorem 5.14 (Nondefinability Theorem) *1. There is no set $N(x)$ of formulas of L_0 such that, for an arbitrary group G and arbitrary element $g \in G$, $N(g)$ holds if and only if g is nonprojectible.*

2. There is no set $T(x)$ of formulas of L_0 such that, for an arbitrary group G and arbitrary element $g \in G$, $T(g)$ holds if and only if g is a test element.

Theorem 5.15 (Nonaxiomatizability Theorem) *The class of Turner groups is not the model class of any set of sentences of* L_0.

5.2 Conjugacy Separability of Elementary Free Groups

A group G is *conjugacy separable* if given any two elements $g_1, g_2 \in G$ either g_1 is conjugate to g_2 or there exists a homomorphism $\rho : G \to H$ where H is a finite group and in which $\rho(g_1)$ is not conjugate to $\rho(g_2)$. It is known that all free groups are conjugacy separable. Here we next prove that all finitely generated elementary free groups are conjugacy separable.

Theorem 5.16 *Let G be a finitely generated elementary free group. Then G is conjugacy separable.*

Proof Suppose G is a finitely generated elementary free group and g_1, g_2 are two nonconjugate elements of G. Since free groups are conjugacy separable to show that G is conjugacy separable it suffices to show that there is a free homomorphic image of G in which the images of g_1 and g_2 are nonconjugate.

Suppose there is no free homomorphic image of G in which g_1 is not conjugate to g_2. Note that a finitely generated elementary free group, in fact more generally a finitely generated fully residually free group must be finitely presented (see [52]). Fix a finite presentation for G,

$$\langle a_1, \ldots, a_n; R_1(a_1, \ldots, a_n) = \cdots = R_m(a_1, \ldots, a_n) = 1 \rangle$$

and suppose that $g_i = w_i(a_1, \ldots, a_n)$ for $i = 1, 2$. Then since there are no free homomorphic images of G in which g_1 and g_2 are not conjugate the following universal-existential sentence which we denote by $\langle 1 \rangle$ of L_0 would be true in every nonabelian free group

$$\forall x_1, \ldots, x_n \exists y (\wedge_{i=1}^{m} (R_i(x_1, \ldots, x_n) = 1)) \to (w_2(x_1, \ldots, x_n) = y^{-1} w_1(x_1, \ldots, x_n) y).$$

It follows that $\langle 1 \rangle$ would have to be true in G. But this contradicts the fact that g_1 is not conjugate to g_2 in G. Therefore there must exist a free homomorphic image in which g_1 and g_2 are not conjugate and hence G is conjugacy separable. □

5.3 The Genus Question

Let f lie in the derived group $[F, F]$ of the free group F. We define the *genus* of f, written $genus_F(f)$, as follows. If $f = 1$, then $genus_F(f) = 0$. Otherwise, $genus_F(f) = g$ where g is the least positive integer n such that f can be expressed in F as the product of n commutators. The following question was posed in [43].

The Genus Question For each ordered pair (g, k) of positive integers, must there exist a finite bound $n(g, k)$ on

$$\{genus_F(f) : genus_F(f^k) \leq g\}?$$

The question was answered in the negative in [54]. We present here a variant of their argument. Let $X = \langle x_1, x_2, x_3, x_4; \rangle$ and $F = \langle b_1, b_2; \rangle$. By a solution in F to

the equation $x_1^2x_2^2x_3^2x_4^2 = 1$ we mean an element $(f_1, f_1, f_3, f_4) \in F^4$ such that the homomorphism ϕ given by $x_i \mapsto f_i,\ \ i = 1, 2, 3, 4$, maps $x_1^2x_2^2x_3^2x_4^2$ to 1. Let V be the solution set of $x_1^2x_2^2x_3^2x_4^2 = 1$ in F^4. Note that V is nonempty since, for example, $(b_1, b_2, b_2^{-1}, b_1^{-1}) \in V$. Now let x abbreviate the element $(x_1x_2x_3x_4)^2(x_1^2x_2^2x_3^2x_4^2)^{-1}$ of X so that clearly x lies in its derived group $[X, X]$. Let $g = genus_X(x)$ be its genus. From this it follows that, for all $f_1, f_2, f_3, f_4 \in V$ we would have $genus_F(f_1f_2f_3f_4)^2 \leq g$. If there were a finite bound $n = n(g, 2)$ on

$$\{genus_F(f) : genus_F(f^2) \leq g\}$$

then the following universal-existential sentence (call it ϕ) would hold in the nonabelian free groups.

$$\forall x_1, x_2, x_3, x_4 \exists y_1, \ldots, y_{2n}((x_2^2x_2^2x_3^2x_4^2 = 1) \rightarrow (x_1x_2x_3x_4 = [y_1, y_2] \ldots [y_{2n-1}, y_{2n}]))$$

(Note that if $x_1x_2x_3x_4$ is expressible as the product of $m < n$ commutators then we can append $n - m$ trivial factors $[1, 1]$ to the product.)

Since the nonorientable surface group N_4 of genus 4 is elementary free we would have that the sentence ϕ holds in N_4. Let $N_4 = \langle a_1, a_2, a_3, a_4; a_1^2a_2^2a_3^2a_4^2 = 1\rangle$. Then $a_1^2a_2^2a_3^2a_4^2 = 1$ whereas $a_1a_2a_3a_4$ maps to an element of order 2 modulo the derived group $[N_4, N_4]$. Hence $a_1a_2a_3a_4$ does not lie in the derived group N_4 let alone be the product of n commutators in N_4. This contradiction shows that there can be no finite bound on the genus.

5.4 The Other Direction

Sometimes the solution of the Tarski problem and the fact that surface groups are elementary free can be used to prove results in free groups that are very difficult or impossible to prove directly within nonabelian free groups. As an example we use the negative solution to the genus question. The following sentence is clearly true in N_4, the nonorientable surface group of genus 4, where n is a large natural number;

$$\exists x_1, x_2, x_3, x_4 \forall y_1, y_2, \ldots, y_{2n}((x_1^2x_2^2x_3^2x_4^2) = 1) \wedge (x_1x_2x_3x_4 \neq [y_1, y_2] \cdots [y_{2n-1}y_{2n}]))$$

Since it is first-order and true in N_4 and N_4 is elementary free it follows that it is true in all nonabelian free groups. However this is extremely difficult to prove directly alhtough it is an extension of the well-known result in free groups that a commutator cannot be a square.

5.5 Tame Automorphisms of Elementary Free Groups

As part of the proof of the Tarski theorems, both Kharlampovich-Myasnikov and Sela completely described the structure of finitely generated fully residually free groups in terms of what is called the JSJ-decomposition. These structure results can be used to both solve the isomorphism problem for limit groups and to prove that the automorphism group of a finitely generated fully residually free group is tame. It follows that the automorphism group of an elementary free group is also tame. We explain these concepts.

A minimal finite presentation of a finitely presented group G is a presentation that is minimal with respect to the number of generators. Hence a presentation $G = \langle x_1, \ldots, x_n; r_1, \ldots, r_m \rangle$ is a minimal finite presentation for G if $n = \text{rank}(G)$, the minimal number of generators necessary to present G. Now suppose that $G = \langle x_1, \ldots, x_n; r_1, \ldots, r_m \rangle$ with $1 \leq n, m < \infty$ is minimal finite presentation of G. Let $F = \langle x_1, \ldots, x_n; \rangle$ be the free group of rank n on $\{x_1, \ldots, x_n\}$. An automorphism $\alpha : G \to G$ is *tame* if it is induced by or lifts to an automorphism on F (considered as free on the generators of G). If each automorphism of G is tame we say that the automorphism group $Aut(G)$ is *tame*. In [84] Shpilrain gives a survey of some of the known general results on tame automorphisms and tame automorphism groups. If G is a surface group a result of Zieschang [88] and improved upon by Rosenberger [77] shows that G has only one Nielsen class of minimal generating systems. An easy consequence of this is that that $Aut(G)$ is tame. Rosenberger (see [67] or [76]) uses the term *almost quasifree* for a finitely presented group which has a tame automorphism group. If G is almost quasifree, $G = \langle x_1, \ldots, x_n; r_1, \ldots, r_m \rangle, 1 \leq n, m < \infty$ a minimal finite presentation of G and, in addition, each automorphism of $F = \langle x_1, \ldots, x_n; \rangle$ induces an automorphism of G, G is called *quasifree*. Rosenberger observed that a non-cyclic, non-free one-relator group is quasifree only if it has a presentation $\langle a, b; [a, b]^n = 1 \rangle$ for $n \geq 1$. This is a Fuchsian group if $n \geq 2$ and isomorphic to a free abelian group of rank 2 if $n = 1$.

JSJ decompositions were introduced by Rips and Sela [74]. A JSJ-decomposition of a group G is a graph of groups decomposition of G with abelian edge groups that encodes all other graph of groups decompositions of G. Any finitely generated fully residually free group has a JSJ decomposition with cyclic edge groups and vertex groups of specific types if it is not abelian or a surface group. We refer to the relevant papers for further discussions of these but mention that Bumagin, Kharlampovich and Myasnikov [14] used the JSJ decomposition to describe the automorphism group of a limit group. See [14] for a decsription of the canonical automorphisms.

Theorem 5.17 ([14]) *Let G be a finitely generated fully residually free group not abelian or a surface group and let Γ be a cyclic JSJ-decomposition for G. Let $Out_\Gamma(G)$ be the outer automorphism group of G generated by the canonical automorphisms. Then $Out_\Gamma(G)$ has finite index in $Out(G)$.*

As an application Bumagin, Kharlampovich and Myasnikov [14] were further able to prove that the isomorphism problem is solvable for finitely generated fully residually free groups. This is actually part of the algorithmic study of this class of groups (see [60, 61]).

Theorem 5.18 ([14]) *The isomorphism problem is solvable in the class of finitely generated fully residually free groups. That is given two finite presentations that are known to define fully residually free groups there is an effective algorithm to determine if the defined groups are isomorphic.*

As an additional consequence of the JSJ decomposition of a fully residually free group and the work of Bumagin, Kharlampovich and Mysasnikov, the tameness of $Aut(G)$ for a limit group was proved by Fine, Kharlmapovich, Myasnikov, Rosenberger and Remeslennikov [26].

Theorem 5.19 ([26]) *The automorphism group $Aut(G)$ of a finitely generated freely indecomposable fully residually free group G is tame with respect to a presentation for the JSJ decomposition for G.*

Since each finitely generated elementary free group is universally free and hence fully residually free the proof of the corollary is immediate.

Corollary 5.20 *The automorphism group of a finitely generated freely indecomposable elementary free group G is tame.*

We note that the converse of this corollary is false. That is there do exist groups (in fact hyperbolic groups) where every automorphism is tame but which are not fully residually free. As an example the groups

$$G = \langle a_1, \ldots, a_n; a_1^{\alpha_1} \cdots a_n^{\alpha_n} \rangle, \text{ with } n \geq 4, 2 \leq \alpha_1, \ldots, \alpha_n,$$

and

$$H = \langle s_1, \ldots, s_n; s_1^2, \ldots, s_{n-1}^2, s_n^{2k+1}, s_1 s_2 \cdots s_n \rangle \text{ with } n = 2\ell, n \geq 4 \text{ even and } k \geq 1,$$

are all hyperbolic. Further every automorphism is tame (see [75] and [39]). However not all of these groups are fully residually free.

5.6 C-test Words in Free Groups and Surface Groups

We now give a much more difficult but more technical example. We thank A. Myasnikov for bringing this to our attention. V. Shpilrain (see [64]) posed the following problem:

Problem Are there two elements u_1, u_2 in a free group F_m of rank $m \geq 2$ such that any endomorphism ϕ of F_m with non-cyclic image is uniquely determined by $\phi(u_1)$ and $\phi(u_2)$?

Ivanov [47] answered this in the affirmative in the case where ϕ is a monomorphism. To do this he introduced C-test words.

Definition 5.21 A nonempty word $v(x_1, \ldots, x_n)$ is a C-test word in n letters for F_m if for any two n-tuples $(X_1, \ldots, X_n), (Y_1, \ldots, Y_n)$ of elements of F_m the equality $v(X_1, \ldots, X_n) = v(Y_1, \ldots, Y_n) \neq 1$ implies the existence of an element $S \in F_m$ such that $Y_i = SX_iS^{-1}$ for all $i = 1, \ldots, n$.

Donghi Lee [64] extends Ivanov's work and uses C-test words to provide a positive solution to Shpilrain's question in the case where the endomorphism has non-cyclic image. For this extension Lee proves the following result:

Theorem 5.22 *For every $n \geq 2$ there exists a C-test word $v_n(x_1, \ldots, x_n)$ in n letters for F_m with the additional property that $v_n(x_1, \ldots, x_n) = 1$ if and only if the subgroup $\langle X_1, \ldots, X_n \rangle$ of F_m generated by $X_1, ..., X_n$ is cyclic.*

This theorem can be given by a sequence of elementary sentences of the following form. For each n we assume that we have a fixed C-test word $v_n(x_1, \ldots, x_n)$ satisfying D. Lee's Theorem. The integer n is fixed in each sentence.

$$(S_n) \qquad \forall x_1, \ldots, x_n, y_1, \ldots, y_n (v_n(x_1, \ldots, x_n) = v_n(y_1, \ldots, y_n))$$
$$\wedge (v_n(x_1, \ldots, x_n) \neq 1) \rightarrow \exists T \left(\bigwedge_{i=1}^{n} T x_i T^{-1} = y_i \right)$$
$$\wedge \forall x_1, \ldots, x_n ((v_n(x_1, \ldots, x_n) = 1) \leftrightarrow \bigwedge_{1 \leq i < j \leq n} ([x_i, x_j] = 1)$$

For a general group G we define a C-test word exactly as for a free group.

Definition 5.23 Let G be a group. Then a nonempty word $v(x_1, \ldots, x_n)$ is a C-test word in n letters for G if for any two n-tuples $(X_1, \ldots, X_n), (Y_1, \ldots, Y_n)$ of elements of G the equality $v(X_1, \ldots, X_n) = v(Y_1, \ldots, Y_n) \neq 1$ implies the existence of an element $S \in G$ such that $Y_i = S^{-1} X_i S$ for all $i = 1, \ldots, n$.

From D. Lee's theorem each first order sentence (S_n) holds in any nonabelian free groups and hence each (S_n) must hold in any elementary free groups. Therefore in an elementary free group $v_n(x_1, \ldots, x_n) = 1$ forces $\langle x_1, \ldots, x_n \rangle$ to be abelian and in particular in any finitely generated elementary free group $v_n(x_1, \ldots, x_n) = 1$ forces $\langle x_1, \ldots, x_n \rangle$ to be cyclic.

It follows that D. Lee's result is true in orientable surface groups of of genus $g \geq 2$ and nonorientable surface groups of genus $g \geq 4$. Hence we get the following result.

Theorem 5.24 *Let G be either an orientable surface group of genus $g \geq 2$ or a nonorientable surface group of genus $g \geq 4$. Then for every $n \geq 2$ there exists a C-test word $v_n(x_1, \ldots, x_n)$ in n letters for G with the additional property that $v_n(x_1, \ldots, x_n) = 1$ if and only if the subgroup $\langle X_1, \ldots, X_n \rangle$ of G generated by $X_1, \ldots, X_n$ is cyclic.*

For the remainder of this paper we will be considering the second type of question. We first consider groups satisfying certain quadratic properties that we call *Lyndon properties* and show that the class of groups satisfying these are closed under many amalgam constructions. We next discuss a class of groups that generalize a theorem of B. Baumslag [2] and then generalized by Gaglione and Spellman [42] and independently Remeslennikov [73]. All elementary free groups satisfy these theorems and we show that the classes of groups satisfying these results are fairly extensive and beyond the class of elementary free groups.

6 Cyclically Pinched and Conjugacy Pinched Constructions

The algebraic generalization of the one-relator presentation type of a surface group presentation leads to *cyclically pinched one-relator groups*. These groups have the same general form of a surface group and have proved to be quite amenable to study. In particular a *cyclically pinched one-relator group* is a one-relator group of the following form

$$G = \langle a_1, \ldots, a_p, a_{p+1}, \ldots, a_n; U = V \rangle$$

where $1 \neq U = U(a_1, \ldots, a_p)$ is a cyclically reduced, non-primitive (not part of a free basis) word in the free group F_1 on $a_1, \ldots, a_p$ and $1 \neq V = V(a_{p+1}, \ldots, a_n)$ is a cyclically reduced, non-primitive word in the free group F_2 on $a_{p+1}, \ldots, a_n$.

Clearly such a group is the free product of the free groups on $a_1, \ldots, a_p$ and $a_{p+1}, \ldots, a_n$ respectively amalgamated over the cyclic subgroups generated by U and V. More generally if G_1 and G_2 are groups and U, V are elements of G_1, G_2 respectively then the cyclically pinched construction from these is a free product with amalgamation

$$G = G_1 \underset{\{U=V\}}{\star} G_2.$$

Notice that from the Poincare presentation that most finitely generated Fuchsian groups are cyclically pinched constructions from free products of cyclics.

Cyclically pinched one-relator groups have been shown to be extremely similar to surface groups. We summarize many of the most important results.

Theorem 6.1 *Let G be a cyclically pinched one-relator group. Then*

1. *G is residually finite (G. Baumslag [3]).*
2. *G has a solvable conjugacy problem (S. Lipschutz [65]) and is conjugacy separable (J. Dyer [22]).*
3. *G is subgroup separable (Brunner, Burns and Solitar [13]).*
4. *If neither U nor V is a proper power then G has a faithful representation over some commutative field (Wehrfritz [87]).*
5. *If neither U nor V is a proper power then G has a faithful representation in $PSL_2(\mathbb{C})$ (Fine-Rosenberger [34]). In fact G has a faithful representation in $PSL(2, \mathbb{R})$ ([27]).*
6. *If neither U nor V is a proper power then G is hyperbolic ([9, 48, 53]).*
7. *If neither U nor V is in the commutator subgroup of its respective factor then G is free-by-cyclic ([4]).*
8. *The isomorphism problem for any cyclically pinched one-relator group is solvable; given a cyclically pinched one-relator group G there is an algorithm to decide in finitely many steps whether an arbitrary one-relator group is isomorphic or not to G ([76]).*

The HNN analogs of cyclically pinched one-relator groups are called *conjugacy pinched one-relator groups* and are also motivated by the structure of orientable surface groups. A *conjugacy pinched one-relator group* is a one-relator group of the form

$$G = \langle a_1, \ldots, a_n, t; tUt^{-1} = V \rangle$$

where $1 \neq U = U(a_1, \ldots, a_n)$ and $1 \neq V = V(a_1, \ldots, a_n)$ are cyclically reduced in the free group F on $a_1, \ldots, a_n$.

Structurally such a group is an HNN extension of the free group F on $a_1, \ldots, a_n$ with cyclic associated subgroups generated by U and V and is hence the HNN analog of a cyclically pinched one-relator group.

Groups of this type arise in many different contexts and share many of the general properties of the cyclically pinched case. However many of the proofs become tremendously more complicated in the conjugacy pinched case than the cyclically pinched

case. Further in most cases additional conditions on the associated elements U and V are necessary. In [30], see also [33], a partial solution to the isomorphism problem for conjugacy pinched one-relator groups was given.

An extremely important conjugacy pinched construction is an *extension of centralizers*. Recall that $H \subset G$ is *malnormal* if $xHx^{-1} \cap H = \{1\}$ for $x \notin H$. A CSA group is a group G in which maximal abelian subgroups are malnormal. CSA groups can be shown to be commutative transitive. Let B be a CSA group so that the centralizer of an element is abelian. Let $U \in B$ not a proper power then the rank one extension of centralzier is the conjugacy pinched construction

$$G = \langle t, B; rel(B), t^{-1}Ut = U \rangle$$

Myasnikov and Remeslennikov proved that a finitely generated fully residually free group is embeddable as a subgroup in an iterated extension of centralizers starting with free groups (see [51]–[59])

Theorem 6.2 ([73]) *Any finitely generated fully residually free group can be embedded as a subgroup in a finite iterated extension of centralizers starting with free groups.*

This result has been used effectively to prove results about finitely generated fully residually free groups also called limit groups. In particular in [24] a complete classification of limit groups of rank three or less was given. Using both this theorem and a characterization of limit groups in terms of nonstandard free groups Fine and Rosenberger [35, 36] proved that any limit group has an effective faithful representation in $PSL(2, \mathbb{C})$ generalizing what is known for surface groups.

Theorem 6.3 ([35, 36]) *Let G be a limit group. Then G has a faithful representation $\rho : G \to PSL(2, \mathbb{C})$. Further a faithful representation can be effectively constructed given information on the graph of groups decomposition of G.*

Related to the cyclically pinched and conjugacy pinched constructions is the general concept of n-free groups. A group G is n-free for some natural number n if any subgroup generated by n or fewer elements must be a free group. If S_g is an orientable surface group of genus g then S_g is $(2g-1)$-free.

G. Baumslag [3] proved that a cyclically pinched one-relator group with the property that U and V are not proper powers is 2-free. This was generalized by G. Rosenberger to show that such groups are 3-free. We extended these results in several directions:

Theorem 6.4 ([75]) *Let G be a cyclically pinched one-relator group with free factors F_1, F_2 and amalgamated elements U and V. Suppose that U and V are not proper powers in the respective free groups on the generators which they involve. Then*

1. *G is 3-free.*
2. *Let $H \subset G$ be a subgroup of rank 4. Then one of the following two cases occurs:*
   (i) H is free of rank 4.

(ii) If $\{x_1, \ldots, x_4\}$ is a generating system for H then there is a Nielsen transformation from $\{x_1, \ldots, x_4\}$ to $\{y_1, \ldots, y_4\}$ with $y_1, y_2 \in zF_1z^{-1}$, $y_3, y_4 \in zF_2z^{-1}$ for a suitable $z \in G$. Further there is a one-relator presentation for H on $\{x_1, ..., x_4\}$.

In conjunction with a study on the universal theory of non-abelian free groups the freeness part of the above results was extended Fine, Gaglione, Rosenberger and Spellman [25], using Nielsen reduction techniques.

7 The Basic Lyndon Properties

We now return to the question of what first-order free group results are true beyond the class of elementary free groups.

Vaught proposed the question whether in a free group a solution of the equation $x^2y^2z^2 = 1$ must generate an abelian (and hence cyclic) subgroup. This was proved by Lyndon and Schutzenberger and then generalized by Baumslag. Based on Lyndon's result in any group G we define the following *Lyndon Properties*.

Definition 7.1 The following are called *Lyndon properties*. Let G be a group. Then G satisfies Property

1. *LZ* if whenever $x^2y^2z^2 = 1$ for $x, y, z \in G$ then $\langle x, y, z\rangle$ is cyclic;
2. *LA* if whenever $x^2y^2z^2 = 1$ for $x, y, z \in G$ then $\langle x, y, z\rangle$ is abelian;
3. *LPZ* if whenever $x^py^qz^r = 1$ for $x, y, z \in G$ with $2 \leq p, q, r \in \mathbb{N}$ then $\langle x, y, z\rangle$ is cyclic;
4. *LPA* if whenever $x^py^qz^r = 1$ for $x, y, z \in G$ with $2 \leq p, q, r \in \mathbb{N}$ then $\langle x, y, z\rangle$ is abelian;
5. *LCZ* if whenever $[x^p, y^q]z^r = 1$ for $x, y, z \in G$ with $1 \leq p, q \in \mathbb{N}$, $2 \leq r \in \mathbb{N}$ then $\langle x, y, z\rangle$ is cyclic;
6. *LCA* if whenever $[x^p, y^q]z^r = 1$ for $x, y, z \in G$ with $1 \leq p, q \in \mathbb{N}$, $2 \leq r \in \mathbb{N}$ then $\langle x, y, z\rangle$ is abelian.

All of these properties hold in free groups and Properties LA, LPA and LCA are given by universal sentences. Hence these hold in any fully residually free group. They are all first-order so they hold in an elementary free group. Note that it is an open question as to when a free product with amalgamation of fully residually free groups is still fully residually free.

We show that the Lyndon properties extend beyond the class of elementary free groups by showing that the property is preserved under some general group amalgams.

7.1 Lyndon Properties in Amalgams and One-Relator Groups

Using Nielsen cancellation methods it can be proved that several of the Lyndon properties are preserved under special free product with amalgamation constructions (see [31]). Since these constructions are not always fully residually free it makes the class of groups satisfying the Lyndon properties wider than the class of limit groups. The following results are from [31].

Theorem 7.2 *Suppose that H_1 and H_2 are groups with no elements of order 2 and that G is the amalgamated product $G = H_1 \star_A H_2$ with $H_1 \neq A \neq H_2$ and A is malnormal in both H_1 and H_2. Then*

1. *if both H_1 and H_2 satisfy Property LZ then G also satisfies Property LZ;*
2. *if both H_1 and H_2 satisfy Property LA then G also satisfies Property LA.*

Theorem 7.3 *Suppose that H_1 and H_2 are torsion-free groups and that G is the amalgamated product $G = H_1 \star_A H_2$ with $H_1 \neq A \neq H_2$ and A is malnormal in both H_1 and H_2. Then*

1. *if both H_1 and H_2 satisfy Property LPZ then G also satisfies Property LPZ;*
2. *if both H_1 and H_2 satisfy Property LPA then G also satisfies Property LPA;*
3. *if both H_1 and H_2 satisfy Property LCZ then G also satisfies Property LCZ;*
4. *if both H_1 and H_2 satisfy Property LCA then G also satisfies Property LCA.*

In particular if $x, y \in G$ with $[x, y] \neq 1$ then if both H_1, H_2 have property LCZ or LCA then $[x, y]$ is never a proper power.

We note that the malnormality condition in both theorems is essential. For example in the nonorientable surface group of genus 3

$$G = \langle a, b, c; a^2b^2c^2 = 1\rangle$$

we have an equation $x^2y^2z^2 = 1$ such that $\langle x, y, z\rangle$ is not abelian.

Theorem 7.4 *Suppose that G is a cyclically pinched one-relator group*

$$G = \langle a_1, \ldots, a_p, b_1, \ldots, b_q; WV = 1\rangle$$

where $p \geq 2$, $q \geq 2$, $1 \neq W = W(a_1, \ldots, a_p)$ is not a proper power nor a primitive element in the free group on $a_1, \ldots, a_p$ and $1 \neq V = V(b_1, \ldots, b_q)$ is not a proper power nor a primitive element in the free group on $b_1, \ldots, b_q$. Then G has properties LZ, LPZ, and LCZ.

The key idea in the above theorem is that a cyclically pinched one-relator group of the above form is 3-free, that is any subgroup generated by 3 or fewer elements must be a free group. The following is then immediate.

Lemma 7.5 *Let G be a 3-free group. Then G satisfies properties LZ, LA, LPZ, LPA, LCZ and LCA.*

In [31] a theorem on 3-freeness was further extended.

Theorem 7.6 *Suppose that $G = H_1 \star_A H_2$ is an amalgamated free product with $H_1 \neq A \neq H_2$. Suppose further that A is malnormal in both H_1 and H_2 and that both H_1 and H_2 are 3-free. Then G is 3-free.*

Combining this result with the lemma we get.

Corollary 7.7 *Suppose that $G = H_1 \star_A H_2$ is an amalgamated free product with $H_1 \neq A \neq H_2$. Suppose further that A is malnormal in both H_1 and H_2 and that both H_1 and H_2 are 3-free. Then G satisfies LZ, LA, LPZ, LPA, LCZ, and LCA.*

7.2 The Lyndon Properties and HNN Constructions

The situation for HNN groups is much more complicated.

Theorem 7.8 ([31]) *Suppose that G is an HNN extension of the base B so that G has the form*

$$G = \langle B, t; rel(B), t^{-1}K_1t = K_2\rangle.$$

Suppose further that K_1 and K_2 are both malnormal in B and that B does not contain an element of order 2. Suppose further that B satisfies the basic Lyndon Property LZ. Then if $x^2y^2z^2 = 1$ in G and $\mathcal{U} = \{x, y, z\}$ is regular then $\langle x, y, z\rangle$ is cyclic.

Let $G = \langle B, t; rel(B), t^{-1}K_1t = K_2\rangle$ be an HNN group with base group B, stable letter t and associated subgroups K_1, K_2. An ordered set $\mathcal{U} = \{u_1, \ldots, u_n\} \subset G$ is *regular* if there is no Nielsen transformation from $\mathcal{U}$ to a system $\mathcal{U}' = \{u'_1, \ldots, u'_n\}$ in which one of the elements is conjugate to an element of K_1 or K_2. However the theorem does not necessarily hold for HNN group when $\mathcal{U} = \{x, y, z\}$ is not regular.

Let $G = \langle x, y, z; x^2y^2z^2 = 1\rangle$ be the nonorientable surface group of genus $g = 3$. In G the equation $x^2y^2z^2 = 1$ holds trivially and $\langle x, y, z\rangle$ is nonabelian and hence noncyclic. G can be written as an amalgamated free product $G = H_1 \star_A H_2$ with

$$H_1 = \langle x, y; \rangle, H_2 = \langle z; \rangle \text{ and } A = \langle x^2y^2\rangle = \langle z^{-2}\rangle.$$

However here G does not contradict any of our results since A is not malnormal in H_2. This again shows that malnormality is essential in the amalgamated free product case.

On the other hand using straightforward Tietze transformations ($t = z$, $v = zxz^{-1}$, $u = yz$) the nonorientable surface group $G = \langle x, y, z : x^2y^2z^2 = 1\rangle$ can also be written as an HNN group

$$G = \langle H, t; t^{-1}ut = v^2u^{-1}\rangle$$

with $H = \langle u, v; \rangle$. The element u is not conjugate in the base H to v^2u^{-1} and both associated subgroups are malnormal in the base. However the system $\{x, y, z\}$ is not regular (see the definition above) showing that in the HNN case regularity is essential.

7.3 The Lyndon Properties in Certain One-Relator Groups

A large subclass of the class of one-relator groups satisfies the Lyndon properties since cyclically pinched one-relator groups do. The following due to Fine, Rosenberger and Rosenberger extends the class even further to some one-relator groups with torsion.

Theorem 7.9 ([32]) *Let G be the one-relator group*

$$G = \langle a, b, c, \ldots; R^m\rangle$$

with $m \geq 3$ and m odd and R a cyclically reduced word, not a proper power in the free group on $a, b, c, \ldots$. Let $w(x_1, x_2, x_3)$ be a regular quadratic word in the free group F on x_1, x_2, x_3 and let $\phi : F \to G$ be a homomorphism from F into G with $\phi(x_i) = u_i$ for $i = 1, 2, 3$. If $w(u_1, u_2, u_3) = 1$ in G then the subgroup $\langle u_1, u_2, u_3\rangle$ is cyclic.

In particular G satisfies the Lyndon properties LZ and LCZ.

Recall that the quadratic word $w(x_1, x_3, x_3)$ is regular if there is no automorphism $\alpha : F \to F$ such that $w' = \alpha(w)$, as a word in x_1, x_2, x_3, contains fewer of the generators x_1, x_2, x_3 than w itself, that is w is of maximal rank.

7.4 The Lyndon Properties and Tree-free Groups

N. Brady, L. Ciobanu, A. Martino and S. O'Rourke [12] considered the Lyndon properties in groups acting freely on Λ-trees. Using the concept of translation length in such groups they were able to prove the following.

Theorem 7.10 ([12]) *Let Λ be an ordered abelian group and let G act freely on a Λ-tree. Then if $x^p y^q = z^r$ with $p, q, r \geq 4$ it follows that x, y and z commute. That is such groups satisfy LCA.*

A study was initiated by Ciobanu, Fine and Rosenberger [16] to consider the smaller cases. It was shown that for small cases you can have tree-free groups that do not satisfy the Lyndon properties.

Theorem 7.11 *Let F be a finitely generated non-cyclic free group, and let U and V be elements in F which are not proper powers. Let $G = \langle F, t; tUt^{-1} = V\rangle$ and $r \geq 2$ be a given integer. Then for particular choices of U and V there exist non-commuting elements $a, b, c \in G$ such that $a^2b^2c^r = 1$.*

Corollary 7.12 *There exist Λ-free groups in which $a^2b^2c^r = 1$ holds for non-commuting $a, b, c \in G$, and $r \geq 2$.*

Theorem 7.13 *Let F be a finitely generated free group, U and V elements in F that are not proper powers and U is not conugate to V^{-1}, and $G = \langle F, t; tUt^{-1} = V\rangle$. Then if for $a, b, c \in G$ and $p \geq 2$, $q \geq 3$, $r \geq 3$ the equality $a^pb^qc^r = 1$ holds, the elements a, b, c must commute.*

8 The Class of $B\mathcal{X}$-Groups

Let $\mathcal{X}$ be a class of groups. Then a group G is *residually* $\mathcal{X}$ if given any nontrivial element $g \in G$ there is a homomorphism $\phi : G \to H$ where H is a group in $\mathcal{X}$ such that $\phi(g) \neq 1$. A group G is *fully residually* $\mathcal{X}$ if given finitely many nontrivial elements $g_1, \ldots, g_n$ in G there is a homomorphism $\phi : G \to H$, where H is a group in $\mathcal{X}$, such that $\phi(g_i) \neq 1$ for all $i = 1, \ldots, n$. Fully residually free groups have played a crucial role in the study of equations and first-order formulas over free groups. Recall that *universal theory* of a group G consists of all universal sentences true in G. All nonabelian free groups share the same universal theory and a group G is called *universally free* if it shares the same universal theory as the class of nonabelian free groups. We recall two additonal concepts that are needed. A group G is *commutative transitive* or *CT* if commutativity is transitive on the set of nontrivial elements of G. That is if $[x, y] = 1$ and $[y, z] = 1$ for nontrivial elements $x, y, z \in G$ then $[x, z] = 1$. A group G is *CSA* if maximal abelian subgroups are malnormal. CSA implies commutative transitivity but there exist CT groups that are not CSA. For example it can be shown that a noncyclic one-relator group G with torsion is CT but not CSA if G has elements of order 2 (see [28]). Another example of a CT group that is not CSA is the infinite dihedral group $G = \langle a, b; a^2 = b^2 = 1\rangle$. It is straightforward that free products of abelian groups are CT and hence G is CT. On the other hand the commutator subgroup G' is the cyclic subgroup of G generated by ab. A nonabelian

CSA group cannot have a nontrivial abelian normal subgroup and hence G is not CSA.

Remeslennikov [73] and independently Gaglione and Spellman [42] proved the following remarkable theorem which became one of the cornerstones in the proof of the Tarski problems (see [51] and [78]).

Theorem 8.1 *Suppose G is nonabelian and residually free. Then the following are equivalent:*

1. *G is fully residually free,*
2. *G is commutative transitive,*
3. *G is universally free.*

Therefore the class of nonabelian fully residually free groups coincides with the class of residually free universally free groups. The equivalence of (1) and (2) in the theorem above was proved originally by Benjamin Baumslag [2], where he introduced the concept of fully residually free. Any residually free elementary free group being universally free must satisfy this theorem and hence be fully residually free.

In [18] classes of groups $\mathcal{X}$ were studied for which being fully residually $\mathcal{X}$ is equivalent to being residually $\mathcal{X}$ and commutative transitive, thus extending Baumslag's result.

Definition 8.2 A class of groups $\mathcal{X}$ satisfies $B\mathcal{X}$ if a group G is fully residually $\mathcal{X}$ if and only if G is residually $\mathcal{X}$ and CT.

With this definition B. Baumslag's original theorem says that the class of free groups $\mathcal{F}$ satisfies $B\mathcal{F}$. In [18] it was shown that classes of $B\mathcal{X}$ groups are fairly extensive.

Theorem 8.3 ([18]) *Let $\mathcal{X}$ be a class of groups such that each nonabelian $H \in \mathcal{X}$ is CSA. Let G be a nonabelian and residually $\mathcal{X}$ group. Then the following are equivalent:*

1. *G is fully residually $\mathcal{X}$.*
2. *G is CSA.*
3. *G is CT.*

Therefore the class $\mathcal{X}$ has the property $B\mathcal{X}$.

Hence a class of groups $\mathcal{X}$ satisfies $B\mathcal{X}$ if each nonabelian $H \in \mathcal{X}$ is CSA. Examples of $B\mathcal{X}$ classes abound. In particular we list the following.

Theorem 8.4 ([18]) *Each of the following classes satisfies $B\mathcal{X}$:*

1. *The class of nonabelian free groups.*
2. *The class of noncyclic torsion-free hyperbolic groups (see [33]).*
3. *The class of noncyclic one-relator groups with only odd torsion (see [33]).*
4. *The class of cocompact Fuchsian groups with only odd torsion.*
5. *The class of noncyclic groups acting freely on Λ-trees where Λ is an ordered abelian group (see [46]).*

6. *The class of noncylic free products of cyclics with only odd torsion.*
7. *The class of noncyclic torsion-free RG-groups (see [28] and [1]).*
8. *The class of conjugacy pinched one-relator groups of the following form*

$$G = \langle F, t; tut^{-1} = v \rangle$$

where F is a free group of rank $n \geq 1$ and u, v are nontrivial elements of F that are not proper powers in F and for which $\langle u \rangle \cap x\langle v \rangle x^{-1} = \{1\}$ for all $x \in F$.

The theorem follows from the fact that each of these classes has the property that each nonabelian group in them is CSA.

Since CSA always implies CT we have the following corollary.

Corollary 8.5 *Let $\mathcal{X}$ be a class of CSA groups. Then if G is a nonabelian residually $\mathcal{X}$ group then CT is equivalent to CSA.*

Commutative transitivity (CT) has been shown to be equivalent to many other properties (see [1]) under the additional condition that abelian subgroups are locally cyclic (ALC). A group G is *power commutative* if $[x, y^n] = 1$ implies that $[x, y] = 1$ whenever $y^n \neq 1$. Two elements $a, b \in G$ are in *power relation* to each other if there exists an $x \in G \setminus \{1\}$ with $a = x^n$, $b = x^m$ for some $n, m \in \mathbb{Z}$. G is *power transitive* or PT if this relation is transitive on nontrivial elements. Hence we get the corollary.

Corollary 8.6 *Let $\mathcal{X}$ be a class of groups such that each nonabelian $H \in \mathcal{X}$ is CSA. Let $\mathcal{Y}$ be the subclass of $\mathcal{X}$ consisting of those groups in $\mathcal{X}$ which are ALC. Let G be a nonabelian residually $\mathcal{Y}$ group which is ALC and has trivial center. Then the following are equivalent.*

1. *G is fully residually $\mathcal{Y}$.*
2. *G is CSA.*
3. *G is CT.*
4. *G is power commutative*
5. *G is power transitive.*

This follows directly from the equivalences given in [1].

8.1 Big Powers Groups and Univeral Freeness

The results of the previous section showed the equivalence of fully residually-$\mathcal{X}$ and commutative transitivity for any class $\mathcal{X}$ of CSA groups. To prove an equivalence with universally-$\mathcal{X}$ groups in [18] the *big powers condition* was used. This was introduced originally by G. Baumslag in [3].

Definition 8.7 Let G be a group and $u = (u_1, \ldots, u_k)$ be a sequence of nontrivial elements of G. Then

1. u is *generic* if neighboring elements in u do not commute, that is $[u_i, u_{i+1}] \neq 1$ for every $i \in \{1, \ldots, k-1\}$;
2. u is *independent* if there exists an $n = n(u) \in \mathbb{N}$ such that for any $\alpha_1, \ldots, \alpha_k \geq n$ we have $u_1^{\alpha_1} \cdots u_k^{\alpha_k} \neq 1$;

3. a group satisfies the *big powers condition* or *BP* if every generic sequence in G is independent. We call such groups BP-groups.

G. Baumslag proved that free groups are BP-groups [3] while Olshansky [70] showed that torsion-free hyperbolic groups are BP-groups. For BP groups the following results are known.

Lemma 8.8 ([60]) *A subgroup of a BP-group is itself a BP-group.*

Lemma 8.9 ([70]) *Every torsion-free hyperbolic group is a BP-group.*

A stronger version of this lemma for relatively hyperbolic groups is given in [60].

Lemma 8.10 *A free product of CSA BP-groups is also a BP-group.*

Lemma 8.11 *Let $G = F_1 \underset{U=V}{\star} F_2$ where F_1, F_2 are finitely generated free groups and U, V are nontrivial elements of F_1, F_2 respectively with not both proper powers. Then G is a CSA and BP-group.*

If G and H are groups then we say that G is an H-group or H-domain if G contains an isomorphic copy of H. Being an H-group is crucial for considering the next equivalence. We consider a class of groups $\mathcal{Z}$ in which each finitely generated nonabelian group H in $\mathcal{Z}$ is CSA and BP. Reinterpreting a result in [6] and [5] (see also [60]) and following the same proof the next theorem proved in [18].

Theorem 8.12 ([18]) *Let $\mathcal{Z}$ be a class of finitely presented groups such that each nonabelian $H \in \mathcal{Z}$ is CSA and BP. Let $H \in \mathcal{Z}$ and G a finitely presented nonabelian H-group. Then the following are equivalent:*

1. *G is fully residually H,*
2. *G is universally equivalent to H.*

Note that being an H-group was not necessary in the case of the class of free groups since a nonabelian free group and a nonabelian fully residually free group contain copies of free groups of all countable ranks. It was noticed by D. Spellman that while the BP and CSA conditions were necessary in [6] for embedding a given hyperbolic group into its Lyndon completion and then a modification of this proof with the given conditions was used in the proof of Theorem 8.12 in [18] they were not really necessary for universal equivalence.

In alternative language if a group G is fully residually H then we say that H *discriminates* G. Further if we append to the basic language L_0 appropriate for group theory constants from the group H then we say that G is *H-universally equivalent* to G if G and H have the same universal theory in this extended language. We actually have the following theorem which says that if G is an H-group then H discriminating G is equivalent to H being universally equivalent to G. Further if H is finitely generated then G is H-universally equivalent to G if and only if there is a discriminating family of retractions from G onto H.

Theorem 8.13 *Let G be a finitely presented H-group. Then the following are equivalent:*

1. *G is fully residually H, that is H disciminates G.*
2. *G is universally equivalent to H.*

Further if H is finitely generated then G is H-universally equivalent to H if and only if there is a discriminating family of retractions $G \to H$.

Proof Suppose first that G is an H-group and that G is fully residually H. We show that G is universally equivalent to H. To show this we prove that any universal sentence true in H is also true in G. Hence the universal theory of H is contained in the universal theory of G. However H is a subgroup of G so the universal theory of G is contained in the universal theory of H. The equivalence then follows.

To show that that every universal sentence true in H is also true in G we show that every existential sentence true in G must also be true in H. Suppose the following existential sentence, which we label $(*)$, and whose matrix is written in disjunctive normal form, is true in G:

$$(*) \qquad \exists x_1, \ldots, x_n \left(\bigvee_i \left(\bigwedge_j (u_{ij}(x_1, \ldots, x_n) = 1) \right) \wedge \left(\bigwedge_k (w_{ik}(x_1, \ldots, x_n) \neq 1) \right) \right)$$

The sentence $(*)$ is equivalent to the sentence below which we label $(**)$:

$$(**) \qquad \bigvee_i \exists x_1, \ldots, x_n \left(\left(\bigwedge_j (u_{ij}(x_1, \ldots, x_n) = 1) \right) \wedge \left(\bigwedge_k (w_{ik}(x_1, \ldots, x_n) \neq 1) \right) \right)$$

Since $(**)$ holds in G it follows that at least one disjunct must be true in G. Suppose that

$$(***) \qquad \exists x_1, \ldots, x_n \left(\left(\bigwedge_{j=1}^{r} (u_{i_0 j}(x_1, \ldots, x_n) = 1) \right) \wedge \left(\bigwedge_{k=1}^{q} (w_{i_0,k}(x_1, \ldots, x_n) \neq 1) \right) \right)$$

holds in G. Let $(g_1, \ldots, g_n) \in G^n$ be an n-tuple such that

$$u_{i_0 1}(g_1, \ldots, g_n) = \cdots = u_{i_0 r}(g_1, \ldots, g_n) = 1 \\ \wedge\, w_{i_0 1}(g_1, \ldots, g_n) \neq 1 \wedge \cdots \wedge w_{i_0 q}(g_1, \ldots, g_n) \neq 1.$$

Since G is fully residually H there is a map $\phi : G \to H$ such that

$$\phi(w_{i_0 k}(g_1, \ldots, g_n)) = w_{i_0 k}(\phi(g_1), \ldots, \phi(g_n)) \neq 1 \quad \text{for all } k = 1, \ldots, q.$$

Further clearly

$$u_{i_0 j}(\phi(g_1), \ldots, \phi(g_n)) = \phi(u_{i_0 j}(g_1, \ldots, g_n)) = \phi(1) = 1 \quad \text{for all } j = 1, \ldots, r.$$

Therefore $(***)$ is true in H and working backwards it follows that $(*)$ holds in H.

Therefore every existential sentence true in G is also true in H and hence they are universally equivalent. Further if the discrimination is by retractions onto H then the result holds in the extended language where the elements of H are appended as constants.

Now we show that G being fully residually H is necessary for universal equivalence. Let us assume than that G and H are universally equivalent and we wish to show that H discriminates G.

Suppose that the finitely presented H-group G is universally equivalent to H and let $\langle a_1, \ldots, a_m; R_1(a_1, \ldots, a_m) = \cdots = R_n(a_1, \ldots, a_m) = 1\rangle$ be a finite presentation for G. Let $g_j = w_j(a_1, \ldots, a_m),\ \ j = 1, \ldots, k$ be nontrivial elements in G. Then the following existential sentence that we denote by $(\star)$ is true in G.

$$(\star)\ \ \exists x_1, \ldots, x_n \left(\left(\bigwedge_{i=1}^{n} (R_i(x_1, \ldots, x_n) = 1) \right) \wedge \left(\bigwedge_{j=1}^{k} (w_j(x_1, \ldots, x_n) \neq 1) \right) \right)$$

Therefore $(\star)$ is also true in H. Let $(h_1, \ldots, h_m) \in H^m$ be such that

$$R_1(h_1, \ldots, h_m) = \cdots = R_n(h_1, \ldots, h_m) = 1$$

and

$$w_1(h_1, \ldots, h_m) \neq 1 \wedge \cdots \wedge w_k(h_1, \ldots, h_m) \neq 1.$$

Then since the relations are preserved the map $a_\nu \mapsto h_\nu,\ \ \nu = 1, \ldots, m$, extends to a homomorphism $\phi : G \to H$. Then

$$\phi(g_j) = \phi(w_j(a_1, \ldots, a_m)) = w_j(\phi(a_1), \ldots, \phi(a_m)) = w_j(h_1, \ldots, h_m) \neq 1$$

for all $j = 1, \ldots, k$. Therefore G is fully residually H and the first set of equivalences are completed.

Now we consider H to be finitely generated and we want to consider the extended language where we adjoin the elements of H as constants. From the comments after the first part of the proof we know that if there is a family of discriminating retractions then G is H-universally equivalent to G. Now we assume that G is H-universally equivalent to G and we show that that there is a discriminating family of retractions from G onto H.

Let $a_1, \ldots, a_p$ be a set of generators for H and these extend to a finite set a_1, …, a_p, b_1, …, b_q of generators of G. Let

$$\langle a_1, \ldots, a_p, b_1, \ldots, b_q; R_1(a_1, \ldots, a_p, b_1, \ldots, b_q) = \cdots = R_n(a_1, \ldots, a_p, b_1, \ldots, b_q) = 1\rangle$$

be a finite presentation for G.

Suppose $g_j = w_j(a_1, \ldots, a_p, b_1, \ldots, b_q) \neq 1$ for $j = 1, \ldots, k$ are nontirvial elements of G. The the following existential sentence is true in G:

$$\exists x_1, \ldots, x_q \left(\left(\bigwedge_{i=1}^{n} (R_i(a_1, \ldots, a_p, x_1, \ldots, x_q) = 1) \right) \wedge \left(\bigwedge_{j+1}^{k} (w_j(a_1, \ldots, a_p, x_1, \ldots, x_q) \neq 1) \right) \right).$$

Since G and H are assumed to H-universally equivalent this must also hold in H.

Let $(h_1, \ldots, h_q) \in H^q$ be such that

$$R_1(a_1, \ldots, a_p, h_1, \ldots, h_q) = \cdots = R_n(a_1, \ldots, a_p, h_1, \ldots, h_q) = 1$$

and

$$w_1(a_1, \ldots, a_p, h_1, \ldots, h_q) \neq 1 \wedge \cdots \wedge w_k(a_1, \ldots, a_p, h_1, \ldots, h_q) \neq 1.$$

Then since the relations are preserved the maps $a_\nu \mapsto a_\nu$, $\nu = 1, \ldots, p$, $b_\mu \mapsto h_\mu$, $\mu = 1, \ldots, q$, extend to a retraction $\phi : G \to H$. Furthermore, for all $j = 1, \ldots, k$, we have

$$\begin{aligned}\phi(g_j) = \phi(w_j(a_1, \ldots, a_p, b_1, \ldots, b_q)) &= w_j(\phi(a_1), \ldots, \phi(a_p), \phi(b_1), \ldots, \phi(b_q)) \\ &= w_j(a_1, \ldots, a_p, h_1, \ldots, h_q) \neq 1.\end{aligned}$$

Therefore G is discriminated by retractions completing the proof. □

Summarizing our results:

Theorem 8.14 *Let $\mathcal{Z}$ be a class of finitely presented groups such that each nonabelian $H \in \mathcal{Z}$ is CSA. Let G be a finitely presented nonabelian residually $\mathcal{Z}$ group. Then the following are equivalent:*

1. *G is fully residually $\mathcal{Z}$,*
2. *G is CSA,*
3. *G is CT.*

If in addition G is an H-group for some $H \in \mathcal{Z}$ then the following are equivalent.

(a) *G is fully residually H,*

(b) *G is universally equivalent to H.*

References

[1] P. Ackermann, V. Große Rebel & G. Rosenberger, On power- and commutation transitive, power commutative and restricted Gromov groups, *Group theory, statistics, and cryptography*, 1–4, Contemp. Math. 360, Amer. Math. Soc., 2004.

[2] B. Baumslag, Residually free groups, *Proc. London Math. Soc. (3)* **17** (1967), 635–645.

[3] G. Baumslag, On generalised free products, *Math. Z.* **78** (1962), 423–438.

[4] G. Baumslag, B. Fine, C. Miller & D. Troeger, Virtual properties of cyclically pinched one-relator groups, *Internat. J. Algebra Comput.* 19 (2009), 1–15.

[5] G. Baumslag, A. Myasnikov & V. Remeslennikov, Algebraic geometry over groups I. Algebraic sets and ideal theory, *J. Algebra* **219** (1999), 16–79.

[6] G. Baumslag, A. Myasnikov & V. Remeslennikov, Discriminating completions of hyperbolic groups, *Geom. Dedicata* **92** (2002), 115–143.

[7] G. Baumslag & P. Shalen, Amalgamated products and finitely presented groups, *Comment. Math. Helv.* **65** (1990), 243–254.

[8] J.L. Bell & A.B. Slomson, *Models and Ultraproducts: An Introduction*, Second Revised Printing, North-Holland, Amsterdam, 1971.

[9] M. Bestvina & M. Feighn, A combination theorem for negatively curved groups, *J. Diff. Geom.* **35** (1992), 85–101.

[10] O. Bogopolski, A surface groups analogue of a theorem of Magnus, *Geometric methods in group theory*, 59–69, Contemp. Math. 372, Amer. Math. Soc., 2005.

[11] O. Bogopolski & K. Sviridov, A Magnus theorem for some one-relator groups, *The Zieschang Gedenkschrift*, 63–73, Geometry & Topology Monographs 14, 2008.
[12] N. Brady, L. Ciobanu, A. Martino & S. O'Rourke, The equation $x^p y^q = z^r$ and groups that act freely on Λ-trees, *Trans. Amer. Math. Soc.* **361** (2009), 223–236
[13] A.M. Brunner, R.G. Burns & D. Solitar, The subgroup separability of free products of two free groups with cyclic amalgamation, *Contributions to group theory*, 90–115, Contemp. Math. 33, Amer. Math. Soc., 1984.
[14] I. Bumagin, O. Kharlampovich, A. Myasnikov, The isomorphism problem for finitely generated fully residually free groups, *J. Pure Appl. Algebra* **208** (2007), 961–977.
[15] C.C. Chang & H.J. Keisler, *Model Theory, Second Edition*, North-Holland, Amsterdam, 1977.
[16] L. Ciobanu, B. Fine & G. Rosenberger, On Lyndon's equation in some Λ-free groups and HNN extensions, *J. Group Theory* **14** (2011), 333–339.
[17] L. Ciobanu, B. Fine & G. Rosenberger, The surface group conjecture: cyclically pinched and conjugacy pinched one-relator groups, *Results Math.* **64** (2013), 175–184.
[18] L. Ciobanu, B. Fine & G. Rosenberger, Classes of groups generalizing a theorem of Benjamin Baumslag, *Comm. Algebra* (to appear).
[19] D. Collins & H. Zieschang, Combinatorial group theory; applications to geometry, *Algebra VII*, Springer, 1993.
[20] L. Comerford, C. Edmunds, G. Rosenberger, Commutators as powers in free products, *Proc. Amer. Math. Soc.* **122** (1994), 47–52.
[21] P. Csorgo, B. Fine & G. Rosenberger, On certain equations in free groups, *Acta Sci. Math.* **68** (2002) 895–905.
[22] J.L. Dyer, Separating conjugates in amalgamated free products and HNN extensions, *J. Austral. Math. Soc. Ser. A* **29** (1980), 35–51.
[23] B. Fine, A. Gaglione, S. Lipschutz & D. Spellman, Turner's Theorem is not first-order, in press.
[24] B. Fine, A. Gaglione, A. Myasnikov, G. Rosenberger & D. Spellman, A classification of fully residually free groups of rank three or less, *J. Algebra* **200** (1998), 571–605.
[25] B. Fine, A. Gaglione, G. Rosenberger & D. Spellman, n-Free groups and questions about universally free groups, *Groups '93 Galway/St Andrews*, 191–204, London Math. Soc. Lecture Notes Ser. 211, 1996.
[26] B. Fine, O. Kharlampovich, A. Myasnikov, V. Remeslennikov & G. Rosenberger, On the surface group conjecture, *Sci. Ser. A Math. Sci.* **15** (2007), 1–15.
[27] B. Fine, M. Kreuzer & G. Rosenberger, Real representations of pinched one-relator groups, to appear.
[28] B. Fine, A. Myasnikov, V. große Rebel & G. Rosenberger, A classification of conjugately separated abelian, commutative transitive and restricted Gromov one-relator groups, *Results Math.* **50** (2007), 183–193.
[29] B. Fine, F. Roehl & G. Rosenberger, Two generator subgroups of certain HNN groups, *Combinatorial group theory*, 19–23, Contemp. Math. 109, 1988.
[30] B. Fine, F. Roehl & G. Rosenberger, On HNN groups whose three-generator subgroups are free, *Infinite Groups and Group Rings*, 13–37, World Scientific Press, 1993.
[31] B. Fine, A. Rosenberger, G. Rosenberger, Quadratic properties in group amalgams, *J. Group Theory* **14** (2011), 657–671.
[32] B. Fine, A. Rosenberger, G. Rosenberger, A note on Lyndon properties in one relator groups, *Results Math.* **59** (2011), 239–250.
[33] B. Fine & G. Rosenberger, *Algebraic Generalizations of Discrete Groups*, Marcel-Dekker, 1999.
[34] B. Fine & G. Rosenberger, On restricted Gromov groups, *Comm. Algebra* **20** (1992), 2171–2182.
[35] B. Fine & G. Rosenberger, Surface groups within Baumslag doubles, *Proc. Edinb. Math.*

Soc. **54** (2011), 91–97.
[36] B. Fine & G. Rosenberger, A note on faithful representations of hyperbolic limit groups, *Groups Complex. Cryptol.* **3** (2011), 349–355.
[37] B. Fine & G. Rosenberger, Faithful representations of limit groups 2, *Groups Complex. Cryptol.*, to appear.
[38] B. Fine, G. Rosenberger, D. Spellman & M. Stille, Test elements, generic elements and almost primitivity in free groups, *Pacific J. Math.* **190** (1999), 277–297.
[39] B. Fine, G. Rosenberger & M. Stille, Nielsen transformations and applications: a survey, *Groups Korea '94*, 69–105, DeGruyter, 1995.
[40] B. Fine, G. Rosenberger & M. Stille, Conjugacy pinched and cyclically pinched one-relator groups, *Rev. Mat. Univ. Complut. Madrid* **10** (1997), 207–227.
[41] A. Gaglione, S. Lipschutz & D. Spellman, Almost locally free groups and a theorem of Magnus, *Groups, Complex. Cryptol.* **1** (2009), 181–198.
[42] A. Gaglione & D. Spellman, Even more model theory of free groups, *Infinite Groups and Group Rings*, 37–40, World Scientific Press, 1993.
[43] A. Gaglione & D. Spellman, Almost locally free groups and the genus quesions, *Comm. Algebra* **261** (1998), 2821–2836.
[44] D. Gildenhuys, O. Kharlampovich & A. Myasnikov, CSA groups and separated free constructions, *Bull. Austral. Math. Soc.* **52** (1995), 63–84.
[45] C. Gordon & H. Wilton, Surface subgroups of doubles of free groups, to appear.
[46] J. Howie, Some results on one-relator surface groups, *Bol. Soc. Mat. Mexicana* **10** (2004), 255–262.
[47] S.V. Ivanov, On certain elements of free groups, *J. Algebra* **204** (1998), 394–405.
[48] A. Juhasz & G. Rosenberger, On the combinatorial curvature of groups of F-type and other one-relator products of cyclics, *The mathematical legacy of Wilhelm Magnus: groups, geometry and special functions*, 373–384, Contemp. Math. 169, 1994.
[49] I. Kapovich, P. Schupp & V. Shpilrain, Generic properties of Whitehead's Algorithm and isomorphism rigidity of random one-relator groups, *Pacific J. Math.* **223** (2006), 113–140.
[50] K. Kearnes, Private e-mail communication.
[51] O. Kharlamapovich & A. Myasnikov, Irreducible affine varieties over a free group: I. Irreducibility of quadratic equations and Nullstellensatz, *J. Algebra* **200** (1998), 472–516.
[52] O. Kharlamapovich & A. Myasnikov, Irreducible affine varieties over a free group: II. Systems in triangular quasi-quadratic form and a description of residually free groups, *J. Algebra* **200** (1998), 517–569.
[53] O. Kharlamapovich & A. Myasnikov, Hyperbolic groups and free constructions, *Trans. Amer. Math. Soc.* **350** (1998), 571–613.
[54] O. Kharlampovich & A.G. Myasnikov, Implicit function theorem over free groups and genus problem, *Knots, braids, and mapping class groups*, 77–83, AMS/IP Stud. Adv. Math. 24, 2001.
[55] O. Kharlamapovich & A. Myasnikov, The Implicit Function Theorem over free groups, *J. Algebra* **290** (2005), 1–203.
[56] O. Kharlamapovich & A. Myasnikov, Effective JSJ Decompositions, *Groups, languages, algorithms*, 87–212, Contemp. Math. 378, 2005.
[57] O. Kharlamapovich & A. Myasnikov, Algebraic geometry over free groups: lifting solutions into generic Points, *Groups, languages, algorithms*, 213–318, Contemp. Math. 378, 2005.
[58] O. Kharlamapovich & A. Myasnikov, Elementary theory of free nonabelian groups, *J. Algebra* **302** (2006), 451–552.
[59] O. Kharlamapovich & A.Myasnikov, Algebraic geometry over free groups, to appear.
[60] O. Kharlamapovich, A. Myasnikov, V. Remeslennikov & D. Serbin, Subgroups of fully residually free groups: algorithmic problems, *Trans. Amer. Math. Soc.* **364** (2012), 2847–

2882.
[61] O. Kharlamapovich, A. Myasnikov, V. Remeslennikov & D. Serbin, Groups with regular length functions in $\mathbb{Z}^n$, Contemp. Math. 360.
[62] S. Kim & H. Wilton, Surface subgroups of doubles of free groups, to appear.
[63] S. Kim & S. Oum, Hyperbolic surface subgroups of one-ended doubles of free groups, preprint.
[64] D. Lee, On certain C-test words for free groups, *J. Algebra* **247** (2002), 509–540.
[65] S. Lipschutz, The conjugacy problem and cyclic amalgamation, *Bull. Amer. Math. Soc.* **81** (1975), 114–116.
[66] R.C. Lyndon, The equation $a^2b^2 = c^2$ in free groups, *Michigan Math. J.* **6** (1959), 155–164.
[67] R.C. Lyndon & P.E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, 1977.
[68] W. Magnus, A. Karrass & D. Solitar, *Combinatorial group theory: Presentations of groups in terms of generators and relations*, Wiley, 1966.
[69] D.I. Moldavanskii, Certain subgroups of groups with one defining relation, *Sibirsk. Mat. Ž.* **8** (1967), 1370–1384.
[70] A. Olshanskii, On relatively hyperbolic and G-subgroups of hyperbolic groups, *Internat. J. Algebra Comput.* **3** (1993), 365–409.
[71] J.C. O'Neill & E.C Turner, Test elements and the retract theorem in hyperbolic groups, *New York J. Math* **6** (2000), 107-117.
[72] N. Peczynski & W. Reiwer, On cancellations in HNN groups, *Math. Z.* **158** (1978), 79–86.
[73] V.N. Remeslennikov, $\exists$-free groups, *Siberian Math. J.* **30** (1989), 998–1001.
[74] E. Rips & Z. Sela, Cyclic splittings of finitely presented groups and the canonical JSJ decomposition, *Ann. of Math. (2)* **146** (1997), 53–109.
[75] G. Rosenberger, On one-relaor groups that are free products of two free groups with cyclic amalgamation, *Groups St Andrews 1981*, 328–344, CUP, 1982.
[76] G. Rosenberger, The isomorphism problem for cyclically pinched one-relator groups, *J. Pure Appl. Algebra* **95** (1994), 75–86.
[77] G. Rosenberger, Zum Isomorphieproblem für Gruppen mit einer definierenden Relation, *Illinois J. Math.* **20** (1976), 614–621.
[78] Z. Sela, The isomorphism problem for hyperbolic groups. I, *Ann. of Math. (2)* **141** (1995), 217–283.
[79] Z. Sela, Diophantine geometry over groups. I: Makanin-Razborov diagrams, *Publ. Math. Inst. Hautes Études Sci.* **93** (2001), 31–105.
[80] Z. Sela, Diophantine geometry over groups. II: Completions, closures and formal solutions, *Israel J. Math.* **134** (2003), 173–254.
[81] Z. Sela, Diophantine geometry over groups. IV: An iterative procedure for validation of a sentence, *Israel J. Math.* **143** (2004), 1–71.
[82] Z. Sela, Diophantine geometry over groups. III: Rigid and solid solutions, *Israel J. Math.* **147**, 2005, 1–73.
[83] Z. Sela, Diophantine geometry over groups. V: Quantifier elimination, *Israel J. Math.* **150** (2005), 1–197.
[84] V. Shpilrain, Recognizing automorphisms of the free groups, *Arch. Math.* **62** (1994), 385–392.
[85] W. Szmielew, Elementary properties of abelian groups, *Fund. Math.* **41** (1955), 203–271.
[86] E.C. Turner, Test words for automorphisms of free groups, *Bull. London Math. Soc.* **28** (1996), 255–263.
[87] B.A.F. Wehfritz, Generalized free products of linear groups, *Proc. London Math. Soc.* **27** (1973), 402–424.
[88] H. Zieschang, Über Automorphismen ebener discontinuerlicher Gruppen, *Math. Ann.* **166** (1966), 148–167.

THE GROUPS OF PROJECTIVITIES IN FINITE PLANES

THEO GRUNDHÖFER

Institut für Mathematik, Universität Würzburg, Emil-Fischer-Str. 30, 97074 Würzburg, Germany
Email: grundh@mathematik.uni-wuerzburg.de

Abstract

In every finite non-desarguesian projective plane, the group of all projectivities of a line onto itself is the alternating group or the symmetric group. Similar results hold for affine planes.

1 Introduction

The group of projectivities of a geometry reflects the complexity of the geometry: one obtains a rather large group if the geometry is not a classical geometry; this is in marked contrast to automorphism groups. See [36] for more information on the role of projectivities in geometry.

In finite projective or affine planes, there are only very few possibilities for the groups of projectivities, as the results 3.2, 4.1 and 4.2 below show.

2 Sharply 2-transitive sets of permutations

A set S of permutations of $\Omega = \{1, \dots, n\}$ is said to be *sharply 2-transitive* if the following holds: given $i, j, i', j' \in \Omega$ with $i \neq j$ and $i' \neq j'$, there exists exactly one element $s \in S$ with $s(i) = i'$ and $s(j) = j'$.

The following result is contained in Müller–Nagy [32, Lemma 1 and Theorem 3]. This elementary lemma is useful in the proofs for Theorems 3.2 and 4.3.

Lemma 2.1 *Let S be a sharply 2-transitive subset of the symmetric group S_n. Then $|S \setminus A_n|$ is odd if, and only if, $n \equiv 2, 3 \pmod 4$.*

In particular, for $n \equiv 2, 3 \pmod 4$ the alternating group A_n does not contain any sharply 2-transitive subset.

Proof There exist precisely $\binom{n}{2}^2$ triples (s, i, j) with $s \in S$, $1 \leq i < j \leq n$ and $s(i) > s(j)$, because such a triple is uniquely determined by the two sets $\{i, j\}$ and $\{s(i), s(j)\}$. A second counting of these triples yields

$$\binom{n}{2}^2 = \sum_{s \in S} |\{(i,j) \mid 1 \leq i < j \leq n,\ s(i) > s(j)\}| \ \equiv\ |S \setminus A_n| \pmod 2,$$

and $\binom{n}{2}$ is odd precisely if $n \equiv 2, 3 \pmod 4$. □

Let S be a sharply 2-transitive set of permutations of a finite set Ω with $|\Omega| > 1$. According to Witt [39, p. 267], compare [5, 3.2.4(b), 3.2.6] or [34, 12.1.1], we obtain

an affine plane as follows: the points are the elements of Ω^2, and the lines are the elements $s \in S$ (note that $s \subseteq \Omega^2$ by the set-theoretic definition of mappings) together with the horizontal lines $\Omega \times \{i\}$ and the vertical lines $\{i\} \times \Omega$ where $i \in \Omega$. Conversely, every affine plane arises in this fashion. For example, the classical affine plane over a (skew) field F is obtained from the sharply 2-transitive group $S = \mathrm{AGL}_1 F := \{x \mapsto xa + b \mid a, b \in F, a \neq 0\}$.

The celebrated Bruck–Ryser theorem may be stated as follows: the symmetric group S_n does not contain any sharply 2-transitive subset if $n \equiv 1, 2 \pmod 4$ and n is not a sum of two squares; see, e.g., [5, 3.2.13 and 3.2.6] or [21, III.6].

3 Projective planes

Let L, M be lines of a projective plane $\mathcal{P}$. We consider lines as sets of points. Every point $p \notin L \cup M$ gives rise to the bijection $[L, p, M] : L \to M : x \mapsto (xp) \cap M$. Projectivities are concatenations of bijections of this type. The projectivities of L onto itself form a triply transitive group of permutations of L (by Remark 3.1 below); choosing another line in $\mathcal{P}$ leads to an isomorphic permutation group (see [5, p. 160]). We denote this permutation group by $G(\mathcal{P})$ and call it the *group of projectivities* of the projective plane $\mathcal{P}$. The dual plane of $\mathcal{P}$ has the same group of projectivities, up to an isomorphism of permutation groups.

If $\mathcal{P}$ is the desarguesian projective plane coordinatized by a skew field F, then $G(\mathcal{P}) = \mathrm{PGL}_2 F$ in its natural action on the projective line $F \cup \{\infty\}$. Adriano Barlotti [1, 2] has started the study of groups of projectivities in non-desarguesian planes: he showed that $G(\mathcal{P}) = S_{10}$ if $\mathcal{P}$ is the Hughes plane of order 9 or the nearfield plane of order 9, and that $G(\mathcal{P}) = A_{17}$ for the Hall plane $\mathcal{P}$ of order 16.

Remark 3.1 Consider a line L of a projective plane $\mathcal{P}$ and distinct points $u, v \notin L$. Then the set

$$S := \{\, [L, u, X][X, v, L] \mid X \text{ is a line with } u, v \notin X \,\}$$

consists of projectivities that fix the point $\infty := L \cap uv$, and $S \subseteq G(\mathcal{P})_\infty$ is sharply 2-transitive on $L \setminus \{\infty\}$. This well-known fact is a direct consequence of the axioms for projective planes.

The following theorem was conjectured by Dembowski [5, p. 160] and proved by Müller–Nagy [31], with a computer search to remove the largest Mathieu group M_{24} from [13, Theorem 2]; this computer search was replaced by Lemma 2.1 in Müller–Nagy [32].

Theorem 3.2 *Let $\mathcal{P}$ be a finite non-desarguesian projective plane of order n. Then $G(\mathcal{P}) \in \{A_{n+1}, S_{n+1}\}$, and $G(\mathcal{P}) = S_{n+1}$ if $n \equiv 2, 3 \pmod 4$.*

Proof $G := G(\mathcal{P})$ is a triply transitive subgroup of S_{n+1} and $n > 3$. If G has a regular (i.e., sharply transitive) normal subgroup, then $n = 2^d - 1$ with $d > 2$ and the stabilizer G_∞ is a subgroup of the linear group $\mathrm{GL}_d \mathbb{F}_2$, compare [4, Theorem 1.6] or [6, Theorem 7.2A]; the simple group $\mathrm{GL}_d \mathbb{F}_2$ is contained in A_n and $n \equiv 3 \pmod 4$, which yields a contradiction to Remark 3.1 and Lemma 2.1. Therefore G has no

regular normal subgroup; for another argument see [5, 3.4.10(c)] or [9, Korollar 3.6]. By a result of Burnside [3, § 154, Theorem XIII], compare [4, Section 4.8] or [6, Theorem 7.2E], the group G is almost simple, i.e., $N \leq G \leq \operatorname{Aut} N$ for some simple group N.

If $A_{n+1} \leq G$, then we obtain the assertion using Remark 3.1 and Lemma 2.1. If $\mathrm{PSL}_2 \mathbb{F}_n \leq G \leq \mathrm{P\Gamma L}_2 \mathbb{F}_n$ with a prime power n, then the stabilizer G_∞ has a regular normal subgroup, and $\mathcal{P}$ is desarguesian by the Lüneburg–Yaqub theorem; see [27, Theorem 3] and [40], or Schleiermacher [38].

In the remaining cases, the triply transitive group G is a simple Mathieu group of degree $n+1 \in \{11, 12, 22, 23, 24\}$, or $n+1 = 22$ and $G = \operatorname{Aut} M_{22}$; this is a consequence of the classification of finite simple groups, see [4, Section 7.4]. We have $n \neq 21$ (and $n \neq 22$) by the Bruck–Ryser theorem mentioned above, thus $n \equiv 2, 3 \pmod 4$. The simple group G is contained in A_{n+1}, hence $G_\infty \leq A_n$; this is a contradiction to Remark 3.1 and Lemma 2.1. (The two small Mathieu groups are excluded already by [37], see also [7] or [35, p. 13].) □

Theorem 3.2 and the results in the sections below depend on the classification of all finite permutation groups that are triply (or doubly) transitive, hence on the classification of finite simple groups. This appears to be unavoidable, because an explicit classification of all finite non-desarguesian planes is out of reach. If the order n in Theorem 3.2 is a power of 2, then the classification of finite simple groups can be replaced by a deep group-theoretic result of Holt [20, Corollary 1].

In the situation of Theorem 3.2, one might expect that $G(\mathcal{P}) = A_{n+1}$ if n is even, and $G(\mathcal{P}) = S_{n+1}$ if n is odd. This is true for all André planes ([19] and [10]), in particular for Hall planes, for planes over commutative semifields ([23] and [14]), and for semifield planes and nearfield planes of odd order ([14] and [10]). However, Kilmer [23] describes semifield planes $\mathcal{P}$ with orders $n = 16$, 32 and 64 such that $G(\mathcal{P}) = S_{n+1}$.

Question 3.3 Are there finite projective planes $\mathcal{P}$ of order $n \equiv 1 \pmod 4$ with $G(\mathcal{P}) = A_{n+1}$? Examples will have order $n \geq 13$, because the planes of order 5 and 9 are known (see [5, 3.2.15] and [24]) and covered by the results of Barlotti mentioned above.

4 Affine planes

Let L, M be lines of an affine plane $\mathcal{A}$; again we consider lines as sets of points. Every point p at infinity (i.e., every parallel class of lines) defines a parallel projection $L \to M$ in the direction of p. The affine projectivities are the concatenations of bijections of this type. The affine projectivities of L onto itself form a doubly transitive group of permutations of L (by Remark 3.1, with points u, v at infinity); choosing another line in $\mathcal{A}$ leads to an isomorphic permutation group. We denote this permutation group by $G^{\mathrm{aff}}(\mathcal{A})$ and call it the *group of affine projectivities* of the affine plane $\mathcal{A}$. (This is the group Π^W in [5, p. 161].)

If $\mathcal{A}$ is the desarguesian affine plane coordinatized by a skew field F, then $G^{\mathrm{aff}}(\mathcal{A}) = \mathrm{AGL}_1 F$ in its natural action on F.

By Schleiermacher [38, Satz 2] (see also [35, Proposition 7]), the permutation group $G^{\mathrm{aff}}(\mathcal{A})$ has a regular normal subgroup if, and only if, the affine plane $\mathcal{A}$ is a *translation plane*; this means that the point set of $\mathcal{A}$ is a vector space, each line is an affine subspace, and each translation $v \mapsto v + a$ of the vector space is an automorphism (a collineation) of $\mathcal{A}$. The *kernel* of $\mathcal{A}$ is then the largest skew field affording such a vector space structure; see [29, p. 3] or [21, VII]. This result of Schleiermacher splits the study of the groups $G^{\mathrm{aff}}(\mathcal{A})$ into two quite different cases; see Theorems 4.1 and 4.2 below.

The affine group $\mathrm{AGL}_d\,\mathbb{F}_q := \{v \mapsto Av + a \mid A \in \mathrm{GL}_d\,\mathbb{F}_q,\, a \in \mathbb{F}_q^d\}$ has a normal subgroup $\mathrm{ASL}_d\,\mathbb{F}_q$ defined by the condition $\det A = 1$.

Theorem 4.1 *Let $\mathcal{A}$ be a finite translation plane of order $n = q^d$ with kernel $\mathbb{F}_q$. Then $\mathrm{ASL}_d\,\mathbb{F}_q \le G^{\mathrm{aff}}(\mathcal{A}) \le \mathrm{AGL}_d\,\mathbb{F}_q$.*

This is proved in [11, 3.2], relying on the list of all subgroups of $\mathrm{GL}_d\,\mathbb{F}_q$ that act transitively on the set of non-zero vectors; see Hering [17, 18] and Liebeck [25, Appendix 1]; cp. also Malle [30, Satz 5.1]. The completeness of this list depends on the classification of finite simple groups. Often, perhaps always, one has $G^{\mathrm{aff}}(\mathcal{A}) = \mathrm{AGL}_d\,\mathbb{F}_q$; this holds for all nearfield planes and for all André planes by [10], and for many semifield planes by [14].

An analogous result for infinite topological translation planes homeomorphic to $\mathbb{R}^n$ is proved in [16].

Theorem 4.2 *Let $\mathcal{A}$ be a finite affine plane of order n which is not a translation plane. Then $G^{\mathrm{aff}}(\mathcal{A}) \in \{A_n, S_n\}$, or $n = 24$ and $G^{\mathrm{aff}}(\mathcal{A}) = M_{24}$ is the largest Mathieu group. Moreover $G^{\mathrm{aff}}(\mathcal{A}) = S_n$ if $n \equiv 2, 3 \pmod 4$.*

These groups $G^{\mathrm{aff}}(\mathcal{A})$ have no regular normal subgroups, hence they are almost simple (compare the proof of Theorem 3.2). Therefore Theorem 4.2 is a consequence of Remark 3.1 (with points u, v at infinity) and the following group-theoretic result.

Theorem 4.3 *If an almost simple permutation group $G \le S_n$ contains a sharply 2-transitive subset, then $G \in \{A_n, S_n\}$, or $n = 24$ and $G = M_{24}$.*

This is obtained by combining [15, Theorem 1.9] with Lemma 2.1. In fact, the proof of [15, Theorem 1.9] can be shortened using that Lemma and [32, Section 3]; this proof considers all doubly transitive actions of almost simple groups and extends earlier results of Lorimer [26] and O'Nan [33].

Problem 4.4 One would like to eliminate the unlikely possibility $n = 24$ with the Mathieu group M_{24} in Theorems 4.2 and 4.3. The results in [32, Section 4] suggest that this requires new ideas.

In the projective closure $\mathcal{P}$ of an affine plane $\mathcal{A}$, the parallel projection $L \to M$ in the direction of p is just the projectivity $[L, p, M]$ considered in the previous section (if we ignore the points at infinity of L and M). Thus $G^{\mathrm{aff}}(\mathcal{A})$ is a subgroup of the stabilizer $G(\mathcal{P})_\infty$, where ∞ is a point at infinity.

Fix a parallel class p of $\mathcal{A}$ and consider concatenations of central projections $[L, c, M]$ where $L, M \in p$ and c is a point of $\mathcal{P}$ with $c \notin L \cup M$. The concatenations

mapping a fixed line $L \in p$ onto itself form another 'group of affine projectivities' of $\mathcal{A}$, but this permutation group is isomorphic to $G^{\mathrm{aff}}(\mathcal{A}')$ where the affine plane $\mathcal{A}'$ is obtained by removing the line that contains the set p from the dual plane of $\mathcal{P}$.

Theorem 3.2 is a consequence of Theorem 4.2: every finite non-desarguesian projective plane $\mathcal{P}$ has at most one translation line (see [21, Theorems 6.18 and 6.20]), hence some affine part $\mathcal{A}$ of $\mathcal{P}$ is not a translation plane, we have $G^{\mathrm{aff}}(\mathcal{A}) \leq G(\mathcal{P})_\infty$, and the Mathieu group M_{24} is maximal in A_{24} and has no transitive extension (see [28, 13.2, p. 94]).

5 Locally finite planes

A structure (a group, a field, a projective or affine plane) is called *locally finite*, if every finite subset is contained in a finite substructure. A subplane of an affine plane is required to have the induced parallelity relation, as in [5, p. 118].

Let $\mathcal{P}$ be a locally finite projective (or affine) plane. Every projectivity of $\mathcal{P}$ is determined by finitely many points and lines, hence every orbit of every finitely generated subgroup of $G(\mathcal{P})$ (or $G^{\mathrm{aff}}(\mathcal{P})$) is finite. This entails that these groups of projectivities are locally residually finite; they are not always locally finite, not even periodic, as the affine generalized André planes described in [8, V.3.8] or [22, Section 3] show.

Every projectivity of a subplane of $\mathcal{P}$ extends to a projectivity of $\mathcal{P}$. Therefore the following two results are consequences of Theorems 3.2 and 4.2.

Corollary 5.1 *Let $\mathcal{P}$ be an infinite, locally finite projective plane which is not desarguesian. Then $G(\mathcal{P})$ is t-transitive for every $t \in \mathbb{N}$.*

Corollary 5.2 *Let $\mathcal{A}$ be an infinite, locally finite affine plane which is not a translation plane. Then $G^{\mathrm{aff}}(\mathcal{A})$ is t-transitive for every $t \in \mathbb{N}$.*

For translation planes, the following result is proved in [12, Theorem 2]; note that finite-dimensionality is not a consequence of local finiteness.

Theorem 5.3 *Let $\mathcal{A}$ be a locally finite translation plane of finite dimension $2d$ over its kernel K. Then $\mathrm{ASL}_d K \leq G^{\mathrm{aff}}(\mathcal{A}) \leq \mathrm{AGL}_d K$.*

The equation $G^{\mathrm{aff}}(\mathcal{A}) = \mathrm{AGL}_d K$ holds if $\mathcal{A}$ is an André plane or a nearfield plane.

Question 5.4 Is every locally finite projective plane countable? This holds if the plane is desarguesian, since every locally finite (skew) field is algebraic over its prime field, and therefore countable.

References

[1] Adriano Barlotti, La determinazione del gruppo delle proiettività di una retta in sè in alcuni particolari piani grafici finiti non desarguesiani, *Boll. Unione Mat. Ital., III. Ser.* **14** (1959), 543–547.

[2] Adriano Barlotti, Il gruppo delle proiettività di una retta in sè in un particolare piano non desarguesiano di ordine sedici, *Matematiche* **19** (1964), 89–95.

[3] W. Burnside, *Theory of groups of finite order, second edition* (CUP, Cambridge 1911).

[4] Peter J. Cameron, *Permutation groups* (CUP, Cambridge 1999).
[5] P. Dembowski, *Finite geometries* (Springer, Berlin 1968).
[6] John D. Dixon and Brian Mortimer, *Permutation Groups* (Springer, New York 1996).
[7] Rudolf Fritsch, Ein 'affiner' Beweis des Satzes von v. Staudt-Schleiermacher, *Monatsh. Math.* **86** (1978), 177–184.
[8] Theo Grundhöfer, Projektivitätengruppen von endlichen und lokal endlichen ebenen Geometrien, PhD thesis, Univ. Kaiserslautern 1981.
[9] Theo Grundhöfer, Über Projektivitätengruppen affiner und projektiver Ebenen unter besonderer Berücksichtigung von Moufangebenen, *Geometriae Dedicata* **13** (1983), 435–458.
[10] Theo Grundhöfer, Projektivitätengruppen von Translationsebenen, *Results Math.* **6** (1983), 163–182.
[11] Theo Grundhöfer, Die Projektivitätengruppen der endlichen Translationsebenen, *J. Geom.* **20** (1983), 74–85.
[12] Theo Grundhöfer, Finite subplanes and affine projectivities of translation planes, *Mitt. Math. Sem. Giessen* **164** (1984), 179–184.
[13] Theo Grundhöfer, The groups of projectivities of finite projective and affine planes, *Ars Combin.* **25(A)** (1988), 269–275.
[14] Theo Grundhöfer, Projektivitätengruppen von Ebenen über endlichen Semikörpern, *J. Geom.* **40** (1991), 74–76.
[15] Theo Grundhöfer and Peter Müller, Sharply 2-transitive sets of permutations and groups of affine projectivities, *Beiträge Algebra Geom.* **50** (2009), 143–154.
[16] Theo Grundhöfer and Karl Strambach, Die affinen Projektivitätengruppen der lokalkompakten zusammenhängenden Translationsebenen, *Arch. Math. (Basel)* **47** (1986), 274–278.
[17] Christoph Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order, *Geometriae Dedicata* **2** (1974), 425–460.
[18] Christoph Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order. II, *J. Algebra* **93** (1985), 151–164.
[19] Armin Herzer, Die Gruppe $\Pi(g)$ in den endlichen André-Ebenen gerader Ordnung, *Geometriae Dedicata* **3** (1974), 241–249.
[20] D. F. Holt, Triply-transitive permutation groups in which an involution central in a Sylow 2-subgroup fixes a unique point, *J. London Math. Soc. (2)* **15** (1977), 55–65.
[21] Daniel R. Hughes and Fred C. Piper, *Projective Planes* (Springer, New York 1973).
[22] Hubert Kiechle, Lokal endliche André-Systeme, *J. Geom.* **41** (1991), 79–93.
[23] Dirk Kilmer, Über die Projektivitätengruppen von Semikörperebenen gerader Ordnung, *J. Geom.* **35** (1989), 108–119.
[24] C. W. H. Lam, G. Kolesova, L. Thiel, A computer search for finite projective planes of order 9, *Discrete Math.* **92** (1991), 187–195.
[25] Martin W. Liebeck, The affine permutation groups of rank three, *Proc. London Math. Soc. (3)* **54** (1987), 477–516.
[26] Peter Lorimer, Finite projective planes and sharply 2-transitive subsets of finite groups, in *Proc. Second Internat. Conf. Theory of Groups* (M. F. Newman ed.), Canberra 1973, Lecture Notes in Math. **372** (Springer, Berlin 1974), 432–436.
[27] Heinz Lüneburg, An axiomatic treatment of ratios in an affine plane, *Arch. Math. (Basel)* **18** (1967), 444–448.
[28] Heinz Lüneburg, *Transitive Erweiterungen endlicher Permutationsgruppen*, Lecture Notes in Math. **84** (Springer, Berlin 1969).
[29] Heinz Lüneburg, *Translation planes* (Springer, Berlin 1980).
[30] Gunter Malle, Fast-einfache Gruppen mit langen Bahnen in absolut irreduzibler Operation, *J. Algebra* **300** (2006), 655–672.
[31] Peter Müller and Gábor P. Nagy, A note on the group of projectivities of finite projective

planes, *Innov. Inc. Geom.* **6/7** (2007/08), 291–294.
[32] Peter Müller and Gábor P. Nagy, On the non-existence of sharply transitive sets of permutations in certain finite permutation groups, *Adv. Math. Commun.* **5** (2011), 303–308.
[33] M. E. O'Nan, Sharply 2-transitive sets of permutations, in *Proc. Rutgers group theory year, 1983–1984* (M. Aschbacher et al., eds.), Rutgers University, New Brunswick, N.J. (CUP, Cambridge 1985), 63–67.
[34] Günter Pickert, *Projektive Ebenen, zweite Auflage* (Springer, Berlin 1975).
[35] Günter Pickert, Projectivities in projective planes, in [36, pp. 1–49].
[36] Peter Plaumann and Karl Strambach (eds), *Geometry — von Staudt's point of view*, Proc. Bad Windsheim 1980 (Reidel, Dordrecht 1981).
[37] Adolf Schleiermacher, Bemerkungen zum Fundamentalsatz der projektiven Geometrie, *Math. Z.* **99** (1967), 299–304.
[38] Adolf Schleiermacher, Reguläre Normalteiler in der Gruppe der Projektivitäten bei projektiven und affinen Ebenen, *Math. Z.* **114** (1970), 313–320.
[39] Ernst Witt, Über Steinersche Systeme, *Abh. Math. Sem. Univ. Hamburg* **12** (1938), 265–275.
[40] Jill C. D. S. Yaqub, On the group of projectivities on a line in a finite projective plane, *Math. Z.* **104** (1968), 247–248.

ON THE RELATION GAP AND RELATION LIFTING PROBLEM

JENS HARLANDER

Boise State University, Boise, Idaho, USA
Email: jensharlander@boisestate.edu

Abstract

This article surveys results in connection with the relation gap problem, the relation lifting problem, and the geometric realization problem. These three problems lie in the intersection of combinatorial group theory and 2-dimensional homotopy theory. We present key examples that lie at the heart of these problems.

1 Three problems

Let G be a group. A generating set $\mathcal{G} = \{g_1, \ldots, g_n\}$ defines an epimorphism $\phi_{\mathcal{G}} : F(x_1, \ldots, x_n) \to G$ that sends x_i to g_i, $i = 1, \ldots, n$. Let $N_{\mathcal{G}}$ be the kernel of $\phi_{\mathcal{G}}$. A fundamental problem in combinatorial group theory is to determine a minimal normal generating set for $N_{\mathcal{G}}$. If it is clear which generating set is used we will drop the subscript and simply write N instead of $N_{\mathcal{G}}$. When we say

$$F/N = \langle\, x_1, \ldots, x_n \mid r_1 = 1, \ldots, r_m = 1 \,\rangle$$

is a presentation for a group G, we mean that F is a free group on $\{x_1, \ldots, x_n\}$, N is the normal closure of $r_1, \ldots, r_m$ in F, and G is isomorphic to F/N. The conjugation action of F on N provides a $\mathbb{Z}G$-module structure on $N_{ab} = N/[N, N]$. This module is the relation module of the presentation F/N of G. We have

$$d_F(N) \geq d_G(N/[N, N]) \geq d(N/[F, N]) = n - tfr(H_1(G)) + d(H_2(G)).$$

Here $d_F(-)$ denotes the minimal number of F-group generators, $d_G(-)$ denotes the minimal number of G-module generators, $d(-)$ denotes the minimal number of generators, and $tfr(-)$ denotes the torsion free rank. The chain of inequalities follows from the exact sequence

$$H_2(G) \to H_2(Q) \to \mathbb{Z} \otimes_{\mathbb{Z}G} H_1(H) \to H_1(G) \to H_1(Q) \to 0,$$

associated with an exact sequence of groups $1 \to H \to G \to Q \to 1$ (see Brown [5]). We say the presentation F/N has a relation gap if

$$d_F(N) - d_G(N/[N, N]) > 0.$$

The presentation is called efficient if $d_F(N) = d(N/[F, N])$.

Relation gap problem: Does there exist a finite presentation F/N with a relation gap?

Infinite relation gaps are known to exists for finitely generated groups. See Bestvina and Brady [2].

Relation lifting problem: Given set $s_1[N,N],\ldots,s_m[N,N]$ of relation module generators, do there exists defining relators $r_1,\ldots,r_m$ (i.e., elements that normally generate N), such that $r_i[N,N]=s_i[N,N]$, $i=1,\ldots,m$?

We refer to the elements r_i as lifts of the generators $s_i[N,N]$, $i=1,\ldots,m$. The relation lifting problem arose in work of C. T. C. Wall [31]. M. Dunwoody [8] provided an example where lifting is not possible. We will provide more details on Dunwoody's construction in a later section.

Let $\mathcal{G}$ be a finite generating set of a group G and $F/N_{\mathcal{G}}$ be the associated presentation. Since the relation module $N_{\mathcal{G}ab}$ is isomorphic to $H_1(\Gamma_{\mathcal{G}})$, where $\Gamma_{\mathcal{G}}$ is the Cayley graph of G associated with the generating set $\mathcal{G}$, a choice of relation module generators $\mathcal{M}$ gives rise to an partial resolution $\mathcal{K}_{\mathcal{M}}$ of the trivial $\mathbb{Z}G$-module $\mathbb{Z}$:

$$\mathbb{Z}G^m \xrightarrow{\partial_2} \mathbb{Z}G^n \xrightarrow{\partial_1} \mathbb{Z}G \xrightarrow{\epsilon} \mathbb{Z} \to 0,$$

where m is the number of elements in $\mathcal{M}$. We call such a partial resolution $\mathcal{K}_{\mathcal{M}}$ an algebraic 2-complex for G. We say an algebraic 2-complex $\mathcal{K}$ for G is geometrically realizable if it is chain homotopy equivalent to an algebraic 2-complex $\mathcal{C}(\tilde{K})$ that arises as the augmented chain complex of the universal covering $\tilde{K}$ of a 2-complex K with fundamental group G.

Geometric realization problem: Does there exist an algebraic 2-complex that is not geometrically realizable?

A relation gap in a finite presentation F/N would certainly provide a set of relation module generators that can not be lifted. If in addition $d_F(N)-d(F)$ is minimal among all finite presentations of G one can construct an algebraic 2-complex that is not realizable. Indeed, if $\mathcal{M}$ is a minimal generating set for the relation module $N/[N,N]$, then the Euler characteristic of $\mathcal{K}_{\mathcal{M}}$ is smaller that the Euler characteristic of any finite 2-complex with fundamental group G. Geometric realization has been studied by F. E. A. Johnson and his students in connection with the $D(2)$-problem. See [25].

2 Examples concerning the relation gap question

Example 2.1 Consider the presentation

$$F/N=\langle\, a,b,c,d \mid [a,b]=1, a^m=1, [c,d]=1, c^n=1\,\rangle$$

of the group $G=(\mathbb{Z}_m\oplus\mathbb{Z})\star(\mathbb{Z}_n\oplus\mathbb{Z})$, where m and n are relatively prime. Note that $d(N/[F,N])=4-2+1=3$. Epstein asked [10] if this presentation is efficient. Gruenberg and Linnell showed [12] that $d_G(N/[N,N])=3$.

Example 2.2 The following examples were constructed by Bridson and Tweedale [3]. They are very much in the spirit of the Epstein example above, but an unexpectedly small set of relation module generators can be seen more easily. Let Q_n be the group defined by $\langle\, a,b,x \mid a^n=1, b^n=1, [a,b]=1, xax^{-1}=b\,\rangle$. This groups is an HNN-extension of $\mathbb{Z}_n\times\mathbb{Z}_n$ where the stable letter x conjugates one factor

into the other. Note that $\langle\, a, x \mid a^n = 1, [a, xax^{-1}] = 1 \,\rangle$ also presents Q_n. Let $\rho_n(a, x) = [xax^{-1}, a]a^{-n}$ and let $q_n = (n+1)^n - 1$. Then

$$F/N = \langle\, a, x, b, y \mid a^m = 1, [a, xax^{-1}] = 1, b^n = 1, [b, yby^{-1}] = 1 \,\rangle$$

is a presentation of the free product $Q_m \star Q_n$. Bridson and Tweedale show that the relation module N_{ab} is generated by $\rho_m(a, x)[N, N]$, $\rho_n(b, y)[N, N]$, and $a^m b^{-n}[N, N]$. Other articles by Bridson and Tweedale that address the relation gap are [4] and [13].

Example 2.3 The following construction first appeared in [16]. Let F_1/N_1 and F_2/N_2 be finite presentations of groups G_1 and G_2, respectively. Let H be a finitely generated subgroup of both G_1 and G_2 and let F/N be the standard presentation of the amalgamated product $G = G_1 \star_H G_2$. One can show that

$$d_G(N_{ab}) \leq d_{G_1}(N_{1ab}) + d_{G_2}(N_{2ab}) + d_H(IH),$$

where IH is the augmentation ideal of H. Denote by H^n the n-fold direct product $H \times \cdots \times H$. Cossey, Gruenberg, and Kovacs [7] showed that $d_{H^n}(IH^n) = d_H(IH)$ in case H is a finite perfect group. Since $d(H^n) \to \infty$ as $n \to \infty$ one can produce arbitrarily large generation gaps $d(H^n) - d_{H^n}(IH^n)$. This leads to unexpectedly small generating sets for the relation module N_{ab} for for presentations F/N of $G_1 \star_{H^n} G_1$. The hope is that the amalgamated product shifts the generation gap into a relation gap.

3 Dunwoody's counter example to relation lifting [8]

Consider the presentation $F/N = \langle\, a, b \mid a^5 = 1 \,\rangle$ of the group $G = \mathbb{Z}_5 \star \mathbb{Z}$. Note that $(1 - a + a^2)(a + a^2 - a^4) = 1$, so $1 - a + a^2$ is a non-trivial unit in $\mathbb{Z}G$. Hence so is $\alpha = (1 - a + a^2)b$. It follows that

$$\alpha \cdot a^5[N, N] = (ba^5b^{-1})(aba^{-5}b^{-1}a^{-1})(a^2ba^5b^{-1}a^{-2})[N, N] = s[N, N]$$

is a generator for the relation module. This generator can not be lifted. For suppose that $\langle\langle a^5 \rangle\rangle = \langle\langle r \rangle\rangle$ and $r[N, N] = s[N, N]$. One relator group theory implies that $r = wa^{\pm 5}w^{-1}$, for some $w \in F$. Using the well known resolution for $\mathbb{Z}_5$ (see Brown [5], Chapter I, Section 6) one can show that $N_{ab} \cong \mathbb{Z}G/\mathbb{Z}G\langle a - 1 \rangle \cong A$, where A is the free abelian group with basis the elements of G that end in $b^{\pm 1}$ (this can also be seen by directly inspecting the Cayley graph Γ for the generating set $\{a, b\}$.) The isomorphism $N_{ab} \to A$ sends $s[N, N]$ to $b - ab + a^2b$ and sends $r[N, N]$ to a single basis element in A. Hence $r[N, N] \neq s[N, N]$.

Dunwoody's construction uses non-trivial units in $\mathbb{Z}G$ and one relator group theory. We do not know an example of a torsion free group where relation lifting fails. Also note that the algebraic 2-complex $\mathcal{K}$ that one obtains from the relation module generator $s[N, N]$ is geometrically realizable. In fact it is chain homotopically equivalent to the cellular chain complex of the universal covering of the standard 2-complex K built from the presentation $F/N = \langle\, a, b \mid a^5 = 1 \,\rangle$ of G. Thus the example does not provide a negative answer to the geometric realization question.

It is easy to give examples of relation module generators that are not defining relators but can be lifted. Let

$$F/J = \langle\, a, b, c, d, x \mid a = 1, b = 1, c = 1, d = 1 \,\rangle$$

be a presentation of the infinite cyclic group. Let $s_1 = b[b, a]$, $s_2 = c[c, b]$, $s_3 = d[d, c]$, $s_4 = a[a, b]$. The elements $s_i[J, J]$, $i = 1, 2, 3, 4$, generate the relation module J_{ab}, but the s_i, $i = 1, 2, 3, 4$, do not normally generate J. Indeed, notice that $s_1 = 1$ can be rewritten as $aba^{-1} = b^2$. Rewriting the other relations $s_i = 1$ in a similar way the presentation $F/N = \langle\, a, b, c, d, x \mid s_1 = 1, s_2 = 1, s_3 = 1, s_4 = 1 \,\rangle$ turns into the presentation

$$F/N = \langle\, a, b, c, d, x \mid aba^{-1} = b^2, bcb^{-1} = c^2, cdc^{-1} = d^2, dad^{-1} = a^2 \,\rangle,$$

which presents the free product $H \star \mathbb{Z}$, where $H = \langle a, b, c, d\rangle$ is the Higman group on four generators. Higman groups H_n on $n \geq 2$ generators are defined similarly [19]. It is known that H_n is trivial for $n = 2, 3$ and infinite (in fact aspherical, see Gersten [11]) for $n \geq 4$. All H_n are perfect and do not have proper subgroups of finite index. In particular F/N does not present the infinite cyclic group. Since $s_1[J, J] = b[J, J]$, $s_2[J, J] = c[J, J]$, $s_3[J, J] = d[J, J]$, $s_4[J, J] = a[J, J]$, the relation module generators $s_i[J, J]$, $i = 1, 2, 3, 4$ can be lifted.

The following observation is often useful when constructing examples of presentations that are interesting in view of relation lifting.

Lemma 3.1 *Let $F/N = \langle\, x_1, \ldots, x_n \mid s_1, \ldots, s_m \,\rangle$ be a presentation of a group G and suppose $P = J/N$ is a perfect normal subgroup. Let Q be the quotient G/P. Then F/J is a presentation for Q and the relation module J_{ab} is generated by $s_1[J, J], \ldots, s_m[J, J]$.*

Proof Since $P_{ab} = J/N[J, J] = 1$ we have $J = N[J, J]$. The result follows. □

4 Stabilization

All examples encountered so far are free products, except for Example 2.3 in Section 2, where the group is a free product amalgamated over a finite group. Free products can have unexpectedly small presentations. See Hog-Angeloni, Metzler, Lustig [21]. In his work on distinguishing homotopy and simply homotopy type for 2-complexes, Metzler [27] showed that Whitehead torsion elements can be topologically realized if one allows stabilization using copies of the complex $K = \langle\, a, b \mid [a, b], a^2, b^4 \,\rangle$. See also Hog-Angeloni, Metzler [22] and [24], Chapter XII. These stabilization techniques can also be applied to closing the relation gap. See [14], [15], and also [16].

Theorem 4.1 *Given a finite presentation F/N. Then there exists $k \geq 0$ such that*

$$F/N \star \langle\, a, b \mid a^2 = 1, b^2 = 1, [a, b] = 1 \,\rangle \star \cdots \star \langle\, a, b \mid a^2 = 1, b^2 = 1, [a, b] = 1 \,\rangle$$

(k copies) does not have a relation gap.

Here is the main idea of the proof. Suppose the relation module is generated by $r_1[N,N],\ldots,r_m[N,N]$. Then N is normally generated by $r_1,\ldots,r_m$ together with a finite set of elements of the form $[s,t]$, where $s,t\in N$. For simplicity assume that $N=\langle\!\langle r_1,\ldots,r_m,[s,t]\rangle\!\rangle$. Now note that

$$\langle\, x_1,\ldots,x_n,a,b \mid r_1=1,\ldots,r_m=1,[s,t]=1,a^2=1,b^2=1,[a,b]=1\,\rangle$$

presents the same groups as

$$\langle\, x_1,\ldots,x_n,a,b \mid r_1=1,\ldots,r_m=1,s=a^2,t=b^2,[a,b]=1\,\rangle.$$

Indeed, since a and b commute, the squares a^2 and b^2 also commute, hence s and t commute. So the relation $[s,t]=1$ holds. Since $N=\langle\!\langle r_1,\ldots,r_m,[s,t]\rangle\!\rangle$, we get $s=1$ and $t=1$, and hence $a^2=1$ and $b^2=1$. The commutator relation $[s,t]=1$ got "absorbed", the second presentation makes due with one less relator than the first presentation.

We will illustrate this method further by working through Dunwoody's example considered in Section 3. For economical reasons we use the notation $x*y=xyx^{-1}$. Let $r=a^5$ and let $s=(b{*}r)(ab{*}r^{-1})(a^2b{*}r)\in N=\langle\!\langle r\rangle\!\rangle$. We know that $\langle\, a,b \mid s=1\,\rangle$ is not a presentation for $G=\mathbb{Z}_5*\mathbb{Z}$. But since s generates the relation module, we can add commutators of relators to obtain a presentation. We need to add enough commutators to be able to do the calculation $(a+a^2-a^4)\cdot s[N,N]=b\cdot r[N,N]$(based on $(1-a+a^2)(a+a^2-a^4)b=b$) in N. We claim that

$$\langle\, a,b \mid s=1,[r,b*r]=1,[b*r,ab*r]=1,[b*r,a^2b*r]=1\,\rangle$$

does present G. Note that $(a*s)(a^2*s)(a^4*s^{-1})$ is

$$(ab*r)(a^2b*r^{-1})(a^3b*r)(a^2b*r)(a^3b*r^{-1})(a^4b*r)(a^6b*r^{-1})(a^5b*r)(a^4b*r^{-1}).$$

Since $b{*}r$ commutes with $ab{*}r$, it follows that $a^ib{*}r$ commutes with $a^{i+1}b{*}r$. And since $b{*}r$ commutes with $a^2b{*}r$ it follows that $a^ib{*}r$ commutes with $a^{i+2}b{*}r$. Thus the above expression becomes $(ab{*}r)(a^6b{*}r^{-1})(a^5b{*}r)$. Now note that $a^5b{*}r=r*(b{*}r)=b{*}r$ because r commutes with $b*r$. It follows that $a^6b*r=a*(a^5*r)=ab*r$. Thus our expresion becomes $(ab*r)(ab*r^{-1})(b*r)=b*r$. Thus $b*r$ and hence r defines the trivial element in the group defined by the presentation above. Hence we have a presentation of $G=\mathbb{Z}_5*\mathbb{Z}$ as claimed. It follows that the presentation of generators a,b,c,d,e,f,g,h and relations $s=1$, $c^2=r$, $d^2=b*r$, $e^2=b*r$, $f^2=ab*r$, $g^2=b*r$, $h^2=a^2b*r$, $[c,d]=[e,f]=[g,h]=1$ is a presentation for $(Z_5\star\mathbb{Z})\star(\mathbb{Z}_2\times\mathbb{Z}_2)\star(\mathbb{Z}_2\times\mathbb{Z}_2)\star(\mathbb{Z}_2\times\mathbb{Z}_2)$.

5 One relator groups and another example of Dunwoody's

Let $F/N=\langle\, x_1,\ldots,x_n \mid r\,\rangle$ be a presentation of a torsion-free one-relator group G. Then the relation module N_{ab} is isomorphic to $\mathbb{Z}G$. Let α and β be left module generators of $\mathbb{Z}G$. Let $\partial_{\alpha,\beta}:\mathbb{Z}G\oplus\mathbb{Z}G\to N_{ab}$ be the homomorphism that sends $e_1=(1,0)$ to $\alpha\cdot r[N,N]$ and $e_2=(0,1)$ to $\beta\cdot r[N,N]$. Since the relation module

N_{ab} is isomorphic to the kernel of $\partial_1 : \mathbb{Z}G^n \to \mathbb{Z}G$ that sends the basis element e_i to $x_i - 1$, $i = 1, \ldots, n$, we obtain an algebraic 2-complex $\mathcal{K}_{\alpha,\beta}$

$$\mathbb{Z}G \oplus \mathbb{Z}G \stackrel{\partial_{\alpha,\beta}}{\to} \mathbb{Z}G \oplus \mathbb{Z}G \stackrel{\partial_1}{\to} \mathbb{Z}G \stackrel{\epsilon}{\to} \mathbb{Z} \to 0.$$

This construction provides easy access to examples relevant for relation lifting and geometric realization.

Example 5.1 (M. Dunwoody [9]) Let G be the trefoil group presented by $\langle\, a, b \mid a^2 = b^3 \,\rangle$. Then $\alpha = 1 + a + a^2$ and $\beta = 1 + b + b^2 + b^3$ generate the left module $\mathbb{Z}G$. In order to see this observe that $(a-1)\alpha = a^3 - 1$ and $(b-1)\beta = b^4 - 1$. Since a^3 and b^4 generate G (simply note that $(a^3)^3(b^4)^{-3} = a$ and $(a^3)^2(b^4)^{-2} = b$), the elements $(a-1)\alpha$ and $(b-1)\beta$ generate the augmentation ideal. Thus $(a-1)\alpha$, $(b-1)\beta$, $\beta - \alpha$, and hence α and β, generate $\mathbb{Z}G$. It follows that $\alpha \cdot r[N,N] = (r)(ara^{-1})(a^2ra^{-2})[N,N]$ and $\beta \cdot r[N,N] = (r)(brb^{-1})(b^2rb^{-2})(b^3rb^{-3})[N,N]$ generate the relation module, where $r = a^2b^{-3}$. One obtains an algebraic 2-complex $\mathcal{K}_{\alpha,\beta}$. Dunwoody shows in [9] that $H_2(\mathcal{K}_{\alpha,\beta})$ is stably-free but not free. In particular $\mathcal{K}_{\alpha,\beta}$ is not chain homotopically equivalent to the chain complex of the universal covering of $\langle\, a, b \mid a^2b^{-3} \,\rangle \vee S^2$. The algebraic 2-complex $\mathcal{K}_{\alpha,\beta}$ is geometrically realizable. Dunwoody shows that

$$\langle\, a, b \mid (r)(ara^{-1})(a^2ra^{-2}) = 1, (r)(brb^{-1})(b^2rb^{-2})(b^3rb^{-3}) = 1 \,\rangle$$

is indeed a presentation of G. This provided the first example of different homotopy types of 2-complexes K and L with the same fundamental group G and Euler characteristic $\chi(K) = \chi(L) = \chi_{min}(G) + 1$. For finite groups different homotopy types can occur only at the minimal Euler characteristic level. See [24], Chapter III. Other examples similar to Dunwoody's are known [17]. Later M. Lustig [23] showed that there are infinitely many distinct homotopy types for G on the Euler characteristic $\chi_{min}(G) + 1$.

6 Algebraic 2-complexes for the Klein bottle group

The homotopy classification of 2-complexes with fundamental group G is complete in case G is free of rank n, or G is free abelian of rank 2. In the first case $(S^1 \vee \cdots \vee S^1) \vee S^2 \vee \cdots \vee S^2$ (n copies of S^1) is a complete list, and in the second case $(S^1 \times S^1) \vee S^2 \cdots \vee S^2$ is a complete list. This follows from the fact that the homotopy type of a 2-complexe K is assembled from $\pi_1(K)$, $\pi_2(K)$, and the k-invariant $\kappa \in H^3(\pi_1(K), \pi_2(K))$. See Chapter II in [24]. For G free or $G = \mathbb{Z} \times \mathbb{Z}$, the cohomology group $H^3(G, M) = 0$ for all $\mathbb{Z}G$-modules M. The second homotopy module $\pi_2(K)$ is stably free since the cohomological dimension of G is less or equal to two, and hence free by results of Cohn [6] for free groups (see also Hog-Angeloni [20] for a short topological proof) and Quillen [29] (independently, Suslin) for free abelian groups. It follows that the homotopy type is determined by the Euler characteristic.

The situation is more complicated for the Klein bottle group G. Let $F/N = \langle a, b \mid aba^{-1} = b^{-1} \rangle$ be the standard presentation of G. Then N_{ab} is isomorphic to $\mathbb{Z}G$. Let $p(b) \in \mathbb{Z}G$ be a polynomial in b and let $q(b) = p(b^{-1})$. The elements

$\alpha = a + q(b)$ and $\beta = p(b)$ generate $\mathbb{Z}G$ as a left module. Indeed,

$$\begin{aligned}(a - p(b))\alpha + p(b^{-1})\beta &= (a - p(b))(a + p(b^{-1}) + p(b^{-1})p(b) \\ &= a^2 + ap(b^{-1}) - p(b)a - p(b)p(b^{-1}) + p(b^{-1})p(b) = a^2\end{aligned}$$

Artamonov [1] and Stafford [30] studied the K-theory of solvable groups. Their results can be used to construct exotic alsgebraic 2-complexes for G.

Theorem 6.1 *Let $p_n(b) = 1+nb+nb^3$ and $q_n(b) = p(b^{-1})$, $n \in \mathbb{N}$. Then the elements $\alpha_n = a+q_n(b)$ and $\beta_n = p_n(b)$ generate $\mathbb{Z}G$ as a left module and the set $\{\mathcal{K}_{\alpha_n.\beta_n}\}_{n\in\mathbb{N}}$ contains infinitely many distinct homotopy types of algebraic 2-complexes for G, each of Euler characteristic one.*

The algebraic 2-complexes $\mathcal{K}_{\alpha_n,\beta_n}$ are studied in [18]. We do not know if the relation module generators

$$\begin{aligned}\alpha_n \cdot r[N,N] &= (ara^{-1})(r)(b^{-1}r^nb)(b^{-3}r^nb^3)[N,N], \\ \beta_n \cdot r[N,N] &= (r)(br^nb^{-1})(b^3r^nb^{-3})[N,N],\end{aligned}$$

where $r = aba^{-1}b^{-2}$, can be lifted, or if any of the complexes $\mathcal{K}_{\alpha_n,\beta_n}$ are geometrically realizable.

We conclude this article by providing some details on the work of Artamonov [1] and Stafford [30]. Given a Noetherian domain R and an automorphism $\sigma : R \to R$, one can defined the skewed Laurent-polynomial ring $S = R[x, x^{-1}, \sigma]$, where $xr = \sigma(r)x$. Stafford shows that given two elements r_1, r_2 that satisfy the properties

1. $S = Sr_1 + S(x + r_2)$,
2. $\sigma(r_1)r_2 \notin Rr_1$,

then the left ideal $K = \{\, s \in S \mid sr_1 \in S(x+r_2) \,\}$ is not generated by a single element. Note that K is isomorphic to the kernel of the S-module homomorphism $S \oplus S \to S$, sending e_1 to r_1 and e_2 to $x + r_2$. Hence $K \oplus S \cong S \oplus S$. Since K is not cyclic, it is not free.

If G is the Klein bottle group as above, then $\mathbb{Z}G = \mathbb{Z}H[a, a^{-1}, \sigma]$, where $H = \langle b \rangle$. Now one can take $r_1 = 1 + nb + nb^3$ and $r_2 = 1 + nb^{-1} + nb^{-3}$. By Stafford K_n is not free. Note that K_n is isomorphic to $H_2(\mathcal{K}_{\alpha_n,\beta_n})$, where $\mathcal{K}_{\alpha_n,\beta_n}$ is the algebraic 2-complex defined above. Artamonov shows that the set $\{K_n\}$ contains infinitely many distinct isomorphism types. His reasoning is as follows. First construct a set of primes Q such that if $p < q$ and both p and q are in Q, then $q = 1$ modulo p. Let $K_{n,p} = K_n/pK_n$. Using Stafford's construction one can show that $K_{q,p}$ is not free, but $K_{q,q}$ is free. Thus if $q_1 < q_2$, then K_{q_1} and K_{q_2} are not isomorphic because K_{q_1,q_1} is free but K_{q_2,q_1} is not free. It follows that the set $\{H_2(\mathcal{K}_{\alpha_n.\beta_n})\}_{n\in\mathbb{N}}$ contains infinitely many isomorphism types and hence the set $\{\mathcal{K}_{\alpha_n.\beta_n}\}_{n\in\mathbb{N}}$ contains infinitely many algebraic homotopy types.

References

[1] V.A. Artamonov, Projective, nonfree modules over group rings of solvable groups, *Math. USSR Sbornik.* **116** (1981), 232–244.

[2] M. Bestvina & N. Brady, Morse theory and finiteness properties of groups, *Invent. Math.* **129** (1997), 445–470.

[3] M.R. Bridson & M. Tweedale, Deficiency and abelianized deficiency of some virtually free groups, *Math. Proc. Camb. Phil. Soc.* **143** (2007), 257–264.
[4] M.R. Bridson & M. Tweedale, Constructing presentations of subgroups of right-angled Artin groups, *Geom. Dedicata* **169** (2014), 1–14.
[5] K.S. Brown, *Cohomology of Groups*, Graduate Texts Math. 87, Springer Verlag, 1982.
[6] P.M. Cohn, Free ideal rings, *J. Algebra* **1** (1964), 47–69.
[7] J. Cossey, K.W. Gruenberg & L.G. Kovacs, The presentation rank of a direct product of finite groups, *J. Algebra* **28** (1974), 597–603.
[8] M.J. Dunwoody, Relation modules, *Bull. London Math. Soc.* **4** (1972), 151–155.
[9] M.J. Dunwoody, The homotopy type of a two-dimensional complex, *Bull. London Math. Soc.* **8** (1976), 282–285.
[10] D.B.A. Epstein, Finite presentations of groups and 3-manifolds, *Quart. J. Math. Oxford Ser. (2)* **12** (1961), 205–212.
[11] S.M. Gersten, Reducible diagrams and equations over groups, *Essays in group theory*, 15–73, Math. Sci. Res. Inst. Publ. 8, Springer, 1987.
[12] K. Gruenberg & P. Linnell, Generation gaps and abelianized defects of free products, *J. Group Theory* **11** (2008), 587–608.
[13] Guido's Book of Conjectures, collected by Indira Chatterji, Monographies de L'Enseignement Mathématique 40, 2008.
[14] J. Harlander, Closing the relation gap by direct product stabilization, *J. Algebra* **182** (1996), 511–521.
[15] J. Harlander, Embeddings into efficient groups, *Proc. Edinburgh Math. Soc.* **40** (1997), 317–324.
[16] J. Harlander, Some aspects of efficiency, *Groups—Korea '98*, 165–180, de Gruyter, 2000.
[17] J. Harlander & J.A. Jensen, Exotic relation modules and homotopy types for certain 1-relator groups, *Algebr. Geom. Topol.* **6** (2006), 2163–2173.
[18] J. Harlander & A. Misseldine, On the K-theory and homotopy theory of the Klein bottle group, *Homology Homotopy Appl.* **12** (2010), 1–10.
[19] G. Higman, A finitely generated infinite simple group, *J. London Math. Soc.* **26** (1951), 61–64.
[20] C. Hog-Angeloni, A short topological proof of Cohn's theorem, *Topology and Combinatorial Group Theory*, 90–95, Lecture Notes Math. 1440, Springer, 1990.
[21] C. Hog-Angeloni, M. Lustig & W. Metzler, Presentation classes, 3-manifolds and free products, *Geometry and Topology*, 154–167, Lecture Notes Math. 1167, Springer, 1985.
[22] C. Hog-Angeloni & W. Metzler, Stabilization by free products giving rise to Andrews-Curtis equivalences, Note Mat. **10** (1990), 305–314.
[23] M. Lustig, Infinitely many pairwise homotopy inequivalent 2-complexes K_i with fixed $\pi_1(K_i)$ and $\chi(K_i)$, *J. Pure Appl. Algebra* **88** (1993), 173–175.
[24] *Two-dimensional Homotopy and Combinatorial Group Theory*, eds. C. Hog-Angeloni, W. Metzler & A.J. Sieradski, London Math. Soc. Lecture Note Ser. 197, CUP, 1993.
[25] F.E.A. Johnson, *Stable Modules and the $D(2)$-Problem*, London Math. Soc. Lecture Note. Ser. 301, CUP, 2003.
[26] W.H. Mannan, A commutative version of the group ring, *J. Algebra* **379** (2013), 113–143.
[27] W. Metzler, Die Unterscheidung von Homotopietyp und einfachem Homotopietyp bei 2-dimensionalen Komplexen, *J. Reine Angew. Math.* **403** (1990), 201–219.
[28] D. Osin & A. Thom, Normal generation and ℓ^2-betti numbers of groups, *Math. Ann.* **355** (2013), 1331–1347.
[29] D. Quillen, Projective modules over polynomial rings, *Invent. Math.* **36** (1976), 167–171.
[30] J.T. Stafford, Stably free, projective right ideals, *Compositio Math.* **54** (1985), 63–78.
[31] C.T.C. Wall, Finiteness conditions for CW complexes, *Ann. of Math. (2)* **81** (1965), 56–69.

SOME RESULTS ON PRODUCTS OF FINITE SUBSETS IN GROUPS

MARCEL HERZOG*, PATRIZIA LONGOBARDI† and MERCEDE MAJ†

*School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel
Email: herzogm@post.tau.ac.il

†Dipartimento di Matematica, Università di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (Salerno), Italy
Email: plongobardi@unisa.it, mmaj@unisa.it

Abstract

In this paper we survey some recent results concerning products of finite subsets in groups, some of which are still under preparation for publication.

1 Introduction

If X, Y are subsets of a group G, we write

$$X^{-1} = \{x^{-1} \mid x \in X\}, \quad X^2 = \{x_1x_2 \mid x_1, x_2 \in X\} \text{ and } XY = \{xy \mid x \in X, y \in Y\}.$$

If G is an additive group, then we write

$$X + Y = \{x + y \mid x \in X, y \in Y\} \text{ and } 2X = \{x_1 + x_2 \mid x_1, x_2 \in X\}.$$

In this survey we are interested in two topics related to products of finite subsets in groups. Let X be a finite subset of a group G. The first topic deals with the size of the product set X^2, while the second one is concerned with those finite subsets X of G which satisfy the equation $|XX^{-1}| = |X^{-1}X|$.

Results related to the first topic appear in our paper [8], joint with G.A. Freiman, and in the following joint papers with G.A. Freiman and Y.V. Stanchescu: [10] and [9], [11] (both under preparation). The second topic was investigated in the paper [14], joint with G. Kaplan.

Our aim in this survey is to be of interest to group theorists who are experts in the described area of research, as well as to those who wish only to take a glimpse into the area. Therefore, in addition to definitions and referenced results, we provide a brief overview of subjects under consideration.

We continue with a more detailed description of our topics of interest. As mentioned above, our *first major topic* deals with the size of the product set X^2, where X is a finite subset of a group G.

More precisely, we are concerned with the following problem.

Problem 1. *Let S be a finite subset of a group G of size $|S| = k$. Determine the* structure *of S if it satisfies the following restriction: $|S^2| \leq f(k)$ for some function f of k.*

Problems of this type are called "*inverse*" problems.

In particular, we shall be interested in "*small doubling*" problems, dealing with the structure of finite subsets S of G, of order k, satisfying the inequality $|S^2| \leq \alpha k + \beta$ for some small $\alpha \geq 1$ and small $|\beta|$.

For example, if X is a finite *subgroup* of a group G of order $|X| = k$, then $|X^2| = k$. This is a *direct* (trivial) result. The corresponding "*ordinary*" inverse problem is:

Problem 2. *Let S be a finite subset of a group G of order $|S| = k$ and suppose that $|S^2| = k$. What is the structure of S?*

If $1 \in S$, then S is a subgroup of G, since by our assumptions clearly $S^2 = S$. But without any assumptions, S does not need to be a subgroup of G. Indeed, if H is a normal subgroup of G of order $|H| = k$ and $a \in G$, then the coset $S = aH$ also satisfies $|S| = k$ and $|S^2| = |aHaH| = |a^2H| = |H| = k$.

What happens in general? It turns out that S is *always* a coset of a subgroup of G. This follows from a more general "*extended*" inverse theorem of G.A. Freiman.

Theorem 1.1 ([6]) *Let A be a finite subset of a group G and suppose that $|A^2| < \frac{3}{2}|A|$. Then A^2 is a coset of a subgroup of G.*

The first version of the paper [6] was published in 1951.

How does our claim follow from Freiman's theorem? If $|S^2| = |S|$, then clearly $S^2 = xS$ for every $x \in S$. By Theorem 1.1, $xS = S^2 = uH$ for some subgroup H of G and some $u \in G$. It follows that $S = x^{-1}uH$, as claimed.

Why is Theorem 1.1 called an "extended" inverse theorem? We call Problem 2 an "ordinary" inverse problem, since the assumption is that the minimal bound on $|S^2|$ holds. We call Theorem 1.1 an "extended" inverse result, because the assumed bound for $|S^2|$ is a bit higher than the minimal one, still enabling us to say something about the structure of S.

The 1951 version of [6] was the beginning of what is now called the "Freiman's structural theory of set addition". The foundations for this theory were laid in his book "Foundations of a structural theory of set addition" (see [4]). For us, the following result from [4] is the most important.

Theorem 1.2 ([4, Theorem 1.9, page 11]) *Let K be a finite set of* integers *of size k and suppose that $|K^2| = 2k - 1 + b$, where $0 \leq b < k - 2$. Then K is a subset of an arithmetic progression*

$$P = \{a, a + q, a + 2q, \ldots, a + (k + b - 1)q\},$$

where a and q are integers with $q > 0$. In particular, $|P| \leq 2k - 3$.

By now, Freiman's theory had been extended tremendously. Freiman and many other mathematicians showed that problems in various fields may be looked at and treated as Structure Theory problems. In [5], Freiman listed, among others, the following areas and problems, whose solution was influenced by the ideas and methods of Structure Theory.

Additive Number Theory. The paper [7] of G.A. Freiman, H. Halberstam and I.Z. Ruzsa confronts the problem of showing that given a set of integers A, the set $rA = \{ra \mid a \in A\}$ contains an arithmetic progression of length L and difference d.

Combinatorial Number Theory. The paper [19] of Lipkin confronts the problem of determining the value of the *critical number* $c(G)$ of a finite abelian group G. The critical number $c(G)$ is the minimal integer n, such that if A is any subset of G of size at least n, then the set of all sums of subsets of A equals G.

Group Theory. In the paper [2], L. Brailovsky proved the following theorem.

Theorem 1.3 ([2, Lemma 1]) *A group G (finite or infinite) is central-by-finite if and only if there exists a positive integer k such that $|K^2| \leq k^2 - k$ for each subset K of G of size $|K| = k$.*

For more details, see [5], Sections 21–27. Freiman lists there also problems from additional areas: *Integer Programming*, *Probability Theory*, *Coding Theory* and *Mathematical Statistics*. This paper of Freiman served as an introductory paper to the 258th volume of the journal *Astérisque*, which was entirely dedicated to Freiman's "Structure Theory of Set Addition".

Since 1999, Freiman's theory has continued to flourish. In particular, various results of Freiman concerning subsets of integers were extended to subsets of other groups. We shall conclude this glimpse into Freiman's theory by quoting another of his basic results and its extension to abelian groups by Ben Green and Imre Z. Ruzsa.

In order to state these results, we need the following definitions. A subset P of $\mathbb{Z}$ is called a *proper d-dimensional progression* if

$$P = \{v_0 + l_1 v_1 + l_2 v_2 + \cdots + l_d v_d \mid l_i = 0, 1, \ldots, L_i - 1,\ i = 1, 2, \ldots, d\}$$

and $|P| = \prod_1^d L_i$, where v_i are integers and L_i are positive integers. A *proper d-dimensional progression* of an abelian group G is defined similarly, the only difference being that in this case $v_0, v_1, \ldots, v_d$ are elements of G. Finally, a subset C of G is called a *coset progression* if $C = P + H$, where H is a subgroup of G, P is a proper d-dimensional progression of G and if $p, p' \in P$ and $h, h' \in H$, then $p + h = p' + h'$ implies that $p = p'$ and $h = h'$. The size of C is defined as $|P|$.

In [4], Freiman proved the following theorem:

Theorem 1.4 ([4, Theorem 2.8], more generally stated) *If A is a finite subset of $\mathbb{Z}$ and $|A + A| \leq K|A|$, for a constant K, then A is contained in a proper $d(K)$-dimensional progression of size at most $f(K)|A|$, for some functions d and f.*

In 2007, Green and Ruzsa proved an analogous theorem for abelian groups.

Theorem 1.5 ([13, Theorem 1.1]) *If A is a finite subset of an abelian group G (finite or infinite) and $|A + A| \leq K|A|$, for a constant K, then A is contained in a coset progression of dimension at most $d(K)$ and size at most $f(K)|A|$, for some functions d and f.*

Many results concerning the possible orders of magnitude of the functions d and f were obtained in the literature. While the assumptions in Theorems 1.4 and 1.5 are much more general that those in Theorem 1.2 and in our analogous Theorem 3.1 for finite subsets in ordered groups (see Section 3), their conclusions concerning the structure of the set A are much less precise.

More results concerning the recent advances in the "Structure theory of set addition" may be found in T. Sanders' survey article [23].

Our joint papers [8], [9], [10] and [11] deal with problems which belong to the first major topic which was described above. Our aim in papers [8] and [9] was to extend the results of Freiman's Theorem 1.2 to finite subsets in *ordered groups* G, which are, like $\mathbb{Z}$, torsion-free and totally ordered. Assuming that a finite subset S of G satisfies the inequality $|S^2| \leq \alpha|S| + \beta$ for some α close to 3 and some small $|\beta|$, we tried to determine the *precise* structure of S. Our results apply in particular to the class of *torsion-free nilpotent groups*, which were shown to be orderable groups by K. Iwasawa, A.I. Mal'cev and B.H. Neumann (see Section 2).

The papers [10] and [11] deal with finite subsets of Baumlag-Solitar groups satisfying similar inequalities.

We continue with a survey of the main results in the above-mentioned four papers.

In order to describe the results of [8], we assume that G denotes an ordered group and S is a finite subset of G of size $k \geq 2$. It is easy to see that, like in $\mathbb{Z}$, $|S^2| \geq 2k-1$ and if $|S^2| = 2k-1$, then S is an abelian geometric progression (see Proposition 2.2). This result raises the basic question: what is the maximal upper bound on $|S^2|$ which implies that the set S is abelian? We solved this problem by proving that if $|S^2| \leq 3k-3$, then S is abelian and this bound is the best possible (see Theorem 2.6 and Example 2.7). Moreover, we extended Freiman's Theorem 1.2 to ordered groups (see Theorem 2.5).

The results of [8] are presented in Section 2.

In [9] we considered *non-abelian* subsets X of ordered *nilpotent groups G of class* 2. If $|X| = k$, then by Theorem 2.6 we must have $|X^2| \geq 3k-2$. We succeeded in determining the possible structures of X in the two smallest possible cases: $|X^2| = 3k-2$ and $|X^2| = 3k-1$ (see Section 3).

In [10] we dealt with inverse problems concerning finite subsets in *Baumslag-Solitar groups*. Baumslag-Solitar groups are denoted by $BS(n,m)$, where m and n are integers, and they are defined as follows:

$$BS(m,n) = \langle\, a, b \mid\ a^m b = b a^n \,\rangle.$$

We concentrated our attention to Baumslag-Solitar groups with $m = 1$: $BS(1,n) = \langle\, a, b \mid\ ab = ba^n \,\rangle$.

We noticed in [10] that inverse small doubling problems in these groups are related to similar problems concerning sums of dilates in Additive Number Theory. A dilate is a subset $r * A$ of $\mathbb{Z}$, where A is a finite subset of $\mathbb{Z}$, r is a positive integer and

$$r * A = \{ra \mid\ a \in A\}.$$

Finding bounds for sizes of sums of dilates of the form

$$r * A + s * A = \{ra + sb \mid\ a, b \in A\}$$

is a popular subject in Additive Number Theory. For example, it was shown in [3] that $|A+2*A| \geq 3|A|-2$ and $|A+2*A| = 3|A|-2$ if and only if A is an arithmetic progression. For more information, see M. Nathanson's book [21].

In [10], our contribution to this area consisted of two new results. First, we proved that if $|A| \geq 3$ and $|A+2*A| < 4|A|-4$, then A is contained in an arithmetic progression of size $\leq 2|A|-3$ (see Theorem 4.2). Our second result is that if $r \geq 3$, then $|A+r*A| \geq 4|A|-4$ (see Theorem 4.1).

We return now to the Baumslag-Solitar groups $BS(1,n)$. Let S be a finite subset of $BS(1,n)$ of size k contained in the coset $b^r\langle a\rangle$ of $\langle a\rangle$, where r is a positive integer. Then

$$S = \{b^r a^{x_0}, b^r a^{x_1}, \ldots, b^r a^{x_{k-1}}\},$$

where $A = \{x_0, x_1, \ldots, x_{k-1}\}$ is a set of integers. We introduce now the notation

$$S = \{b^r a^x : x \in A\} = b^r a^A.$$

Thus $|S| = |A|$. We proved the following equality: $|S^2| = |n^r * A + A|$ (see Theorem 4.3). This result served as the major means for investigating $|S^2|$ for $S \in BS(1,n)$, using information about sizes of sums of dilates.

Here is a partial list of our results. Suppose, first, that $S = ba^A$ is a subset of $BS(1,2)$. Then:

1. $|S^2| \geq 3|S|-2$ (see Proposition 4.4).
2. $|S^2| = 3|S|-2$ if and only if A is an arithmetic progression (see Proposition 4.4).
3. If $|S^2| < 4|S|-4$, then A is a subset of an arithmetic progression of size $\leq 2|A|-3$ (see Theorem 4.5).

Suppose, next, that $S = b^m a^A$ is a subset of $BS(1,2)$ with $m \geq 2$. Then:

4. $|S^2| \geq 4|S|-4$ (see Theorem 4.6).

Finally suppose that $S = ba^A$ is a subset of $BS(1,r)$ with $r \geq 3$. Then:

5. $|S^2| \geq 4|S|-4$ (see Theorem 4.7).

The results of [10] are presented in the first part of Section 4.

Finally, in [11], we proved an extended inverse theorem concerning *arbitrary* finite subsets of the *monoid* $BS^+(1,2)$. This monoid is a subset of the Baumslag-Solitar group $BS(1,2)$ and it is defined by:

$$BS^+(1,2) = \{\, g = b^m a^x \in BS(1,2) \mid x, m \text{ are integers, } m \geq 0 \,\}.$$

In particular, the set $BS^+(1,2)$ is closed with respect to multiplication and all elements can be uniquely represented by a word of the form $b^m a^x$. This property is basic for our arguments.

We proved the following result. Suppose that S is a finite non-abelian subset of $BS^+(1,2)$. Then:

6. If $|S^2| < \frac{7}{2}|S|-4$, then $S = ba^A$, where A is contained in an arithmetic progression of length less than $\frac{3}{2}|A|$ (see Theorem 4.8)

This result is best possible. Its proof required rather complicated arguments. For more details, see the second part of Section 4.

Our *second major topic* deals with the size of products XY, where X and Y are finite subsets of a group G. In the joint paper [14] with G. Kaplan, we investigated the relation between the sizes $|XY|$ and $|YX|$, and in particular we concentrated on the comparison of $|XX^{-1}|$ and $|X^{-1}X|$. As far as we know, the problems and results of our paper [14] had not been dealt with before.

Following the notation in [14], we call a finite subset X of a group G "*good*" if $|XX^{-1}| = |X^{-1}X|$ and "bad" otherwise. In Section 5 we present some results concerning "good" and "bad" subsets of G. Then in Section 6 we describe our classification of all groups in which all finite subsets are "good".

In more detail, if k is a positive integer, we say that a group G is a *$\mathcal{P}$-group* (*$\mathcal{P}_k$-group*) if each finite subset X of G (each subset X of G of size $|X| \leq k$) satisfies

$$|XX^{-1}| = |X^{-1}X|,$$

i.e., if each finite subset (finite subset of size $\leq k$) of G is "good". We shall also write G is in $\mathcal{P}$ (G is in $\mathcal{P}_k$) or $G \in \mathcal{P}$ ($G \in \mathcal{P}_k$). In [14], the $\mathcal{P}$-groups were completely classified. They are either abelian, or Hamiltonian 2-groups, or isomorphic to one of seven fixed finite groups of order at most 20 (see Theorem 6.14).

In Section 7 we describe some results which were obtained in [14], concerning $\mathcal{P}_k$-groups, where k is a positive integer. It is easy to see that every group is a $\mathcal{P}_3$-group, but there exist groups that are not in $\mathcal{P}_4$ (see Proposition 5.3). In [14], $\mathcal{P}_4$-groups were completely described. They are of one of the following three types: (i) groups in which every involution is central, (ii) groups of the form $A \rtimes \langle t \rangle$, where t is an involution and $a^t = a^{-1}$ for any $a \in A$, or (iii) some particular finite 2-groups of exponent 4 (see Theorem 7.1).

We also showed that $\mathcal{P}_6 \subset \mathcal{P}_5 \subset \mathcal{P}_4$ (see the discussion in Section 7). We have not been able to find a $\mathcal{P}$-group which is not a $\mathcal{P}_6$-group. This raises the following question:

Is every $\mathcal{P}_6$-group a $\mathcal{P}$-group?

Finally, in Section 8 we use our classification of $\mathcal{P}$-groups for the characterization of the following two related families of groups: (i) groups in which $XX^{-1} = X^{-1}X$ for *any* subset X (see Theorem 8.1), and (ii) groups satisfying $|PQ| = |QP|$ for all finite subsets P and Q (see Theorem 8.6).

The authors of this survey are grateful to the referee for his constructive remarks.

2 Products of subsets in ordered groups

We say that $(G, \leq)$ is an *ordered group* if G is a group and $\leq$ is a total order relation defined on the set G, satisfying the following condition:

$$\text{for all } a, b, x, y \in G, \qquad a \leq b \qquad \text{implies that} \qquad xay \leq xby.$$

A group G is called *orderable* if there exists a total order relation $\leq$ on the set G, such that $(G, \leq)$ is an ordered group.

It is easy to prove that an orderable group is torsion free. The converse of this statement is also true.

Theorem 2.1 (see [15], [18], [20] and [22]) *A nilpotent group is orderable if and only if it is torsion-free.*

More information about ordered groups may be found, for example, in [1] and in [12].

First we state some basic facts concerning $|S^2|$. For $G = \mathbb{Z}$, the result of next Proposition is part of the folklore of the additive number theory. The proof for ordered groups is, as a matter of fact, almost identical.

Proposition 2.2 ([8, Theorem 1.1]) *Let $(G, \leq)$ be an ordered group and let $S = \{x_1, x_2, \ldots, x_k\}$ be a finite subset of G of size k, with $x_1 < x_2 < \cdots < x_k$. Then the following statements hold:*

1. *$|S^2| \geq 2k - 1$.*
2. *If $|S^2| = 2k - 1$, then S is an abelian geometric progression.*

Our next proposition is basic for our arguments and it is also of independent interest.

Proposition 2.3 ([8, Proposition 2.3]) *Let $(G, \leq)$ be an ordered group and let S be a finite subset of G of size k. If $y \in G \setminus C_G(S)$, then*

$$|yS \cup Sy| \geq k + 1.$$

In particular, there exist $x_i, x_j \in S$ such that $yx_i \notin Sy$ and $x_j y \notin yS$.

This result implies the following corollary concerning normalizers of finite subsets in ordered groups.

Corollary 2.4 ([8, Corollary 2.4]) *Let G be an ordered group and let S be a finite subset of G. Then $N_G(S) = C_G(S)$.*

Recall that a finite subset X of G is said to satisfy the *small doubling property* if $|X^2| \leq \alpha|S| + \beta$, where α and β denote real numbers, α and $|\beta|$ small and $\alpha \geq 1$.

In [8], a joint paper with G.A. Freiman, we studied subsets of ordered groups G satisfying the small doubling property for $\alpha = 3$ and for small $|\beta|$. Our first aim was to extend Freiman's Theorem 1.2 dealing with subsets of $\mathbb{Z}$ to subsets of ordered groups. Indeed, we proved the following theorem:

Theorem 2.5 ([8, Corollary 3.4]) *Let $(G, \leq)$ be an ordered group and let $S = \{x_1, x_2, \ldots, x_k\}$ be a finite subset of G of size $k \geq 3$, with $x_1 < x_2 < \cdots < x_k$. Assume that $t = |S^2| \leq 3k - 4$.*

Then S is abelian and there exists $g \in G$, $g > 1$, such that $gx_1 = x_1 g$ and S is a subset of $\{x_1, x_1 g, x_1 g^2, \ldots, x_1 g^{t-k}\}$.

Since, by Theorem 2.5, subsets S of ordered groups satisfying $|S^2| \leq 3|S| - 4$ are abelian, we considered the ensuing question: what is the maximal upper bound on $|S^2|$ which guarantees that S is abelian? It turned out that $3|S| - 3$ is such a bound. We proved the following theorem:

Theorem 2.6 ([8, Theorem 3.2]) *Let $(G, \leq)$ be an ordered group and let S be a finite subset of G of size $k \geq 2$. Suppose that $|S^2| \leq 3k-3$. Then S is abelian.*

This result is best possible. In fact, there exist ordered groups G and finite non-abelian subsets S of G such that $|S| = k \geq 2$ and $|S^2| = 3k-2$.

Example 2.7 ([8, in the proof of Theorem 3.2]) Let $G = A \rtimes \langle b \rangle$ be a semidirect product of an abelian subgroup A, isomorphic to the additive rational group $(\mathbb{Q}, +)$, by an infinite cyclic group $\langle b \rangle$, such that $a^b = a^2$ for each $a \in A$. Then G is torsion-free and it is orderable (see [1] and [16]). Take $a \in A \setminus \{1\}$ and let $S = \{b, ba, ba^2, \ldots, ba^{k-1}\}$. Since $ab = ba^2$, it is easy to see that

$$S^2 = \{b^2, b^2a, b^2a^2, b^2a^3, \ldots, b^2a^{3k-3}\}.$$

Thus S is non-abelian and $|S^2| = 3k-2$, as required.

3 Small doubling in torsion-free nilpotent groups of class 2

Let G be a torsion-free nilpotent group of class 2. Then G is orderable and the results of the previous section apply. In particular, it follows from Theorems 2.5 and 2.6 that if $S = \{x_1, x_2, \ldots, x_k\}$ is a finite subset of G of size $k \geq 2$ with $x_1 < x_2 < \cdots < x_k$ and $|S^2| \leq 3k-3$, then S is abelian, and if $t = |S^2| \leq 3k-4$, then S is a subset of the geometric progression

$$\{x_1, x_1g, x_1g^2, \cdots, x_1g^{t-k}\}$$

for some $g \in G$, with $g > 1$ and $gx_1 = x_1g$.

In [9], a joint paper with G.A. Freiman and Y.V. Stanchescu which is under preparation, we considered the following question: what is the structure of a finite subset S of G of size k, if S is non-abelian and either $|S^2| = 3k-2$ or $|S^2| = 3k-1$?

If $|S^2| = 3k-2$ and $k > 2$, then the following two results supply a complete answer. First we consider the general case: $k \geq 4$.

Theorem 3.1 *Let G be a torsion-free nilpotent group of class 2 and let S be a non-abelian subset of G of size $k \geq 4$. Then $|S^2| = 3k-2$ if and only if*

$$S = \{a, ac, \ldots, ac^i, b, bc, bc^2, \ldots, bc^j\},$$

with $1 + i + 1 + j = k$ and $ab = bac$, $c > 1$.

The remaining case: $k = 3$ is dealt with in the following proposition:

Proposition 3.2 *Let G be a torsion-free nilpotent group of class 2 and let S be a non-abelian subset of G of size $k = 3$. Then $|S^2| = 3k-2 = 7$ if and only if one of the following holds:*

(i) $S \cap Z(\langle S \rangle) \neq \emptyset$;

(ii) $S = \{a, ac, b\}$, *with $c > 1$ and either $ab = bac$ or $ba = abc$; in particular, $c \in Z(G)$.*

If $|S^2| = 3k-1$ and $k > 3$, then again we state separately the results for $k \geq 5$ and for $k = 4$.

Theorem 3.3 *Let G be a torsion-free nilpotent group of class 2 and let S be a non-abelian subset of G of size $k \geq 5$. Then $|S^2| = 3k-1$ if and only if one of the following holds:*

(i) $S = \{a, ac, \ldots, ac^{i-1}, b, bc, \ldots, bc^{j-1}\}$, *with* $ab = bac^2$, $c > 1$;

(ii) $S = \{a, ac^2, b, bc, \ldots, bc^j\}$, $j \geq 2$, *with either* $ab = bac$ *or* $ba = abc$, $c > 1$.

In the case when $k = 4$, the situation is more complicated and more cases arise. In fact, we have:

Theorem 3.4 *Let G be a torsion-free nilpotent group of class 2 and let S be a non-abelian subset of G of size $k = 4$. Then $|S^2| = 3k - 1 = 11$ if and only if one of the following holds:*

(i) *There exist* $s, t \in S \cap Z(\langle S \rangle)$, $s \neq t$;

(ii) $S = \{a, ac, b, bc, \}$, *with* $ab = bac^2$, $c > 1$;

(iii) $S = \{a, ac^2, b, bc\}$, *with* $ba = bac$ *or* $ab = bac$, $c > 1$;

(iv) $S = \{a, ac, ac^2, b\}$, *with either* $ba = bac^2$ *or* $ab = ba^2c$, $c > 1$;

(v) $S = \{a, ac, b, x\}$, *with either* $ba = bac$ *or* $ab = bac$, $c > 1$, $ax = xa$ *and* $bx = xb$;

(vi) $S = \{a, ac, ac^2, x\}$, *with* $ac = ca$ *and there exists exactly one* $i \in \{0, 1, 2\}$ *such that* $ac^i x = xac^i$;

(vii) $S = \{a, ac, b, x\}$, *with either* $ab = bac, xa = axc$, $c > 1$ *or* $ba = abc, ax = xac, c > 1$, *and* $x = b^{-1}a^2 = a^2b^{-1}c^2$.

The following Lemma was very useful in our proofs.

Lemma 3.5 *Let G be a torsion-free nilpotent group of class 2 and let $S = \{x_1, \ldots, x_k\}$ be a subset of G satisfying $x_1 < x_2 < \cdots < x_k$ and $x_k x_{k-1} \neq x_{k-1} x_k$. Let $T = \{x_1, \ldots, x_{k-1}\}$. Then: $|S^2| \geq |T^2| + 4$. In particular, if T is non-abelian, then $|S^2| \geq 3k - 1$.*

We end this section with a consequence of Lemma 3.5, concerning completely non-abelian subsets of G.

We say that a subset S of a group G is completely non-abelian ($S \in CNA$) if $ab \neq ba$ for all $a, b \in S$, $a \neq b$.

Proposition 3.6 *Let S be a CNA-subset of a torsion-free nilpotent group of class 2 of size $|S| = k$. Then $|S^2| \geq 4k - 4$.*

This result raises the following question:

Is Proposition 3.6 true for all ordered groups?

4 Inverse problems in Baumslag-Solitar groups

We noticed in [10], a joint paper with G.A. Freiman and Y.V. Stanchescu, that there exists an important connection between results on sums of dilates in Additive Number Theory and some small doubling problems in the Baumslag-Solitar groups. So we begin this section with a short introduction into the theory of sums of dilates, including two new results which we proved in [10].

Let A be a finite set of integers and let r be positive integers. Define $r * A = \{ra \mid a \in A\}$. Such sets are called *dilates* and sums of dilates of the form

$$r * A + s * A = \{ra + sb \mid \ a, b \in A\}$$

have been studied recently by several authors. For example, it was shown in [3] that

$$|A + 2 * A| \geq 3|A| - 2$$

and $|A + 2 * A| = 3|A| - 2$ if and only if A is an arithmetic progression.

In [10], we proved the following two new results in this area:

Theorem 4.1 ([10, Theorem 6]) *Let A be a finite set of integers and let r be an integer satistying $r \geq 3$. The the following inequality holds:*

$$|A + r * A| \geq 4|A| - 4.$$

Theorem 4.2 ([10, Theorem 4]) *Let A be a finite set of integers of size $|A| \geq 3$ and suppose that $|A+2*A| < 4|A|-4$. Then A is a subset of an arithmetic progression of size $\leq 2|A| - 3$.*

We continue with the definition of *Baumslag-Solitar groups* $BS(m.n)$. These are two-generated groups with one relation which are defined as follows:

$$BS(m, n) = \langle\, a, b \mid a^m b = ba^n \,\rangle,$$

where m and n are integers. We shall concentrate our attention to the groups

$$BS(1, n) = \langle\, a, b \mid ab = ba^n \,\rangle.$$

We shall describe now the connection between small doubling problems in the Baumslag-Solitar groups $BS(1, n)$ and sums of dilates.

Let S be a finite subset of $BS(1, n)$ of size k_1 contained in the coset $b^r\langle a\rangle$ for some positive integer r and let T be a finite subset of $BS(1, n)$ of size k_2 contained in the coset $b^s\langle a\rangle$ for some positive integer s. Then

$$S = \{b^r a^{x_0}, b^r a^{x_1}, \ldots, b^r a^{x_{k_1-1}}\},$$

where $A = \{x_0, x_1, \ldots, x_{k_1-1}\}$ is a set of integers. We introduce now the notation

$$S = \{b^r a^x : x \in A\} = b^r a^A.$$

Thus $|S| = |A|$.

Similarly, $T = b^s a^B$ for some set of integers $B = \{y_0, y_1, \ldots, y_{k_2-1}\}$. Since $ab = ba^n$, it follows that $a^{-1}b = ba^{-n}$ and

$$a^x b^t = b^t a^{n^t x} \qquad \text{for each } x \in \mathbb{Z} \text{ and } t \in \mathbb{N}. \tag{4.1}$$

In particular, $a^x b = ba^{nx}$ for each $x \in \mathbb{Z}$. Equation (4.1) implies that

$$(b^r a^x)(b^s a^y) = b^r (a^x b^s) a^y = b^r (b^s a^{n^s x}) a^y = b^{r+s} a^{n^s x + y}$$

for each $x, y \in \mathbb{Z}$ and for each $r, s \in \mathbb{N}$. Therefore the *product set* $ST = \{vw \mid v \in S,\ w \in T\}$ can be written as

$$\begin{aligned} ST &= \{(b^r a^{x_i})(b^s a^{y_j}) \mid\ i \in \{0, 1, \ldots, k_1 - 1\},\ j \in \{0, 1, \ldots, k_2 - 1\}\} \\ &= \{b^{r+s} a^{n^s x_i + y_j} \mid\ i \in \{0, 1, \ldots, k - 1\}\},\ j \in \{0, 1, \ldots, k_2 - 1\}\} \\ &= b^{r+s} a^{n^s * A + B} \end{aligned} \tag{4.2}$$

and $|ST| = |n^s * A + B|$.

We proved the following basic theorem.

Theorem 4.3 ([10, Theorem 1]) *Suppose that $S = b^r a^A \subseteq BS(1, n)$, $T = b^s a^B \subseteq BS(1, n)$ where $r, s \in \mathbb{N}$ and A, B are finite subsets of $\mathbb{Z}$. Then*

$$ST = b^{r+s} a^{n^s * A + B} \qquad \textit{and} \qquad |ST| = |n^s * A + B|.$$

In particular,

$$S^2 = b^{2r} a^{n^r * A + A} \qquad \textit{and} \qquad |S^2| = |n^r * A + A|.$$

This result served us as the major means for investigating $|ST|$, and in particular $|S^2|$, using known information about sizes of sums of dilates. For example, Theorem 4.3 and the above mentioned results concerning $|A + 2 * A|$ yield the following proposition:

Proposition 4.4 ([10, Theorem 2(a)]) *Let A be a finite set of integers and let $S = ba^A \subseteq BS(1, 2)$. Then*

$$|S^2| = |A + 2 * A| \geq 3|S| - 2.$$

Moreover, $|S^2| = 3|S| - 2$ if and only if A is an arithmetic progression.

Moreover, Theorem 4.3 and Theorems 4.1 and 4.2 yield the following additional results:

Theorem 4.5 ([10, Theorem 4]) *Let A be a finite set of integers of size k and let $S = ba^A \subseteq BS(1, 2)$. Suppose that $|S^2| < 4k - 4$. Then A is a subset of an arithmetic progression of size $\leq 2k - 3$.*

Proof By Theorem 4.3, $|S^2| = |A + 2 * A| < 4k - 4$ and Theorem 4.2 implies that A is a subset of an arithmetic progression of size $\leq 2k - 3$. □

Theorem 4.6 ([10, Theorem 6]) *Let A be a finite set of integers of size k and let $S = b^m a^A \subseteq BS(1,2)$, where $m \geq 2$. Then $|S^2| \geq 4k-4$.*

Proof By Theorem 4.3 $|S^2| = |A + 2^m * A|$. Since $m \geq 2$, $2^m \geq 4 > 3$ and Theorem 4.1 implies that $|S^2| \geq 4k-4$. □

Theorem 4.7 ([10, Corollary 1]) *Let A be a finite set of integers of size k and let $S = ba^A \subseteq BS(1,r)$ for $r \geq 3$. Then $|S^2| \geq 4k-4$.*

Proof By Theorem 4.3 $|S^2| = |A + r * A|$ and Theorem 4.1 implies that $|S^2| \geq 4k-4$. □

Until now, we have obtained inverse results concerning subsets of $BS(1,2)$ which were contained in one coset of $\langle a \rangle$. In [11], a joint paper with G A. Freiman and Y.V. Stanchescu under preparation, we dealt with an inverse problem concerning *arbitrary* non-abelian finite sets S contained in the following subset of $BS(1,2)$:

$$BS^+(1,2) = \{g = b^m a^x \in BS(1,2) \mid \ x, m \text{ are integers, } m \geq 0\}.$$

This subset of $BS(1,2)$ is closed with respect to multiplication, so it constitutes a monoid.

We assumed that a finite subset S of $BS^+(1,2)$ satisfies the following small doubling condition: $|S^2| < \frac{7}{2}k - 4$. Using rather complicated arguments, we proved the following theorem.

Theorem 4.8 *Let S be a finite non-abelian subset of $BS^+(1,2)$ of size k and suppose that $|S^2| < \frac{7}{2}k - 4$. Then $S = ba^A$, where A is a set of integers of size k which is contained in an arithmetic progression of size less than $\frac{3}{2}k$.*

The result is best possible. In fact, there exist non-abelian subsets S of $BS^+(1,2)$ satisfying $|S^2| = \frac{7}{2}k - 4$, which are not contained in one coset of the cyclic subgroup $\langle a \rangle$ of $BS^+(1,2)$.

Example 4.9 Theorem 4.8 is *optimal* in view of the following example:

$$S = a^{A_0} \cup \{b\} \subset BS^+(1,2),$$

where $A_0 = \{0, 1, 2, \ldots, k-2\}$ and $k > 2$ is even. The set S is clearly non-abelian and not contained in one coset of $\langle a \rangle$. Moreover,

$$S^2 = a^{A_0} a^{A_0} \cup b a^{A_0} \cup a^{A_0} b \cup \{b^2\}.$$

Using $a^{A_0} b = b a^{2*A_0}$, we get

$$S^2 = a^{A_0 + A_0} \cup (b a^{A_0} \cup b a^{2*A_0}) \cup \{b^2\} = a^{A_0+A_0} \cup b a^{A_0 \cup 2*A_0} \cup \{b^2\}.$$

Since $a^{A_0+A_0} \subseteq a^{\mathbb{Z}}$, $ba^{A_0 \cup 2*A_0} \subseteq ba^{\mathbb{Z}}$, $\{b^2\} \subseteq b^2 a^{\mathbb{Z}}$, the three components of S^2 are disjoint in pairs and hence

$$|S^2| = |A_0 + A_0| + |A_0 \cup 2 * A_0| + 1 = (2k-3) + (\tfrac{3}{2}k - 2) + 1 = \tfrac{7}{2}k - 4.$$

Remark 4.10 We noticed that the Baumslag-Solitar groups $BS(1,n)$ are metabelian. They are semidirect products of the normal closure of $\langle a\rangle$, which is isomorphic to the n-adic rational group, by $\langle b\rangle$, which acts by raising to n-th power (see, for example, the Encyclopedia of Mathematics, under "Baumslag-Solitar group"). Since the n-adic rational group is torsion-free and abelian, it is orderable, and using a theorem of Kargapolov in [16] (see also [8], Theorem K), it follows that this order can be extended to an order in $BS(1,n)$. Hence *the groups* $BS(1,n)$ *are orderable.*

5 Subsets X satisfying $|XX^{-1}| = |X^{-1}X|$

Let G be a group and let X, Y denote *finite subsets* of G. In this paragraph we describe our results in [14], a joint paper with G. Kaplan, considering the relationship between the orders $|XY|$ and $|YX|$. In particular, we are interested in comparison of $|XX^{-1}|$ and $|X^{-1}X|$.

Let G be a group and let X be a finite subset of G.

Question 5.1 Is $|XX^{-1}| = |X^{-1}X|$?

The answer is clearly "yes" if X is a coset gH of a finite subgroup H of G. In fact, in this case $X^{-1} = Hg^{-1}$, $XX^{-1} = gHg^{-1}$ and $X^{-1}X = H$. Thus $|XX^{-1}| = |H| = |X^{-1}X|$.

If X is a union of two cosets of H, say, $X = H \cup gH$ for some $g \in G \setminus H$, then the equality is not always true. In fact, we have:

Example 5.2 ([14, in the proof of Proposition 2.4])
Consider the semi-dihedral group of order 16:

$$G = \langle a, b \mid a^8 = b^2 = 1, a^b = a^3\rangle$$

and let $X = \langle b\rangle \cup a^{-1}\langle b\rangle = \{1, b, a^{-1}, a^{-1}b\}$. Then $X^{-1} = \{1, b, a, b^{-1}a\}$,

$$X^{-1}X = \{1, b, a, a^3, a^5, a^7, a^3b, ab, a^5b, a^7b\}$$

and

$$XX^{-1} = \{1, b, a, a^7, a^3b, a^2b, a^7b\}.$$

Thus $X^{-1}X$ is of size 10, while XX^{-1} is of size 7.

In this section we present some basic results concerning XX^{-1} and $X^{-1}X$. We call a finite subset X of a group G "*good*" if $|XX^{-1}| = |X^{-1}X|$ and "*bad*" otherwise. We have:

Proposition 5.3 ([14, Proposition 2.4]) *Let X be a finite subset of a group G and suppose that one of the following holds:*

1. $|X| \leq 3$;
2. *G is finite and $|X| > |G|/2$.*

Then X is "good".

Moreover, if $|X| = 4$, then X may be "bad".

We also proved:

Theorem 5.4 ([14, Theorem 2.1(3)]) *If G is a finite group of order n and X is a subset of G satisfying $|X| > n/k$ for some integer $k \geq 2$, then*

$$|XX^{-1}| \geq \frac{n}{k-1} \quad \textit{and} \quad |X^{-1}X| \geq \frac{n}{k-1}.$$

Furthermore, we mention the following two results of T. Tao.

Theorem 5.5 ([25]) *Let X be a finite subset of a group G. Then $|X^2| < 2|X|$ implies $XX^{-1} = X^{-1}X$.*

Theorem 5.6 ([24]) *Let X be a finite subset of a group G and suppose that $|XX^{-1}| < \frac{3}{2}|X|$. Then XX^{-1} and $X^{-1}X$ are conjugate subgroups of G.*

6 Groups with all finite subsets X satisfying $|XX^{-1}| = |X^{-1}X|$

In this section we continue to present results from [14].

In the previous section we have seen that a finite subset of a group G can be "bad". So another question arises:

Question 6.1 Which groups G satisfy $|XX^{-1}| = |X^{-1}X|$ for *all* finite subsets X of G?

We recall our definition of *$\mathcal{P}$-groups*.

Definition 6.2 Let k denote a positive integer. A group G is called a *$\mathcal{P}$-group* (*$\mathcal{P}_k$-group*) if each finite subset X of G (each subset X of G of size $|X| \leq k$) satisfies

$$|XX^{-1}| = |X^{-1}X|.$$

We shall also write G is in $\mathcal{P}$ (G is in $\mathcal{P}_k$) or $G \in \mathcal{P}$ ($G \in \mathcal{P}_k$).

Remark 6.3 If G is an abelian group, then obviously $XX^{-1} = X^{-1}X$ for *each* subset X of G. Therefore, in particular, all abelian groups are $\mathcal{P}$-groups.

Other examples of $\mathcal{P}$-groups are the Dedekind 2-groups. They also satisfy the stronger condition: $XX^{-1} = X^{-1}X$ for *each* subset X.

Proposition 6.4 ([14, Proposition 6.1]) *Let G be a Dedekind 2-group. Then $XX^{-1} = X^{-1}X$ for* each *subset X of G.*

Therefore, in particular, all Dedekind 2-groups are $\mathcal{P}$-groups.

Notice that a Dedekind group need not be a $\mathcal{P}$-group. For instance, $C_3 \times Q_8 \notin \mathcal{P}$. Furthermore, there are non-abelian $\mathcal{P}$-groups which are not 2-groups. In fact we have:

Example 6.5 ([14, Proposition 5.3]) Let $G = \langle a \rangle \rtimes \langle c \rangle$, where $|a| \in \{3, 5\}$, $|c| \in \{2, 4\}$ and $a^c = a^{-1}$. Then G is a $\mathcal{P}$-group.

There are also non-Dedekind 2-groups which are $\mathcal{P}$-groups.

Example 6.6 ([14, Lemma 6.5]) The following 2-groups are $\mathcal{P}$-groups: the dihedral group D_8, the quaternion group Q_{16} and the group $G = \langle a \rangle \rtimes \langle b \rangle$, where $|a| = |b| = 4$ and $a^b = a^{-1}$.

Moreover, the following 2-group is not a $\mathcal{P}$-group.

Example 6.7 Let $G = (\langle a \rangle \times \langle s \rangle) \rtimes \langle b \rangle$, where $|a| = 4$, $|s| = |b| = 2$, $a^b = a^{-1}$ and $s^b = a^2 s$. Then G is not a $\mathcal{P}$-group.

If G is a $\mathcal{P}$-group ($\mathcal{P}_k$-group) and H is a subgroup of G, then also H is a $\mathcal{P}$-group ($\mathcal{P}_k$-group). Thus the classes $\mathcal{P}$ and $\mathcal{P}_k$ are closed with respect to the subgroup operation, but they are not closed with respect to the quotient operation. Nevertheless, we were able to prove the following result.

Proposition 6.8 ([14, Corollary 3.2]) *Let G be a $\mathcal{P}$-group, H a subgroup of G and suppose that $\langle h \rangle$ is normal in G for each $h \in H$. Then G/H is a $\mathcal{P}$-group.*

Our main result in [14] was the classification of the $\mathcal{P}$-groups. We dealt separately with the following three complementary cases: finite groups, infinite periodic groups and non-periodic groups. Our classification consists of three theorems, each corresponding to one of these cases.

Theorem 6.9 ([14, Theorem 7.15]) *Let G be a finite group. Then G is a $\mathcal{P}$-group if and only if one of the following holds:*

(i) G is abelian;

(ii) G is a finite Hamiltonian 2-group;

(iii) G is isomorphic to one of the following seven groups:

$$D_6, \quad D_8, \quad D_{10}, \quad Q_{16}, \quad \langle a_3 \rangle \rtimes \langle b \rangle, \quad \langle a_4 \rangle \rtimes \langle b \rangle, \quad \langle a_5 \rangle \rtimes \langle b \rangle,$$

where $|a_3| = 3$, $|a_4| = 4$, $|a_5| = 5$, $|b| = 4$ and $a_i^b = a_i^{-1}$ for $i = 3, 4, 5$.

In the infinite case, we start with the study of periodic $\mathcal{P}$-groups. First we have:

Proposition 6.10 ([14, Proposition 8.1]) *Let G be a periodic $\mathcal{P}$-group. Then G is locally-finite.*

Now it is not difficult to obtain the following complete classification of infinite periodic $\mathcal{P}$-groups.

Theorem 6.11 ([14, Theorem 8.2]) *Let G be an infinite periodic group. Then G is a $\mathcal{P}$-group if and only if G is either an abelian group or a Hamiltonian 2-group.*

In order to study the non-periodic $\mathcal{P}$-groups, we first consider torsion-free groups. In this case, we proved the following results.

Proposition 6.12 ([14, Proposition 9.3]) *Let G be a torsion-free $\mathcal{P}$-group. Then $[\langle b \rangle : \langle b \rangle \cap \langle b \rangle^a]$ is finite for any $a, b \in G$.*

Groups G satisfying the property that $[\langle b\rangle : \langle b\rangle \cap \langle b\rangle^a]$ is finite for any $a, b \in G$ had been called *C_2-groups* and have been studied by Lennox, Longobardi, Maj, Smith, Wiegold in [17]. They proved that *if a C_2-group G is torsion-free and finitely generated, then $G/Z(G)$ is periodic.* Using this result, we are able to prove the following classification of non-periodic $\mathcal{P}$-groups.

Theorem 6.13 ([14, Theorem 9.1]) *Let G be a non-periodic group. Then G is a $\mathcal{P}$-group if and only if G is abelian.*

To summarize, we have established the following complete classification of $\mathcal{P}$-groups.

Theorem 6.14 ([14, Theorems 7.15, 8.2 and 9.1]) *The group G is a $\mathcal{P}$-group if and only if one of the following statements holds:*

1. *G is abelian;*
2. *G is a Hamiltonian 2-group;*
3. *G is one of the following seven fixed non-abelian finite groups of orders* 6, 8, 10, 12, 16, 16 *and* 20*:*

$$D_6, \quad D_8, \quad D_{10}, \quad \langle a_3\rangle \rtimes \langle b\rangle, \quad Q_{16}, \quad \langle a_4\rangle \rtimes \langle b\rangle, \quad \langle a_5\rangle \rtimes \langle b\rangle,$$

where $|a_3| = 3$, $|a_4| = 4$, $|a_5| = 5$, $|b| = 4$ and $a_i^b = a_i^{-1}$ for $i = 3, 4, 5$.

7 $\mathcal{P}_k$-groups

Obviously $\mathcal{P}_1$ is the class of all groups. Moreover, by Proposition 5.3, $\mathcal{P}_2 = \mathcal{P}_3$ are again classes of all groups and there exist groups that are not $\mathcal{P}_4$-groups. Therefore the following problem arises:

Determine the class of $\mathcal{P}_4$-groups.

In [14], a joint paper with G. Kaplan, we characterized these groups. In fact, we proved the following theorem:

Theorem 7.1 ([14, Theorem 10.4]) *A group G is a $\mathcal{P}_4$-group if and only if one of the following holds:*

(i) every involution is central in G;

(ii) $G = A \rtimes \langle t\rangle$, with A abelian, $t^2 = 1$ and $a^t = a^{-1}$ for each $a \in A$;

(iii) G is a 2-group of exponent 4 and $G = E \times H$, where E is elementary abelian and either $H = V$, an extraspecial group, or H is the central product of a cyclic group $\langle s\rangle$ of order 4 and an extraspecial group V, with $\langle s^2\rangle$ and V^2 amalgamated.

In our proof we use the following result.

Theorem 7.2 ([14, Theorem 10.2]) *Let G be a group and let X be a subset of G such that $1 \in X$ and $|X| = 4$. Then $|XX^{-1}| \neq |X^{-1}X|$ if and only if $X = \{1, y, t, z\}$, where $|y| > 2$, t is an involution, $z \in \{yt, ty\}$ and $y^t \notin \{y, y^{-1}\}$.*

This theorem gives rise to another characterization of $\mathcal{P}_4$-groups.

Theorem 7.3 ([14, Theorem 10.3]) *Let G be a group. Then $G \in \mathcal{P}_4$ if and only if for any $y, t \in G$ with $|y| > 2$ and t an involution, we have $y^t \in \{y, y^{-1}\}$.*

Next we notice that $\mathcal{P}_5 \subset \mathcal{P}_4$. In fact we have:

Proposition 7.4 ([14, Proposition 4.10]) *Let $G = A\langle b\rangle$, where $A \leq G$, $b \in G$, A is not an elementary abelian 2-group, $a^b = a^{-1}$ for each $a \in A$ and $b^2 \notin A$ if $|b| = 4$. If L is any group which is not an elementary abelian 2-group, then $G \times L \notin \mathcal{P}_5$.*

Thus, if every involution in G is central and L is of odd order, then $G \times L \in \mathcal{P}_4$ by Theorem 7.1, but $G \times L \notin \mathcal{P}_5$ by Proposition 7.4.

So we may ask:

Question 7.5 Which groups belong to $\mathcal{P}_5$?

Moreover, by definition

$$\mathcal{P} \subseteq \cdots \subseteq \mathcal{P}_{m+1} \subseteq \mathcal{P}_m \subseteq \cdots \subseteq \mathcal{P}_3 \subseteq \mathcal{P}_2 \subseteq \mathcal{P}_1.$$

So we may also ask:

Question 7.6 Is $\mathcal{P} = \mathcal{P}_m$ for some integer m? In other words, does there exist an integer m such that every finite group G satisfying $G \in \mathcal{P}_m$, satisfies also $G \in \mathcal{P}$?

Notice that $\mathcal{P}_5 \subset \mathcal{P}_4 \subset \mathcal{P}_3 = \mathcal{P}_2 = \mathcal{P}_1 =$ class of all groups.

We also have $\mathcal{P}_6 \subset \mathcal{P}_5$. In fact we have:

Proposition 7.7 ([14, Proposition 11.1]) *Let $G = A \rtimes \langle b\rangle$, where $A \leq G$, $|b| = 2$ and $a^b = a^{-1}$ for each $a \in A$. Then $G \in \mathcal{P}_5$.*

On the other hand, if $G = \langle a\rangle \rtimes \langle b\rangle$, where $a, b \in G$, $|a| > 5$, $|b| = 2$ and $a^b = a^{-1}$, then $G \notin \mathcal{P}_6$. Indeed, if $S = \{1, a, a^3\}$ and $X = S \cup Sb$, one can easily verify that $|XX^{-1}| \leq 13$, while $|X^{-1}X| = 15$.

Therefore, if $\mathcal{P} = \mathcal{P}_m$ for some integer m, then $m \geqslant 6$.

Some of our results led us to the following *conjecture*:

G is a $\mathcal{P}$-group if and only if G is a $\mathcal{P}_6$-group.

Our conjecture is also supported by the following new result.

Theorem 7.8 *Let G be a periodic group. If G is a $\mathcal{P}_6$-group, then every element of G of odd order generates a normal subgroup of G.*

Indeed, let $G \in \mathcal{P}_6$ be a periodic group without elements of even order. Then, by Theorem 7.8, every element of G generates a normal subgroup of G. Thus G is a Dedekind group without elements of order 2, hence abelian. In particular, G is a $\mathcal{P}$-group. We state this fact as a corollary.

Corollary 7.9 *Let $G \in \mathcal{P}_6$ be a periodic group without elements of even order. Then G is a $\mathcal{P}$-group.*

We conclude this section with a proof of Theorem 7.8.

Proof Suppose that $G \in \mathcal{P}_6$. We need to show that $\langle b\rangle$ is normal in G for any element $b \in G$ of odd order. We clearly may assume that G is not a 2-group.

First suppose that $|b| = 3$ and let $H = \langle b\rangle$ and $X = H \cup a^{-1}H$. As $|X| = 6$ and $G \in \mathcal{P}_6$, it follows that $|XX^{-1}| = |X^{-1}X|$ and by Theorem 2.1 of [14], $|H|$ is even, a contradiction.

So suppose that $|b| > 3$. Since G is a $\mathcal{P}_4$-group and it is not a 2-group, it follows by Theorem 7.1 that either $G = A\langle t\rangle$, with A abelian, t an involution and $c^t = c^{-1}$ for any $c \in A$, or every involution is central in G. In the first case obviously $\langle b\rangle$ is normal in G. So we may assume that every involution is central in G.

Suppose that there exists $a \in G \setminus N_G(\langle b\rangle)$. Write $S = \{1, b, b^2\}$ and $X = S \cup aS$. As $|X| = 6$ and $G \in \mathcal{P}_6$, it follows that $|XX^{-1}| = |X^{-1}X|$. Now $SS^{-1} = S^{-1}S = \{1, b, b^{-1}, b^2, b^{-2}\}$ and $|SS^{-1}| = 5$. Furthermore, since $a \notin \langle b\rangle$, we have:

$$XX^{-1} = (SS^{-1} \cup aSS^{-1}a^{-1}) \,\dot\cup\, (aSS^{-1} \cup SS^{-1}a^{-1})$$

and

$$X^{-1}X = S^{-1}S \,\dot\cup\, (S^{-1}aS \cup S^{-1}a^{-1}S).$$

Therefore $|XX^{-1}| \leq 9 + 5 + 5 = 19$.

Next we show that $|X^{-1}X| = 23$, a contradiction.

Our first claim is that $S^{-1}aS \cap S^{-1}a^{-1}S = \emptyset$. Assume, to the contrary, that there exist integers i, j, h, k such that $(i, j) \neq (h, k)$ and $b^{-i}ab^j = b^{-h}a^{-1}b^k$. Then $ab^{h-i}a = b^{k-j}$ and since also $a^{-1} \in G \setminus N_G(\langle b\rangle)$, we may assume that $i \neq h$. Since $(b^{h-i}a)^2 = b^{k-j+h-i}$, b is of odd order and the element $b^{h-i}a$ is non-trivial as $a \notin \langle b\rangle$, it follows that $b^{h-i}a$ has order $2d$, where d is an odd integer. Thus $(b^{h-i}a)^d$ is an involution and hence $(b^{h-i}a)^d \in Z(G)$. As $(b^{h-i}a)^2 \in C_G(\langle b\rangle)$, it follows that $[b^{h-i}a, b] = 1$, so $[a, b] = 1$ and $a \in N_G(\langle b\rangle)$, a contradiction. This proves our claim.

Next we claim that $|S^{-1}aS| = 9$. Assume, to the contrary, that there exist $i, j, h, k \in \{0, 1, 2\}$, such that $(i, j) \neq (h, k)$ and $b^{-i}ab^j = b^{-h}ab^k$. Then $a^{-1}b^{h-i}a = b^{k-j}$ and arguing as in the previous paragraph, we may assume that $h > i$. Thus $0 < h - i \leq 2$ and b being of odd order implies that $\langle b\rangle = \langle b^2\rangle$. Consequently $a^{-1}\langle b\rangle a \subseteq \langle b\rangle$ and $a^{-1}\langle b\rangle a = \langle b\rangle$ since $\langle b\rangle$ is finite. So $a \in N_G(\langle b\rangle)$, a contradiction. This proves our second claim.

Hence $|S^{-1}aS| = |S^{-1}a^{-1}S| = 9$ and $|X^{-1}X| = 5 + 9 + 9 = 23$, as claimed.

It follows from this contradiction that $a \in G \setminus N_G(\langle b\rangle)$ does not exist and hence $\langle b\rangle$ is normal in G, as required. □

8 Two applications

Using Proposition 6.4 and Theorem 6.14, we can prove two interesting theorems. First of all, we have the following new result.

Theorem 8.1 *The class of groups G satisfying*

$$XX^{-1} = X^{-1}X \tag{8.1}$$

for all finite subsets X of G coincides with the class of groups G satisfying (8.1) *for all (finite or infinite) subsets X of G and consists of two families of groups: the abelian groups and the Hamiltonian 2-groups.*

Proof It is clear that groups G satisfying (8.1) either for all finite subsets X of G or for all subsets X of G, are $\mathcal{P}$-groups. Now, the class of $\mathcal{P}$-groups consists, by Theorem 6.14, of the abelian groups, the Hamiltonian 2-groups and of the seven non-abelian groups: D_6, D_8, D_{10}, Q_{16}, $\langle a_3\rangle \rtimes \langle b\rangle$, $\langle a_4\rangle \rtimes \langle b\rangle$, $\langle a_5\rangle \rtimes \langle b\rangle$, $|a_i| = i$, $|b| = 4$, $a_i^b = a_i^{-1}$. Moreover, by Proposition 6.4, the abelian groups and the Hamiltonian 2-groups G satisfy (8.1) for all subsets X of G and in particular for all finite subsets X of G.

Thus it follows that the two classes coincide and consist of the abelian groups and the Hamiltonian 2-groups, provided that each of the seven groups mentioned above does not satisfy (8.1) for all finite subsets X.

Now, each of these seven groups contains two non-commuting elements a, b such that $a^2 \neq b^2$. Consider $X = \{a, b\}$. Then: $XX^{-1} = \{1, ab^{-1}, ba^{-1}\}$ and $X^{-1}X = \{1, a^{-1}b, b^{-1}a\}$. But, $ab^{-1} \neq 1$ and $ab^{-1} \neq b^{-1}a$ since a, b are non-commuting. Moreover, also $ab^{-1} \neq a^{-1}b$ since $a^2 \neq b^2$. Hence $ab^{-1} \notin X^{-1}X$ and $XX^{-1} \neq X^{-1}X$, as required. The proof of Theorem 8.1 is complete. □

The second application of our classification of $\mathcal{P}$-groups was proved in [14]. It concerns the following class of groups.

Definition 8.2 A group G is called a $\mathcal{Q}$-group if $|AB| = |BA|$ for all finite subsets A and B of G.

Example 8.3 Abelian groups are $\mathcal{Q}$-groups.

Remark 8.4 $Q_8 \notin \mathcal{Q}$. For, if $A = \{1, i, -i, j\}$ and $B = \{i, j, k\}$, then $AB = Q_8 \setminus \{-i\}$, while $BA = Q_8$.

Using Theorem 6.14 we are able to prove the following result.

Proposition 8.5 ([14, Proposition 12.2]) *Let G be a finite $\mathcal{Q}$-group. Then G is abelian.*

The following general theorem follows rather easily from Proposition 8.5.

Theorem 8.6 ([14, Theorem 12.1]) *If G is a $\mathcal{Q}$-group, then G is abelian.*

References

[1] R. Botto Mura & A. Rhemtulla, *Orderable groups*, Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, Inc., New York and Basel, 1977.

[2] L. Brailovsky, Combinatorial conditions forcing commutativity of an infinite group, *J. Algebra* **165** (1994), 394–400.

[3] J. Cilleruelo, M. Silva & C. Vinuesa, A sumset problem, *J. Comb. Number Theory* **2** (2010), 79–89.
[4] G.A. Freiman, *Foundations of a structural theory of set addition*, Translations of mathematical monographs **37**, Amer. Math. Soc., Providence, 1973.
[5] G.A. Freiman, Structure theory of set addition, *Astérisque* **258** (1999), 1–33.
[6] G.A. Freiman, On finite subsets of non-abelian groups with small doubling, *Proc. Amer. Math. Soc.* **140** (2012), 2997–3002.
[7] G.A. Freiman, H. Halberstam & I.Z. Ruzsa, Integer sum sets containing long arithmetic progressions, *J. London Math. Soc.* **46** (1992), 193–201.
[8] G.A. Freiman, M. Herzog, P. Longobardi & M. Maj, Small doubling in ordered groups, *J. Aust. Math. Soc.* **96** (2014), 316–325.
[9] G.A. Freiman, M. Herzog, P. Longobardi, M. Maj & Y.V. Stanchescu, Small doubling in ordered nilpotent groups of class 2, in preparation.
[10] G.A. Freiman, M. Herzog, P. Longobardi, M. Maj, Y.V. Stanchescu, Direct and inverse problems in additive number theory and in non-abelian group theory, *European J. Combin.* **40** (2014), 42–54.
[11] G.A. Freiman, M. Herzog, P. Longobardi, M. Maj & Y.V. Stanchescu, A small doubling structure theorem in a Baumslag-Solitar group, in preparation.
[12] A.M.W. Glass, *Partially ordered groups*, Series in Algebra 7, World Scientific Publishing Co., 1999.
[13] B.J. Green & I.Z. Ruzsa, Freiman's theorem in an arbitrary abelian group, *J. London Math. Soc. (2)* **75** (2007), 163–175.
[14] M. Herzog, G. Kaplan, P. Longobardi & M. Maj, Products of subsets of groups by their inverses, *Beitr. Algebra Geom.*, to appear.
[15] K. Iwasawa, On linearly ordered groups, *J. Math. Soc. Japan* **1** (1948), 1–9.
[16] M.I. Kargapolov, Completely orderable groups, *Algebra i Logica Sem.* **1** (1962), 16–21.
[17] J.C. Lennox, P. Longobardi, M. Maj, H. Smith & J. Wiegold, Some finiteness conditions concerning intersections of conjugates of subgroups, *Glasgow Math. J.* **37** (1995), 327–335.
[18] F.W. Levi, Arithmetische Gesetze im Gebiete diskreter Gruppen, *Rend. Circ. Mat. Palermo* **35** (1913), 225–236.
[19] E. Lipkin, Subset sums of sets of residues, *Astérisque* **258** (1999), 187–193.
[20] A.I. Mal'cev, On ordered groups, *Izv. Akad. Nauk. SSSR Ser. Mat.* **13** (1948), 473–482.
[21] M.B. Nathanson, *Additive number theory: Inverse problems and geometry of sumsets*, Springer, New York, 1996.
[22] B.H. Neumann, On ordered groups, *Amer. J. Math.* **71** (1949), 1–18.
[23] T. Sanders, The structure theory of set addition revisited, *Bull. Amer. Math. Soc.* **50** (2013), 93–127.
[24] T.C. Tao, Product set estimates for non-commutative groups, *Combinatorica* **28** (2008), 547–594.
[25] T.C. Tao, An elementary non-commutative Freiman theorem, http://terrytao.wordpress.com/2009/11/10/an-elementary-non-commutative-freiman-theorem (2009), 1–6.

FORMAL LANGUAGES AND GROUP THEORY

SAM A. M. JONES and RICHARD M. THOMAS

Department of Computer Science, University of Leicester, Leicester LE1 7RH, U.K.

1 Introduction

Our aim is to explore some connections between word problems of groups and formal language theory. One question is whether any finitely presented group has a recursive word problem, i.e., if there is an algorithm to decide if a given word in the generators of such a group represents the identity; this was shown not to be the case by Novikov and Boone independently [37, 6]. A finitely presented group has a recursively enumerable word problem however, i.e., there is a process listing the words representing the identity; the process will not terminate (there are infinitely many such words) but any word representing the identity will eventually appear.

We are interested in relating the complexity of the word problem (as a formal language) to the algebraic structure of the group. With regards to the classes of languages we have just mentioned, there is the beautiful Higman embedding theorem [16] which says that a finitely generated group has a recursively enumerable word problem if and only if it can be embedded in a finitely presented group. For recursive languages it was shown by Boone and Higman [7] that a finitely generated group has a recursive word problem if and only if it can be embedded in a simple group which can, in turn, be embedded in a finitely presented group. This was strengthened by Thompson [44] who showed that a finitely generated group has a recursive word problem if and only if it can be embedded in a finitely generated simple group which can, in turn, be embedded in a finitely presented group. There is a natural (and seemingly difficult) question (attributed to Higman) which asks if we can strengthen this further by proving that every finitely generated group with a recursive word problem can be embedded in a finitely presented simple group.

We will survey some work concerning groups whose word problem is a simpler type of language. We will concentrate on the class of context-free languages, some subclasses of the context-free languages and some other related classes (such as intersections and complements of context-free languages); there are many other interesting classes of languages we have not discussed such as the the class of real-time languages (see [18, 20, 23] for example) the class of growing context-sensitive languages (see [22, 29]) and the class of context-sensitive languages (see [31, 41]).

To make the paper reasonably self-contained (from a group theorist's perspective) we summarize the concepts from formal language theory we need; we introduce the basic definitions in Section 2 and then discuss some operations on languages and the closure properties of classes of languages in Sections 3 and 4 respectively. In Section 5 we discuss some general issues concerning the connections between formal languages and word problems of groups and then turn in Section 6 to some characterizations of groups whose word problem lies in some specified class of languages. We finish the paper with a discussion of some decidability issues in Section 7.

2 Formal language theory

In this section we will survey some of the concepts, notation and results we need from formal language theory; for further information see [4, 24, 26].

If Σ is a finite set (or *alphabet*), let Σ^* denote the set of all *words*, i.e., finite strings of symbols from Σ, including the *empty word* ϵ, and Σ^+ the set of all non-empty finite strings of symbols from Σ. A subset of Σ^* is called a *language* (or, if we want to stress the set Σ, a *language over* Σ). If α is a word $a_1 a_2 \dots a_n$ in Σ^* with $n \geqslant 1$ (where $a_i \in \Sigma$ for each i) we denote the length n of α by $|\alpha|$ and the number of occurrences of the symbol x in α (where $x \in \Sigma$) by $|\alpha|_x$ (we define $|\epsilon|$ and $|\epsilon|_x$ to be 0 for any $x \in \Sigma$). A *factor* of $\alpha = a_1 a_2 \dots a_n$ is a sequence $a_i a_{i+1} \dots a_{j-1} a_j$ of consecutive characters from α for some $1 \leqslant i \leqslant j \leqslant n$ (the only factor of ϵ is ϵ).

We introduce the first of our notions of some sort of abstract "machine":

Definition 2.1 A (*nondeterministic*) *finite automaton* (or NFA) M is a quintuple (Q, Σ, τ, s, A), where Q and Σ are non-empty finite sets, τ is a subset of $Q \times \Sigma \times Q$, s is a designated element of Q and A is a subset of Q.

Here Q is the set of *states* of M, Σ is the set of *inputs*, τ is the *transition relation*, s is the *start state* and A is the set of *accept states*. Some definitions also allow the possibility of *empty moves* as well, i.e., moves of the form (q, ϵ, r) where $q, r \in Q$.

Given an NFA (Q, Σ, τ, s, A), we may extend τ to a subset of $Q \times \Sigma^* \times Q$ by composition. We first define (q, ϵ, r) to be in τ if $q = r$; if we have allowed empty moves, we can also have $(q, \epsilon, r) \in \tau$ for some $q, r \in Q$ with $q \neq r$. We then inductively define $(q, a\beta, r)$ to be in τ (where $a \in \Sigma$, $\beta \in \Sigma^*$) whenever $(q, a, p) \in \tau$ and $(p, \beta, r) \in \tau$ for some $p \in Q$. Given this, we make the following definition:

Definition 2.2 An NFA $M = (Q, \Sigma, \tau, s, A)$ *accepts* a word $\alpha \in \Sigma^*$ if $(s, \alpha, f) \in \tau$ for some $f \in A$, and $L(M)$ (the *language accepted by* M) is the set of all words accepted by M. A word in Σ^* that is not accepted by M is *rejected* by M.

A language $L \subseteq \Sigma^*$ is called *regular* if $L = L(M)$ for some NFA M and $\mathcal{R}eg$ denotes the class of all regular languages. A finite automaton is said to be a *deterministic finite automaton* (DFA) if τ is a function from $Q \times \Sigma$ to Q (some definitions allow τ to be a partial function); note that a DFA is an example of an NFA. It is a standard result that, for every NFA, there is a DFA accepting the same language.

We extend our automaton by adding a *stack*, i.e., a memory device where we may store a sequence of elements but where we may only access the topmost element. Each move may read an input (we allow empty moves), "pops" off the topmost element and "pushes" a (possibly empty) sequence of elements onto the stack:

Definition 2.3 A (*nondeterministic*) *pushdown automaton* (or NPDA) M is a sextuple $(Q, \Sigma, \Gamma, \tau, s, A)$, where Q, Σ and Γ are non-empty finite sets with $\perp$ a designated element of Γ, τ a finite subset of $Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \times Q \times \Gamma^*$, s a designated element of Q and A a subset of Q. If $\Gamma_1 = \Gamma - \{\perp\}$, we insist that:

$$\begin{aligned} (q, a, \perp, r, \gamma) \in \tau &\implies \gamma \in \Gamma_1^* \{\perp\}; \\ (q, a, g, r, \gamma) \in \tau,\ g \in \Gamma_1 &\implies \gamma \in \Gamma_1^*. \end{aligned}$$

Here Q is the set of *states*, Σ the set of *inputs*, Γ the set of *stack symbols*, s the start state, A the set of accept states and τ the transition relation. A *configuration* of M is an element of $Q \times \Sigma^* \times \Gamma_1^*\{\bot\}$; this records the current state of M, the input remaining to be read and the contents of the stack (from top to bottom). We write $(q, a\beta, g\gamma) \models (r, \beta, \theta\gamma)$ if $a \in \Sigma$ and $(q, a, g, r, \theta) \in \tau$, and $(q, \beta, g\gamma) \models (r, \beta, \theta\gamma)$ if $(q, \epsilon, g, r, \theta) \in \tau$. If $\models^*$ denotes the reflexive and transitive closure of $\models$, then M *accepts* $\alpha \in \Sigma^*$ if $(s, \alpha, \bot) \models^* (f, \epsilon, \gamma\bot)$ for some $f \in A$ and $\gamma \in \Gamma_1^*$. We let $L(M)$ denote the set of all words accepted by M. We say that L is a *context-free language* (CFL) if $L = L(M)$ for some NPDA M and let $\mathcal{CF}$ denote the class of CFLs.

Some definitions of an NPDA do not include the bottom symbol $\bot$; this is not significant as the classes of languages accepted are the same. There are other notions of acceptance in NPDAs, such as acceptance by empty stack, but these different notions of acceptance all give rise to the same class of languages.

If, given any configuration (q, α, γ) of an NPDA M, there is at most one (r, β, θ) such that $(q, \alpha, \gamma) \models (r, \beta, \theta)$ we say that M is a *deterministic pushdown automaton* (DPDA). If $L = L(M)$ for some DPDA M then L is a *deterministic context-free language* (DCFL). Let $\mathcal{DCF}$ denote the class of all DCFLs. Unlike finite automata, insisting on determinism does make a difference, in that there are languages that are context-free but not deterministic context-free. We also have to be a little more careful with our notions of acceptance; for example, there are languages accepted by a DPDA as described here, where we accept by accept state, but which are not accepted by any DPDA that always has the stack empty when accepting a word.

We now consider a restricted type of NPDA. A *nondeterministic one-counter automaton* (NOCA) is an NPDA $(Q, \Sigma, \Gamma, \tau, s, A)$ where $|\Gamma| = 2$, $\Gamma = \{\bot, g\}$ say. At any stage the stack of an NOCA contains $g^n\bot$ for some $n \geqslant 0$ and so can be described by a single natural number n; hence the title "one-counter". If the NOCA is deterministic we have a *deterministic one-counter automaton* (DOCA). L is a *one-counter language* (OCL) if $L = L(M)$ for some NOCA M and a *deterministic one-counter language* (DOCL) if $L = L(N)$ for some DOCA N. We let $\mathcal{OC}$ and $\mathcal{DOC}$ denote the classes of one-counter and deterministic one-counter languages.

Clearly $\mathcal{R}eg \subseteq \mathcal{DOC} \subseteq \mathcal{OC} \subseteq \mathcal{CF}$; in fact we have that $\mathcal{R}eg \subset \mathcal{DOC} \subset \mathcal{OC} \subset \mathcal{CF}$. It is clear that $\mathcal{DOC} \subseteq \mathcal{OC} \cap \mathcal{DCF}$ and, in fact, $\mathcal{DOC} \subset \mathcal{OC} \cap \mathcal{DCF}$. We also have that $\mathcal{DCF}$ and $\mathcal{OC}$ are incomparable (i.e., that $\mathcal{DCF} \not\subseteq \mathcal{OC}$ and $\mathcal{OC} \not\subseteq \mathcal{DCF}$).

Another approach to languages is via grammars. We have the following:

Definition 2.4 A *grammar* G is a quadruple (N, Σ, P, S) where N is a finite set of *nonterminals*, Σ is a finite set of *terminals* (where we insist that $N \cap \Sigma = \emptyset$), $P \subseteq (V^* - \Sigma^*) \times V^*$ is a finite set of *productions* (where $V = N \cup \Sigma$) and the *sentence symbol* S is a designated element of N.

We write $\alpha \Rightarrow \beta$ if α and β are of the form $\gamma\rho\delta$, $\gamma\sigma\delta$ respectively where $(\rho, \sigma) \in P$ and $\gamma, \delta \in V^*$. We let $\stackrel{*}{\Rightarrow}$ denote the reflexive and transitive closure of $\Rightarrow$ and then define the *language* $L(G)$ generated by the grammar G to be $\{\alpha \in \Sigma^* : S \stackrel{*}{\Rightarrow} \alpha\}$.

Definition 2.4 allows a wider range of languages than just CFLs: if we only insist that $P \subseteq (V^* - \Sigma^*) \times V^*$ we get the class of recursively enumerable languages. However, if every rule in P is of the form (A, α) with $A \in N$ (i.e., $P \subseteq N \times V^*$), we have the *context-free grammars* (CFGs) which generate precisely the CFLs.

There are many interesting subclasses of $\mathcal{CF}$ obtained by imposing further restrictions on the types of production allowed in a CFG. For example, for $\mathcal{R}eg$ (defined above as the languages accepted by NFAs) we have *regular grammars* which are CFGs where each production is either of the form (A, xB) or (A, ϵ) with $A, B \in N$ and $x \in \Sigma$ (there are other definitions used for regular grammars but they are equivalent in that they all give rise to the same class of languages).

An interesting class of languages is that of the *NTS languages* (see Remark 6.3) which are the languages generated by CFGs with the property that, if A and B are nonterminals such that $A \stackrel{*}{\Rightarrow} \beta$ and $B \stackrel{*}{\Rightarrow} \alpha\beta\gamma$, then $B \stackrel{*}{\Rightarrow} \alpha A\gamma$. The class of *linear languages* (see Remarks 5.6 and 6.8) is the subset of $\mathcal{CF}$ generated by CFGs with the property that each production has at most one nonterminal on the right-hand side, i.e., such that $P \subseteq N \times (\Sigma^* \cup \Sigma^* N \Sigma^*)$; there are various definitions used for linear grammars but they all give rise to the same class of languages. We can also define the class of linear languages by restricting the range of the possible NPDAs (where we insist that all the push operations must precede any pop operations).

There are also grammars generating classes of languages that contain $\mathcal{CF}$ but which are still proper subsets of the class of recursively enumerable languages. One such class is that of the *indexed languages* which are defined by extending our notion of a grammar (although they can be generated by grammars as in Definition 2.4). There are some variations in the definition of an indexed grammar (though not in the class of languages they define); we follow the approach in [24].

In an *indexed grammar* (N, Σ, I, P, S) we have the set N of nonterminals, the set Σ of terminals, the set P of productions and a designated element S of N as before; however, we also have a finite set I of *indices*. As before we let V denote the set $N \cup \Sigma$. Each production rule in P is then of one of the following types:

(T1)	$A \to \alpha$	$A \in N$, $\alpha \in V^*$;
(T2)	$A \to Bf$	$A, B \in N$, $f \in I$;
(T3)	$Af \to \alpha$	$A \in N$, $f \in I$, $\alpha \in V^*$.

A derivation is similar to one in a CFG except that a nonterminal may be followed by a sequence of indices; so our derived strings are elements of $(NI^* \cup \Sigma)^*$. We define $\Rightarrow$ (where $\beta, \gamma \in (NI^* \cup \Sigma)^*$, $\delta \in I^*$ and $X_1, X_2, \ldots, X_k \in V$) as follows:

(i) if $(A, X_1X_2 \ldots X_k) \in P$ is of type (T1) then $\beta A\delta\gamma \Rightarrow \beta X_1\delta_1 X_2\delta_2 \ldots X_k\delta_k\gamma$ where $\delta_i = \delta$ if $X_i \in N$ and $\delta_i = \epsilon$ if $X_i \in \Sigma$;

(ii) if $(A, Bf) \in P$ is of type (T2) then $\beta A\delta\gamma \Rightarrow \beta Bf\delta\gamma$;

(iii) if $(Af, X_1X_2 \ldots X_k) \in P$ is of type (T3) then $\beta Af\delta\gamma \Rightarrow \beta X_1\delta_1 X_2\delta_2 \ldots X_k\delta_k\gamma$ where $\delta_i = \delta$ if $X_i \in N$ and $\delta_i = \epsilon$ if $X_i \in \Sigma$.

As before, let $\stackrel{*}{\Rightarrow}$ denote the reflexive and transitive closure of $\Rightarrow$ and define the language $L(G)$ generated by G to be $\{\alpha \in \Sigma^* : S \stackrel{*}{\Rightarrow} \alpha\}$. There is an alternative approach based on an extension of NPDAs known as *nested stack automata* but we will not go into that here. The class of indexed languages properly contains $\mathcal{CF}$.

3 Operations on languages

We now turn our attention to various operations on languages. There is the operation of concatenation on words (we denote the concatenation of α and β by $\alpha\beta$); we can

extend this to languages and define the *concatenation* L_1L_2 of L_1 and L_2 to be $\{\alpha\beta : \alpha \in L_1,\ \beta \in L_2\}$. We define the *Kleene star* of a language L to be

$$L^* = \{\epsilon\} \cup \{\alpha_1\alpha_2 \dots \alpha_n : n > 0,\ \alpha_i \in L \text{ for all } i\}.$$

We now have the notion of a "prefix": For a word $\alpha = x_1x_2 \cdots x_n$ with $x_i \in \Sigma$ for each i, a *prefix* of α is a word of the form $x_1x_2 \cdots x_m$ where $0 \leqslant m \leqslant n$. Note that, for any word α, the words ϵ and α are both prefixes of α (and the only prefix of ϵ is ϵ). We can extend this to an operation on languages: for any language L we denote the set of all the prefixes of the words in L by $\mathrm{Pref}(L)$, so that

$$\mathrm{Pref}(L) = \{\alpha \in \Sigma^* : \text{there exists } \beta \in \Sigma^* \text{ such that } \alpha\beta \in L\}.$$

We call $\mathrm{Pref}(L)$ the *prefix closure* of the language L.

As languages are sets the operations $\cup$, $\cap$ and $-$ are naturally defined on languages. As Σ^* (under concatenation) is a monoid with identity element ϵ for any set Σ we can also form languages via monoid homomorphism: if φ is a monoid homomorphism from Σ^* to Ω^* for some finite sets Σ and Ω and if $K \subseteq \Sigma^*$, then $K\varphi \subseteq \Omega^*$. We similarly have inverse homomorphisms: if $L \subseteq \Omega^*$ then $L\varphi^{-1} \subseteq \Sigma^*$.

A generalization of a homomorphism is a "gsm-mapping" which can be thought of as a mapping computed by a finite automaton with output. A *generalized sequential machine* (gsm) is a sextuple $(Q, \Sigma, \Delta, \tau, s, A)$ where Q is a finite set of *states*, Σ a finite set of *inputs*, Δs a finite set of *outputs*, τ a function from $Q \times \Sigma$ to finite subsets of $Q \times \Delta^*$, s the *start state* and A the set of *accept states*.

If M is a generalized sequential machine, then we can extend τ to a function from $Q \times \Sigma^*$ to finite subsets of $Q \times \Delta^*$ as follows:

(i) if $q \in Q$ then $\tau(q, \epsilon) = \{(q, \epsilon)\}$;

(ii) if $q \in Q$, $x \in \Sigma$ and $\alpha \in \Sigma^*$ then $\tau(q, \alpha x)$ is defined to be:

$$\{(p, \beta) : \exists \gamma, \delta \in \Sigma^*,\ r \in Q \text{ such that } \beta = \gamma\delta \text{ with } (r, \gamma) \in \tau(q, \alpha),\ (p, \delta) \in \tau(r, x)\}.$$

Given this extended definition of τ, if $\alpha \in \Sigma^*$ we define $M(\alpha)$ to be

$$\{\beta \in \Delta^* : \text{there exists } f \in A \text{ such that } (f, \beta) \in \tau(s, \alpha)\},$$

in other words $M(\alpha)$ is the set of words β such that M can output β and finish in an accept state when given α as input. If $L \subseteq \Sigma^*$ then we define

$$M(L) = \bigcup \{M(\alpha) : \alpha \in L\}.$$

We say that $M(L)$ is a *gsm-mapping*. In a similar vein, if $\beta \in \Delta^*$, we can define $M^{-1}(\beta) = \{\alpha \in \Sigma^* : \beta \in M(\alpha)\}$, and then, if $K \subseteq \Delta^*$, we define

$$M^{-1}(K) = \{\alpha \in \Sigma^* : \beta \in M(\alpha) \text{ for some } \beta \in K\}.$$

We say that $M^{-1}(K)$ is an *inverse gsm-mapping*. Note that this is not a true inverse as we do not necessarily have that $M(M^{-1}(K)) = K$ or that $M^{-1}(M(L)) = L$.

We have one last operation to define, that of "shuffle". The *shuffle* $L_1 \leftrightarrow L_2$ of $L_1 \subseteq \Sigma^*$ with $L_2 \subseteq \Delta^*$ is defined to be the language over $\Sigma \cup \Delta$ defined by:

$$\{\alpha_1\beta_1\alpha_2\beta_2 \cdots \alpha_n\beta_n : n \geqslant 0,\ \alpha_1\alpha_2 \cdots \alpha_n \in L_1, \beta_1\beta_2 \cdots \beta_n \in L_2,\ \alpha_i \in \Sigma^*,\ \beta_i \in \Delta^*\}.$$

4 Closure properties

We now come to the topic of closure properties of classes $\mathcal{F}$ of languages.

(i) If u is a unary operation on languages (i.e., if $L \subseteq \Sigma^*$ then $u(L)$ is a uniquely defined subset of Ω^* for some Ω) we say that $\mathcal{F}$ is *closed under* u if, whenever $L \in \mathcal{F}$, we have that $u(L) \in \mathcal{F}$.

(ii) If b is a binary operation on languages (i.e., if $L_1 \subseteq \Sigma^*$ and $L_2 \subseteq \Delta^*$ then $b(L_1, L_2)$ is a uniquely defined subset of Ω^* for some Ω) we say that $\mathcal{F}$ is *closed under* b if, whenever $L_1, L_2 \in \mathcal{F}$, we have that $b(L_1, L_2) \in \mathcal{F}$.

We could extend this to n-ary operations for $n \geqslant 3$ but we won't need such operations here. Examples of closure under unary operations include closure under complementation, Kleene star, prefix closure, homomorphism, inverse homomorphism, gsm-mapping and inverse gsm-mapping; closure under binary operations includes closure under union, intersection, concatenation and shuffle.

Going on from this, we say that a class of languages $\mathcal{F}$ is *closed under intersection with regular languages* if

$$L_1 \in \mathcal{F},\ L_2 \in \mathcal{R}eg,\ L_1 \subseteq \Sigma^*,\ L_2 \subseteq \Sigma^* \implies L_1 \cap L_2 \in \mathcal{F},$$

and that $\mathcal{F}$ is *closed under union with regular languages* if

$$L_1 \in \mathcal{F},\ L_2 \in \mathcal{R}eg,\ L_1 \subseteq \Sigma^*,\ L_2 \subseteq \Sigma^* \implies L_1 \cup L_2 \in \mathcal{F}.$$

There are variants such as closure under intersection and union with CFLs.

Three of the most important closure properties are combined in the notion of a *cone* which is a class of languages closed under homomorphism, inverse homomorphism and intersection with regular languages. We have taken the term "cone" from [4]; there are other names in use such as "full trio" (see [24]). If we have these three properties then others must hold as well; for example any cone is closed under gsm and inverse gsm mappings (see [24]); as a result, there are other (equivalent) definitions of a cone used. The classes of languages we will be most concerned with in this paper, namely $\mathcal{R}eg$, $\mathcal{OC}$ and $\mathcal{CF}$, are all cones. The classes $\mathcal{DOC}$ and $\mathcal{DCF}$ are not cones, however, as they are not closed under homomorphism.

5 Formal languages and word problems

We will now explain how word problems are related to formal language theory.

If G is a group and Σ is a finite subset of G such that every element of G can be expressed in the form $a_1 a_2 \dots a_n$ for some $a_i \in \Sigma$ and $n \geqslant 0$ (so that G is finitely generated) then Σ generates G (as a monoid). We then have a natural (monoid) homomorphism φ from Σ^* onto G. For each element $a \in \Sigma$ let $\overline{a}$ be an element of Σ^* such that $\overline{a}\varphi = (a\varphi)^{-1}$. We have that $a_1 a_2 \dots a_n = b_1 b_2 \dots b_m$ in G (where $a_i, b_j \in \Sigma$) if and only if $a_1 a_2 \dots a_n \overline{b_m}\, \overline{b_{m-1}} \dots \overline{b_1}$ represents the identity element of G; so we focus on the set of the words in Σ^* representing the identity and refer to this as the *word problem* $W(G, \Sigma)$ of G with respect to the generating set Σ.

Given a group G and a surjective homomorphism $\varphi : \Sigma^* \to G$ for some finite set Σ, the word problem of G is completely specified (given such a homomorphism φ we can

consider Σ as a subset of G by identifying a and $a\varphi$ for each $a \in \Sigma$). Conversely, given a language $L \subseteq \Sigma^*$ that is the word problem of some group, then it turns out that the group in question is completely specified by the language L. One way of thinking about this is to consider the "syntactic monoid" of L:

If Σ is a finite set and $L \subseteq \Sigma^*$ then the *syntactic congruence* $\approx_L$ is the congruence on Σ^* defined by:

$$\alpha \approx_L \beta \iff (\gamma\alpha\delta \in L \Leftrightarrow \gamma\beta\delta \in L \text{ for all } \gamma, \delta \in \Sigma^*).$$

Another way to think of $\approx_L$ is that it is the coarsest congruence on Σ^* such that L is a union of congruence classes. The *syntactic monoid* M_L of L is then the quotient $\Sigma^*/\approx_L$. If L is a language that is the word problem of a group G, then G is the syntactic monoid of L (see [15] for example).

Remark 5.1 There are many interesting results on syntactic monoids; one fundamental result is that the syntactic monoid M_L is finite if and only if L is regular. We note one feature of this connection that will be useful later in this paper.

Given an NFA accepting L, there is a procedure for constructing the minimal (with respect to the number of states) DFA P accepting L (and this DFA is then uniquely determined up to the labelling of its states). If Q is the set of states of P then each element of Σ^* induces a function from Q to Q; the set of all such functions forms a monoid under composition known as the *transition monoid* of P. One way to then think of the syntactic monoid of L is as this transition monoid.

Let φ be the natural map from Σ^* to $M = M_L$, so that $L = S\varphi^{-1}$ for some $S \subseteq M$ (given that L is a union of congruence classes). For each $x \in M$ we can test whether $x \in S$ by checking whether $x\varphi^{-1} \subseteq L$ (we can pick any element α of Σ^* with $\alpha\varphi = x$, i.e., any element of Σ^* inducing the same function from Q to Q as x, and then test whether $\alpha \in L$) and so we may effectively determine S.

One natural question is which languages are word problems of groups; the following characterization was given in [38]:

Theorem 5.2 *A language L over an alphabet Σ is the word problem of a group with generating set Σ if and only if L satisfies the following two conditions:*

(W1) *for all $\alpha \in \Sigma^*$ there exists $\beta \in \Sigma^*$ such that $\alpha\beta \in L$;*

(W2) *if $\alpha\delta\beta \in L$ and $\delta \in L$ then $\alpha\beta \in L$.*

The condition (W1) in Theorem 5.2 says that the prefix closure $\mathrm{Pref}(L)$ of L is Σ^*; this is referred to as saying that L has the *universal prefix closure property*. The condition (W2) is referred to as saying that the language L is *deletion closed*.

It is clear that conditions (W1) and (W2) in Theorem 5.2 are necessary for L to be the word problem of a group but it is perhaps a little surprising that they are also sufficient. One obvious further condition that word problems of groups must satisfy is the following:

(W3) if $\alpha\beta \in L$ and $\delta \in L$ then $\alpha\delta\beta \in L$.

Property (W3) is sometimes referred to as saying that L is *insertion closed*.

These notions give rise to further closure operations: "insertion closure" and "deletion closure". Suppose that L is a language over Σ. Given that the intersection of

a collection of insertion closed languages is insertion closed, the intersection I of all the insertion closed languages over Σ containing L is itself insertion closed and is the minimal insertion closed language over Σ containing L. We call I the *insertion closure* of L (in Σ^*) and define the *deletion closure* of L similarly. It is intriguing that we have a connection between word problems of groups and natural formal language conditions such as insertion and deletion closure as studied in [27].

The proof of Theorem 5.2 in [38] essentially shows that L is the set of words in the syntactic monoid M_L that represent the identity element of M_L and then that M_L is a group (which has L as its word problem). Let us leave groups for a moment and consider this issue about representing the identity element of a monoid.

If we consider monoids then the set of words representing the identity clearly does not determine when two words represent the same element of the monoid. Indeed, this set of words could tell us very little; for example, consider an arbitrary semigroup S with a (semigroup) homomorphism $\varphi : \Sigma^+ \to S$ and then add an identity element 1 to S to form a monoid M. We can extend φ to a (monoid) homomorphism $\varphi : \Sigma^* \to M$ by defining $\epsilon\varphi$ to be 1 but the only word representing the identity is then ϵ. Notwithstanding all this, we note the following fact (compare Proposition 2.1 of [30]):

Proposition 5.3 *If Σ is an alphabet and $\emptyset \neq L \subseteq \Sigma^*$ then there is a monoid M and a monoid homomorphism $\varphi : \Sigma^* \to M$ with $L = \{1\}\varphi^{-1}$ if and only if L satisfies conditions* (W2) *and* (W3) *above, i.e., if and only if*

$$\text{if } \delta \in L \text{ then, for all } \alpha, \beta \in \Sigma^*, \text{ we have that } (\alpha\beta \in L \Longleftrightarrow \alpha\delta\beta \in L).$$

Proof It is easy to check that, if there is a monoid M and a monoid homomorphism $\varphi : \Sigma^* \to M$ with $L = \{1\}\varphi^{-1}$, then L must satisfy conditions (W2) and (W3): if $\delta \in L$, i.e., $\delta\varphi = 1$, then $(\alpha\delta\beta)\varphi = (\alpha\varphi)(\delta\varphi)(\beta\varphi) = (\alpha\varphi)(\beta\varphi) = (\alpha\beta)\varphi$, and so $(\alpha\delta\beta)\varphi = 1$ if and only if $(\alpha\beta)\varphi = 1$ as required. It remains to prove the converse.

Suppose that L satisfies (W2) and (W3). As $L \neq \emptyset$, we must have that $\epsilon \in L$ (using (W2) with $\alpha = \beta = \epsilon$ and δ any element of L). Let M be the syntactic monoid M_L of L and then let φ be the natural homomorphism from Σ^* onto M.

If $\delta \in L$ then, given (W2) and (W3), we have that $\alpha\delta\beta \in L$ if and only if $\alpha\epsilon\beta \in L$, and so $\delta \approx_L \epsilon$ for all $\delta \in L$, i.e., L forms a single congruence class under $\approx_L$. So $L = \{m\}\varphi^{-1}$ for some $m \in M$ and, since $\epsilon \in L$, we have $m = 1$ as required. □

If we want to address the word problem in semigroups, we do have to consider pairs of words (α, β) such that α and β represent the same element and it is not immediate how to link this with formal language theory. One natural approach is taken in [11]; if S is a semigroup generated by a finite set Σ then consider the set

$$W(S, \Sigma) = \{\alpha\#\beta^{rev} : \alpha, \beta \in \Sigma^+, \alpha \text{ and } \beta \text{ represent the same element of } S\} \quad (*)$$

and think of $W(S, \Sigma)$ as being the word problem of S with respect to Σ.

Here $\#$ is some new symbol (i.e., $\# \notin \Sigma$) and β^{rev} denotes the *reversal* of the word β (i.e., if β is the word $x_1x_2 \ldots x_{n-1}x_n$ then β^{rev} is $x_nx_{n-1} \ldots x_2x_1$). This is a natural extension of the way we approached the word problem of groups where we considered $\alpha\beta^{-1}$; if we restrict the concept $W(S, \Sigma)$ back to groups then we can think of β^{rev} playing the role of β^{-1} and the $\#$ symbol as indicating that we are operating

with inverse symbols after that point (or, if we take a monoid generating set, as we have here, with words representing the inverses of the symbols in β).

One satisfying aspect of this approach is that the word problem of a finitely generated group G (the set of words representing the identity) lies in a class $\mathcal{F}$ of languages if and only if the word problem $W(G, \Sigma)$ of G (as a semigroup) defined in $(*)$ lies in $\mathcal{F}$, and so we have a natural extension. Semigroups with a one-counter word problem were investigated in [19] and with a context-free word problem in [17].

An interesting alternative approach was taken in [36] (see also [9]). Instead of considering words of the form $\alpha\#\beta^{rev}$, take a finite automaton with two tapes, one of which contains α and the other β; the tapes can be read synchronously (in which case the shorter of the words α and β needs to be padded with some symbol at the end so that the resulting words are of the same length) or asynchronously. The family of semigroups whose word problem can be recognized by such a machine is intriguing and they are not the same as those having a regular word problem in the sense of [11]: the latter is the family of finite semigroups whereas there are infinite semigroups, such finitely generated free semigroups, whose word problems are accepted by two-tape synchronous finite automata.

Returning to groups, when we examine finitely generated groups based on their word problem as a formal language it is quite natural to try to classify groups based on what class of languages their word problem lies in. However, a group will have many different finite generating sets and there is no guarantee, in general, that the word problem of the group will necessarily lie in the same class $\mathcal{F}$ of languages for different generating sets. However the following result (see [15] for example) shows that, under certain mild assumptions on the class $\mathcal{F}$, this is not a problem:

Theorem 5.4 *If a class of languages $\mathcal{F}$ is closed under inverse homomorphism and the word problem of a group G with respect to some finite generating set lies in $\mathcal{F}$ then the word problem of G with respect to any finite generating set lies in $\mathcal{F}$.*

Given Theorem 5.4, if $\mathcal{F}$ is closed under inverse homomorphism, then we say that a finitely generated group G is an $\mathcal{F}$-group if the word problem of G lies in $\mathcal{F}$ with respect to some finite generating set. If we assume some further closure properties of the class $\mathcal{F}$, then we can say something about the family of $\mathcal{F}$-groups:

Proposition 5.5 *Let $\mathcal{F}$ be a class of languages which is closed under inverse homomorphism.*

(i) If $\mathcal{F}$ is closed under intersection with regular languages then the family of $\mathcal{F}$-groups is closed under taking finitely generated subgroups.

(ii) If $\mathcal{F}$ is closed under union with regular languages and inverse gsm-mappings then the family of $\mathcal{F}$-groups is closed under passing to finite index overgroups.

(iii) If $\mathcal{F}$ is closed under shuffle then the family of $\mathcal{F}$-groups is closed under taking direct products.

(iv) If $\mathcal{F}$ is closed under union and insertion closure then the family of $\mathcal{F}$-groups is closed under taking free products.

Proof For the proofs of (i) and (ii) see [21].

For (iii) let A and B be $\mathcal{F}$-groups with finite generating sets Σ and Ω respectively

and consider the direct product $A \times B$ with generating set $\Sigma \cup \Omega$. Let J and K be the word problems of A and B with respect to Σ and Ω respectively. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are words over Σ and $\beta_1, \beta_2, \ldots, \beta_n$ are words over Ω then $\alpha_1\beta_1\alpha_2\beta_2 \ldots \alpha_n\beta_n$ represents the identity of $A \times B$ if and only if $\alpha_1\alpha_2 \ldots \alpha_n$ represents the identity of A and $\beta_1\beta_2 \ldots \beta_n$ represents the identity of B. So the word problem of $A \times B$ is

$$\{\alpha_1\beta_1\alpha_2\beta_2 \ldots \alpha_n\beta_n : \alpha_1\alpha_2 \ldots \alpha_n \in J \text{ and } \beta_1\beta_2 \ldots \beta_n \in K\} = J \leftrightarrow K,$$

and, as $\mathcal{F}$ is closed under shuffle, we have that $A \times B$ is an $\mathcal{F}$-group as required.

For (iv), let A and B be $\mathcal{F}$-groups with finite generating sets Σ and Ω respectively and consider the free product $A * B$ with generating set $\Sigma \cup \Omega$. Let J and K again be the word problems of A and B with respect to Σ and Ω. Let L be the insertion closure of $J \cup K$; by assumption we have that $L \in \mathcal{F}$. Clearly any word in $J \cup K$ represents the identity in $A * B$ and, if we insert a word representing the identity in $A * B$ into another such word, then the resulting word still represents the identity in $A * B$. So L is contained in the word problem W of $A * B$ with respect to $\Sigma \cup \Omega$.

Suppose that $L \neq W$ and let $\theta = \alpha_1\beta_1 \ldots \alpha_n\beta_n$ $(n \geqslant 1)$ be a word in $W - L$ where the α_i are words over Σ, the β_i are words over Ω and the free product length n of θ is minimal over all such words in $W - L$. We must have that at least one of the factors α_i or β_i represents the identity in A or B respectively and that the word η that results from deleting this factor ζ from θ still represents the identity in $A * B$ but with smaller free product length (if we have deleted α_1 or β_n we need to take a cyclic permutation of the word η, moving β_1 to the end or α_n to the start of η as appropriate, to reduce the free product length). However, by the minimality of n, we would have that $\eta \in L$ and, as θ can be formed by inserting ζ into η with $\zeta \in L$, that $\theta \in L$, a contradiction. So $L = W$ as required and $A * B$ is an $\mathcal{F}$-group. □

Remark 5.6 The assumptions in Proposition 5.5 are sufficient but are not necessary; in particular, it could be that the restriction on being a word problem of a group means that we do not need the full force of the closure properties of the class $\mathcal{F}$ of languages to get the corresponding property of the family of $\mathcal{F}$-groups.

As an example, let $\mathcal{F}$ be the class of linear languages. We comment below (Remark 6.8) that the family of $\mathcal{F}$-groups is actually the family of finite groups, and so is closed under direct products. However, the linear languages are not closed under shuffle. For example, if $L_1 = \{u^m v^m : m \in \mathbb{N}\}$ and $L_2 = \{x^n y^n : n \in \mathbb{N}\}$, then L_1 and L_2 are both linear. However, if J is the shuffle $L_1 \leftrightarrow L_2$ and K is the regular language $\{u^a x^b v^c y^d : a, b, c, d \in \mathbb{N}\}$, then $J \cap K = \{u^m x^n v^m y^n : m, n \in \mathbb{N}\}$. This language $J \cap K$ is not linear. As the class of linear languages is closed under intersection with regular languages, the shuffle $J = L_1 \leftrightarrow L_2$ is not linear.

6 Groups, word problems and algebraic characterizations

In the previous section we discussed some general properties which families of groups whose word problem lies in some designated class of languages might have. Let us now turn our attention to what is known for some specific classes. We start with the groups with a regular word problem which were classified by Anisimov [1]:

Theorem 6.1 *A finitely generated group G has a regular word problem if and only if it is a finite group.*

Further work was done by Muller and Schupp [35] (see also [2, 34, 42]) and their work, modulo a subsequent result of Dunwoody [12], characterized the groups with a context-free word problem. Before we state the result recall that, if $\mathcal{P}$ is a property of groups (such as being abelian or free) we say that a group is *virtually-$\mathcal{P}$* if it has a subgroup of finite index with property $\mathcal{P}$. Given that, we have the following:

Theorem 6.2 *A finitely generated group G has a context-free word problem if and only if it is a virtually free group.*

This is a beautiful result and uses many deep theorems from group theory (in addition to Dunwoody's theorem) such as Stallings' characterization [43] of groups with more than one end. The following will be of relevance in what follows:

Remark 6.3 As a consequence of the classification in Theorem 6.2, one can show that any CFL that is the word problem of a group is accepted by a DPDA with the stack empty on acceptance; indeed, by [3], it is even an NTS language.

In the light of this we have the following consequence of Theorems 5.2 and 6.2:

Corollary 6.4 *If L is a CFL that satisfies the following two conditions:*

(W1) *for all $\alpha \in \Sigma^*$ there exists $\beta \in \Sigma^*$ such that $\alpha\beta \in L$;*

(W2) *if $\alpha\delta\beta \in L$ and $\delta \in L$ then $\alpha\beta \in L$.*

then L is a DCFL.

Indeed, given Remark 6.3, one could strengthen Corollary 6.4 to deduce that L is an NTS-language. The only proof of Corollary 6.4 we are aware of uses Theorem 6.2, and hence we have a simple fact about formal languages deduced using deep results from the theory of groups. It is tempting, therefore, to ask the following:

Question 6.5 Is there a natural proof of Corollary 6.4 that avoids the use of Theorem 6.2?

We accept that the word "natural" in Question 6.5 is a little vague but we are thinking about a proof couched purely in terms of CFGs or NPDAs which proves Corollary 6.4 without any need to appeal to results from group theory. In addition, there are variants of Question 6.5 which are still open. If we consider the class of indexed languages one can ask whether a finitely generated group with an indexed word problem must have a context-free word problem. This is true if we impose some extra hypotheses on the type of indexed language (see [13] for example) but is open in general. It is an intriguing question which is equivalent to the following:

Question 6.6 If L is an indexed language that satisfies:

(W1) for all $\alpha \in \Sigma^*$ there exists $\beta \in \Sigma^*$ such that $\alpha\beta \in L$;

(W2) if $\alpha\delta\beta \in L$ and $\delta \in L$ then $\alpha\beta \in L$.

does it follow that L is necessarily context-free?

Returning to Theorems 6.1 and 6.2 one might ask what other classes $\mathcal{F}$ of languages contained in $\mathcal{CF}$ give rise to interesting families of groups. Herbst [14] showed that, if $\mathcal{F}$ is a cone, then there are not many possibilities:

Theorem 6.7 *If $\mathcal{F}$ is a cone which is a subset of $\mathcal{CF}$ then the family of finitely generated groups whose word problem lies in $\mathcal{F}$ is the family of groups with a regular word problem, the family of groups with a one-counter word problem or the family of groups with a context-free word problem.*

Remark 6.8 Theorem 6.7 is not saying that $\mathcal{R}eg$, $\mathcal{OC}$ and $\mathcal{CF}$ are the only cones contained in $\mathcal{CF}$; there are others such as the cone of all linear languages. What it says is that, if $\mathcal{F}$ is a cone contained in $\mathcal{CF}$, then the set of $\mathcal{F}$-groups is one of the three families of groups mentioned. If we take the cone of linear languages, for example, then this properly contains $\mathcal{R}eg$ but the family of groups with a linear word problem is the same as the family of groups with a regular word problem, i.e., the finite groups by Theorem 6.1. This gives another result analogous to Corollary 6.4, in that a linear language satisfying (W1) and (W2) is necessarily regular.

In the light of Theorems 6.1, 6.2 and 6.7 it is natural to ask which groups have a one-counter word problem. Herbst characterized these groups in [14]:

Theorem 6.9 *A finitely generated group G has a one-counter word problem if and only if it is a virtually cyclic group.*

The proof in [14] uses Theorem 6.2; see [19] for a proof that avoids doing so. One consequence of Theorem 6.9 is that any OCL which is the word problem of a group is also a DOCL (and so we have a version of Question 6.5 for OCLs).

Following on from Theorem 6.9, the following generalization was proved in [19]:

Theorem 6.10 *If $n \geqslant 1$ the following are equivalent for a finitely generated group G:*

(i) The word problem of G is the intersection of n OCLs.
(ii) The word problem of G is the intersection of n DOCLs.
(iii) G is virtually abelian of free abelian rank at most n.

This theme was continued by Brough who investigated groups whose word problem is an intersection of finitely many CFLs; such a language is said to be *poly-context-free* or *poly-$\mathcal{CF}$*. If we want to specify that a language is the intersection of n CFLs for some specific value of n, then we say that it is *n-context-free* or *n-$\mathcal{CF}$*.

It is pointed out in [8] that, for any $n \geqslant 1$, the class of n-$\mathcal{CF}$ languages is closed under inverse homomorphisms, inverse gsm-mappings, union with CFLs and intersection with regular languages. The class of *poly-$\mathcal{CF}$* languages is closed under all these operations and also under intersection and union. Given this, if we fix $n \geqslant 1$, the family of groups whose word problem is an n-$\mathcal{CF}$ language is closed under taking finitely generated subgroups and finite index overgroups by Proposition 5.5, and the same holds for groups whose word problem is a *poly-$\mathcal{CF}$* language.

It is also shown in [8] that the direct product of a group whose word problem is m-$\mathcal{CF}$ with a group whose word problem is n-$\mathcal{CF}$ results in a group whose word problem is $(m+n)$-$\mathcal{CF}$, so that the family of groups whose word problem is *poly-$\mathcal{CF}$*

is closed under direct products. The proof in [8] essentially follows the line we have taken in Proposition 5.5 above given the fact that the class of *poly-$\mathcal{CF}$* languages is closed under shuffle. Given all this, one has the following:

Proposition 6.11 *If G is a finitely generated group which has a subgroup H of finite index such that H is a finitely generated subgroup of a direct product of free groups then G has a* poly-$\mathcal{CF}$ *word problem.*

We have included the assumption that H is finitely generated for clarity but it follows from the fact that H has finite index in a finitely generated group. In the light of Proposition 6.11, Brough makes the following intriguing conjecture:

Conjecture 6.12 *A finitely generated group G has a poly-$\mathcal{CF}$ word problem if and only if G has a subgroup H of finite index such that H is a subgroup of a direct product of free groups.*

Some evidence for this conjecture is provided in [8]. For example, it is shown that a polycyclic group or a finitely generated nilpotent group G has a *poly-$\mathcal{CF}$* word problem if and only if G is virtually abelian. Going on from this, we have from [8] the following restriction of Conjecture 6.12 to soluble groups:

Conjecture 6.13 *A finitely generated soluble group G has a poly-$\mathcal{CF}$ word problem if and only if G is virtually abelian.*

Some further evidence for Conjecture 6.13 is provided by the following result [8]:

Theorem 6.14 *If G is a finitely generated soluble group with a poly-$\mathcal{CF}$ word problem, then one of the following must hold:*

(i) G is virtually abelian; or (possibly)

(ii) G has a finitely generated subgroup H with an infinite normal torsion subgroup U such that H/U is either free abelian or isomorphic to a Gc-group which is not virtually abelian.

The notion of a *Gc*-group was defined in [10]. For soluble groups Conjectures 6.12 and 6.13 imply that case (ii) in Theorem 6.14 cannot occur.

The study of groups with a poly-$\mathcal{CF}$ word problem is potentially related to groups whose word problem is the complement of a CFL (see below). If $\mathcal{F}$ is a class of languages let *co*$\mathcal{F}$ denote the class of languages that are complements of languages in $\mathcal{F}$. The following will be convenient in what follows:

Definition 6.15 If G is a group generated by a finite set Σ then the *co-word problem* of G (with respect to Σ) is the set of elements in Σ^* that do not represent the identity element of G.

Groups with a context-free co-word problem were initially studied in [21]. Studying such groups may seem a little strange but the original interest sprang from the following observation. If the word problem of a group is a CFL, then it is a DCFL (Remark 6.3) and, unlike $\mathcal{CF}$, the class $\mathcal{DCF}$ is closed under complementation; so we have that $\mathcal{DCF} = co\mathcal{DCF} \subseteq \mathcal{CF} \cap co\mathcal{CF}$. Given this, we have the following:

Proposition 6.16 *If G is a finitely generated group such that word problem for G is a CFL then the word problem for G is co-context-free.*

Like $\mathcal{CF}$, *co*$\mathcal{CF}$ is closed under inverse homomorphism, and so we do not need to refer to the particular choice of generating set in Proposition 6.16 by Theorem 5.4.

One might expect the converse of Proposition 6.16 to be true and that the families of groups with context-free and co-context-free word problems to coincide, but this is not the case. as the following example shows.

Example 6.17 Consider the group $\mathbb{Z} \times \mathbb{Z}$ generated by a and b and let Σ be the (monoid) generating set $\{a, b, A, B\}$ where A represents a^{-1} and B represents b^{-1}; then the co-word problem of G with respect to Σ is the set

$$\{\alpha \in \Sigma^* : |\alpha|_a \neq |\alpha|_A \text{ or } |\alpha|_b \neq |\alpha|_B\}$$

which is context-free; so the word problem is co-context-free. However, the word problem of G is $\{\alpha \in \Sigma^* : |\alpha|_a = |\alpha|_A \text{ and } |\alpha|_b = |\alpha|_B\}$ which is not a CFL.

Some information about groups with a co-context-free word problem comes from [21]:

Proposition 6.18

(i) All finitely generated free groups are co$\mathcal{CF}$-groups.

(ii) All finitely generated abelian groups are co$\mathcal{CF}$-groups.

(iii) The family of co$\mathcal{CF}$-groups is closed under passing to finite index overgroups.

(iv) The family of co$\mathcal{CF}$-groups is closed under taking finitely generated subgroups.

(v) The family of co$\mathcal{CF}$-groups is closed under taking direct products.

(vi) If G is a co$\mathcal{CF}$-group and H is a $\mathcal{CF}$-group then the restricted standard wreath product $G \wr H$ of G with H is a co$\mathcal{CF}$-group.

It is not clear whether the hypothesis in Proposition 6.18 (vi) that H is a $\mathcal{CF}$-group (i.e., that H is virtually free by Theorem 6.2) is necessary; it is conjectured in [21] that this is the case, i.e., that, if $G \wr H$ is a *co*$\mathcal{CF}$-group, then H is a $\mathcal{CF}$-group (we have that G and H are *co*$\mathcal{CF}$-groups by part (iv) of Proposition 6.18 however).

Conjecture 6.12 stated that a finitely generated group G has a poly-$\mathcal{CF}$ word problem if and only if G has a subgroup H of finite index such that H is a subgroup of a direct product of free groups. We see from Proposition 6.18 that any such group would have a co-context free word problem; one wonders if one could establish this fact in general (that any word problem for a group in poly-$\mathcal{CF}$ must lie in *co*$\mathcal{CF}$) without proving Conjecture 6.12 in full. If this were so then it must depend on the fact that we are considering word problems of groups as we do not have that poly-$\mathcal{CF}$ is contained in *co*$\mathcal{CF}$ (it is not even the case that $\mathcal{CF}$ is contained in *co*$\mathcal{CF}$).

One might wonder if Proposition 6.18 describes all *co*$\mathcal{CF}$-groups, i.e., any group with a *co*$\mathcal{CF}$ word problem can be constructed from free groups and abelian groups via a sequence of the operations described in parts (iii), (iv), (v) and (vi); however this is not the case. If all *co*$\mathcal{CF}$-groups could be constructed in this way then, for any particular *co*$\mathcal{CF}$-group G, there would only be finitely many natural numbers n such that G contains an element of order n. However, in [33], Lehnert and Schweitzer

showed that the co-word problem for the Higman-Thompson group $G_{n,r}$ is context-free and the set of elements of finite order in $G_{n,r}$ does not have this property.

Another interesting question about the closure properties of the $co\mathcal{CF}$-groups is that of free products. We saw in Proposition 5.5 (iv) that, if $\mathcal{F}$ is a class of languages closed under inverse homomorphism, union and insertion closure, then the family of $\mathcal{F}$-groups is closed under free products. These closure properties hold in the $\mathcal{CF}$ for example (so we could deduce that the family of $\mathcal{CF}$-groups is closed under taking free products without recourse to Theorem 6.2); on the other hand, $co\mathcal{CF}$ is not closed under union. However, as pointed out in Remark 5.6, we are not claiming that the hypotheses in Proposition 5.5 are necessary.

As a specific example, there is the question as to whether the co-word problem of $G = (\mathbb{Z} \times \mathbb{Z}) * \mathbb{Z}$ is context-free. One approach might be to show that G embeds in a group with context-free co-word problem and use Proposition 6.18 (iv). A possible candidate for such a group was Thompson's group $V = G_{2,1}$ but it was shown by Bleak and Salazar-Díaz in [5] that G does not embed in V.

In [14] Herbst asked the following: if L is a DCFL and the syntactic monoid of L is a group G, does it follow that G has context-free word problem, i.e., that G is virtually free? This was answered by Röver in [39]. In an analogous result to Proposition 6.18 (vi), he showed that the family of groups which are syntactic monoids of DCFLs is closed under taking restricted standard wreath products with virtually free top groups. So a group such as a restricted standard wreath product $\mathbb{Z} \wr F$, where F is finite, is the syntactic monoid of a DCFL but is not virtually free. An interesting question from [39] is the following: suppose that the syntactic monoid of a DCFL is a group G; is G is necessarily a $co\mathcal{CF}$-group?

7 Decidability

In this section we turn our attention to some questions of decidability concentrating on the concepts and characterizations introduced above. One natural question is that of asking, given a language L (specified by some means such as an automaton or grammar), whether L is the word problem of a group. Related to this is the decidability of properties of languages such as (W1), (W2) and (W3) introduced in Section 5. In particular, if properties (W1) and (W2) were both decidable for some class $\mathcal{F}$ of languages, then the question of deciding whether a language in $\mathcal{F}$ is the word problem of a group would also be decidable by Theorem 5.2.

In general such questions are decidable for $\mathcal{R}eg$. One approach is to consider the syntactic monoid. Taking the notation used in Remark 5.1, (W1) is equivalent to

$$\text{for all } x \in M \text{ there exists } y \in M \text{ such that } xy \in S.$$

Since M is a finite monoid, this is easily seen to be decidable. In a similar fashion, conditions (W2) and (W3) are equivalent to

> if $xuy \in S$ and $u \in S$ then $xy \in S$, and
> if $xy \in S$ and $u \in S$ then $xuy \in S$,

respectively; again, these are easily seen to be decidable. As a result, we can decide whether a regular language L is the word problem of a group by Theorem 5.2.

If we are only interested in deciding whether a regular language L is the word problem of a group, and not in whether L satisfies (W1), (W2) and (W3) individually, we could check this without using Theorem 5.2: we simply calculate the transition monoid M of the minimal DFA of L (as in Remark 5.1) and then see if M is a group with L as the pre-image of the identity element of M.

The situation changes for $\mathcal{CF}$ however; in fact, even for $\mathcal{OC}$, all the properties (W1), (W2) and (W3) are undecidable. For (W3) the argument is straightforward. It is known [25] that the problem of deciding whether a OCL L over an alphabet Σ is equal to Σ^* is undecidable. Now suppose we could decide (W3) for OCLs. We first check whether ϵ and the elements of Σ are in L (we can do this as the membership problem for OCLs is decidable); if not we know that $L \neq \Sigma^*$. Assuming that ϵ and the elements of Σ are in L, we next see whether L satisfies (W3), i.e., whether L is insertion closed. If L is not insertion closed, then $L \neq \Sigma^*$ and, if L is insertion closed then, as $\Sigma \subseteq L$, we have that $\Sigma^+ \subseteq L$ and then, as $\epsilon \in L$, $\Sigma^* \subseteq L$, and hence $L = \Sigma^*$. So we would be able to decide whether $L = \Sigma^*$, a contradiction.

The arguments that (W1) and (W2) are undecidable for OCLs are more involved and may be found in [28]. We also cannot decide whether an arbitrary OCL is the word problem of a group. This does not follow from the undecidability of (W1) and (W2), as it is possible to have two undecidable properties whose conjunction is decidable, but it was shown in [32] that the problem of deciding whether a CFL is the word problem is undecidable and this was generalized to OCLs in [28].

However, given a DCFL L, it is decidable whether L is the word problem of a group [28]. At first sight this may seem rather strange, in that it is undecidable whether a context-free language is the word problem of a group and any CFL that is such a word problem is necessarily a DCFL by Remark 6.3. We have to be careful how we express this problem. The important point is that, when we say "given a DCFL L", we are insisting that it is given as a DCFL, i.e., that we are given a DPDA accepting L; so we are considering the following problem:

Input: a DPDA $M = (Q, \Sigma, \Gamma, \tau, s, A)$.
Output: "yes" if $L(M)$ is the word problem of a group; "no" otherwise.

This changes the problem; a critical component of the proof in [28] of this decidability result is the fact that, while the equivalence problem (given M_1 and M_2 do we have $L(M_1) = L(M_2)$) is undecidable for NPDAs, it is decidable for DPDAs [40].

The fact that our CFL L would have to be a DCFL in order for L to be the word problem of a group does not help as the problem of deciding whether a CFL is a DCFL is undecidable. Even if we are given the promise that L is a DCFL, we cannot find a DPDA accepting L, i.e., there is no algorithm solving the following:

Input: an NPDA M such that $L(M)$ is deterministic context-free.
Output: a DPDA N with $L(N) = L(M)$.

Acknowledgements The second author would like to thank Hilary Craig for all her help and encouragement.

References

[1] A. V. Anisimov, Group languages, *Kibernetika* **4** (1971), 18–24.

[2] A. V. Anisimov, Certain algorithmic questions for groups and context-free languages, *Kibernetika* **8** (1972), 4–11.
[3] J. M. Autebert, L. Boasson & G. Sénizergues, Groups and NTS languages, *J. Comput. System Sci.* **35** (1987), 243–267.
[4] J. Berstel, *Transductions and Context-free Languages* (Leitfäden der Angewandten Mathematik und Mechanik **38**, Teubner, Stuttgart, 1979).
[5] C. Bleak & O. Salazar-Díaz, Free products in R. Thompson's group V, *Trans. Amer. Math. Soc.* **365** (2013), 5967–5997.
[6] W. W. Boone, The word problem, *Ann. of Math.* **70** (1959), 207–265.
[7] W. W. Boone & G. Higman, An algebraic characterization of groups with soluble word problem, *J. Austral. Math. Soc.* **18** (1974), 41–53.
[8] T. Brough, Groups with poly-context-free word problem, *Groups Complex. Cryptol.* **6** (2014), 9–29.
[9] T. Brough, Inverse semigroups with rational word problem are finite, *arXiv: math/ 9812028*.
[10] T. Brough & D. F. Holt, Finitely generated soluble groups and their subgroups, *Comm. Algebra* **41** (2013), 1790–1799.
[11] A. Duncan & R. H. Gilman, Word hyperbolic semigroups, *Math. Proc. Cambridge Philos. Soc.* **136** (2004), 513–524.
[12] M. J. Dunwoody, The accessibility of finitely presented groups, *Invent. Math.* **81** (1985), 449–457.
[13] R. H. Gilman & M. Shapiro, On groups whose word problem is solved by a nested stack automaton, *arXiv: math/ 9812028*.
[14] T. Herbst, On a subclass of context-free groups *RAIRO Inform. Théor. Appl.* **25** (1991), 255–272.
[15] T. Herbst & R. M. Thomas, Group presentations, formal languages and characterizations of one-counter groups, *Theoret. Comput. Sci.* **112** (1993), 187–213.
[16] G. Higman, Subgroups of finitely presented groups, *Proc. Roy. Soc. Ser. A* **262** (1961), 455–475.
[17] M. Hoffmann, D. F. Holt, M. D. Owens & R. M. Thomas, Semigroups with a context-free word problem, *in* H-C. Yen & O. H. Ibarra (eds), *Developments in Language Theory* (LNCS **7410**, Springer, Heidelberg, 2012), 97–108.
[18] D. F. Holt, Word-hyperbolic groups have real-time word problem, *Internat. J. Algebra Comput.* **10** (2000), 221–227.
[19] D. F. Holt, M. D. Owens & R. M. Thomas, Groups and semigroups with a one-counter word problem, *J. Austral. Math. Soc.* **85** (2008), 197–209.
[20] D. F. Holt & S. E. Rees, Solving the word problem in real time, *J. London Math. Soc.* **63** (2001), 623–639.
[21] D. F. Holt, S. E. Rees, C. E. Röver & R. M. Thomas, Groups with a context-free co-word problem, *J. London Math. Soc.* **71** (2005), 643–657.
[22] D. F. Holt, S. E. Rees & M. Shapiro, Groups that do and do not have growing context-sensitive word problem, *Internat. J. Algebra Comput.* **18** (2008), 1179–1191.
[23] D. F. Holt & C. E. Röver, On real-time word problems, *J. London Math. Soc.* **67** (2003), 289–301.
[24] J. E. Hopcroft & J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, (Addison-Wesley Publishing Company, Reading, Massachusetts, 1979).
[25] O. H. Ibarra, Restricted one-counter machines with undecidable universe problems, *Math. Systems Theory* **13** (1979/80), 181–186.
[26] M. Ito, *Algebraic Theory of Automata and Languages* (World Scientific Publishing Co., River Edge, New Jersey, 2004).
[27] M. Ito, L. Kari & G. Thierren, Insertion and deletion closure of languages, *Theoret. Comput. Sci.* **183** (1997), 3–19.

[28] S. A. M. Jones & R. M. Thomas, Formal languages, word problems of groups and decidability, *in* P. A. Abdulla & I. Potapov (eds), *Reachability Problems* (LNCS **8169**, Springer, Berlin-Heidelberg, 2013), 146–158.

[29] M. Kambites & F. Otto, Church-Rosser groups and growing context-sensitive groups, *J. Autom. Lang. Comb.* **13** (2008), 249–267.

[30] L. Kari & G, Thierrin, Languages and monoids with disjunctive identity, *Collect. Math.* **6** (1995), 97–107

[31] S. R. Lakin & R. M. Thomas, Context-sensitive decision problems in groups, *in* C. S. Calude, E. Calude & M. J. Dinneen (eds), *Developments in Language Theory* (LNCS **3340**, Springer, Berlin, 2004), 296–307.

[32] S. R. Lakin & R. M. Thomas, Complexity classes and word problems of groups, *Groups Complex. Cryptol.* **1** (2009), 261–273.

[33] J. Lehnert & P. Schweitzer, The co-word problem for the Higman-Thompson group is context-free, *Bull. Lond. Math. Soc.* **39** (2007), 235–241.

[34] A. A. Letičevskiĭ & L. B. Smikun, A certain class of groups with a decidable problem of automata equivalence, *Dokl. Akad. Nauk SSSR* **227** (1976), 36–38.

[35] D. E. Muller & P. E. Schupp, Groups, the theory of ends, and context-free languages, *J. Comput. System Sci.* **26** (1983), 295–310.

[36] M. Neunhöffer, M. Pfeiffer & N. Ruškuc, Deciding word problems of semigroups using finite state automata, *arXiv: 1206.1714 [cs.FL]*.

[37] P. S. Novikov, On the algorithmic unsolvability of the word problem in group theory, *Trudy Mat. Inst. im. Steklov.* **44** (1955), 143 pages.

[38] D. W. Parkes & R. M. Thomas, Groups with context-free reduced word problem, *Comm. Algebra* **30** (2002), 3143–3156.

[39] C. E. Röver, On groups which are syntactic monoids of deterministic context-free languages, *Internat. J. Algebra Comput.* **14** (2004), 499–504.

[40] G. Sénizergues, $L(A) = L(B)$? Decidability results from complete formal systems, *Theoret. Comput. Sci.* **251** (2001), 1–166.

[41] M. Shapiro, A note on context-sensitive languages and word problems, *Internat. J. Algebra Comput.* **4** (1994), 493–497.

[42] L. B. Smikun, A connection between context-free groups and groups with a solvable problem of automata equivalence, *Kibernetica* **5** (1976), 33–37.

[43] J. Stallings, *Group Theory and Three-dimensional Manifolds* (Yale Mathematical Monographs **4**, Yale University Press, New Haven-Connecticut-London, 1971).

[44] R. J. Thompson, Embeddings into finitely generated groups which preserve the word problem, *in* S. I. Adian, W. W. Boone & G. Higman (eds), *Word problems, II* (Stud. Logic Foundations Math. **95**, North Holland, Amsterdam-New York, 1980), 401–441.

ON THE CASTELNUOVO-MUMFORD REGULARITY OF THE COHOMOLOGY OF FUSION SYSTEMS AND OF THE HOCHSCHILD COHOMOLOGY OF BLOCK ALGEBRAS

RADHA KESSAR and MARKUS LINCKELMANN

Department of Mathematical Sciences, City University London, Northampton Square, London EC1V 0HB, U.K.
Email: radha.kessar.1@city.ac.uk, markus.linckelmann.1@city.ac.uk

Abstract

Symonds' proof of Benson's regularity conjecture implies that the regularity of the cohomology of a fusion system and that of the Hochschild cohomology of a p-block of a finite group is at most zero. Using results of Benson, Greenlees, and Symonds, we show that in both cases the regularity is equal to zero.

Let p be a prime and k an algebraically closed field of characteristic p. Given a finite group G, a *block algebra of* kG is an indecomposable direct factor B of kG as a k-algebra. A *defect group of a block algebra* B *of* kG is a minimal subgroup P of G such that B is isomorphic to a direct summand of $B \otimes_{kP} B$ as a B-B-bimodule. The defect groups of B form a G-conjugacy class of p-subgroups of G. The Hochschild cohomology of B is the algebra $HH^*(B) = \mathrm{Ext}^*_{B \otimes_k B^{\mathrm{op}}}(B)$, where B^{op} is the opposite algebra of B, and where B is regarded as a $B \otimes_k B^{\mathrm{op}}$-module via left and right multiplication. By a result of Gerstenhaber, the algebra $HH^*(B)$ is graded-commutative; that is, for homogeneous elements $\zeta \in HH^m(B)$ and $\eta \in HH^m(B)$ we have $\eta\zeta = (-1)^{nm}\zeta\eta$, where m, n are nonnegative integers. In particular, if $p = 2$, then $HH^*(B)$ is commutative, and if p is odd, then the even part $HH^{\mathrm{ev}}(B) = \bigoplus_{n \geq 0} HH^{2n}(B)$ is commutative and all homogeneous elements in odd degrees square to zero. The extension of the Castelnuovo-Mumford regularity to graded-commutative rings with generators in arbitrary positive degrees is due to Benson [2, §4]. We follow the notational conventions in Symonds [18]. In particular, if p is odd and $T = \bigoplus_{n \geq 0} T^n$ is a finitely generated graded-commutative k-algebra and M a finitely generated graded T-module, we denote by $\mathrm{reg}(T, M)$ the Castelnuovo-Mumford regularity of M as a graded T^{ev}-module, where $T^{\mathrm{ev}} = \bigoplus_{n \geq 0} T^{2n}$ is the even part of T. We set $\mathrm{reg}(T) = \mathrm{reg}(T, T)$; that is, $\mathrm{reg}(T)$ is the Castelnuovo-Mumford regularity of T as a graded T^{ev}-module. See also [3] and [8] for more background material and references. We note that Benson's definition of regularity uses the ring T instead of T^{ev}, but the two definitions are equivalent. This can be seen by noting that [18, Proposition 1.1] also holds for finitely generated graded commutative k-algebras.

Theorem 1 *Let G be a finite group and B a block algebra of kG. We have*

$$\mathrm{reg}(HH^*(B)) = 0.$$

This will be shown as a consequence of a statement on Scott modules. Given a finite group G and a p-subgroup P of G, there is up to isomorphism a unique indecomposable kG-module $Sc(G; P)$ with vertex P and trivial source having a quotient

(or equivalently, a submodule) isomorphic to the trivial kG-module k. The module $Sc(G;P)$ is called the *Scott module of kG with vertex P*. It is constructed as follows: Frobenius reciprocity implies that $\mathrm{Hom}_{kG}(\mathrm{Ind}_P^G(k),k) \cong \mathrm{Hom}_{kP}(k,k) \cong k$, and hence $\mathrm{Ind}_P^G(k)$ has up to isomorphism a unique direct summand $Sc(G;P)$ having k as a quotient. Since $\mathrm{Ind}_P^G(k)$ is selfdual, the uniqueness of $Sc(G;P)$ implies that $Sc(G;P)$ is also selfdual, and hence $Sc(G;P)$ can also be characterised as the unique summand, up to isomorphism, of $\mathrm{Ind}_P^G(k)$ having a nonzero trivial submodule. Moreover, it is not difficult to see that $Sc(G;P)$ has P has a vertex. See [7] for more details on Scott modules, as well as [11] for connections between Scott modules and fusion systems. For a finitely generated graded module X over $H^*(G;k)$ we denote by $H^{*,*}_m(X)$ the local cohomology with respect to the maximal ideal of $H^*(G;k)$ generated by all elements in positive degree. The first grading is here the local cohomological grading, and the second is induced by the grading of X.

Theorem 2 *Let G be a finite group and P a p-subgroup of G. We have*

$$\mathrm{reg}(H^*(G;k);H^*(G;Sc(G;P))) = 0.$$

Remark 3 Using Benson's reinterpretation in [1, §4], of the 'last survivor' from [5, §7], applied to the Scott module instead of the trivial module, one can show more precisely that

$$H^{r,-r}_m(H^*(G;Sc(G,P))) \neq \{0\},$$

where r is the rank of P. It is not clear whether this property, or even the property of having cohomology with regularity zero, characterises Scott modules amongst trivial source modules.

For $\mathcal{F}$ a saturated fusion system on a finite p-group P, we denote by $H^*(P;k)^{\mathcal{F}}$ the graded subalgebra of $H^*(P;k)$ consisting of all elements ζ satisfying $\mathrm{Res}^P_Q(\zeta) = \mathrm{Res}_\varphi(\zeta)$ for any subgroup Q of P and any morphism $\varphi : Q \to P$ in $\mathcal{F}$. If $\mathcal{F}$ is the fusion system of a finite group G on one of its Sylow-p-subgroups P, then $H^*(P;k)^{\mathcal{F}}$ is isomorphic to $H^*(G;k)$ through the restriction map Res^G_P, by the characterisation of $H^*(G;k)$ in terms of stable elements due to Cartan and Eilenberg. In that case we have $\mathrm{reg}(H^*(P;k)^{\mathcal{F}}) = 0$ by [18, Corollary 0.2]. If $\mathcal{F}$ is the fusion system of a block algebra B of kG on a defect group P, then $H^*(P;k)^{\mathcal{F}}$ is the block cohomology $H^*(B)$ as defined in [14, Definition 5.1]. It is not known whether all block fusion systems arise as fusion systems of finite groups. There are examples of fusion systems which arise neither from finite groups nor from blocks; see [10], [13].

Theorem 4 *Let $\mathcal{F}$ be a saturated fusion system on a finite p-group P. We have*

$$\mathrm{reg}(H^*(P;k)^{\mathcal{F}}) = 0.$$

The key ingredients for proving the above results are Greenlees' local cohomology spectral sequence [9, Theorem 2.1], results and techniques in work of Benson [1], [2], [4], and Symonds' proof in [18] of Benson's regularity conjecture. We use the properties of the regularity from [18, §1] and [19, §2].

Lemma 5 *Let G be a finite group and V an indecomposable trivial source kG-module. Then* $\mathrm{reg}(H^*(G;k);H^*(G;V)) \leq 0$.

Proof Since V is a direct summand of $\mathrm{Ind}_P^G(k)$, we have

$$\mathrm{reg}(H^*(G;k);H^*(G;V)) \leq \mathrm{reg}(H^*(G;k);H^*(G;\mathrm{Ind}_P^G(k)).$$

By [12, Lemma 4], the right side is equal to $\mathrm{reg}(H^*(P;k))$, hence zero by [18, Corollary 0.2]. □

Lemma 6 *Let G be a finite group and V a finitely generated kG-module. If $H_0(G;V) \neq \{0\}$, then* $\mathrm{reg}(H^*(G;k);H^*(G;V)) \geq 0$.

Proof It follows from the assumption $H_0(G;V) \neq \{0\}$ and Greenlees' spectral sequence [9, Theorem 2.1] that there is an integer s such that $H_m^{s,-s}(H^*(G;V)) \neq \{0\}$, which implies the result. □

Proof of Theorem 2 Set $V = Sc(G;P)$. By Lemma 5 we have

$$\mathrm{reg}(H^*(G;k);\mathrm{Ext}_{kG}^*(k;V)) \leq 0.$$

Since V has a nonzero trivial submodule, we have $H_0(G;V) \neq \{0\}$, and hence the other inequality follows from Lemma 6. □

Theorem 1 will be a consequence of Theorem 2 and the following well-known observation (for which we include a proof for the convenience of the reader; the block theoretic background material can be found in [20]).

Lemma 7 *Let G be a finite group, B a block algebra of kG and P a defect group of B. As a module over kG with respect to the conjugation action of G on B, the kG-module B has an indecomposable direct summand isomorphic to the Scott module $Sc(G;P)$.*

Proof Since the conjugation action of G on B induces the trivial action on $Z(B)$ and since $Z(B) \neq \{0\}$, it follows that the kG-module B has a nonzero trivial submodule. Moreover, B is a direct summand of kG, hence B is a p-permutation kG-module, and the vertices of the indecomposable direct summands of B are conjugate to subgroups of P. Thus B has a Scott module with a vertex contained in P as a direct summand. Since $Z(B)$ is not contained in the kernel of the Brauer homomorphism Br_P, it follows that B has a direct summand isomorphic to the Scott module $Sc(G;P)$. □

Proof of Theorem 1 By [12, Proposition 5] we have $\mathrm{reg}(HH^*(B)) \leq 0$. Recall that $HH^*(kG)$ is an $H^*(G;k)$-module via the diagonal induction map, and we have a canonical graded isomorphism $HH^*(B) \cong H^*(G;B)$ as $H^*(G;B)$-modules where G acts on B by conjugation; see, e.g., [17, (3.2)]. It follows from [12, Lemma 4] that

$$\mathrm{reg}(HH^*(B)) = \mathrm{reg}(H^*(G;k);H^*(G;B)).$$

By Lemma 7, the kG-module B has a direct summand isomorphic to $V = Sc(G;P)$, where P is a defect group of B. Thus as an $H^*(G;k)$-module, $H^*(G;B)$ has a direct summand isomorphic to $H^*(G;V)$. It follows that

$$\mathrm{reg}(HH^*(B)) \geq \mathrm{reg}(H^*(G;k);H^*(G;V)) = 0,$$

where the last equality is from Theorem 2. This completes the proof of Theorem 1. □

Remark 8 The above proof can be adapted to show that the regularity of the stable quotient $\overline{HH^*}(B)$ of $HH^*(B)$ also equals zero. Recall that $\overline{HH^*}(B)$ is the quotient of $HH^*(B)$ by the ideal $Z^{\rm pr}(B) = {\rm Tr}_1^G(B)$ of $Z(B) \cong HH^0(B)$. Note that $Z^{\rm pr}(B)$ is concentrated in degree 0. Alternatively, $\overline{HH^*}(B)$ may be defined as the non-negative part of the Tate Hochschild cohomology of B. Our interest in $\overline{HH^*}(B)$ comes from the fact that Tate Hochschild cohomology of symmetric algebras is an invariant of stable equivalence of Morita type. We briefly indicate how the regularity of $\overline{HH^*}(B)$ may be calculated. Let $B = \bigoplus_i M_i$ be a decomposition of B into a direct sum of indecomposable kG-modules M_i, where G acts by conjugation on B. The canonical graded $H^*(G;k)$-module isomorphism $HH^*(B) \cong H^*(G;B)$ induces an isomorphism

$$HH^0(B) \cong H^0(G;B) = \bigoplus_i H^0(G;M_i)$$

in degree zero. Composing this with the the canonical isomorphisms $Z(B) \cong HH^0(B)$ and $H^0(G;M_i) \cong M_i^G$, it is easy to check that the image of $Z^{\rm pr}(B)$ in $\bigoplus_i M_i^G$ is $\bigoplus_i {\rm Tr}_1^G(M_i)$. Since B is a p-permutation kG-module, ${\rm Tr}_1^G(M_i)$ is non-zero precisely if M_i is isomorphic to the Scott module $Sc(G;1)$ (which is a projective cover of the trivial kG-module). Let M' denote the sum of all M_i's in the above decomposition which are isomorphic to $Sc(G,1)$ and let M'' be the complement of M' in B with respect to the above decomposition. Since $Z^{pr}(B)$ is concentrated in degree zero, we have a direct sum decomposition $HH^*(B) \cong \oplus H^*(G;M'') \oplus Z^{\rm pr}(B)$ as $H^*(G;k)$-modules. In particular,

$$\mathrm{reg}(H^*(G;k);HH^*(B)) = \max\{\mathrm{reg}(H^*(G;k);H^*(G;M'')), \mathrm{reg}(H^*(G;k);Z^{pr}(B))\}.$$

We may assume that a defect group P of B is non-trivial. By Lemma 7, M'' contains a direct summand isomorphic to $Sc(G;P)$. Hence $\mathrm{reg}(H^*(G;k);H^*(G;M'')) \geq 0$, by Theorem 2. It follows from Theorem 1 and the above displayed equation that $\overline{HH^*}(B) \cong H^*(G;M'')$ has regularity zero.

Proof of Theorem 4 By [18, Proposition 6.1] we have $\mathrm{reg}(H^*(P;k)^{\mathcal{F}}) \leq 0$. For the other inequality we follow the arguments in [1, §3, §4], applied to transfer maps using fusion stable bisets. For Q a subgroup of P and $\varphi : Q \to P$ an injective group homomorphism, we denote by $P \times_{(Q,\varphi)} P$ the P-P-biset of equivalence classes in $P \times P$ with respect to the relation $(uw, v) \sim (u, \varphi(w)v)$, where $u, v \in P$, and $w \in Q$. The kP-kP-bimodule having $P \times_{(Q,\varphi)} P$ as a k-basis is canonically isomorphic to $kP \otimes_{kQ} ({}_\varphi kP)$. This biset gives rise to a transfer map ${\rm tr}_{P\times_{(Q,\varphi)}P}$ on $H^*(P;k)$ obtained by composing the restriction map ${\rm res}^P_{\varphi(Q)} : H^*(P;k) \to H^*(\varphi(Q);k)$, the isomorphism $H^*(\varphi(Q);k) \cong H^*(Q;k)$ induced by φ, and the transfer map ${\rm tr}^P_Q : H^*(Q;k) \to H^*(P;k)$. Let X be an $\mathcal{F}$-stable P-P-biset satisfying the conclusions of [6, Proposition 5.5]. That is, every transitive subbiset of X is isomorphic to $P \times_{(Q,\varphi)} P$ for some subgroup Q of P and some group homomorphism $\varphi : Q \to P$ belonging to $\mathcal{F}$, the integer $|X|/|P|$ is prime to p, and for any subgroup Q of P and any group homomorphism $\varphi : Q \to P$ in $\mathcal{F}$, the Q-P-bisets ${}_\varphi X$ and ${}_Q X$ (resp., the P-Q-bisets X_Q and X_φ) are isomorphic. By taking the sum, over the transitive subbisets $P \times_{(Q,\varphi)} P$, of the transfer maps ${\rm tr}_{P\times_{(Q,\varphi)}P}$, we obtain a transfer map ${\rm tr}_X$ on $H^*(P;k)$.

Following [15, Proposition 3.2], the map tr_X acts as multiplication by $|X|/|P|$ on $H^*(P;k)^{\mathcal{F}}$, hence $\Im(\mathrm{tr}_X) = H^*(P;k)^{\mathcal{F}}$, and we have a direct sum decomposition

$$H^*(P;k) = H^*(P;k)^{\mathcal{F}} \oplus \ker(\mathrm{tr}_X)$$

as $H^*(P;k)^{\mathcal{F}}$-modules. A similar decomposition holds for Tate cohomology, and for homology (using either the canonical duality $H_n(P;k) \cong H^n(P;k)^\vee$ or the isomorphism $H_n(P;k) \cong \hat{H}^{-n-1}(P;k)$ obtained from composing the previous duality with Tate duality). By [1, Equation (4.1)], the transfer map tr_Q^P induces a homomorphism of Greenlees' local cohomology spectral sequences

$$\begin{array}{ccc} H_m^{i,j}H^*(Q,k) & \Longrightarrow & H_{-i-j}(Q;k) \\ {\scriptstyle (\mathrm{tr}_Q^P)_*}\downarrow & & \downarrow{\scriptstyle (\mathrm{res}_Q^P)_*} \\ H_m^{i,j}H^*(P;k) & \Longrightarrow & H_{-i-j}(P;k) \end{array}$$

where $(\mathrm{tr}_Q^P)_*$ and $(\mathrm{res}_Q^P)_*$ are the maps induced by tr_Q^P and the inclusion $Q \to P$, respectively. The isomorphism $\varphi : Q \to \varphi(Q)$ induces an obvious isomorphism of spectral sequences

$$\begin{array}{ccc} H_m^{i,j}H^*(\varphi(Q),k) & \Longrightarrow & H_{-i-j}(\varphi(Q);k) \\ {\scriptstyle \cong}\downarrow & & \downarrow{\scriptstyle \cong} \\ H_m^{i,j}H^*(Q;k) & \Longrightarrow & H_{-i-j}(Q;k) \end{array}$$

Restriction and transfer on Tate cohomology are dual to each other under Tate duality, and hence the dual version of [1, Equation (4.1)] implies that the restriction $\mathrm{res}^P_{\varphi(Q)}$ induces a homomorphism of spectral sequences

$$\begin{array}{ccc} H_m^{i,j}H^*(P,k) & \Longrightarrow & H_{-i-j}(P;k) \\ {\scriptstyle (\mathrm{res}^P_{\varphi(Q)})_*}\downarrow & & \downarrow{\scriptstyle (\mathrm{tr}^P_{\varphi(Q)})_*} \\ H_m^{i,j}H^*(\varphi(Q);k) & \Longrightarrow & H_{-i-j}(\varphi(Q);k) \end{array}$$

Composing the three diagrams above yields a homomorphism induced by $\mathrm{tr}_{P\times_{(Q,\varphi)}P}$ on the spectral sequence for P, and taking the sum over all transitive subbisets of X yields a homomorphism of spectral sequences

$$\begin{array}{ccc} H_m^{i,j}H^*(P,k) & \Longrightarrow & H_{-i-j}(P;k) \\ {\scriptstyle (\mathrm{tr}_X)_*}\downarrow & & \downarrow{\scriptstyle (\mathrm{tr}_{X^\vee})_*} \\ H_m^{i,j}H^*(P;k) & \Longrightarrow & H_{-i-j}(P;k) \end{array}$$

where $X^\vee$ is the P-P-biset X with the opposite action $u \cdot x \cdot v = v^{-1}xu^{-1}$ for all $u, v \in P$ and $x \in X$. One easily checks that $X^\vee$ is isomorphic to a dual basis of X in the dual bimodule $\mathrm{Hom}_k(kX, k)$. By [6, Proposition 5.2], $H^*(P;k)$ is finitely generated as a module over $H^*(P;k)^{\mathcal{F}}$. Thus the local cohomology spaces $H_m^{i,j}H^*(P;k)$ can

be calculated using for m the maximal ideal of positive degree elements in $H^*(P;k)^{\mathcal{F}}$ instead of $H^*(P;k)$. It follows that tr_X induces a homomorphism of spectral sequences

$$\begin{array}{ccc} H_m^{i,j}H^*(P,k) & \Longrightarrow & H_{-i-j}(P;k) \\ {\scriptstyle (\mathrm{tr}_X)_*}\downarrow & & \downarrow{\scriptstyle (\mathrm{tr}_{X^\vee})_*} \\ H_m^{i,j}H^*(P;k)^{\mathcal{F}} & \Longrightarrow & H_{-i-j}(P;k)^{\mathcal{F}} \end{array}$$

For $i=-j=r$, where r is the rank of P, the edge homomorphism yields a commutative diagram of the form

$$\begin{array}{ccccc} H_m^{r,-r}H^*(P;k) & \xrightarrow{\gamma_P} & H_0(P;k) & \xrightarrow{\cong} & k \\ {\scriptstyle (\mathrm{tr}_X)_*}\downarrow & & {\scriptstyle (\mathrm{tr}_{X^\vee})_*}\downarrow & & \downarrow{\scriptstyle \cdot|X|/|P|} \\ H_m^{r,-r}H^*(P;k)^{\mathcal{F}} & \xrightarrow[\delta_{\mathcal{F}}]{} & H_0(P;k)^{\mathcal{F}} & \xrightarrow[\cong]{} & k \end{array}$$

where the right vertical map is multiplication on k by $|X|/|P|$. By [1, Theorem 4.1], the map γ_P is surjective, and hence so is the map $\delta_{\mathcal{F}}$. In particular, $H_m^{r,-r}H^*(P;k)^{\mathcal{F}} \neq \{0\}$, whence the result. □

Remark 9 The fact that transfer and restriction on Tate cohomology are dual to each other under Tate duality can be deduced from a more general duality for transfer maps on Tate-Hochschild cohomology of symmetric algebras induced by bimodules which are finitely generated projective as left and right modules (cf. [16]).

References

[1] D. Benson, Modules with injective cohomology, and local duality for a finite group, *New York J. Math.* **7** (2001), 201–215.

[2] D. J. Benson, Dickson invariants, regularity and computation in group cohomology, *Illinois J. Math.* **48** (2004), 171–197.

[3] D. J. Benson, Commutative algebra in the cohomology of groups, *Trends in Commutative Algebra*, MSRI Publication **51**, (2004.

[4] D. J. Benson, On the regularity conjecture for the cohomology of finite groups, *Proc. Edinburgh Math. Soc.* **51** (2008), 273–284.

[5] D. J. Benson and J. F. Carlson, Projective resolutions and Poincaré duality complexes, *Trans. Amer. Math. Soc.* **342** (1994), 447–488

[6] C. Broto, R. Levi, and B. Oliver, The homotopy theory of fusion systems, *J. Amer. Math. Soc.* **16** (2003), 779–856.

[7] M. Broué, On Scott modules and p-permutation modules: an approach through the Brauer morphism, *Proc. Amer. Math. Soc.* **93** (1985), 401–408.

[8] D. Eisenbud, *The Geometry of Syzygies.* Graduate Texts Math. **229**, Springer-Verlag, New York, 2005.

[9] J. P. C. Greenlees, Commutative algebra in group cohomology, *J. Pure Appl. Algebra* **98** (1995), 151–162.

[10] R. Kessar, The Solomon system $\mathcal{F}_{\mathrm{Sol}}(3)$ does not occur as fusion system of a 2-block, *J. Algebra* **296** (2006), 409–425.

[11] R. Kessar, N. Kunugi, and N. Mitsuhashi, On saturated fusion systems and Brauer indecomposability of Scott modules, *J. Algebra* **340** (2011), 90–103.

[12] R. Kessar and M. Linckelmann, On the Hilbert series of Hochschild cohomology of block algebras, *J. Algebra* **371** (2012), 457–461.
[13] R. Kessar and R. Stancu, A reduction theorem for fusion systems of blocks, *J. Algebra* **319** (2008), 806–823.
[14] M. Linckelmann, Transfer in Hochschild cohomology of blocks of finite groups, *Algebr. Represent. Theory* **2** (1999), 107–135.
[15] M. Linckelmann, A Note on the Schur multiplier of fusion systems, *J. Algebra* **296** (2006), 402–408.
[16] M. Linckelmann, Tate duality and transfer in Hochschild cohomology, *J. Pure Appl. Algebra* **217** (2013), 2387–2399.
[17] S. F. Siegel and S. Witherspoon, The Hochschild cohomology ring of a group algebra, *Proc. London Math. Soc.* **79** (1999), 131–157.
[18] P. Symonds, On the Castelnuovo-Mumford regularity of the cohomology ring of a group, *J. Amer. Math. Soc.* **23** (2010), 1159–1173.
[19] P. Symonds, On the Castelnuovo-Mumford regularity of rings of polynomial invariants, *Ann. of Math. (2)* **174** (2011), 499–517.
[20] J. Thévenaz, *G-Algebras and Modular Representation Theory*, Oxford Science Publications, Clarendon Press, Oxford, 1995.

RECENT ADVANCES ON TORSION SUBGROUPS OF INTEGRAL GROUP RINGS

WOLFGANG KIMMERLE* and ALEXANDER KONOVALOV†

*Fachbereich Mathematik. Universität Stuttgart, Pfaffenwaldring 57, D - 70550 Stuttgart, Germany

Email: kimmerle@mathematik.uni-stuttgart.de

†School of Computer Science, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SX, U.K.

Email: alexk@mcs.st-andrews.ac.uk

Abstract

This survey reports on recent progress made on finite subgroups of the unit group of integral group rings of finite groups. We show that the Gruenberg–Kegel graph of $\mathbb{Z}G$ coincides with that one of G provided $|G|$ is divisible by at most three primes and give an outline how such a result may be obtained with the aid of computational algebra. In the last section we discuss this question for sporadic simple groups and their automorphism groups.

1 Introduction

Let G be a finite group, and let $\mathbb{Z}G$ denotes its integral group ring. The map $\varepsilon : \mathbb{Z}G \longrightarrow \mathbb{Z}$ defined by $\sum_{g\in G} z_g g \mapsto \sum_{g\in G} z_g$ is called, as usual, the *augmentation map*, and its kernel is the augmentation ideal $I(G)$. The unit group is denoted by $U(\mathbb{Z}G)$, and $V(\mathbb{Z}G)$ denotes its subgroup consisting of units with augmentation 1, which is called the *normalised unit group*, or the *group of normalised units*.

A fundamental question in the theory of integral group rings is which properties of G are reflected by $\mathbb{Z}G$. One may expect that the most informative answers to this question come from considering $U(\mathbb{Z}G)$.

For a long time, starting with G. Higman's thesis [41] in 1940 (cf. also [58]), the focus was on the question whether any torsion subgroup of $V(\mathbb{Z}G)$ is isomorphic to a subgroup of the group G. In 1997, M. Hertweck showed with his counterexample to the isomorphism problem that this is not always the case [32]. Nevertheless for many groups G the answer is affirmative, in particular if G is a p-group [57, 63].

Thus it makes sense to rephrase the question in the following way:

Question 1.1 Classify the finite groups H with the property that whenever H occurs as subgroup in the unit group $V(\mathbb{Z}G)$ of the integral group ring $\mathbb{Z}G$ of a finite group G then H is isomorphic to a subgroup of G.

Z. Marciniak posed this question in the special case when H is the Klein four-group on a satellite conference of the ICM at Granada 2006. One can consider the rephrased question as the subgroup isomorphism problem (SIP), and we say that (SIP) holds for a finite group H if this question has an affirmative answer for H. In this article we focus especially on recent results on the following classes of subgroups.

Question 1.2 Has (SIP) an affirmative answer when H is cyclic?

Question 1.3 Does (SIP) hold when H is a p-group?

The second question clearly points into the direction of a Sylow-like theorem, cf. Section 3. The first question circulates around the Zassenhaus Conjecture (ZC) and the prime graph question (PQ), cf. Sections 2 and 4.

Clearly we could consider mainly the case when G is soluble. But the research done in the last years tends also in the direction of arbitrary finite groups or at least of groups whose nonabelian composition factors have small Lie rank or are of specific nature. Note that in order to get a positive result on (SIP) one has to establish such a result also for integral group rings of the nonabelian simple groups.

Integral group rings of soluble groups may be handled in terms of successive abelian extensions of the integral group ring of the groups C_p of prime order p. G. Higman showed in his thesis [41] that for a finite abelian group A the torsion subgroups of $V(\mathbb{Z}A)$ are just the subgroups of A. However with respect to nonabelian simple groups very little is known concerning the torsion subgroups of their integral group rings, so this gives the first big contrast to the soluble case. The second difficulty which arises in handling arbitrary finite groups is that in addition to abelian extensions one has also to deal with perfect extensions. Nevertheless within the last ten years some progress has been made also concerning several classes of insoluble groups.

2 Around the Zassenhaus Conjecture

H. Zassenhaus posed in [66] the following conjecture for a finite group G.

(ZC) Every torsion unit $u \in V(\mathbb{Z}G)$ is conjugate within $\mathbb{Q}G$ to an element in G.

As reported in [60, p.205], H.Zassenhaus stated concerning torsion subgroups even stronger conjectures than (ZC). The strongest one, the so-called (ZC 3), says that each torsion subgroup of $V(\mathbb{Z}G)$ is conjugate within $\mathbb{Q}G$ to a subgroup of G.

In general the (ZC 3) conjecture fails. A finite subgroup H of $V(\mathbb{Z}G)$ which has the same order as G is called a group basis. K.W. Roggenkamp and L.L. Scott constructed a metabelian group G such that $V(\mathbb{Z}G)$ has isomorphic group bases which are not conjugate within $\mathbb{Q}G$ [49, 56, 59]. Of course the counterexample of M.Hertweck to the isomorphism problem provides even stronger counterexample to (ZC 3). In the meantime, the smallest counterexamples to (ZC 3) with group bases which are not conjugate within $\mathbb{Q}G$ are groups of order 96 [5, 39].

The conjecture (ZC) however is still open. The most far reaching results are the following.

Theorem 2.1 (A. Weiss, [64]) *The conjecture (ZC) holds provided G is nilpotent. Moreover, in this case even (ZC 3) is valid.*

For a long time, it was an open question whether (ZC) is valid for metacyclic groups. Finally, in 2007 M. Hertweck gave a positive answer [36]. M. Hertweck proved that (ZC) holds provided G has a cyclic normal subgroup C which has an

abelian complement. The following result shows that it suffices to assume that G/C is abelian.

Theorem 2.2 (M. Caicedo, A. del Rio and L. Margolis [22]) *The conjecture (ZC) holds provided G has a cyclic normal subgroup C such that G/C is abelian.*

Theorem 2.2 is the best result on (ZC) obtained so far concerning metabelian groups. As pointed out before, (ZC 3) is not valid for all metabelian groups. But if G is metabelian, it is known that each torsion subgroup of $V(\mathbb{Z}G)$ is isomorphic to a subgroup of G. This follows immediately from the result of Z. Marciniak and S.K. Sehgal that the units of the form

$$1 + I(G)I(A),$$

where A is an abelian normal subgroup of the finite group G, form a torsion-free normal subgroup of $V(\mathbb{Z}G)$ [53]. Note that these units are just the kernel of the normalised unit group under the projection from $\mathbb{Z}G$ onto the so called *small group ring* $S(G, A) := \mathbb{Z}G/I(G)I(A)$.

As the results on (ZC 3) show, it might be worth to check systematically groups of small order. In [42] it was shown that (ZC) holds for all groups of order $|G| \leq 71$. From Theorem 2.2 it follows that groups of order 72 as well can not provide a counterexample to (ZC). The next critical small group order which should be considered is certainly 96.

In order to get evidence concerning (ZC) one might ask whether at least torsion units of prime order in $V(\mathbb{Z}G)$ are conjugate to an element of G. This is unknown and stands of course in a big contrast to the celebrated results on p-groups and nilpotent groups mentioned above. From this point of view one could certainly say that for arbitrary finite groups we still know nothing about (ZC). There is just one general result in this context.

Proposition 2.3 ([37]) *Suppose that the finite group G has precisely one conjugacy class of elements of prime order p. Then all elements of order p of $V(\mathbb{Z}G)$ are conjugate within $\mathbb{Q}G$ to an element of G.*

Thus, it is certainly justified to look first in the context of (SIP) for an isomorphism of cyclic torsion subgroups. Note that we get then precisely [60, Research Problem 8]. Indeed, we know in this situation at least a bit more.

Proposition 2.4 ([23]) *Let G be a finite group. The cyclic subgroups of prime power order of $V(\mathbb{Z}G)$ are isomorphic to subgroups of G.*

A result due to D.S. Passman [55] shows that this holds also in the case when G is infinite, see also [61, 3.10, 3.14]. With respect to non-cyclic subgroups the following is known.

Proposition 2.5 *Let G be a finite group.*

(a) [43, pp.3169–3170] *Suppose that $V(\mathbb{Z}G)$ has a Klein four-group $C_2 \times C_2$ as subgroup. Then G has a subgroup isomorphic to $C_2 \times C_2$.*

(b) [33] *Let p be an odd prime. Suppose that $C_p \times C_p$ occurs as subgroup of $V(\mathbb{Z}G)$, then as well it occurs as subgroup of G.*

Clearly, Proposition 2.5(a) answers the question of Z. Marciniak raised in the introduction. It is worth to note that the proofs of parts (a) and (b) of Proposition 2.5 are substantially different. The proof of part (a) relies on the Brauer–Suzuki theorem, whereas for part (b) elementary arguments from ordinary character theory suffice. Before we give a proof of Proposition 2.5(a), we explain partial augmentations which provide one of the fundamental methods for the analysis of torsion units.

Let $g \in G$ and $u = \sum_{x \in G} u_x x \in \mathbb{Z}G$. Then

$$\nu_g(u) = \sum_{h \in g^G} u_h$$

is called the partial augmentation of u with respect to the conjugacy class g^G.

Theorem 2.6 *Let G be a finite group.*

(a) [54, Theorem 2.5],[51] *A unit $u \in V(\mathbb{Z}G)$ is rationally conjugate to a trivial unit $g \in G$ if and only if $\nu_x(v) \geq 0$ for every $v \in \langle u \rangle$ and every $x \in G$.*

(b) [38, Theorem 2.3] *Let $u \in V(\mathbb{Z}G)$ be of finite order $o(u)$. Then*

$$\nu_g(u) \neq 0 \Longrightarrow o(g) \text{ divides } o(u).$$

Proof of 2.5(a) Let G be a counterexample of minimal order. By [25, Lemma 2.1] we may assume that G has no normal subgroup of odd order. Because G is a counterexample, Sylow 2-subgroups of G have precisely one involution. Thus Sylow 2-subgroups of G are either cyclic or generalized quaternion groups. In the second case it follows from the Brauer–Suzuki theorem that G has precisely one involution t which has to be central. In the first case we get the same conclusion by Burnside's transfer theorem. By [23] (or by Theorem 2.6(b)) for each involution u of $V(\mathbb{Z}G)$ there exists an involution $j \in G$ such that $\nu_j(u) \neq 0$. Thus each involution of $V(\mathbb{Z}G)$ has a non-vanishing partial augmentation on t. If $u \neq t$ we get via $u \cdot t$ a non-trival torsion unit of $V(\mathbb{Z}G)$ with 1-coefficient $\neq 0$. Now the classical result of Berman and Higman [4, 41], shows that $u \cdot t = 1$.

The next case to be considered is certainly the one of arbitrary finite cyclic subgroups of $V(\mathbb{Z}G)$.

Proposition 2.7 ([34]) *Let G be a finite soluble group. Then finite cyclic subgroups of $V(\mathbb{Z}G)$ are isomorphic to subgroups of G.*

However with respect to a general finite group much less is known. In the case when the subgroup has order $p \cdot q$ where p and q are different primes some general statements can be made. We report on this in the section on the Gruenberg–Kegel graph of $\mathbb{Z}G$.

There are also positive results on (ZC) for some specific insoluble groups. Most of them will be presented in the Gruenberg–Kegel graph section.

3 Sylow-type results

The previous sections certainly support the question whether a Sylow-like theorem may hold in $V(\mathbb{Z}G)$. We say that in $V(\mathbb{Z}G)$ a Sylow-like theorem holds provided for each prime p each finite p-subgroup of $V(\mathbb{Z}G)$ is conjugate within $\mathbb{Q}G$ to a Sylow p-subgroup of G. If for a fixed prime p the p-subgroups of $V(\mathbb{Z}G)$ have the property stated above then we say that in $V(\mathbb{Z}G)$ the p-Sylow-like theorem is valid. Clearly if (ZC 3) holds for $\mathbb{Z}G$ then a Sylow-like theorem holds for $V(\mathbb{Z}G)$.

Thus, a Sylow-like theorem holds in $V(\mathbb{Z}G)$ when G is nilpotent by Theorem 2.1. Further evidence is given because in $\mathrm{GL}(n,\mathbb{Z})$ a Sylow-like theorem holds [1].

M. Dokuchaev and S.O. Juriaans showed using the results of A. Weiss the following which covers the supersoluble groups.

Theorem 3.1 ([25, Theorem 2.9]) *Let G be a finite group with a nilpotent normal subgroup N such that G/N is nilpotent. Then the Sylow-like theorem holds in $V(\mathbb{Z}G)$.*

With respect to group bases even more is known.

Theorem 3.2 ([47]) *Let G be a finite soluble group and let H be a group basis of $\mathbb{Z}G$. Let p be a prime. Then each p-subgroup of H is conjugate within $\mathbb{Q}G$ to a subgroup of a Sylow p-subgroup of G.*

If one assumes that Sylow p-subgroups of G have a special structure much more is known. This is especially the case when G has abelian Sylow subgroups.

Proposition 3.3 ([25, Proposition 2.11]) *Assume that G is a finite soluble group. Suppose that G has abelian Sylow p-subgroups. Then a p-subgroup of $V(\mathbb{Z}G)$ is rationally conjugate to a subgroup of G.*

This proposition may easily be generalized to p-constrained case [2, Proposition 3.2]. The same is the case for Theorem 3.2. Nevertheless these are results which still circulate as all other results before in this section around the class of soluble groups. But for abelian Sylow 2-subgroups there are also results which hold for each group which such Sylow subgroups.

If G has elementary abelian Sylow 2-subgroups then by Proposition 2.4 all 2-elements of $V(\mathbb{Z}G)$ are involutions. Thus all 2-subgroups of $V(\mathbb{Z}G)$ are abelian. Because the order of a torsion subgroup of $V(\mathbb{Z}G)$ divides $|G|$ it follows that each 2-subgroup is isomorphic to a subgroup of a Sylow 2-subgroup of G.

For small abelian Sylow 2-subgroups it is known that a Sylow-like theorem holds.

Proposition 3.4 *Let G be a finite group whose Sylow 2-subgroups are abelian of order ≤ 8. Then each 2-subgroup of $V(\mathbb{Z}G)$ is rationally conjugate to a subgroup of G.*

Proof If G has cyclic Sylow 2-subgroups or Sylow 2-subgroups isomorphic to $C_4 \times C_2$ then by Burnside's transfer theorem G is soluble and the result follows from Proposition 3.3. If $S \in Syl_2(G)$ is elementary abelian of order 4 or 8 then [2, Proposition 3.4] completes the proof. □

Note that the proof of [2, Proposition 3.4] uses the classification of the finite simple groups with abelian Sylow 2-subgroups [62]. These simple groups S have all precisely one conjugacy class of involutions. By Proposition 2.3 it follows that in $V(\mathbb{Z}S)$ all involutions are conjugate. This establishes a substantial part for the insoluble part of Proposition 3.4.

We remark further that it is unknown whether a Sylow-like theorem holds in $V(\mathbb{Z}G)$ when G has elementary abelian Sylow 2-subgroups of order 16. In the case when G is soluble it is shown in [25, Theorem 5.3] that a Sylow-like theorem holds in $V(\mathbb{Z}G)$ provided Sylow p-subgroups of G have order dividing p^3.

The following result indicates that a Sylow-like theorem in nonabelian simple groups of small Lie rank may be true.

Theorem 3.5 ([40]) *Let $G = \mathrm{PSL}(2, q)$. Then finite 2-subgroups of $V(\mathbb{Z}G)$ are isomorphic to subgroups of G.*

When q is even, this is clear from the above because $G = \mathrm{PSL}(2, 2^f)$ has elementary abelian subgroups. Thus, the main part of the Theorem is the case of an odd q, cf. [40, Theorem 2.1]. Note that if q is odd then the Sylow 2-subgroups of $\mathrm{PSL}(2, q)$ are dihedral or elementary abelian.

Theorem 3.6 ([8, Theorem 3]) *Let G be a finite Frobenius group. Then each torsion unit of prime power order in $V(\mathbb{Z}G)$ is conjugate within $\mathbb{Q}G$ to an element of G.*

V. Bovdi and M. Hertweck completed the proof of Theorem 3.6 by showing that (ZC) holds for the covering group $\hat{S}_5$ of the symmetric group S_5 which has a unique conjugacy class of involutions[1]. This was the missing piece in earlier work on integral group rings of Frobenius groups done in [25, 26, 45]. Note that by [25, Corollary 5.2] for soluble Frobenius groups G the Sylow-like theorem is valid in $V(\mathbb{Z}G)$.

4 The Gruenberg–Kegel graph of $\mathbb{Z}G$

The prime graph $\Pi(G)$ of a group G is the graph whose vertices are the primes dividing the order of a torsion element of G. Two different vertices p and q are connected by an edge if and only if G has an element g with $o(g) = pq$. By Proposition 2.4 the vertices of the prime graph of G and that one of $V(\mathbb{Z}G)$ coincide.

The interest in the prime graph of a finite group comes from the work of K.W. Gruenberg and K.W. Roggenkamp in 1970s on the decomposition of the integral augmentation ideal $I(G)$ of $\mathbb{Z}G$ [30]. Using the classification of the finite simple groups, J.S. Williams finally established that for a finite group G the augmentation ideal of $\mathbb{Z}G$ decomposes if and only if the prime graph of G is disconnected. The proof uses a purely group-theoretical result of K.W. Gruenberg and O. Kegel on the structure of finite groups with disconnected prime graph [65, Theorem A]. J.S. Williams and finally A.S. Kondratiev described the prime graphs of all simple groups with more than one connected component [50, 65].

[1] $\hat{S}_5$ is the group (240,89) in the GAP Small Groups Library

In the case of the normalised unit group $V(\mathbb{Z}G)$ we call $\Pi(V(\mathbb{Z}G))$ the Gruenberg–Kegel graph of $\mathbb{Z}G$ and denote it by $\Gamma(\mathbb{Z}G)$. It is certainly natural to pose the prime graph question (PQ) asking whether $\Gamma(\mathbb{Z}G) = \Pi(G)$ [46]. In the context of the decomposition of augmentation ideals an affirmative answer just means that the augmentation ideal of a torsion subgroup U of $V(\mathbb{Z}G)$ decomposes if and only if there are primes dividing $|U|$ which belong to different components of $\Pi(G)$. Clearly, if (ZC) is valid, the prime graph question has an affirmative answer. In the context of (SIP), the prime graph question is just the question whether (SIP) is true for cyclic subgroups of order $p \cdot q$, where p and q are different primes.

(PQ) has been intensively studied during recent years. One main method used for it is the so-called HeLP–method. This method has its origin in the proof of (ZC) for the alternating group A_5 given by I.S. Luthar and I.B.S. Passi [51]. It uses the relationship between partial augmentations and the eigenvalues of a torsion unit u on a Wedderburn component of $\mathbb{C}G$. A first algorithmic treatment has been given in the Diplomarbeit of R. Wagner under the supervision of the first author. This is documented in [9, p.293]. Later M. Hertweck noticed that the same arguments may be applied with Brauer characters [38, §3,§4].

For the convenience of the reader we give a short explanation. Let F be a field whose characteristic does not divide the order n of the unit u and which contains a primitive n-th root of unity. Assume that the partial augmentations of u and all its powers are known. Let χ be the character of D and let ξ be a possible eigenvalue of $D(u)$. Denote by $\mu(u, \xi, \chi)$ the multiplicity of ξ as an eigenvalue of $D(u)$. Then

$$\mu(u, \xi, \chi) = \frac{1}{n} \sum_{d|n} \mathrm{Tr}_{\mathbb{Q}(\zeta^d)/\mathbb{Q}}(\chi(u^d)\xi^{-d}), \tag{1}$$

where n is the order of u, is a non-negative integer. Note that the eigenvalues of $D(u)$ and the character values are considered within $\mathbb{Q}(\zeta)$ where ζ is a primitive n-th root of unity, and $\mathrm{Tr}_{\mathbb{Q}(\zeta^d)/\mathbb{Q}}(x)$ denotes the trace.

Clearly the degrees of the (irreducible) F-representations bound the possibilities for $\mu(u, \xi, \chi)$. Note that $\chi(u^d)$ is given by the partial augmentations of u^d. Since by [31] there is a bound for absolute values of partial augmentations given by the condition $\nu_g(u)^2 \leq |g^G|$, there are only finitely many possible partial augmentations of u, and computational tools may be employed nowadays to implement the HeLP–method and systematically investigate more groups. In some situations this method shows that u and its powers have only the trivial partial augmentations, i.e., for $u^m, m \in \mathbb{N}$ there is precisely one conjugacy class C_m of group elements with $\nu_{C_m}(u^m) = 1$ and the partial augmentations of all other classes vanish. By Theorem 2.6(a) it follows then that u is conjugate within $\mathbb{Q}G$ to a group element. Also it is useful to exclude possible orders of torsion units and so permits to attack (PQ) and (ZC) using the ordinary character table and the Brauer tables for the group in question or with generic character tables for series of groups, cf. [9, §6].

In particular sporadic simple groups have been considered that way, mainly by V. Bovdi and the second author. For about a half of the 26 sporadic simple groups the prime graph question could be positively settled, cf. Section 5. The following reduction to almost simple groups demonstrates that these computations do not just play a role of single examples, but may give a rise to a more general result. This

reduction has been obtained for soluble extensions in [46] and with respect to arbitrary extensions in [48].

Proposition 4.1 *Let N be a minimal normal subgroup of the finite group G. Assume that N is abelian or not simple perfect and that $\Gamma(\mathbb{Z}G/N) = \Pi(G/N)$ then*

$$\Gamma(\mathbb{Z}G) = \Pi(G).$$

Because for abelian (simple) groups the only torsion elements of $V(\mathbb{Z}G)$ are the elements of G it follows immediately that the Gruenberg–Kegel graph of $\mathbb{Z}G$ coincides with the prime graph of G provided G is soluble [46].

So the remaining task for a given nonabelian simple group S is to determine the prime graph of $V(\mathbb{Z}H)$ for the almost simple groups H sandwiched between $S = \operatorname{Inn} S$ and $\operatorname{Aut} S$. In Section 5 we give a survey on all such results concerning sporadic simple groups. The rest of this section is devoted to the following result.

Theorem 4.2 *Let G be a finite group whose order is divisible by at most three primes. Then*

$$\Gamma(\mathbb{Z}G) = \Pi(G).$$

Proof By [46, Proposition 4.3] we may assume that G is not soluble and that G has no minimal normal subgroup which is abelian. By the classical Burnside result, groups of order $p^a \cdot q^b$ are soluble. By assumption $|\pi(G)| \leq 3$. Thus if G has more than one minimal normal subgroup which is perfect or one minimal normal subgroup which is not simple then the prime graph of G is a complete graph. Because the vertices of $\Gamma(\mathbb{Z}G)$ and $\Pi(G)$ coincide [23] it follows in this case that $\Gamma(\mathbb{Z}G) = \Pi(G)$.[2] So it remains to consider the almost simple groups whose order is divisible by exactly three primes. By the classification of the finite simple groups (see also [27]), the simple groups S with $|\pi(S)| = 3$ are as given in the table below. We also include in the table the isomorphism type of $\operatorname{Out} S$. In the third column we indicate with (ZC) when the Zassenhaus conjecture is established, with (SIP-C) when the the cyclic subgroups of $V(\mathbb{Z}S)$ are isomorphic to subgroups of S and with (PQ) that the prime graph question has an affirmative answer. The fourth column contains the references.

	$\operatorname{Out} S$		
$\mathrm{A}_5 \cong \mathrm{PSL}(2,5)$	C_2	(ZC)	[51]
$\mathrm{PSL}(2,7)$	C_2	(ZC)	[37]
$\mathrm{PSL}(2,8)$	C_3	(ZC)	[29, 48]
$\mathrm{A}_6 \cong \mathrm{PSL}(2,9)$	$C_2 \times C_2$	(ZC)	[35]
$\mathrm{PSL}(2,17)$	C_2	(ZC)	[29]
$\mathrm{PSL}(3,3)$	C_2	(PQ)	[2]
$\mathrm{PSP}(3,4) \cong \mathrm{U}(4,2)$	C_2	(PQ)	[48]
$\mathrm{U}(3,3)$	C_2	(PQ)	[48]

For the full automorphism groups of these simple groups the results are as follows.

[2] Note we also could have used Proposition 4.1. But for Theorem 4.2 the given direct argument is shorter.

$\mathrm{Aut}\,\mathrm{A}_5 \cong S_5$	(ZC)	[52]
$\mathrm{Aut}\,\mathrm{PSL}(2,7) \cong \mathrm{PGL}(2,7)$	(ZC)	[48]
$\mathrm{Aut}\,\mathrm{PSL}(2,8) \cong \mathrm{P\Gamma L}(2,8)$	(SIP-C)	[48]
$\mathrm{Aut}\,\mathrm{A}_6 \cong \mathrm{P\Gamma L}(2,9)$	(PQ)	[48]
$\mathrm{Aut}\,\mathrm{PSL}(2,17) \cong \mathrm{PGL}(2,17)$	(ZC)	[48]
$\mathrm{Aut}\,\mathrm{PSL}(3,3)$	(PQ)	[48]
$\mathrm{Aut}\,\mathrm{PSP}(3,4)$	(PQ)	[48]
$\mathrm{Aut}\,\mathrm{U}(3,3)$	(PQ)	[48]

(SIP-C) holds for S_6 by [48]. The other two subgroups of $\mathrm{Aut}\,A_6$ of index 2, the groups M_{10} and $PGL(2,9)$ could not be settled with the HeLP–method. The HeLP–method does not answer whether there are elements of order 6 in $V(\mathbb{Z}G)$. Very recently A. Bächle and L. Margolis were able to complete the proof of Theorem 4.2 using additionally special integral representations in order to settle the question on the elements of order 6. So (PQ) holds for these two groups [3]. □

5 Results on sporadic simple groups

For sporadic simple groups, the first result appeared in [10]. Later, a series of papers dedicated to further sporadic simple groups has emerged. At the moment, the following is known:

- (PQ) answered affirmatively for 13 sporadic simple groups:
  - M_{11} [10], M_{12} [19], M_{22} [16], M_{23} [12], M_{24} [15];
  - J_1, J_2, J_3 [7];
  - HS [14], McL [11], He [6], Ru [13], Suz [18]
- For $G = ON$ the prime graph of $V(\mathbb{Z}G)$ is not connected [6];
- For $G = Co_3$, Co_2 and $G = Co_1$, prime graphs of G and $V(\mathbb{Z}G)$ have the same number of components [17].

The technique used for these groups is based on the Hertweck–Luthar–Passi (HeLP) method. The key observation is that even being unable to prove the rational conjugacy for elements of orders that appear in G, one could try to use information about their partial augmentations to eliminate some torsion units in $V(\mathbb{Z}G)$.

Some optimisations of the HeLP–method has been developed, in particular, the notion of (p,q)-constant characters and a hybrid symbolic–numeric approach, see [17] for the latest exposition of these techniques. In particular, in some cases it became possible to answer (PQ) by eliminating torsion units of order pq without preliminary enumeration of all admissible tuples of partial augmentations of elements of orders p and q.

The next table summarises results about possible orders and partial augmentations of normalised torsion units in integral group rings for the 17 sporadic simple groups listed above. It is divided into three cases w.r.t. (PQ) by horizontal lines.

- Column 2 lists orders (of normalised torsion units) for which the rational conjugacy is known, either as an immediate consequence of [37, Proposition 3.1] or using the HeLP–method for orders displayed using the bold face.

- Column 3 lists orders of elements of G with remaining non-trivial tuples of partial augmentations. In each entry of the form M(N), M means the order and N means the number of all (both trivial and non-trivial) admissible tuples of partial augmentations that are produced by the HeLP–method. Here and in columns 5 and 6 numbers in *italics* denote cases where only computer-aided results are available (these were computed using the software by V. Bovdi and the second author developed on the base of the GAP package LAGUNA [20]); for other cases, theoretical proofs are available.
- Column 4 lists orders of elements of G which were omitted as not relevant to (PQ) (for some groups it was possible to cover all or most of orders, though). A dash (—) means that no orders were omitted.
- Column 5 lists orders that do not appear neither in G nor in $V(\mathbb{Z}G)$.
- Column 6 lists some orders with remaining non-trivial tuples of partial augmentations that still have to be eliminated for positive answers on (SIP-C) for M_{11}, M_{22}, J_1 or (PQ) for ON and Conway groups.

Remark 5.1 Infinite number of admissible tuples of partial augmentations for order 57 in ON corresponds to the condition $\nu_{3a} = -18$, $\nu_{19a}+\nu_{19b}+\nu_{19c} = 19$ from [6].

Remark 5.2 Thus, the remaining 9 more sporadic simple groups for which the (PQ) or its weakened variations are not yet settled, are Fi_{22}, HN, Ly, Th, Fi_{23}, J_4, Fi_{24}', B and M. Also, for the Tits group T, which is sometimes referred as the 27^{th} sporadic simple group, the second author and V. Bovdi established the positive answer to (PQ). This result will be reported elsewhere.

G	ZC	order(#) in G	Not considered orders in G	No orders in $V(\mathbb{Z}G)$	order(#) in $V(\mathbb{Z}G)$
1	2	3	4	5	6
M_{11}	2, 3, 5, **11**	4(2), 6(5), 8(4)	—	10, 15, 22, 24, 33, 55	12(2)
M_{12}	5	2(6), 3(5), 10(2), 11(4)	4, 6, 8	15, 22, 33, 55	
M_{22}	2, 3, 5	4(34), 6(15), 7(4), *8(76)*, 11(10)	—	10, 14, 15, 21, 22, 33, 35, 55, 77	*12 (1166)*
M_{23}	2, 3, 5, **23**	4(3), 6(21), 7(4), 8(10), 11(20), 15(6)	14	10, 21, 22, *28*, 33, 35, 46, 55, *56*, 69, 77, 115, 161, 253	
M_{24}	5, 11, **23**	2(6), 3(6), 7(4), 10(11), *15(34)*, *21(21)*	4, 6, 8, 12, 14	22, 33, 35, 46, 55, 69, 77, 115, 161, 253	

G	ZC	order(#) in G	Not considered orders in G	No orders in $V(\mathbb{Z}G)$	order(#) in $V(\mathbb{Z}G)$
1	2	3	4	5	6
J_1	2, 3, 7, 11, **19**	5(4), 6(6), 10(12), 15(4)	—	14, 21, 22, 33, 35, 38, 55, 57, 77, 95, 133, 209	*30(6)*
J_2	7, **15**	2(6), 3(3), 4(15), 5(10), 8(18)	6, 10, 12	14, 21, 35	
J_3	2	3(10), 4(3), 5(8), 8(15), 17(10), 19(10)	6, 9, 10, 12, 15	34, 38, 51, 57, 85, 95, 323	
HS	3, 7	2(6), 5(23), 11(10)	4, 6, 8, 12, 15, 20	14, 21, 22, 33, 35, *55*, 77	
McL	2	3(4), 5(6), 7(174), 11(20)	4, 6, 8, 9, 10, 12, 14, 15, 30	21, 22, 33, 35, 55, 77	
He	5	2(13), 3(10), 17(30)	4, 6, 7, 8, 10, 12, 14, 15, 21, 28	34, 35, 51, 85, 119	
Ru	3, 7, 13	2(22), 5(8), 29(10)	4, 6, 8, 10, 12, 14, 15, 16, 20, 24, 26	21, 35, 39, 58, 65, 87, 91, 145, 203, 377	
Suz	7, 11	2(8), 5(10), 13(18)	4, 6, 8, 9, 10, 12, 14, 15, 18, 20, 21, 24	22, 26, 33, 35, 39, 55, 65, 77, 91, 143	
ON	2, 3, 5, 11	7(26), *19(2145)*, 31(80)	4, 6, 8, 10, 12, 14, 15, 16, 20, 28	21, 22, 35, 38, 55, 62, 77, 93, 95, 133, 155, 209, 217, 341, 589	33(1), 57 (∞)
Co_3	7	2(6), 3(155), *4(510)*, 5(6), 11(24), *14(5)*, 23(12)	6, 8, 9, 10, 12, 15, 18, 20, 21, 22, 24, 30	33, 46, 55, 69, 77, 115, 161, 253	35(2)

G	ZC	order(#) in G	Not considered orders in G	No orders in $V(\mathbb{Z}G)$	order(#) in $V(\mathbb{Z}G)$
1	2	3	4	5	6
Co_2	7, 11	2(48), 3(4), 5(6), 23(66)	4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 24, 28, 30	21, 22, 33, 46, 55, 69, 77, 115, 161, 253	35(2)
Co_1	11, 13	*2(216)*, *3(15239)* *5(1041)*, 7(47), 23(58588)	6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22, 24, 26, 28, 30, 33, 35, 36, 39, 40, 42, 60	*46*, *69*, 77, 91, *115*, 143, 161, 253, 299	*55(36)*, *65(14)*

As shown in Section 3 (cf. [48]), it is important to consider also automorphism groups of sporadic simple groups. Below we will give an outline of the proof that for all 13 sporadic simple groups where (PQ) has a positive answer, is has also a positive answer for their automorphism groups.

Theorem 5.3 *Let G be one of the following sporadic simple groups: M_{12}, M_{22}, J_2, J_3, HS, McL, He, Suz. Then (PQ) holds for* $\operatorname{Aut} G$*, i.e.,*

$$\Gamma(\mathbb{Z}\operatorname{Aut} G) = \Pi(\operatorname{Aut} G).$$

Accordingly to [24], the following 5 groups have trivial $\operatorname{Out} G$: M_{11}, M_{23}, M_{24}, J_1, Ru. Thus we get together with the results of Section 4 the following consequence.

Corollary 5.4 *(PQ) holds for a finite group provided its composition factors S are isomorphic to one of the following 13 sporadic simple groups:*

$$M_{11},\ M_{12},\ M_{22},\ M_{23},\ M_{24},\ J_1,\ J_2,\ J_3,\ HS,\ McL,\ He,\ Ru,\ Suz,$$

or $|\pi(S)| = 3$, or S is of prime order.

Proof of Theorem 5.3. We have to consider 8 groups, namely M_{12}, M_{22}, J_2, J_3, HS, McL, He, Suz. We will give here several examples to demonstrate our methods, and will only give an outline proof for the remaining cases, due to space considerations and the review nature of the paper. The full proof will be published elsewhere.

• Let $G = Aut(M_{12})$. Then $|G| = 2^7 \cdot 3^3 \cdot 5 \cdot 11$, $\exp(G) = 2^3 \cdot 3 \cdot 5 \cdot 11$ and we have to show that there are no units of orders 15, 22, 33 and 55 in $V(\mathbb{Z}G)$.

For this group, we may complete the entire proof by using the technique of (p, q)-constant characters [17], which, when successful, allows to eliminate torsion units of order pq without preliminary consideration of units of orders p and q. We will use the ordinary and p-Brauer character tables of G for $p \in \{2, 3, 5, 11\}$ which can be found using the Character Table Library [21] of the computational algebra system GAP [28], deriving its data from [24, 44]. Therefore, we will use the same notation, including indexation, for characters and conjugacy classes as used in the GAP Character Table Library.

Let $u \in V(\mathbb{Z}G)$ be a unit of order 22. Then $\nu_2 + \nu_{11} = 1$, and all other its partial augmentations are equal to zero by Theorem 2.6. Now consider the ordinary character $\chi = \chi_{12} + \chi_{15}$ (notation $\chi = \chi_{(n_1,\ldots,n_s)_{[p]}}$ indicates that $\chi = \chi_{n_1} + \cdots + \chi_{n_s}$ and p is either 0 when ordinary characters are used, or corresponding p when p-Brauer character tables are used). Then χ is $(2, 11)$-constant character, since $\chi(C_2) = 5$ and $\chi(C_{11}) = 0$. Let ζ be a primitive 22^{nd} root of unity. Using Equation (1) from Section 4, we obtain the following system of constraints on partial augmentations of u:

$$\mu(u, 1, \chi_{(12,15)_{[0]}}) = \tfrac{1}{22}(50\nu_2 + 170) \geq 0;$$
$$\mu(u, \zeta^{11}, \chi_{(12,15)_{[0]}}) = \tfrac{1}{22}(-50\nu_2 + 160) \geq 0,$$

which yields $-3 \leq \nu_2 \leq 3$. Now the condition

$$\mu(u, \zeta^2, \chi_{(12,15)_{[0]}}) = \tfrac{1}{22}(-5\nu_2 + 170) \geq 0$$

eliminates these cases, since the left hand side is not an integer for any $-3 \leq \nu_2 \leq 3$.

The proof for units of orders 15, 33 and 55 can be derived similarly from the table below containing the data for the constraints on partial augmentations ν_p and ν_q for possible orders pq (including the order 22 as well) to write the constraint

$$\mu(u, \zeta^l, \chi) = \tfrac{1}{pq}\,(m_1 + \nu_p m_p + \nu_q m_q) \geq 0, \tag{2}$$

where ζ is the pq^{th} primitive root of unity.

$\lvert u\rvert$	p	q	χ	$\chi(C_p)$	$\chi(C_q)$	l	m_1	m_p	m_q
15	3	5	$\chi = (3,4)_{[0]}$	0	4	0	70	0	32
						5	70	0	16
22	2	11	$\chi = (12,15)_{[0]}$	5	0	0	170	50	0
						2	170	-5	0
						11	160	-50	0
33	3	11	$\chi = (7)_{[0]}$	0	-1	0	44	0	-20
						11	44	0	10
55	5	11	$\chi = (3)_{[0]}$	2	0	0	30	80	0
						1	20	2	0
						11	20	-20	0

Similarly, (p, q)-constant characters technique suffices to show that there are no units of orders 15, 21, 22, 33, 35, 55 and 77 in $V(\mathbb{Z}G)$ for $G = \mathrm{Aut}(M_{22})$; of orders 21 and 35 for $G = \mathrm{Aut}(J_2)$; of orders 38, 51, 57, 85, 95 and 323 for $G = \mathrm{Aut}(J_3)$. For $G = \mathrm{Aut}(HS)$, however, it suffices for orders 21, 33, 55, 77 but we did not manage to find suitable (p, q)-constant characters for orders 22 and 35. However, they were eliminated using a different technique, which we will demonstrate later describing the case of $G = \mathrm{Aut}(Suz)$.

• Let $G = \mathrm{Aut}(McL)$. Then we have to show that there are no units of orders 21, 33, 35, 55 and 77 in $V(\mathbb{Z}G)$. For all orders except 35, the proof may be obtained using (p, q)-constant characters method. For the order 35, we substitute into Equation (1)

in Section 4 partial augmentations of u^5 and u^7 as described in [17] and obtain for the ordinary character χ_1 that

$$\mu(u, 1, \chi_1) = \tfrac{1}{35}(6\nu_{7a_5} + 1) \geq 0$$

where ν_{7a_5} is the partial augmentation of u^5 corresponding to the class $7a$. Since elements of order 7 are rationally conjugate to an element of G, we have $\nu_{7a_5} = 1$ and $\mu(u, 1, \chi_1)$ is not an integer, a contradiction.

• Let $G = \mathrm{Aut}(He)$. Then we have to show that there are no units of orders 34, 35, 51, 85 and 119 in $V(\mathbb{Z}G)$. Here the (p, q)-constant characters method works for all orders except 35. For order 35, we substitute into Equation (1) in Section 4 partial augmentations of u^5 and u^7 and obtain that for the ordinary character χ_1

$$\mu(u, 1, \chi_1) = \tfrac{1}{35}(6(\nu_{7a_5} + \nu_{7b_5} + \nu_{7c_5}) + 1) \geq 0$$

where $\mu(u, 1, \chi_1)$ is not an integer since $\nu_{7a_5} + \nu_{7b_5} + \nu_{7c_5} = 1$ (note that the HeLP–method implementation by V. Bovdi and the second author produces 225 admissible tuples of partial augmentation for units of order 7, but we do not need to check them individually since we obtained a contradiction in a much more elegant way).

• Let $G = \mathrm{Aut}(Suz)$. Then we have to show that there are no units of orders 26, 33, 35, 39, 55, 65, 77, 91 and 143 in $V(\mathbb{Z}G)$.

First, (p, q)-constant characters method succeeds for orders 55, 65, 77, 91, 143.

For order 35, we obtain that for the ordinary character χ_1

$$\mu(u, 1, \chi_1) = \tfrac{1}{35}(6\nu_{7a_5} + 1) \geq 0.$$

Since elements of order 7 are rationally conjugate to elements of G, we have $\nu_{7a_5} = 1$, so $\mu(u, 1, \chi_1)$ is not an integer, a contradiction.

For order 39, we substitute into Equation (1) in Section 4 partial augmentations of u^3 and u^{13} as described in [17] and obtain the following system using ordinary characters χ_3, χ_5 and χ_7 (ζ is the 39^{th} root of unity):

$$\begin{aligned}
\mu(u, 1, \chi_3) &= \tfrac{1}{39}(840\nu_{3a} + 192\nu_{3b} - 24\nu_{3c} + 70\nu_{3a_{13}} + 16\nu_{3b_{13}} - 2\nu_{3c_{13}} + 143) \geq 0;\\
\mu(u, \zeta, \chi_3) &= \tfrac{1}{39}(35\nu_{3a} + 8\nu_{3b} - \nu_{3c} - 35\nu_{3a_{13}} - 8\nu_{3b_{13}} + \nu_{3c_{13}} + 143) \geq 0;\\
\mu(u, \zeta^3, \chi_3) &= \tfrac{1}{39}(-70\nu_{3a} - 16\nu_{3b} + 2\nu_{3c} + 70\nu_{3a_{13}} + 16\nu_{3b_{13}} - 2\nu_{3c_{13}} + 143) \geq 0;\\
\mu(u, \zeta^{13}, \chi_3) &= \tfrac{1}{39}(-420\nu_{3a} - 96\nu_{3b} + 12\nu_{3c} - 35\nu_{3a_{13}} - 8\nu_{3b_{13}} + \nu_{3c_{13}} + 143) \geq 0;\\
\mu(u, 1, \chi_5) &= \tfrac{1}{39}(-336\nu_{3a} + 312\nu_{3b} + 96\nu_{3c} - 28\nu_{3a_{13}} + 26\nu_{3b_{13}} + 8\nu_{3c_{13}} + 364) \geq 0;\\
\mu(u, \zeta, \chi_5) &= \tfrac{1}{39}(-14\nu_{3a} + 13\nu_{3b} + 4\nu_{3c} + 14\nu_{3a_{13}} - 13\nu_{3b_{13}} - 4\nu_{3c_{13}} + 364) \geq 0;\\
\mu(u, \zeta^3, \chi_5) &= \tfrac{1}{39}(28\nu_{3a} - 26\nu_{3b} - 8\nu_{3c} - 28\nu_{3a_{13}} + 26\nu_{3b_{13}} + 8\nu_{3c_{13}} + 364) \geq 0;\\
\mu(u, \zeta^{13}, \chi_5) &= \tfrac{1}{39}(168\nu_{3a} - 156\nu_{3b} - 48\nu_{3c} + 14\nu_{3a_{13}} - 13\nu_{3b_{13}} - 4\nu_{3c_{13}} + 364) \geq 0;\\
\mu(u, 1, \chi_7) &= \tfrac{1}{39}(2520\nu_{3a} - 72\nu_{3b} + 144\nu_{3c} + 210\nu_{3a_{13}} - 6\nu_{3b_{13}} + 12\nu_{3c_{13}} + 780) \geq 0;\\
\mu(u, \zeta^{13}, \chi_7) &= \tfrac{1}{39}(-1260\nu_{3a} + 36\nu_{3b} - 72\nu_{3c} - 105\nu_{3a_{13}} + 3\nu_{3b_{13}} - 6\nu_{3c_{13}} + 780) \geq 0;
\end{aligned}$$

which, as confirmed using the HeLP–method implementation by V. Bovdi and the second author, has no solutions. The same technique works for eliminating orders 26 and 33. This completes the proof.

References

[1] H. Abold and W. Plesken, Ein Sylowsatz für endliche p-Untergruppen von $\mathrm{GL}(n,\mathbb{Z})$, *Math. Ann.* **232** (1978), 183–186.

[2] A. Bächle and W. Kimmerle, On torsion subgroups in integral group rings of finite groups, *J. Algebra* **326** (2011), 34–46.

[3] A. Bächle and L. Margolis, Rational conjugacy of torsion units in integral group rings of non-solvable groups, arXiv:1305.7419v2 [math.RT], 27 February 2014.

[4] S. D. Berman, On certain properties of integral group rings (Russian), *Dokl. Akad. Nauk SSSR (N.S.)* **91** (1953), 7–9.

[5] P. F. Blanchard, Exceptional group ring automorphisms for groups of order 96, *Comm. Algebra* **29** (2001), no.11, 4823–4830.

[6] V. Bovdi, A. Grishkov and A. Konovalov, Kimmerle conjecture for the Held and O'Nan sporadic simple groups. *Sci. Math. Jpn.* **69** (2009), no.3, 353–361.

[7] V. Bovdi, E. Jespers and A. Konovalov, Torsion units in integral group rings of Janko simple groups, *Math. Comp.* **80** (2011), no.273, 593–615.

[8] V. Bovdi and M. Hertweck, Zassenhaus conjecture for central extensions of S_5. *J. Group Theory* **11** (2008), no.1, 63–74.

[9] V. Bovdi, C. Höfert and W. Kimmerle, On the first Zassenhaus conjecture for integral group rings, *Publ. Math. Debrecen* **65/3-4** (2004), 291–303.

[10] V. Bovdi and A. Konovalov, Integral group ring of the first Mathieu simple group, in *Groups St. Andrews 2005, Vol. 1* (C. M. Campbell et al., eds.), London Math. Soc. Lecture Note Ser. **399** (CUP, Cambridge 2007), 237–245.

[11] V. Bovdi and A. Konovalov, Integral group ring of the McLaughlin simple group, *Algebra Discrete Math.* **2** (2007), 43–53.

[12] V. Bovdi and A. Konovalov, Integral group ring of the Mathieu simple group M_{23}, *Comm. Algebra* **36** (2008), no.7, 2670–2680.

[13] V. Bovdi and A. Konovalov, Integral group ring of Rudvalis simple group, *Ukraïn. Mat. Zh.* **61** (2009), no.1, 3–13.

[14] V. Bovdi and A. Konovalov, Torsion units in integral group ring of Higman-Sims simple group, *Studia Sci. Math. Hungar.* **47** (2010), no.1, 1–11.

[15] V. Bovdi and A. Konovalov, Integral group ring of the Mathieu simple group M_{24}, *J. Algebra Appl.* **11** (2012), no.1, 1250016.

[16] V. Bovdi, A. Konovalov and S. Linton, Torsion units in integral group ring of the Mathieu simple group M_{22}, *LMS J. Comput. Math.* **11** (2008), 28–39.

[17] V. Bovdi, A. Konovalov and S. Linton, Torsion units in integral group rings of Conway simple groups, *Internat. J. Algebra Comput.* **21** (2011), no.4, 615–634.

[18] V. Bovdi, A. Konovalov and E.N. Marcos, Integral group ring of the Suzuki sporadic simple group. *Publ. Math. Debrecen* **72** (2008), no.3-4, 487–503.

[19] V. Bovdi, A. Konovalov and S. Siciliano, Integral group ring of the Mathieu simple group M_{12}, *Rend. Circ. Mat. Palermo (2)* **56** (2007), no.1, 125–136.

[20] V. Bovdi, A. Konovalov, R. Rossmanith and Cs. Schneider, *LAGUNA — Lie AlGebras and UNits of group Algebras, Version 3.6.1*; GAP package, 2012, http://www.cs.st-andrews.ac.uk/~alexk/laguna/.

[21] T. Breuer, *The GAP Character Table Library, Version 1.2.1*; GAP package, 2012, http://www.math.rwth-aachen.de/~Thomas.Breuer/ctbllib.

[22] M. Caicedo, A. del Rio and L. Margolis, Zassenhaus conjecture for cyclic-by-abelian groups, *J. London Math. Soc. (2)* **88** (2013), 65 – 78.

[23] J. A. Cohn and D. Livingstone, On the structure of group algebras I, *Canad. J. Math.* **17** (1965), 583–593.

[24] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, *Atlas of Finite Groups* (OUP, Oxford 1985). Maximal subgroups and ordinary characters for simple groups, With computational assistance from J. G. Thackray.

[25] M. A. Dokuchaev and S. O. Juriaans, Finite subgroups in integral group rings, *Canad. J. Math.* **48** (1996), No.6, 1170–1179.
[26] M. A. Dokuchaev, S. O. Juriaans and C. Polcino Milies, Integral group rings of Frobenius groups and the conjectures of H. J. Zassenhaus, *Comm. Algebra* **25** (1997), no.7, 2311–2325.
[27] W. Feit, The current situation in the theory of finite simple groups, *Actes. Congres. Intern. Math.* 1970, Tome 1, 55–93.
[28] The GAP Group, *GAP — Groups, Algorithms, and Programming, Version 4.6.4*; 2013, http://www.gap-system.org.
[29] J. Gildea, Zassenhaus conjecture for integral group ring of simple linear groups, *J. Algebra Appl.* **12** (2013), 1350016.
[30] K. W. Gruenberg and K. W. Roggenkamp, Decomposition of the augmentation ideal and of the relation modules of a finite group, *Proc. London Math. Soc. (3)* **31** (1975), 149–166.
[31] A. W. Hales, I. S. Luthar and I. B. S. Passi, Partial augmentations and Jordan decomposition in group rings. *Comm. Algebra* **18** (1990), no.7, 2327–2341.
[32] M. Hertweck, A counterexample to the isomorphism problem for integral group rings, *Ann. of Math.* **154** (2001), 115–138.
[33] M. Hertweck, Unit groups of integral finite group rings with no noncyclic abelian finite p-subgroups, *Comm. Algebra* **36** (2008), no.9, 3224–3229.
[34] M. Hertweck, The orders of torsion units in integral group rings of finite solvable groups, *Comm. Algebra* **36** (2008), no.10, 3585–3588.
[35] M. Hertweck, Zassenhaus Conjecture for A_6, *Proc. Indian Acad. Sci. Math. Sci.* **118** (2008), no.2, 189–195.
[36] M. Hertweck, Torsion units in integral group rings of certain metabelian groups, *Proc. Edinb. Math. Soc. (2)* **51** (2008), no.2, 363–385.
[37] M. Hertweck, On the torsion units of some integral group rings, *Algebra Colloq.* **13** (2006), no.2, 329–348.
[38] M. Hertweck, Partial augmentations and Brauer character values of torsion units in group rings, Preprint arXiv:math.RA/0612429, 2007.
[39] M. Hertweck, Contributions to the integral representation theory of groups, Elektronische Hochschulschrift, http://elib.uni-stuttgart.de/opus, Habilitationsschrift, Universität Stuttgart, 2004.
[40] M. Hertweck, C. Höfert and W. Kimmerle, Finite groups of units and their composition factors in the integral group rings of the groups PSL$(2,q)$, *J. Group Theory* **12** (2009), no.6, 873–882.
[41] G. Higman, Units in group rings, D.Phil. thesis, Oxford Univ., 1940.
[42] C. Höfert and W. Kimmerle, On Torsion Units of Integral Group Rings of Small Order, in *Groups, Rings and Group Rings* (A. Giambruno, C. P. Milies and S.K.Sehgal, eds.), Lect. Notes in Pure Math., **248** (Chapman & Hall 2006), 243–252.
[43] E. Jespers, W. Kimmerle, G. Nebe and Z. Marciniak, Miniworkshop Arithmetik von Gruppenringen, vol.4 Oberwolfach Reports no.55 (EMS, Zürich 2007), 3149–3179,
[44] C. Jansen, K. Lux, R. Parker and R. Wilson, *Atlas of Brauer Characters*, volume 11 of *London Mathematical Society Monographs (New Series)*, (Clarendon Press Oxford University Press, New York, 1995). Appendix 2 by T. Breuer and S. Norton, Oxford Science Publications.
[45] S.O.Juriaans and C. Polcino Milies, Units of integral group rings of Frobenius groups, *J. Group Theory* **3** (2000), no.3, 277–284.
[46] W. Kimmerle, On the prime graph of the unit group of integral group rings of finite groups, *Contemporary Mathematics* **420** (2006), 215–228.
[47] W. Kimmerle and K. W. Roggenkamp, A Sylowlike theorem for integral group rings of finite solvable groups, *Arch. Math. (Basel)* **60** (1993), no.1, 1–6.

[48] W. Kimmerle and A. Konovalov, On the prime graph of the unit group of integral group rings of finite groups II, *Stuttgarter Mathematische Berichte* 2012–018, http://www.mathematik.uni-stuttgart.de/fachbereich/math-berichte/listen.jsp.

[49] L. Klingler, Construction of a counterexample to a conjecture of Zassenhaus, *Comm. Algebra* **19** (1991), no.8, 2303–2330.

[50] A. S. Kondrat'ev, On prime graph components of finite simple groups (Russian), *Mat. Sb.* **180** (1989), no.6, 787–797,864; translated in *Math. USSR Sb.* **67** (1990).

[51] I. S. Luthar and I. B. S. Passi, Zassenhaus conjecture for A_5, *J. Nat. Acad. Math. India* **99** (1989), 1–5.

[52] I. S. Luthar and Poonam Trama, Zassenhaus conjecture for S_5, *Comm. Algebra* **19** (1991), no.8, 2353–2362

[53] Z. S. Marciniak and S. K. Sehgal, The unit group of $1 + \Delta(G)\Delta(A)$ is torsion-free, *J. Group Theory* **6** (2003), 223–228.

[54] Z. S. Marciniak, J. Ritter, S. K. Sehgal and A. Weiss, Torsion Units in Integral Group Rings of Some Metabelian Groups, II, *J. Number Theory* **25** (1987), 340–352.

[55] D. S. Passman, Isomorphic groups and group rings, *Pacific J. Math.* **15** (1965), 561–583.

[56] K. W.Roggenkamp, Observations to a conjecture of H. Zassenhaus, *Groups St Andrews 1989*, LMS Lecture Note Ser. **160** (1991), 427–444.

[57] K. W. Roggenkamp and L. L. Scott, Isomorphisms of p-adic group rings, *Ann. of Math.* **126** (1987), 593–647.

[58] R. Sandling, Graham Higman's thesis "Units in Group Rings" in Integral Representations and Applications ed. K.W. Roggenkamp, Springer Lecture Notes **882** (1981), 93–116.

[59] L. L. Scott, On a conjecture of Zassenhaus, and beyond, Algebra, Proc. Internat. Conf. Memory A.I. Malcev, Novosibirsk 1989, *Contemporary Mathematics* **131** (1992), no.1, 325–343.

[60] S. K. Sehgal, *Units in integral group rings* (with an appendix of A. Weiss), Pitman Monographs and Surveys in Pure and Applied Mathematics, **69** (Longman Scientific & Technical, Harlow, 1993).

[61] S. K. Sehgal, Group Rings, in *Handbook of Algebra, Vol. 3* (ed. by M.Hazewinkel), (Elsevier Science, 2003), 455–541.

[62] J. H. Walter, The characterization of finite groups with Abelian Sylow 2-subgroups, *Ann. of Math.* **89** (1969), 405–514.

[63] A. Weiss, Rigidity of p-adic p-torsion, *Ann. of Math. (2)* **127** (1987), no.2, 317–332.

[64] A. Weiss, Torsion units in integral group rings, *J. Reine Angew. Math.* **415** (1991), 175–187.

[65] J. S. Williams, Prime graph components of finite groups, *J. Algebra* **21** (1981), 487–513.

[66] H. Zassenhaus, On the torsion units of finite group rings, *Studies in Math. (in honour of A. Almeida Costa)*, (Instituto de Alta Cultura, Lisboa, 1974), 119–126.

ON FINITE GROUPS WITH SMALL PRIME SPECTRUM[1]

ANATOLY S. KONDRATIEV* and IGOR V. KHRAMTSOV†

Institute of Mathematics and Mechanics of UB RAS 16, Ekaterinburg, 620990, Russia
*Email: a.s.kondratiev@imm.uran.ru
†Email: ihramtsov@gmail.com

Abstract

We survey the recent results of the authors on finite groups with a small prime spectrum. The prime spectrum $\pi(G)$ of a finite group G is the set of prime divisors of its order. If $|\pi(G)| = n$ then G is called n-primary. We describe the chief factors of 3-primary groups and the chief factors of commutator subgroups of 4-primary groups whose prime graphs are disconnected. As a corollary, 3-primary finite almost simple groups and 4-primary finite simple groups recognizable by prime graph are determined. The complete irreducibility of $GF(2)A_7$-modules in which an element of order 5 acts fixed-point-freely is proved. Finite groups with the same prime graph as the group $\mathrm{Aut}(J_2)$ or A_{10} are described.

1 Introduction

In finite group theory, many researchers are interested in various problems of recognizability, i.e., in the characterization of a group by a certain set of its parameters up to isomorphism. Examples of such problems are the problems of recognizing finite groups by their spectrum or prime graph. Let G be a finite group. Denote by $\pi(G)$ the set of prime divisors of the order of G. We also call $\pi(G)$ the *prime spectrum*. Denote by $\omega(G)$ the *spectrum* of the group G, i.e., the set of its element orders. The set $\omega(G)$ defines the *prime graph* (the *Gruenberg–Kegel graph*) $\Gamma(G)$ of the group G; in this graph, the vertex set is $\pi(G)$ and two different vertices p and q are connected by an edge if and only if $pq \in \omega(G)$. Denote the number of connected components of $\Gamma(G)$ by $s(G)$ and the set of connected components by $\{\pi_i(G) \mid 1 \leq i \leq s(G)\}$; for a group G of even order, we assume that $2 \in \pi_1(G)$.

The notion of prime graph appeared by the investigation some cohomological problems related to integer representations of finite groups and was found very fruitful.

A group G is called *recognizable by its spectrum* (resp., *prime graph*) if, for any finite group H, the equality $\omega(H) = \omega(G)$ (resp., $\Gamma(H) = \Gamma(G)$) implies a group isomorphism $H \cong G$. Here, the equality of the graphs $\Gamma(H)$ and $\Gamma(G)$ means the coincidence of the sets of their vertices and their edges, respectively. It is clear that the recognizability of a finite group by prime graph implies its recognizability by its

The work is supported by the RFBR (grant no. 13-01-00469), the RFBR–NSFC (grant no. 12-01-91155), the Program of the Division of Mathematical Sciences of the Russian Academy of Sciences (grant no. 12-T-1-1003), and the Programs of the Joint Investigations of the Ural Branch of the Russian Academy of Sciences with Sibirian Division of the Russian Academy of Sciences (grant no. 12-1-10018) and with the Belarussian National Academy of Sciences (grant no. 12-C-1-1009), by the grant of the President of Russian Federation for young scientists (grant no. MK-3395.2012.1) and a grant from the IMM of UB RAS for young scientists in 2013.

spectrum. At present, large progress in the solving the problem of the recognizability of finite simple groups by spectrum has been achieved and the problem of the recognizability of finite simple groups by prime graph has been intensively investigated.

In 1996, Chen [5] proved that any sporadic simple group is recognizable by its order and prime graph up to isomorphism in class of all finite groups. In 2003, Hagie [6] gave the first examples of finite groups recognizable by prime graph, namely the sporadic simple groups J_1, M_{22}, M_{23}, M_{24} and Co_2, and obtained a description of finite groups G such that $\Gamma(G) = \Gamma(S)$, where S is a sporadic simple group. But the description was not complete classification. Further many authors obtained other results in this direction.

The problem of recognition of a group by prime graph is a particular case of more general problem: *study finite groups by the properties of their prime graphs*. In the frame of this general problem, our attention draws first of all a more detailed study of the class of finite groups with disconnected prime graph. In fact, this class generalizes widely the class of finite Frobenius groups as is obvious from the well-known stuctural Gruenberg-Kegel theorem on finite groups with disconnected prime graph (see [40]). And Frobenius groups occupy an absolutely exceptional place in the finite group theory. Note also that the class of finite groups with disconnected prime graph coincides with the class of finite groups having an *isolated subgroup* (i.e., a proper subgroup containing the centralizer of any its nontrivial element) which have been studied without the classification of finite simple groups by many known algebraists (Frobenius, Suzuki, Feit, Thompson, G. Higman, Arad, Chillag, Busarkin, Gorchakov, Podufalov, and others). See, for example, [1]).

The classification of connected components of prime graph for finite simple groups was established in papers of Williams [40], the first author [21], and Iiyori and Yamaki [15]. Lucido [29] extended this classification onto all finite *almost simple groups*, i.e., groups with nonabelian simple socle. The finite simple groups with disconnected prime graph compose a sufficiently restricted class of all finite simple groups, but include many "small" (in various senses) groups which arise often in the investigations. For example, all finite simple groups of exceptional Lie type besides the the groups $E_7(q)$ for $q > 3$, as well as simple groups from the well-known "Atlas of finite groups" [7] besides the group A_{10}, have disconnected prime graphs.

The problem of the study of finite unsoluble groups with disconnected prime graph, which are not almost simple, is solved for several particular cases only, because here some nontrivial problems related with modular representations of finite almost simple groups arise. Let us consider such a problem.

Let G be a finite group with disconnected prime graph, and let G be nonisomorphic to a Frobenius group or a 2-Frobenius group. A 2-*Frobenius group* is a finite group G such that $G = ABC$ where A and AB are normal subgroups in G, and AB and BC are Frobenius groups with kernels A and B and complements B and C, respectively. Then, by the Gruenberg–Kegel theorem, the group $\overline{G} := G/F(G)$ is almost simple and is known by the above mentioned results. Assume that $F(G) \neq 1$. Each connected component $\pi_i(G)$ of the graph $\Gamma(G)$ for $i > 1$ corresponds to a nilpotent isolated $\pi_i(G)$-Hall subgroup $X_i(G)$ of the group G. Any nontrivial element x from $X_i(G)$ $(i > 1)$ acts *fixed-point-freely (freely)* on $F(G)$, i.e., $C_{F(G)}(x) = 1$. Let K and L be two neighboring terms of a chief series of the group G and $K < L \leq F(G))$. Then,

the (chief) factor $V = L/K$ is an elementary abelian p-group for some prime p (we will call it the *p-chief factor* of the group G), and we can consider it as a faithful irreducible $GF(p)\overline{G}$-module (since $C_{G/K}(V) = F(G)/K$). Moreover, any nontrivial element from $X_i(G)$ $(i > 1)$ acts fixed-point-freely on V.

Therefore, the problem of studying the structure of the group G largely reduces to the following problem, which is of independent interest.

Problem 1.1 For the finite simple group G and given prime p, describe all irreducible $GF(p)G$-modules V such that an element of prime order $\neq p$ from G acts on V fixed-point-freely.

Extending and refining Problem 1.1 we obtain the following:

Problem 1.2 Let G be a finite group, Q be a normal nontrivial subgroup from G, $\overline{G} = G/Q$ be a known group and an element of prime order from $G\backslash Q$ acts on Q fixed-points-freely. The following questions arise.

1) What are the chief factors of the group G in Q?
2) What is the structure of the group Q?
3) If Q is elementary abelian group, is the action of $\overline{G}$ on Q completely irreducible?
4) Is the extension of G over Q splittable?

The well-known Thompson's theorem implies that Q is a nilpotent group in this situation.

In spite of importance of the questions 1) – 4), we have few results about them. In general, this important problem is far from being solved.

The first work, devoted to the study of the case when $\overline{G}$ is a simple nonabelian group, was a classical work of G. Higman [12]. If $\overline{G} \cong L_2(2^m)$ for $m \geq 2$ and an element of order 3 from G acts on Q fixed-point-freely then Higman gave affirmative answers on all above-formulated questions. In particular, Q is an elementary abelian 2-group, the action of $\overline{G}$ on Q is completely irreducible and every 2-chief factors of G is isomorphic to the natural $GF(2^m)SL_2(2^m)$-module.

Later Martineau [30, 31] obtained an analogous result for the case when $\overline{G}$ is isomorphic to the Suzuki group $Sz(2^n)$ and an element of order 5 from G acts on Q fixed-point-freely.

Continuing the work of Higman, Stewart [39] showed that $Q = 1$ in the case when $\overline{G} \cong L_2(q)$ for odd $q > 5$ and an element of order 3 from G acts on Q fixed-point-freely.

The papers of Prince [34, 35], Zurek [43], Holt and Plesken [13] were devoted to the study of the case, when $Q = O_2(G)$, $\overline{G} \cong A_5$ and an element of order 5 from G acts on Q fixed-point-freely. This case is difficult, because Q can be a nonabelian group. Prince and Zurek gave affirmative answers on the questions 1), 3) and 4). In particular, Q is a product of $\overline{G}$-invariant subgroups Q_i's, isomorphic to either a homocyclic 2-group of the rank 4, or the special 2-group of order 2^8 with the center of order 2^4 (isomorphic to the unipotent radical some parabolic maximal subgroup in $U_5(2)$). In addition, in the first case every 2-chief factor of G involving in Q_i is isomorphic to the orthogonal (permutational) $GF(2)A_5$-module, and in the second case the group $Z(Q_i)$ is isomorphic to the orthogonal $GF(2)A_5$-module, but $Q_i/Z(Q_i)$ is isomorphic

to the natural $GF(4)SL_2(4)$-module. By an early result of G. Higman [11], a theoretical upper bound of the nilpotency class of Q was 6. Zurek conjectured that such bound must be 2. But later on, Holt and Plesken proved that the nilpotency class of Q is at the most 3 and constructed an example of the group Q of order 2^{28} where this bound is reached. Using a computer, they showed also that this is an example of minimal order.

If $\overline{G} \cong L_2(7)$ and an element of order 7 from G acts on a 2-group Q fixed-point-freely then Holt and Plesken in [13] proved that neither the nilpotency class nor the derived length of Q can be bounded.

Prince [34, 35] proved that if $Q = O_2(G)$, $\overline{G} \cong A_6$ and an element of order 5 from G acts on Q fixed-point-freely then the questions 1) – 4) are solved affirmatively. Dolfi, Jabara and Lucido [8] proved that if $\overline{G} \cong A_6$ and an element of order 5 from G acts on Q fixed-point-freely, then $O(Q)$ is abelian, $O(Q) = O_3(G)$ and 3-chief factors of G are isomorphic to the 4-dimentional permutational $GF(3)\overline{G}$-module. In [8], it is asserted also that if $\overline{G} \cong A_5$ and an element of order 5 from G acts on Q fixed-point-freely, then $O(Q)$ is abelian. But this assertion is found wrong. Recently, Astill, Parker and Waldecker in[3] proved that in this situation $O(Q)$ is a nilpotent group of class at most 2 and, for any odd prime $p \neq 5$, constructed a r-group of class 2 admitting the group A_5 with the mentioned property.

If the socle of the group $\overline{G}$ is a finite simple group of Lie type over a field of a prime characteristic p, then, for the solving the item 1) of Problem 1.2, the classification of Guralnick and Tiep [9] of all unisingular finite simple group of Lie type is useful. A finite simple group S of Lie type over a field of a prime characteristic p is called *unisingular* if any element $s \in S$ has a non-trivial fixed point in any non-trivial finite abelian p-group on which S acts.

Zavarnitsine in [41, 42] found some sufficient conditions for an element of a large prime order in the group $S = L_n^{\pm}(q)$, where q is a power of a prime p, to have non-zero fixed points in S-modules over a field of characteristic p.

The mentioned results of Guralnick–Tiep and Zavarnitsine are useful for the study of reconizability of finite simple groups by spectrum or prime graph.

The considered partial results show that Problem 1.2 is complicated. In general, this important problem is far from being solved.

If the table of irreducible Brauer characters is known (for example, from [16] or [17]) then the following result can be applied for the solving the item 1) of Problem 1.2.

Proposition 1.3 *Suppose that G is a finite quasi-simple group, F is a field of characteristic $p > 0$, V is a faithful absolutely irreducible FG-module, and β is a Brauer character of the module V. If g is an element in G of a prime order coprime to $p|Z(G)|$, then*

$$\dim C_V(g) = (\beta|_{\langle g\rangle}, 1|_{\langle g\rangle}) = \frac{1}{|g|}\sum_{x\in\langle g\rangle}\beta(x).$$

In this paper, we survey the recent author's results on finite groups with small prime spectrum.

Our notation and terminology are mostly standard and can be found in [2, 7, 16]. Denote by G^{∞} the last member in the derived series of a finite group G.

2 Finite 3-primary and 4-primary groups with disconnected prime graph

The authors investigate finite groups whose prime graph is disconnected and has a small number of vertices. First of all, let us consider the trivial cases, when the prime graph of a finite group has one or two vertices. A finite group G is called *n-primary* if $|\pi(G)| = n$. The class of 1-primary groups coincides with the boundless class of all primary groups. Using Gruenberg-Kegel theorem and the properties of solvable complements in finite Frobenius groups, it is not difficult to describe 2-primary (biprimary) groups with disconnected prime graph. They are either Frobenius groups or 2-Frobenius groups of a special form.

In [24, 26], the authors described the chief factors of 3-primary groups with disconnected prime graph. In particular, the following theorem is proved.

Theorem 2.1 *Let G be a finite threeprimary group with disconnected prime graph and $\overline{G} = G/F(G)$. Then, one of the following statements holds:*

(1) *G is a Frobenius group.*

(2) *G is a 2-Frobenius group.*

(3) *$s(G) = 3$ and either G is isomorphic to A_5, A_6, $L_2(7)$, $L_2(8)$, M_{10} or $L_2(17)$, or $G/O_2(G) \cong L_2(2^n)$, where $n \in \{2, 3\}$ and $O_2(G)$ is a direct product of minimal normal subgroups of the order 2^{2n} in G, each of which as $\overline{G}$-module is isomorphic to the natural $GF(2^n)SL_2(2^n)$-module.*

(4) *$s(G) = 2$, $\pi_1(G) = \{2, 5\}$ and $G \cong PGL_2(9)$.*

(5) *$s(G) = 2$, $\pi_1(G) = \{2, 3\}$, $F(G) = O_2(G) \times O_3(G)$, and one of the following statements (i)–(viii) holds:*

   (i) *$\overline{G} \cong A_5$ or S_5, any 2-chief factor of the group G as $GF(2)\overline{G}$-module is isomorphic to one of two 4-dimensional irreducible $GF(2)\overline{G}$-modules, and any 3-chief factor of G as $GF(3)\overline{G}$-module is isomorphic to the 4-dimensional irreducible permutation $GF(3)\overline{G}$-module.*

   (ii) *$\overline{G} \cong A_6$, S_6 or M_{10}, $F(G)$ is the direct product of an elementary abelian 2-group and an abelian 3-group, and $F(G) \neq 1$ for $\overline{G} \cong A_6$ or M_{10}. If $O_2(G) \neq 1$ then $O_2(G)$ is the direct product of G'-invariant subgroups of order 16 that are as $GF(2)\overline{G}'$-module isomorphic to either the 4-dimensional irreducible permutation $GF(2)A_6$-module or conjugated with them by an outer automorphism of S_6. Any 3-chief factor in G' as $GF(3)\overline{G}'$-module is isomorphic to the 4-dimensional irreducible permutation $GF(3)A_6$-module.*

   (iii) *$\overline{G} \cong U_4(2)$ and $F(G) = O_2(G)$ is an elementary abelian 2-group. Any 2-chief factor of the group G as $GF(4)\overline{G}$-module is isomorphic to the natural unitary 4-dimensional $GF(4)SU_4(2)$-module.*

   (iv) *$\overline{G} \cong L_2(8)$ or $Aut(L_2(8))$, $F(G) = O_2(G)$, and $F(G) \neq 1$ for $\overline{G} \cong L_2(8)$. Any 2-chief factor of the group G' as $GF(8)\overline{G}'$-module is isomorphic to the natural 2-dimensional $GF(8)SL_2(8)$-module or 4-dimensional irreducible $GF(8)L_2(8)$-module.*

   (v) *$\overline{G} \cong L_2(7)$ or $PGL_2(7)$, and $F(G) \neq 1$ for $\overline{G} \cong L_2(7)$. Any 2-chief*

factor of the group G' as $GF(2)\overline{G}'$-module is isomorphic to the natural 3-dimensional $GF(2)SL_3(2)$ modul or to the module conjugated with them by an outer involutive automorphism of the group $SL_3(2)$. Any 3-chief factor of the group G' as $\overline{G}'$-module is isomorphic to the 3-dimensional irreducible $GF(9)L_2(7)$-module or the 6-dimensional absolutely irreducible $GF(3)L_2(7)$-module.

(*vi*) $\overline{G} \cong U_3(3)$ *or* $\mathrm{Aut}(U_3(3))(\cong G_2(2))$. *Any 2-chief factor of the group G as $GF(2)\overline{G}$-module is isomorphic to the 6-dimensional absolutely irreducible $GF(2)\overline{G}$-module. Any 3-chief factor of the group G' as $GF(9)\overline{G}'$-module is isomorphic to the natural unitary 3-dimensional $GF(9)U_3(3)$-module or the 6-dimensional $GF(9)U_3(3)$-module.*

(*vii*) $\overline{G} \cong L_3(3)$ *or* $\mathrm{Aut}(L_3(3))$. *Any 2-chief factor of G' is isomorphic as $GF(2)\overline{G}'$-module to the 12-dimensional absolutely irreducible $GF(2)L_3(3)$-module. Any 3-chief factor of the group G' as $GF(3)\overline{G}'$-module is isomorphic to one of the three absolutely irreducible $GF(3)L_3(3)$-modules of the dimensions 3, 6 or 15; for those dimensions up to isomorphism there exists exactly two $GF(3)L_3(3)$-modules that are conjugated by an outer involutive automorphism of the group $L_3(3)$.*

(*viii*) $\overline{G} \cong L_2(17)$ *or* $PGL_2(17)$ *and* $F(G) \neq 1$ *for* $\overline{G} \cong L_2(17)$. *Any 2-chief factor of the group G' as $\overline{G}'$-module is isomorphic either to the 8-dimensional absolutely irreducible $GF(2)L_2(17)$-module, to the module conjugated with them by an outer involutive automorphism of the group $L_2(17)$, to the 16-dimensional absolutely irreducible $GF(2)L_2(17)$-module, or to the 16-dimensional irreducible $GF(8)L_2(17)$-module. Any 3-chief factor of the group G as $GF(3)\overline{G}$-module is isomorphic to the 16-dimensional absolutely irreducible $GF(3)\overline{G}$- module.*

Each item of the theorem is realized.

The proof of Theorem 2.1 uses the well-known description of finite simple 3-primary groups (see, for example, [10]).

As a corollary of Theorem 2.1, the following result is obtained.

Corollary 2.2 *The finite 3-primary almost simple group with disconnected prime graph is recognizable by prime graph if and only if it is isomorphic to $L_2(17)$.*

In [25], the authors described chief factors of commutator subgroups of finite 4-primary groups with disconnected prime graph. In some cases, all possibilities for such chief factors were not determined; however, the existence of at least one possibility was proved. The description is too large so we formulate here only first theorem, which was proved in [25].

Theorem 2.3 *Let G be a finite tetraprimary group with disconnected prime graph, and let $\overline{G} = G/F(G)$. Then, one of the following statements holds:*

(1) *G is a Frobenius group;*

(2) *G is a 2-Frobenius group;*

(3) *$\overline{G}$ is an almost simple triprimary group;*

(4) $\overline{G} \cong L_2(2^m)$, *where* $m \geq 5$, $2^m - 1$, *and* $(2^m + 1)/3$ *are primes;*

(5) $\overline{G} \cong L_2(3^m)$ *or* $PGL_2(3^m)$, *where* m *and* $(3^m - 1)/2$ *are odd primes and* $(3^m + 1)/4$ *is either a prime or* 11^2 *(for* $m = 5$*);*

(6) $\overline{G} \cong L_2(r)$ *or* $PGL_2(r)$, *where* r *is a prime,* $17 \neq r \geq 11$, $r^2 - 1 = 2^a 3^b s^c$, $s > 3$ *is a prime,* $a, b \in \mathbb{N}$, *and* c *is either* 1 *or* 2 *for* $r \in \{97, 577\}$*;*

(7) $\overline{G} \cong A_7$, S_7, A_8, S_8, A_9, $L_2(16)$, $L_2(16)\colon 2$, $\mathrm{Aut}(L_2(16))$, $L_2(25)$, $L_2(25)\colon 2$, $L_2(27)\colon 3$, $L_2(49)$, $L_2(49)\colon 2_1$, $L_2(49)\colon 2_3$, $L_2(81)$, $L_2(81)\colon 2$, $L_2(81)\colon 4$, $L_3(4)$, $L_3(4)\colon 2_1$, $L_3(4)\colon 2_3$, $L_3(5)$, $\mathrm{Aut}(L_3(5))$, $L_3(7)$, $L_3(7)\colon 2$, $L_3(8)$, $L_3(8)\colon 2$, $L_3(8)\colon 3$, $\mathrm{Aut}(L_3(8))$, $L_3(17)$, $\mathrm{Aut}(L_3(17))$, $L_4(3)$, $L_4(3)\colon 2_2$, $L_4(3)\colon 2_3$, $U_3(4)$, $U_3(4)\colon 2$, $\mathrm{Aut}(U_3(4))$, $U_3(5)$, $U_3(5)\colon 2$, $U_3(7)$, $\mathrm{Aut}(U_3(7))$, $U_3(8)$, $U_3(8)\colon 2$, $U_3(8)\colon 3_1$, $U_3(8)\colon 3_3$, $U_3(8)\colon 6$, $U_3(9)$, $U_3(9)\colon 2$, $\mathrm{Aut}(U_3(9))$, $U_4(3)$, $U_4(3)\colon 2_2$, $U_4(3)\colon 2_3$, $U_5(2)$, $\mathrm{Aut}(U_5(2))$, $S_4(4)$, $S_4(4)\colon 2$, $\mathrm{Aut}(S_4(4))$, $S_4(5)$, $S_4(7)$, $S_4(9)$, $S_4(9)\colon 2_1$, $S_4(9)\colon 2_3$, $S_6(2)$, $G_2(3)$, $\mathrm{Aut}(G_2(3))$, $O_8^+(2)$, ${}^3D_4(2)$, $\mathrm{Aut}({}^3D_4(2))$, $Sz(8)$, $Sz(32)$, $\mathrm{Aut}(Sz(32))$, ${}^2F_4(2)'$, ${}^2F_4(2)$, M_{11}, M_{12}, $\mathrm{Aut}(M_{12})$, *or* J_2.

The proof of Theorem 2.3 uses the description of finite simple 4-primary groups obtained in [36, 14, 4]. Shi wrote Question 13.65 in "The Kourovka Notebook" [28]: *Is the number of finite simple tetraprimary groups finite or infinite?* However, Shi's question is still open.

In the proofs of the theorems from [24, 25, 26], computations are carried out by applying the computer system GAP. A program written in the language of this system makes it possible to compute by the formula from Proposition 1.3 the dimension of the centralizer in the vector space of an element of prime order from a finite simple group that acts irreducibly on this space.

As a corollary of Theorems 1–8 from [25], the following result is obtained.

Corollary 2.4 *A finite* 4*-primary simple group is recognizable by prime graph if and only if it is isomorphic to one of the following groups:* A_8, $L_3(4)$, *and* $L_2(q)$, *where* $|\pi(q^2 - 1)| = 3$, $q > 17$, *and either* $q = 3^m$ *and* m *is an odd prime or* q *is a prime and* $q \not\equiv 1 \pmod{12}$ *or* $q \in \{97, 577\}$.

Theorems 5 and 6 from [25] that are concerned with 4-primary sporadic groups M_{11}, M_{12}, and J_2 refine essentially the corresponding Hagie's results [6].

Vasil'ev wrote Problem 16.26 in "The Kourovka Notebook" [28] about the finding the maximal number of pairwise nonisomorphic finite nonabelian simple groups with the same prime graph. It is conjectured that this number equals to 5 and is achieved on the groups J_2, A_9, $C_3(2)$, $D_4(2)$. Theorem 6 from [25] shows that the set $\{J_2, A_9, C_3(2), D_4(2)\}$ is a maximal set of pairwise nonisomorphic finite nonabelian simple groups with the same prime graph.

In [27], it is obtained the positive solution for all items of Problem 2.2 in the case when $Q = O_2(G)$, $\overline{G} \cong A_7$ and an element of order 5 from G acts on Q fixed points freely. The following theorem is proved.

Theorem 2.5 *Let* G *be a finite group with a nontrivial normal* 2*-subgroup* Q *and* $G/Q \cong A_7$. *Suppose that an element of order* 5 *from* G *acts on* Q *fixed points freely. Then the extension* G *over* Q *is split,* Q *is an elementary abelian group and* Q *is the direct product of minimal normal subgroups each of which as* $GF(2)G/Q$*-module*

is isomorphic to one of the two 4*-dimensional irreducible* $GF(2)A_7$*-modules that are conjugated by outer automorphism of the group* A_7*.*

3 Finite groups with the same graph as the group $Aut(J_2)$

Khosravi in [18, 19, 20] obtained a description of a group having the same prime graph as the group $\mathrm{Aut}(S)$ for any sporadic simple group S except for the group J_2. He posed the problem: *describe all groups* G *such that* $\Gamma(G) = \Gamma(\mathrm{Aut}(J_2))$. Note that if S is a sporadic simple group then $|\mathrm{Aut}(S) : S| \leq 2$ and graphs $\Gamma(S)$ and $\Gamma(\mathrm{Aut}(S))$ are disconnected except for the graphs $\Gamma(\mathrm{Aut}(J_2))$ and $\Gamma(\mathrm{Aut}(McL))$. In [22], the first author solved the Khosravi's problem. The following theorem is proved.

Theorem 3.1 *Let* G *be a finite group,* $\Gamma(G) = \Gamma(\mathrm{Aut}(J_2))$ *and* $\overline{G} = G/O_2(G)$*. Then one of the following statements holds:*

(1) G *is soluble, the* 2*-complement in* G *is a Frobenius group, whose core is a* 7*-group and complement* B *is a cyclic* $\{3,5\}$*-group of order divisible by* 15*, the factor-group* $G/O_{\{2,7\}}(G)$ *is isomorphic to a subgroup of order dividing* $8|B|$ *from* $\mathrm{Hol}(C)$*;*

(2) G *is soluble, the* 2*-complement* R *in* G *is a Frobenius group of form* $A : B$*, where* $A = F(R)$ *is a biprimary* $\{3,5\}$*-group, and* B *is a cyclic* 7*-group, the factor-group* $O_{7'}(G)/O_2(G)$ *has the normal* 2*-complement* $AO_2(G)/O_2(G)$*, and the factor-group* $G/O_{7'}(G)$ *is isomorphic to* B *or the dihedral group of order* $2|B|$*;*

(3) G *is soluble, the* 2*-complement* R *in* G *is a* 2*-Frobenius group of form* $A : B : C$*, where* $A = F(R)$ *is a* $\{3,5\}$*-group of order divisible by* 5*,* B *is a cyclic* 7*-group, and* $|C| = 3$*, the factor-group* $O_{7'}(G)/O_2(G)$ *has the normal* 2*-complement* $AO_2(G)/O_2(G)$*, and the factor-group* $G/O_{7'}(G)$ *is isomorphic to a Frobenius group of order* $3|B|$ *or* $6|B|$*;*

(4) $\overline{G}$ *is isomorphic to one of the groups* A_8*,* S_8*,* A_9*,* S_9*,* $S_6(2)$*,* $O_8^+(2)$*,* $O_8^+(2) : 2$*,* J_2 *or* $\mathrm{Aut}(J_2)$*;*

(5) $\overline{G}$ *is isomorphic to an extension of a nontrivial nilpotent* $\{3,5\}$*-group* A *by a group* B *such that* $F^*(B) = O_2(B) \times L$*, where the group* L *is isomorphic to* A_7*, the group* $B/O_2(B)$ *is isomorphic to* A_7 *or* S_7*, the group* L *induces (by conjugation) on any* p*-chief factor of the group* $\overline{G}^\infty$ *the irreducible* 6*-dimensional* $GF(p)A_7$*-module for* $p \in \{3,5\}$*;*

(6) $\overline{G}$ *is isomorphic to an extension of a nilpotent* $\{3,5\}$*-group* A *of order divisible by* 5 *by a group* B *such that* $F^*(B) = O_2(B) \times L$*, where the group* L *is isomorphic to* $U_3(3)$*, the group* $B/O_2(B)$ *is isomorphic to* $U_3(3)$ *or* $G_2(2)$*, the group* L *induces on any* 3*-chief factor of the group* $\overline{G}^\infty$ *the natural unitary* 3*-dimensional* $GF(9)U_3(3)$*-module or the irreducible* 6*-dimensional* $GF(9)U_3(3)$*-module, and on any its* 5*-chief factor the absolutely irreducible* 6*-dimensional* $GF(5)U_3(3)$*-module;*

(7) $\overline{G}$ *is isomorphic to an extension of a nilpotent* $\{3,5\}$*-group* A *of order divisible by* 5 *by a group* B *such that* $F^*(B) = O_2(B) \times L$*, where the group* L *is isomorphic to* $L_2(7)$*, the group* $B/O_2(B)$ *is isomorphic to* $L_2(7)$ *or* $PGL_2(7)$*, the group* L

induces on any p-chief factor of the group $\overline{G}^\infty$ the irreducible 3-dimensional $GF(p^2)L_2(7)$-module or the absolutely irreducible 6-dimensional $GF(p)L_2(7)$-module for $p \in \{3,5\}$;

(8) *$\overline{G}$ is isomorphic to a semidirect product of a nontrivial abelian 3-group A on a group B such that $F^*(B) = O_2(B) \circ L$, where the group L is isomorphic to $2^{\cdot}L_3(4)$ or $2^{\cdot}U_4(3)$, the group $B/F^*(B)$ is isomorphic to a subgroup from D_8, the involution from $Z(L)$ inverts A, and the group L induces on any 3-chief factor of the group AL the faithful irreducible 6-dimensional $GF(3)L$-module;*

(9) *$\overline{G}$ is isomorphic to a semidirect product of a nontrivial abelian 3-group A on a group B such that $F^*(B) = O_2(B) \circ L$, where the group L is isomorphic to $2^2{\cdot}L_3(4)$, the group $B/F^*(B)$ is isomorphic to a subgroup from 2^2, $Z(L)$ is generated by some involutions z_1 and z_2 such that $A = C_A(z_1) \times C_A(z_2)$, and the group L induces on any 3-chief factor of the group AL the faithful irreducible 6-dimensional $GF(3)2^{\cdot}L_3(4)$-module;*

(10) *$\overline{G}$ is isomorphic to a semidirect product of a abelian $\{3,5\}$-group A on a group B such that $F^*(B) = O_2(B) \circ L$, where the group L is isomorphic to $2^{\cdot}J_2$, the group $B/O_2(B)$ is isomorphic to J_2 or $\mathrm{Aut}(J_2)$, the involution from $Z(L)$ inverts A, and the group L induces on any 3-chief factor of the group AL the faithful irreducible 6-dimensional $GF(9)L$-module and on any its 5-chief factor the faithful irreducible 6-dimensional $GF(5)L$-module;*

(11) *$\overline{G}$ is isomorphic to an extension of a nilpotent $\{3,5\}$-group A of order divisible by 5 by a group B such that $F^*(B) = O_2(B) \circ L$, where the group L is isomorphic to $SL_2(7)$, the group $B/O_2(B)$ is isomorphic to $L_2(7)$ or $PGL_2(7)$, and the group L induces on any p-chief factor of the group $\overline{G}^\infty$ for $p \in \{3,5\}$ either a unfaithful irreducible L-module with the core of order 2 (see item (7)), or the faithful irreducible 6-dimensional $GF(p^2)L$-module*

Each from items (1)–(11) of the theorem is realised.

As a corollary of Theorem 3.1, we obtain:

Corollary 3.2 *A finite group G such that $|G| = |\mathrm{Aut}(J_2)|$ and $\Gamma(G) = \Gamma(\mathrm{Aut}(J_2))$ is isomorphic to $\mathrm{Aut}(J_2)$, $2 \times J_2$ or $2^{\cdot}J_2$.*

4 Finite groups with the same graph as the group A_{10}

The group A_{10} is exceptional in many senses. It is the only group with connected prime graph among all finite simple groups from "Atlas of Finite Groups" [7] and also among all 4-primary simple groups. The non-recognizability by spectrum of the group A_{10} was established by Mazurov [32] in 1998. Staroletov in [37, 38] determined the structure of the group G such that $\omega(G) = \omega(A_{10})$ and in particular proved its unsolvability. In [33] it is proved that the group A_{10} is recognizable by its prime graph and order. Extending these results, recently the first author in [23] describe all finite group with the same prime graph as the group A_{10}. The following theorem is proved.

Theorem 4.1 *Let G be a finite group, $\Gamma(G) = \Gamma(A_{10}$ and $\overline{G} = G/O_3(G)$. Then one of the following statements holds:*

(1) *G is soluble, 3-complement R in G is a Frobenius group, whose core is a non-cyclic 7-group and complement B is a biprimary group of form $C : D$, where C is a cyclic 5-group and D is a cyclic or (generalized) quaternion 2-group, the factor-group $G/O(G)$ is isomorphic to D, $SL_2(3)$, or $Q_8.S_3$;*

(2) *G is soluble, 3-complement R in G is a Frobenius group of form $A : B$, where $A = F(R)$ is a biprimary $\{2,5\}$-group and B is a cyclic 7-group, the factor-group $O_{7'}(G)/O_3(G)$ has the normal 3-complement $AO_3(G)/O_3(G)$, and the factor-group $G/O_{7'}(G)$ is isomorphic to B or a Frobenius group of order $3|B|$;*

(3) *G is soluble, 3-complement R in G is a 2-Frobenius group of form $A : B : C$, where $A = F(R)$ is a $\{2,5\}$-group of order divisible by 5, B is a cyclic 7-group, and $|C| = 2$, the factor-group $O_{7'}(G)/O_3(G)$ has the normal 2-complement $AO_3(G)/O_3(G)$, and the factor-group $G/O_{7'}(G)$ is isomorphic to a Frobenius group of order $2|B|$ or $6|B|$;*

(4) *$\overline{G}$ is isomorphic to a semidirect product of a nontrivial abelian 7-group A on a group B such that $F^*(B) = O_3(B) \times L$, where the group L is isomorphic to $SL_2(q)$ for $q \in \{5,9\}$, the group $B/O_3(B)$ is isomorphic to $L_2(q)$ or $PGL_2(q)$, and any 7-chief factor of the group AL as L-module is isomorphic for $q = 5$ to the faithful irreducible 2-dimensional $GF(49)SL_2(5)$-module or the faithful irreducible 4-dimensional $GF(7)SL_2(5)$-module, and for $q = 9$ to one of two quasiequivalent faithful irreducible 4-dimensional $GF(7)SL_2(9)$-modules;*

(5) *$\overline{G}$ is isomorphic to one of the groups S_7, S_8, A_9, A_{10}, $PGL_2(49)$, $L_3(4) : 2_3$, $L_3(4).3.2_3$, $U_3(5)$, $U_3(5) : 2$, $U_3(5) : 3$, $U_3(5) : S_3$, $S_6(2)$, $O_8^+(2)$, $O_8^+(2) : 3$, or J_2;*

(6) *$\overline{G}$ is isomorphic to an extension of a nilpotent $\{3,5\}$-group A of order divisible by 5 by a group B such that $F^*(B) = O_3(B) \times L$, where the group L is isomorphic to $L_2(7)$, the group $B/O_3(B)$ is isomorphic to $L_2(7)$ or $PGL_2(7)$, any 2-chief factor of the group $\overline{G}^\infty$ as L-module is isomorphic to one of two quasiequivalent irreducible 3-dimensional $GF(3)L_2(7)$-modules, and any 5-chief factor of the group $\overline{G}^\infty$ as L-module is isomorphic either to the irreducible 3-dimensional $GF(25)L_2(7)$-module or to the absolutely irreducible 6-dimensional $GF(5)L_2(7)$-module;*

(7) *$\overline{G}$ is isomorphic to an extension of a nontrivial nilpotent $\{2,5\}$-group A by a group B such that $F^*(B) = O_3(B) \times L$, where the group L is isomorphic to A_7 or $U_3(3)$, $|B : F^*(B)| \leq 2$, and any p-chief factor of the group $\overline{G}^\infty$ as L-module is isomorphic to the irreducible 6-dimensional $GF(p)L$-module for $p \in \{2,5\}$;*

(8) *$\overline{G}$ is isomorphic to an extension of a nontrivial nilpotent $\{2,5\}$-group A by a group B such that $F^*(B) = O_3(B) \circ L$, where $L \cong 3 \cdot A_7$, the group $B/O_3(B)$ is isomorphic to L or $Aut(L)$, any p-chief factor of the group $\overline{G}^\infty$ as L-module for for $p \in \{2,5\}$ is isomorphic either to the faithful irreducible 6-dimensional $GF(p^2)L$-module or to the unfaithful irreducible 6-dimensional $GF(p)L$-module with the core of order 3 (see the item (6));*

(9) *$\overline{G}$ is isomorphic to an extension of a nontrivial 5-group A by a group B such that $F^*(B) = O_3(B) \circ L$, where $L \cong SU_3(5)$, the group $B/O_3(B)$ is isomorphic to a subgroup from $Aut(L)$, any 5-chief factor of the group $\overline{G}^\infty$ as L-module is isomorphic to the faithful irreducible 3-dimensional or 6-dimensional $GF(25)L$-module;*

(10) *$\overline{G}$ is isomorphic to an extension of a nontrivial 2-group A by a group B such that $F^*(B) = O_3(B) \circ L$, where the group L is isomorphic to A_8, $S_6(2)$, $3^{\cdot}U_4(3)$ or J_2, the group $B/O_3(B)$ is isomorphic to a subgroup from S_8, $S_6(2)$, $U_4(3).2_{2/3}$ or J_2, respectively, any 2-chief factor of the group $\overline{G}^{\infty}$ as L-module is isomorphic to the faithful irreducible 6-dimensional L-module over the field $GF(2)$ for the first and second cases and over the field $GF(4)$ for the remaining cases;*

(11) *$\overline{G}$ is isomorphic to an extension of a nontrivial 2-group A by a group B such that $F^*(B) = O_3(B) \circ L$, where the group L is isomorphic to $L_3(4)$ or $SL_3(4)$, the group $B/O_3(B)$ is isomorphic to a subgroup from $L_3(4).6$ or $L_3(4).3.2_3$, respectively, any 2-chief factor of the group $\overline{G}^{\infty}$ as L-module is isomorphic either to the natural 3-dimensional $GF(4)SL_3(4)$-module or to one of two quasiequivalent unfaithful irreducible 9-dimensional $GF(2)SL_3(4)$-modules with the core of order 3.*

Each from items (1)–(11) of the theorem is realised.

Since the different prime graphs $\Gamma(\mathrm{Aut}(J_2))$ and $\Gamma(A_{10})$ are isomorphic as abstract graphs, the arguments in the proofs of Theorems 3.1 and 4.1 are similar. It is interesting that $\Gamma(\mathrm{Aut}(J_2)) = \Gamma(2 \times J_2)$ and $\Gamma(A_{10}) = \Gamma(3 \times J_2)$.

References

[1] Z. Arad and W. Herford, Classification of finite groups with a CC-subgroup, *Comm. Algebra* **32** (2004), 2087–2098.

[2] M. Aschbacher, *Finite group theory*, Cambridge University Press, Cambridge, 1986.

[3] S. Astill, C. Parker, R. Waldecker, A note on groups in which the centralizer of every element of order 5 is a 5-group, *Siberian Math. J.* **52** (2012), 967–977.

[4] Y. Bugeaud, Z. Cao, M. Mignotte, On simple K_4-groups, *J. Algebra* **241** (2001), 658–668.

[5] G. Chen, A new characterization of sporadic simple groups, *Algebra Colloq.* **3** 1996, 49-58.

[6] M. Hagie, The prime graph of a sporadic simple group, *Comm. Algebra* **31** (2003), 4405–4424.

[7] J. H. Conway, R. T. Curtis, S. P. Norton, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups*, Clarendon Press, Oxford, 1985.

[8] S. Dolfi, E. Jabara, and M. S. Lucido, $C55$-groups, *Siberian Math. J.* **45** (2004), 1053–1062.

[9] R. M. Guralnik and P. H. Tiep, Finite simple unisingular groups of Lie type, *J. Group Theory* **6** (2003), 271–310.

[10] M. Herzog, On finite simple groups of order divisible by three primes only, *J. Algebra* **10** (1968), 383–388.

[11] G. Higman, Groups and Lie rings having automorphism without non-trivial fixed points, *J. London Math. Soc. (2)* **32** (1957), 321–334.

[12] G. Higman, *Odd characterizations of finite simple groups: lecture notes*, University Michigan, 1968.

[13] D. F. Holt and W. Plesken, A_5-invariant 2-groups with no trivial sections, *Quart. J. Math. Oxford (2)* **37** (1986), 39–47.

[14] B. Huppert and W. Lempken, Simple groups of order divisible by at most four primes, *Proc. F. Scorina. Gomel State University. Problems in Algebra* **3 (16)** (2000), 64–75.

[15] N. Iiyori and H. Yamaki, Prime graph components of the simple groups of Lie type over the fields of even characteristic, *J. Algebra* **155** (1993), 335–343; Corrigenda, *J. Algebra* **181** (1996), 659.

[16] C. Jansen, K. Lux, R. Parker, and R. Wilson, *An atlas of Brauer characters*, Clarendon Press, Oxford, 1995.
[17] The GAP Group, *GAP — Groups, Algorithms, and Programming*, Ver. 4.4.12, 2008 (URL: http://www.gap-system.org).
[18] B. Khosravi, On the prime graph of the automorphism groups of sporadic simple groups, *Arch. Math. (Brno)* **45** (2009), 83–94.
[19] B. Khosravi, A characterization of the automorphism groups of sporadic groups by the set of orders of maximal abelian subgroups, *Kumamoto J. Math.* **22** (2009), 17–34.
[20] B. Khosravi, On the prime graph of a finite group, *London Math. Soc. Lect. Not. Ser.* **388**, (2011), 424–428.
[21] A. S. Kondrat'ev, Prime graph components of finite simple groups, *Math. USSR Sb.* **67** (1990), 235–247.
[22] A. S. Kondratiev, Finite groups with the same prime graph as the group $\mathrm{Aut}(J_2)$, *Trudy Inst. Mat. Mekh. UrO RAN* **18, no.3** (2012), 131–138 (In Russian).
[23] A. S. Kondratiev, *Finite groups with the same prime graph as the group* A_{10}, *Trudy Inst. Mat. Mekh. UrO RAN* **19, no.1** (2013), 136–143.
[24] A. S. Kondrat'ev and I. V. Khramtsov, On finite threeprimary groups, *Trudy Inst. Mat. Mekh. UrO RAN* **16, no.3** (2010), 150–158.
[25] A. S. Kondrat'ev and I.V. Khramtsov, On finite tetraprimary groups, *Proc. of the Steklov Institute of Math.* **279** (2012), Suppl.1, 43–61.
[26] A. S. Kondratiev and I. V. Khramtsov, On finite nonsimple threeprimary groups with disconnected prime graph, *Sib. Electron. Mat. Izv.* **9** (2012), 472–477.
[27] A. S. Kondratiev, I. V. Khramtsov, Complete irreducibility of some $GF(2)A_7$-modules, *Trudy Inst. Mat. Mekh. UrO RAN* **18** (2012), no.3, 139–143.
[28] *The Kourovka notebook. Unsolved problems in group theory*, eds V.D. Mazurov and E.I. Khukhro. 17th ed., Russian Academy of Sciences Siberian Division, Novosibirsk, 2010.
[29] M. S. Lucido, Prime graph components of finite almost simple groups, *Rend. Sem. Mat. Univ. Padova* **102** (1999), 1–22; addendum, **107** (2002), 189–190.
[30] R. P. Martineau, On representations of the Suzuki groups over fields of odd characteristic, *J. London Math. Soc. (2)* **6** (1972), 153–160.
[31] R. P. Martineau, On 2-modular representations of the Suzuki groups, *Amer. J. Math.* **94** (1972), 55–72.
[32] V. D. Mazurov, Recognition of finite groups by a set of orders of their elements, *Algebra and Logic* **37** (1998), 371–379.
[33] A. R. Moghaddamfar and A. R. Zokayi, OD-characterization of certain finite groups having connected prime graphs, *Algebra Colloq.* **17** (2010), 121–130.
[34] A. R. Prince, On 2-groups admitting A_5 or A_6 with an element of order 5 acting fixed point freely, *J. Algebra* **49** (1977), 374–386.
[35] A. R. Prince, An analogue of Maschke's theorem for certain representations of A_6 over $GF(2)$, *Proc. Roy. Soc. Edinburgh Sect. A* **91** (1982), 175–177.
[36] W. J. Shi, On simple K_4-groups, *Chinese Science Bull.* **36** (1991), 1281–1283.
[37] A. M. Staroletov, Unsolubility of finite groups that are isospectral to the alternating group of degree 10, *Sib. Electron. Mat. Izv.* **5** (2008), 20–24.
[38] A. M. Staroletov, Groups isospectral to the alternating group of degree 10, *Siberian Math. J.* **51** (2010), 507–514.
[39] W. B. Stewart, Groups having strongly self-centralizing 3-centralizers, *Proc. London Math. Soc.* **426** (1973), 653-680.
[40] J. S. Williams, Prime graph components of finite groups, *J. Algebra* **69** (1981), 487–513.
[41] A. V. Zavarnitsin, Properties of element orders in coverings of the groups $L_n(q)$ and $U_n(q)$, *Sib. Math. J.* **49** (2008), 246–256.
[42] A. V. Zavarnitsine, Fixed points of large prime-order elements in the equicharacteristic action of linear and unitary groups, *Sib. Electron. Mat. Izv.* **8** (2011), 333–340.
[43] G. Zurek, Über A_5-invariante 2-Gruppen, *Mitt. Math. Sem. Giessen* **155** (1982).

SOLVABILITY CRITERIA FOR FINITE LOOPS AND GROUPS

EMMA LEPPÄLÄ

Department of Mathematical Sciences, University of Oulu, PL 3000, 90014 Oulu, Finland
Email: emma.leppala@oulu.fi

Abstract

We study how the solvability of finite loops follows from certain properties of their multiplication groups and inner mapping groups. In this survey we review the progress of this research problem from the last two decades and the best results achieved so far. We also introduce a few new improvements with proofs in order to give the readers some idea about the basic methods used in this study.

1 Introduction

Let Q be a groupoid with a neutral element e. If the equations $ax = b$ and $ya = b$ have unique solutions x and y in Q for every $a, b \in Q$, then we say that Q is a *loop*. If a loop Q is associative, then it is in fact a group. For each $a \in Q$ we have two permutations L_a (*left translation*) and R_a (*right translation*) on Q, defined by $L_a(x) = ax$ and $R_a(x) = xa$ for every $x \in Q$. These permutations generate a permutation group $M(Q)$, which is called the *multiplication group* of Q. Clearly, $M(Q)$ is a transitive permutation group on Q. The stabilizer of the neutral element e is called the *inner mapping group* of Q and denoted by $I(Q)$. If Q is a group, then $I(Q)$ is just the group of inner automorphisms of Q. The concepts of the multiplication group and the inner mapping group of a loop were defined by Bruck [1] in 1946.

A subloop H of Q is *normal* in Q if $x(yH) = (xy)H$, $(xH)y = x(Hy)$ and $xH = Hx$ for every $x, y \in Q$. As in groups, a subloop H of a loop Q is normal if and only if H is the kernel of some homomorphism of Q. A loop Q is said to be *solvable* if it has a series $1 = Q_0 \subseteq \cdots \subseteq Q_n = Q$, where Q_{i-1} is normal in Q_i and Q_i/Q_{i-1} is an abelian group for each i.

2 Connected transversals

Let G be a group, $H \leq G$ and let A and B be two left transversals to H in G. We say that the two transversals A and B are *H-connected* if $a^{-1}b^{-1}ab \in H$ for every $a \in A$ and $b \in B$. If A and B are H-connected, it follows that also A^g and B^g are left transversals to H in G for every $g \in G$ [13, Lemma 2.1 and Lemma 2.2]. We denote by H_G the core of H in G (the largest normal subgroup of G contained in H).

Let Q be a loop and write $A = \{L_a : a \in Q\}$ and $B = \{R_a : a \in Q\}$. Now these two sets are left (and right) transversals to $I(Q)$ in $M(Q)$ and as $[A, B] \leq I(Q)$, they are $I(Q)$-connected. Moreover, $M(Q) = \langle A, B \rangle$ and the core of $I(Q)$ in $M(Q)$ is trivial. In 1990 Niemenmaa and Kepka proved the following theorem [13, Theorem 4.1], which gives the relation between multiplication groups of loops and connected transversals.

Theorem 2.1 *A group G is isomorphic to the multiplication group of a loop if and only if there exist a subgroup H and H-connected transversals A and B such that $H_G = 1$ and $G = \langle A, B \rangle$.*

In the following results, which are needed later, we assume that $H \leq G$ and A and B are H-connected transversals in G.

Lemma 2.2 *If $H_G = 1$, then $1 \in A \cap B$.*

Proof Let $1 = ah$, where $a \in A$ and $h \in H$. Then $h = a^{-1}$, and $b^{-1}hb = b^{-1}a^{-1}b = b^{-1}a^{-1}bah \in H$ for every $b \in B$. Thus $h \in \bigcap_{b \in B} H^{b^{-1}} = 1$, and hence $1 = a \in A$. In a similar manner we show that $1 \in B$. □

Lemma 2.3 *If $H_G = 1$, then $N_G(H) = H \times Z(G)$.*

Lemma 2.4 *If $C \subseteq A \cup B$ and $K = \langle H, C \rangle$, then $C \subseteq K_G$.*

For the proofs, see [13, Proposition 2.7 and Lemma 2.5].

Lemma 2.5 *Let N be a normal subgroup of G and set $\bar{G} = G/N$. Then $\bar{A}$ and $\bar{B}$ contain $\bar{H}$-connected transversals in $\bar{G}$.*

Proof Now $[aN, bN] = [a, b]N \subseteq HN$, and clearly $ANH = BNH = G$. Thus $[\bar{A}, \bar{B}] \leq \bar{H}$ and $\bar{A}\bar{H} = \bar{B}\bar{H} = \bar{G}$. □

3 Solvability criteria for finite loops

In 1996 Vesanen [15] studied the connection between solvable loops and solvable groups. He was able to prove the following

Theorem 3.1 *Let Q be a finite loop. If $M(Q)$ is a solvable group, then Q is a solvable loop.*

This result opened new possibilities to study the solvability of finite loops. By combining this result with the theorem of Niemenmaa and Kepka, we may create solvability criteria for finite loops in terms of their inner mapping groups. Hence we focus on the following

Question 3.2 Which properties of the inner mapping group $I(Q)$ of a finite loop Q guarantee the solvability of $M(Q)$, and hence that of the loop Q?

In [14], Niemenmaa and Kepka managed to show that $M(Q)$ and Q are solvable provided that $I(Q)$ is abelian. The case where $I(Q)$ is a nonabelian group of order pq was first investigated in [11]. The case where $|I(Q)| = 2p$ was solved by Csörgő and Niemenmaa in [2] and in 2002 Drápal [3] finally managed to solve the entire nonabelian case of order pq.

The following series of results forms the best answers achieved so far. The proofs to the following theorems can be found in [10, Theorem 1], [12, Theorem 3.1], [8, Theorem 2.7, Theorem 3.2 and Theorem 3.3] and [9, Theorem 3.4].

In the following four theorems we assume that there exist H-connected transversals A and B in G.

Theorem 3.3 *Assume that G is a finite group, $H \leq G$ and H is nilpotent. Then G is solvable.*

Theorem 3.4 *Assume that G is a finite group, $H \leq G$ and H is either a dihedral group or a nonabelian group of order pq ($p \neq q$ are odd primes). Then G is solvable.*

Theorem 3.5 *Assume that G is a finite group, $H \leq G$ and $H = S \times L$, where S is either a dihedral group or a nonabelian group of order pq ($p \neq q$ are odd primes), L is abelian and $\gcd(|S|, |L|) = 1$. Then G is solvable.*

Theorem 3.6 *Assume that G is a finite group, $H \leq G$ and $H = D \times S$, where D is a dihedral 2-group and S is a nonabelian group of order pq ($p \neq q$ are odd primes). Then G is solvable.*

Next we state and prove some minor improvements. These proofs give an idea of the structure of the proofs of some of the solvability theorems above.

A group is called a *Dedekind group*, if its every subgroup is normal. A finite Dedekind group is either abelian or of the form $Q_8 \times K \times N$, where Q_8 is the quaternion group, K is an elementary abelian 2-group and N is an abelian group of odd order (for the details, see [6, pp. 308–309]). Clearly, a finite Dedekind group is nilpotent.

Theorem 3.7 *Let G be a group, $H \leq G$ and $H = D \times S$, where D is a finite Dedekind group, S is a nonabelian group of order pq ($p \neq q$ are odd primes) and $\gcd(|D|, |S|) = 1$. If there exist H-connected transversals A and B in G, then G is solvable.*

Proof First, let G be a finite group. Assume that G is a minimal counterexample. If D is abelian, then by Theorem 3.5, G is solvable. Thus we may assume that $D = Q_8 \times K \times N$, where Q_8, K and N are defined as above.

If $H_G > 1$, then we consider G/H_G and its subgroup H/H_G. Here either $H/H_G \cong D'$ or $H/H_G \cong D' \times S$, where D' is a Dedekind group. By Theorem 3.3 or by induction combined with Lemma 2.5, G/H_G is solvable, and hence G is solvable.

Thus we may assume that $H_G = 1$. If H is not maximal in G, then there exists a subgroup T such that $H < T < G$. By Lemma 2.4, $T_G > 1$ and we may consider G/T_G and its subgroup $HT_G/T_G = T/T_G$. Again by Theorem 3.3 or induction, we conclude that G/T_G is solvable. Since T is solvable by induction, T_G is solvable, and we conclude that G is solvable.

Thus we may assume that H is maximal in G. Now $Q_8 \times K$ is a Sylow 2-subgroup of H. If $[G : H]$ is even, then there exists a 2-subgroup R of G such that $[R : Q_8 \times K] = 2$. But then $Q_8 \times K \trianglelefteq \langle H, R \rangle = G$, as H is maximal in G, which is a contradiction, as $H_G = 1$. Thus we may assume that $[G : H]$ is odd, and hence $Q_8 \times K$ is a Sylow 2-subgroup of G.

If $1 < L \leq Q_8 \times K$, then L is normal in H. Since $H_G = 1$ and H is maximal in G, it follows that $N_G(L) = H$. As $C_G(L) \geq N \times S$, $N_G(L)/C_G(L)$ is a 2-group for every $1 < L \leq Q_8 \times K$. By Frobenius normal p-complement theorem, G is then 2-nilpotent. Thus $G = MP$, where M is normal in G, $P \cong Q_8 \times K$ is a Sylow 2-subgroup of G and $\gcd(|M|, |P|) = 1$.

If $1 \neq a \in A$, then $a = yx$, where $y \in P$ and $x \in M$. Then $aM = yM$ and $(aM)^d = M$, where d divides $|P|$. Thus $a^d \in M$, hence $(a^d)^t = 1$, where t divides $|M|$. It follows that $(a^t)^d = 1$, hence $|a^t|$ divides d, and $|P|$. As P is a Sylow 2-subgroup of G, $a^t \in P^g$ for some $g \in G$. Now P^g is a Dedekind group, and hence $\langle a^t \rangle \trianglelefteq \langle H^g, a \rangle = G$, as A is a left transversals to H^g, too. But $(H^g)_G = 1$, and we conclude that $a^t = 1$. As $\gcd(d, t) = 1$, there exist integers m and n such that $md + nt = 1$. Thus $a = a^{md+nt} = (a^d)^m(a^t)^n = (a^d)^m \in M$.

We may conclude that $A \cup B \subseteq M$. Now $G = AH = APNS = MP$, and hence $M = A(NS) = B(NS)$. Here $[A, B] \leq M \cap H = NS$ and thus by Theorem 3.5, M is solvable and hence G is solvable.

Then assume that G is infinite. Let first $G = \langle A, B \rangle$. Let $a \in A$ and $k \in H$ be fixed and write $E(a, k) = \{b \in B : a^{-1}b^{-1}ab = k\}$. If $b, c \in E(a, k)$, then $a^{-1}b^{-1}ab = a^{-1}c^{-1}ac$, and hence $bc^{-1} \in C_G(a)$ and $b \in C_G(a)c$. Thus $E(a, k) \subseteq C_G(a)c_k$, where $c_k \in E(a, k)$ is fixed.

Now $B = \bigcup_{k \in H} E(a, k)$, and hence $G = BH \subseteq C_G(a)\{c_k : k \in H\}H$. Thus $[G : C_G(a)] \leq |H|^2$ is finite for every $a \in A$. Similarly we see that $[G : C_G(b)]$ is finite for every $b \in B$. As $G = \langle A, B \rangle$, it follows that

$$[G : C_G(H)] = [G : \bigcap_{h \in H} C_G(h)] \leq \prod_{h \in H} [G : C_G(h)]$$

is finite, and hence $[G : N_G(H)]$ is finite, too.

By Lemma 2.3, $N_{G/H_G}(H/H_G) = H/H_G \times Z(G/H_G)$. Thus $N_G(H) = HM$, where $M/H_G = Z(G/H_G)$ and M is normal in G. Now $[G : M] = [G : HM][HM : M] = [G : N_G(H)][H : H \cap M]$ is finite, and we may consider G/M and its subgroup $HM/M \cong H/H \cap M$. By the first part of the proof, G/M is solvable. As $M/H_G = Z(G/H_G)$ is abelian, M is solvable, and thus G is solvable, too.

Assume now that $K = \langle A, B \rangle$ is a proper subgroup of G. Then A and B are $K \cap H$-connected transversals in K. By the previous part of the proof, K is solvable. Since $G = KH$, $[G : K]$ is finite. Thus

$$[G : K_G] = [G : \bigcap_{g \in G} K^g] \leq \prod_{g \in G} [G : K^g]$$

is finite, as there are only finitely many conjugates of K in G. Now we consider G/K_G and its subgroup $HK_G/K_G \cong H/H \cap K_G$. Again by the first part of the proof, G/K_G is solvable. As K is solvable, G is also solvable. □

Remark 3.8 In the previous proof, we were able to avoid using the odd order theorem. However, in the proofs of Theorems 4.2 and 4.3 it is necessary.

Remark 3.9 Here we were able to prove the infinite case, too. However, this does not add to the loop theoretical consequence which we receive by applying Theorems 2.1 and 3.1 to Theorem 3.7.

4 Solvability criteria for loops of odd order

As $M(Q)$ is transitive on Q and $I(Q)$ is the stabilizer of the neutral element, $|Q| = [M(Q) : I(Q)]$. Thus, if Q is a loop of odd order and its inner mapping group $I(Q)$

has odd order too, then Q is solvable by Theorem 3.1 and the odd order theorem. However, loops of odd order are not solvable in general.

Left cosets of a normal subloop form a partition of the loop (this is not true for any subloop). Thus the order of a normal subloop must divide the order of the loop. This implies that every non-associative loop of prime order is in fact non-solvable. The smallest example of a non-solvable loop (of odd order) is also the smallest non-associative loop. The loop is given by the following multiplication table.

1	2	3	4	5
2	1	4	5	3
3	5	1	2	4
4	3	5	1	2
5	4	2	3	1

Here $(3 \cdot 4) \cdot 5 \neq 3 \cdot (4 \cdot 5)$, and thus the loop is non-associative.

In some special cases the analogue of the odd order theorem holds for loops. A loop is called a *Moufang loop* if it satisfies the identity $x(y(zy)) = ((xy)z)y$. In 1968, Glauberman [4] was able to show that Moufang loops of odd order are solvable. A loop is an *automorphic loop* if its every inner mapping is an automorphism. Recently, Kinyon, Kunen, Phillips and Vojtěchovský [7] were able to prove that the odd order theorem holds for automorphic loops, too.

In this section, we consider the situation where $|Q|$ is odd and $|I(Q)|$ is even. Which conditions for $I(Q)$ imply the solvability of the loop Q? Naturally, all results in Section 3 hold for loops of odd order, too.

In the proof of our next result, we need the following theorem by Gorenstein and Walter ([5, Theorem 1]).

Theorem 4.1 *Let G be a finite group with dihedral Sylow 2-subgroups. Let $O(G)$ denote the maximal normal subgroup of odd order. Then G satisfies one of the following conditions:*

1. *$G/O(G)$ is isomorphic to a subgroup of $P\Gamma L(2,q)$ containing $PSL(2,q)$, q odd.*
2. *$G/O(G)$ is isomorphic to the alternating group A_7.*
3. *$G/O(G)$ is isomorphic to a Sylow 2-subgroup of G.*

Theorem 4.2 *Let G be a finite group, $H \leq G$ and $H = D \times T$, where D is a dihedral group, T is a group of odd order and $\gcd(|D|, |T|) = 1$. If there exist H-connected transversals A and B in G and $[G : H]$ is odd, then G is solvable.*

Proof Assume that G is a minimal counterexample. If 4 doesn't divide $|G|$, then the Sylow 2-subgroups of G are cyclic, and by Burnside's folklore theorem, G is 2-nilpotent and thus solvable by the odd order theorem. Thus we may assume that 4 divides $|G|$. If $[G : H] = 1$, then $G = H$ and G is solvable. Thus we may assume that $[G : H] > 1$. By the odd order theorem, Burnside's theorem or induction, we conclude that $H_G = 1$ and H is maximal in G (see the proof of Theorem 3.7).

Now the Sylow 2-subgroup E of D is also the Sylow 2-subgroup of G. Since E is dihedral, we may use Theorem 4.1. If $O(G) > 1$ is the maximal normal subgroup of odd order, then $G = O(G)H$ and G is solvable by the odd order theorem.

Thus we may assume that $O(G) = 1$. As $[G : H] > 1$, G is not a 2-group. If $G \cong A_7$, then G cannot have a maximal subgroup which is isomorphic to $H = D \times T$. Thus we are left with the case that G is isomorphic to a subgroup of $P\Gamma L(2, q)$ containing $PSL(2, q)$ (here $q = p^n$ for an odd prime p). Since $P\Gamma L(2, q) \cong PGL(2, q) \rtimes C_n$, it follows that $G = NH$, where $N \cong PSL(2, q)$ is normal in G and G/N is abelian.

Now $N \cap H = (N \cap D) \times (N \cap T)$, where $N \cap D$ is dihedral. If $N \cap T$ is nontrivial, then $N \cap H \leq N \cong PSL(2, q)$ is not possible. Thus $N \cap T = 1$ and $H/N \cap H \cong G/N$ contains a subgroup isomorphic to T. But then T must be abelian, and again G is solvable by Theorem 3.5. □

Theorem 4.3 *Let G be a finite group, $H \leq G$ and $H = D \times T$, where D is a Dedekind group, T is a group of odd order and* $\gcd(|D|, |T|) = 1$. *If there exist H-connected transversals A and B in G and $[G : H]$ is odd, then G is solvable.*

Proof Assume again that G is a minimal counterexample. By induction or the odd order theorem, we conclude that $H_G = 1$ and H is maximal in G (see the proof of Theorem 3.7).

Assume first that D is abelian. If $|D|$ is odd, then G is solvable by the odd order theorem. Let then $|D|$ be even, and as $[G : H]$ is odd, D contains a Sylow 2-subgroup of G. If $1 < L \leq D$ is a 2-group, then L is normal in H, and since $H_G = 1$ and H is maximal in G, $N_G(L) = H$. As $C_G(L) = H$, $N_G(L)/C_G(L)$ is trivial for every 2-group $L \leq D$.

Assume now that D is a nonabelian Dedekind group. Thus $D = Q_8 \times K \times N$, where Q_8 is the quaternion group, K is an elementary abelian 2-group and N is an abelian group of odd order. As $[G : H]$ is odd, $Q_8 \times K$ is a Sylow 2-subgroup of G. If $1 < L \leq Q_8 \times K$, then L is normal in H and again $N_G(L) = H$. As $C_G(L) \geq N \times T$, $N_G(L)/C_G(L)$ is a 2-group for every 2-group $L \leq D$.

In any case, $N_G(L)/C_G(L)$ is a 2-group for every 2-group L of G. By Frobenius normal p-complement theorem, G is 2-nilpotent, and thus solvable by the odd order theorem. □

If $[G : H]$ is odd and H has a cyclic Sylow 2-subgroup, then G is 2-nilpotent, and hence solvable. We introduce the following

Conjecture 4.4 *Let G be a finite group and $H \leq G$, where H is a solvable group with the Klein four-group as a Sylow 2-subgroup. If there exist H-connected transversals A and B in G, $[G : H]$ is odd and $G = \langle A, B \rangle$, then G is solvable.*

Remark 4.5 The last assumption is unavoidable: Let $G = A_5$, $H \leq G$, $H \cong A_4$ and $A = B = \langle x \rangle$, where $|x| = 5$. Then A and B are H-connected transversals in G, $[G : H]$ is odd, H is solvable and the Sylow 2-subgroup of H is the Klein four-group, but G is not solvable.

5 Loop theoretical interpretation

By combining Theorem 2.1 with Theorems 3.3, 3.4, 3.5, and 3.6 and by applying Theorem 3.1, we have

Theorem 5.1 *Let Q be a finite loop. If $I(Q)$ is nilpotent, dihedral or a nonabelian group of order pq ($p \neq q$ are odd primes), then $M(Q)$ is a solvable group and Q is a solvable loop.*

Theorem 5.2 *Let Q be a finite loop. If $I(Q) = S \times L$, where S is either a dihedral group or a nonabelian group of order pq ($p \neq q$ are odd primes), L is abelian and $\gcd(|S|, |L|) = 1$, then $M(Q)$ is a solvable group and Q is a solvable loop.*

Theorem 5.3 *Let Q be a finite loop. If $I(Q) = D \times S$, where D is a dihedral 2-group and S is a nonabelian group of order pq ($p \neq q$ are odd primes), then $M(Q)$ is a solvable group and Q is a solvable loop.*

By combining Theorem 2.1 with Theorems 3.7, 4.2, and 4.3 and by applying Theorem 3.1, we obtain

Theorem 5.4 *Let Q be a finite loop. If $I(Q) = D \times S$, where D is a Dedekind group, S is a nonabelian group of order pq ($p \neq q$ are odd primes) and $\gcd(|D|, |S|) = 1$, then $M(Q)$ is a solvable group and Q is a solvable loop.*

Theorem 5.5 *Let Q be a loop of odd order. If $I(Q) = D \times T$, where D is a dihedral group, T is a group of odd order and $\gcd(|D|, |T|) = 1$, then $M(Q)$ is a solvable group and Q is a solvable loop.*

Theorem 5.6 *Let Q be a loop of odd order. If $I(Q) = D \times T$, where D is a Dedekind group, T is a group of odd order and $\gcd(|D|, |T|) = 1$, then $M(Q)$ is a solvable group and Q is a solvable loop.*

References

[1] R. H. Bruck, Contributions to the theory of loops, *Trans. Amer. Math. Soc.* **60** (1946), 245–354.

[2] P. Csörgő and M. Niemenmaa, Solvability conditions for loops sand groups, *J. Algebra* **232** (2000), 336–342.

[3] A. Drápal, Orbits of inner mapping groups, *Monatsh. Math.* **134** (2002), 191–206.

[4] G. Glauberman, On loops of odd order II, *J. Algebra* **8** (1968), 393–414.

[5] D. Gorenstein and J. H. Walter, The characterization of finite groups with dihedral Sylow 2-subgroups, *J. Algebra* **2** (1965), 85–151; *J. Algebra* **2** (1965), 218–270; *J. Algebra* **2** (1965), 354–393.

[6] B. Huppert, *Endliche Gruppen I* (Springer-Verlag, Berlin/Heidelberg, 1967).

[7] M. K. Kinyon, K. Kunen, J. D. Phillips and P. Vojtěchovský, The structure of automorphic loops, *Trans. Amer. Math. Soc.* (2013), to appear.

[8] E. Leppälä and M. Niemenmaa, On finite loops whose inner mapping groups are direct products of dihedral groups and abelian groups, *Quasigroups and Related Systems* **20** (2012), 257–260.

[9] E. Leppälä and M. Niemenmaa, A solvability criterion for finite loops, *Bull. Austral. Math. Soc.* (2013), to appear.

[10] M. Mazur, Connected transversals to nilpotent groups, *J. Group Theory* **10** (2007), 195–203.

[11] M. Niemenmaa, On connected transversals to subgroups whose order is a product of two primes, *Europ. J. Combinatorics* **18** (1997), 915–919.

[12] M. Niemenmaa, Finite loops with dihedral inner mapping groups are solvable, *J. Algebra* **273** (2004), 288–294.
[13] M. Niemenmaa and T. Kepka, On multiplication groups of loops, *J. Algebra* **135** (1990), 112–122.
[14] M. Niemenmaa and T. Kepka, On connected transversals to abelian subgroups, *Bull. Austral. Math. Soc.* **49** (1994), 121–128.
[15] A. Vesanen, Solvable loops and groups, *J. Algebra* **180** (1996), 862–876.

THE RATIONAL SUBSET MEMBERSHIP PROBLEM FOR GROUPS: A SURVEY

MARKUS LOHREY

Department für Elektrotechnik und Informatik, University of Siegen, Hölderlinstraße 3, D-57076 Siegen, Germany
Email: lohrey@eti.uni-siegen.de

Abstract

The class of rational subsets of a group G is the smallest class that contains all finite subsets of G and that is closed with respect to union, product and taking the monoid generated by a set. The rational subset membership problem for a finitely generated group G is the decision problem, where for a given rational subset of G and a group element g it is asked whether $g \in G$. This paper presents a survey on known decidability and undecidability results for the rational subset membership problem for groups. The membership problems for finitely generated submonoids and finitely generated subgroups will be discussed as well.

1 Introduction

The study of algorithmic problems in group theory has a long tradition. Dehn [13], in his seminal paper from 1911, introduced the word problem (Does a given word over the generators represent the identity?), the conjugacy problem (Are two given group elements conjugate?) and the isomorphism problem (Are two given finitely presented groups isomorphic?), see [38] for general references in combinatorial group theory. Starting with the work of Novikov and Boone from the 1950's, all three problems were shown to be undecidable for finitely presented groups in general. A generalization of the word problem is the *subgroup membership problem* (also known as the *generalized word problem*) for finitely generated groups: Given group elements $g, g_1, \ldots, g_n$, does g belong to the subgroup generated by $g_1, \ldots, g_n$? Explicitly, this problem was introduced by Mihailova [42] in 1959, although Nielsen [47] had already presented an algorithm for the subgroup membership problem for free groups in his paper from 1921.

Motivated partly by automata theory, the subgroup membership problem was further generalized to the *rational subset membership problem*. Assume that the group G is finitely generated by the set X (where $a \in X$ if and only if $a^{-1} \in X$). A finite automaton A with transitions labelled by elements of X defines a subset $L(A) \subseteq G$ in the natural way; such subsets are the rational subsets of G, see Sections 2 and 3 for precise definitions. The rational subset membership problem asks whether a given group element belongs to $L(A)$ for a given finite automaton (in fact, this problem makes sense for any finitely generated monoid). The notion of a rational subset of a monoid can be traced back to the work of Eilenberg and Schützenberger [15] from 1969. The first decidability result for the rational subset membership problem was shown by Benois [5]: Every finitely generated free group has a decidable rational

subset membership problem.

It seems that after Benois' work the rational subset membership problem had been forgotten for a long time. Aspects of rational sets in monoids that are close to classical formal language theory were studied in the 1980s and 1990s, see [7, 18] for surveys. Only in 1999, Grunschlag [19] returned to the rational subset membership problem in his thesis. He proved that the rational subset membership problem is decidable for finitely generated abelian groups and that decidability of the rational subset membership problem is preserved by finite extensions. Also in 1999, Roman'kov presented at a conference a proof, showing that the rational subset membership problem is undecidable for nilpotent groups (even of class 2), see Section 7. The next step was done by Kambites, Silva, and Steinberg [26] in 2006. They proved that the rational subset membership problem is decidable for the fundamental group of a graph of groups, provided that (i) all edge groups are finite and (ii) every vertex group has a decidable rational subset membership problem, see Section 5. Further (un)decidability results on the rational subset membership problem in various classes of groups (right-angled Artin groups, metabelian groups, wreath products) can be found in [33, 36, 37], see Sections 6, 8, and 9. The latter three papers also studied the submonoid membership problem, which sits in between the subgroup membership problem and the rational subset membership problem. The input consists of group elements $g, g_1, \ldots, g_n \in G$ and it is asked whether g belongs to the submonoid of G generated by $g, g_1, \ldots, g_n$. In [35] it was shown that if the group G has at least two ends, then the rational subset membership problem for G is decidable if and only if the submonoid membership problem for G is decidable, see Section 10.

Rational subsets of groups also found applications for the solution of word equations (here, quite often the term rational constraint is used) [14, 31]. In automata theory, rational subsets are tightly related to valence automata: A valence automaton over a monoid M (the term M-automaton is also used) is a finite automaton, where every transition is labeled with an input symbol and an element of M. A word w is accepted by such a valence automaton if there exists a path from the initial state to a final state such that (i) the concatenation of the inputs symbols along this path yields the word w and (ii) the product of the M-elements along the path is the monoid identity. For any group G, the emptiness problem for valence automata over G is decidable if and only if G has a decidable rational subset membership problem. See [10, 16, 24, 26, 60, 61] for details on valence automata.

2 Finite automata

We assume that the reader has some background in computability theory. She or he should be familiar with the concepts of a *decidable problem* (also called *computable problem*) and *undecidable problem* (also called *unsolvable problem* or *insoluble problem*), see, e.g., [51] for background. In Section 4, we present a proof that requires some basic knowledge of complexity theory, in particular the theory of NP-completeness, see [48] for background. Although we give all needed definitions related to finite automata, some background on automata theory (see, e.g., [21]) makes the paper certainly easier to read.

Let X be a finite set of symbols, which is also called an alphabet. With X^* we

denote the set of all finite words $w = a_1 a_2 \cdots a_n$ with $a_1, \ldots, a_n \in X$. If $n = 0$, then w is the empty word, which is also denoted by ε. A subset of X^* is also called a *language*. A finite automaton over X is a tuple $\mathcal{A} = (Q, \Delta, q_0, F)$, where

- $\Delta \subseteq Q \times X \times Q$ is the set of transitions,
- $q_0 \in Q$ is the initial state, and
- $F \subseteq Q$ is the set of final states.

The language accepted by $\mathcal{A}$, denoted by $L(\mathcal{A})$, is the set of all words $w = a_1 a_2 \cdots a_n \in X^*$ for which there exist states $q_1, q_2, \ldots, q_n \in Q$ with $(q_{i-1}, a_i, q_i) \in \Delta$ for $1 \leq i \leq n$ (note that for $i = 1$, $q_{i-1} = q_0$ is the initial state) and $q_n \in F$. Languages of the form $L(\mathcal{A})$ for $\mathcal{A}$ a finite automaton are called *regular*.

A finite automaton over X with ε-transitions is defined as above, except that $\Delta \subseteq Q \times (X \cup \{\varepsilon\}) \times Q$. A transition $(q, \varepsilon, p) \in \Delta$ means that the automaton can move from state q to state p without reading an input symbol. It is well-known that for every finite automaton with ε-transitions there exists an ordinary finite automaton (without ε-transitions) that accepts the same language [21]. Allowing ε-transitions sometimes simplifies technical details in proofs.

3 Rational subsets of groups

We assume that the reader has some background in combinatorial group theory. A classical reference is [38]. Let G be a finitely generated group and X a finite symmetric generating set for G (symmetric means that X is closed under taking inverses). This mean that the canonical morphism $\pi : X^* \to G$ that maps a word $w \in X^*$ to the group element of G represented by w is surjective. Hence, elements of group G can be represented by finite words over the alphabet X. When we say below that the input for a decision problem consists of a group element $g \in G$ (plus possibly some other objects), then we actually mean that the input consists of a finite word $w \in X^*$ that represents the group element g.

Let us fix a monoid M. For a subset $B \subseteq M$ we denote by B^* the *submonoid* of M generated by B. Of course we have to distinguish B^* from the set of all words over B, which is also denoted by B^*. It will be always clear, whether B^* is viewed as the set of all words over B or as the submonoid of M generated by B. In the case M is a group, we denote with $\langle B \rangle$ the subgroup generated by B. The set of *rational subsets* of M is the smallest subset of 2^M that (i) contains all finite subsets of M and (ii) is closed under union, product, and *.

In the following, we will mainly consider rational subsets of a group G. If G is finitely generated by X^* and $\pi : X^* \to G$ is the corresponding canonical homomorphism, then rational subsets of G can be represented by finite automata over X. The following result can be deduced from Kleene's theorem for regular languages, see [18] for a proof:

Proposition 3.1 *Let G be a finitely generated group, let X be a finite generating set for G, and let $\pi : X^* \to G$ be the corresponding canonical homomorphism. A subset $L \subseteq G$ is rational if and only if there is a finite automaton $\mathcal{A}$ over X such that $L = \pi(L(\mathcal{A}))$.*

This characterization of rational subsets is useful since it allows us to represent a rational subset of G by a finite automaton over X.

We consider the following decision problem for a finitely generated group G together with a canonical morphism $\pi : X^* \to G$.

Decision problem 3.2 (Rational subset membership problem for G)

- INPUT: A finite automaton $\mathcal{A}$ over X and an element $g \in G$.
- QUESTION: Does $g \in \pi(L(\mathcal{A}))$ hold?

Note that $g \in L(\mathcal{A})$ if and only if $1 \in L(\mathcal{A})g^{-1}$. Moreover, the set $L(\mathcal{A})g^{-1}$ is rational too and a finite automaton for this set can be constructed from $\mathcal{A}$ and g. Hence, the rational subset membership problem for G is equivalent to the following problem:

- *INPUT: A finite automaton $\mathcal{A}$ over X.*
- *QUESTION: Does $1 \in \pi(L(\mathcal{A}))$ hold?*

Decision problem 3.3 (Submonoid membership problem for G)

- INPUT: Elements $g, g_1, \ldots, g_n \in G$.
- QUESTION: Does $g \in \{g_1, \ldots, g_n\}^*$ hold?

Decision problem 3.4 (Subgroup membership problem for G)

- INPUT: Elements $g, g_1, \ldots, g_n \in G$.
- QUESTION: Does $g \in \langle g_1, \ldots, g_n \rangle$ hold?

The subgroup membership problem for G is also known as the *generalized word problem for G* or as the *occurrence problem for G*.

Strictly speaking, we should speak of the rational subset membership problem for G with respect to $\pi : X^* \to G$, since another choice for the generating set leads to another decision problem. On the other hand, if X and Y are two finite generating sets for G with canonical morphisms $\pi : X^* \to G$ and $\sigma : Y^* \to G$, then the rational subset membership problem for G with respect to $\pi : X^* \to G$ is decidable, if and only if the rational subset membership problem for G with respect to $\sigma : Y^* \to G$ is decidable. For the proof, one chooses a morphism $\rho : Y^* \to X^*$ such that for every $a \in Y$, $\sigma(a) = \pi(\rho(a))$ (clearly, such a morphism exists). Then, for $w \in Y^*$ and a finite automaton $\mathcal{A}$ over Y, we have $\sigma(w) \in \sigma(L(\mathcal{A}))$ if and only if $\pi(\rho(w)) \in \pi(L(\mathcal{B}))$. Here, $\mathcal{B}$ is the automaton over X that results from $\mathcal{A}$ by replacing every a-labelled transition ($a \in Y$) by a chain of transitions that is labelled with the word $\rho(a)$. A similar remark applies to the submonoid membership problem and the subgroup membership problem for G.

Clearly, decidability of the rational subset membership problem for G implies decidability of the submonoid membership problem for G, and the latter implies decidablity of the subgroup membership problem for G.

Note that in the above three decision problems, the input consists of a group element g and a finite description of a subset $Z \subseteq G$, and it is asked whether $g \in Z$. A more restricted setting is obtained by fixing a subset $Z \subseteq G$. For this set Z, we can consider the following decision problem:

Decision problem 3.5 (Membership problem for the set $Z \subseteq G$)

- INPUT: An element $g \in G$.
- QUESTION: Does $g \in Z$ hold?

Problem 3.6 Is there a finitely generated group G with the following properties?

- For every rational subset $R \subseteq G$, the membership problem for R is decidable.
- The rational subset membership problem for G is undecidable.

The same question can be considered for rational subsets replaced by finitely generated submonoids or finitely generated subgroups.

One should note that a positive answer to this problem is conceivable: There might exist a group G for which there is no algorithm that decides the rational subset membership problem for G, but for every rational subset $R \subseteq G$ there is an algorithm A_R that checks whether a given group element belongs to R. These algorithms A_R must be completely unrelated in the sense that they do not follow a uniform scheme.

Of course, one may also generalize Problem 3.2 further, e.g., by considering context-free languages. Given a context-free grammar $\mathcal{G}$ over the symmetric generating set X of the group G and a group element $g \in G$, one can ask whether $g \in \pi(L(\mathcal{G}))$. But this problem is already undecidable for free groups: To see this, take a finitely presented group $G = \mathsf{Gp}\langle X \mid R\rangle$ with an undecidable word problem. Here $R \subseteq X^*$ is a finite set of relators. Then, for a given word $w \in X^*$ we have $w = 1$ in G if and only if in the free group $F(X)$, w belongs to the normal closure of R. But the latter is the canonical image of the context-free language $L = \{crc^{-1} \mid r \in R, c \in X^*\}$. Hence, if the word problem for G is undecidable, then the membership problem for the free group image of the context-free language L is undecidable.

By the last paragraph, the membership problem for (images of) context-free sets is already undecidable for the simplest finitely generated groups[1] (namely free groups). On the other hand, the following sections will show that for the rational subset membership problem we can prove non-trivial decidability results. This is one of the reasons for restricting the membership problem to rational sets in this paper.

4 Classical results

The first decidability result for the rational subset membership problem was shown by Benois [5] in 1969 for free groups:

Theorem 4.1 *Every free group of finite rank has a decidable rational subset membership problem.*

This result can be shown by a simple automata saturation procedure. Consider a free group $F(Y)$, where Y (a finite set) generates $F(Y)$ as a group. Let $X = Y \cup Y^{-1}$. Let $\mathcal{A} = (Q, \Delta, q_0, F)$ be a finite automaton with ε-transitions over the alphabet X. Since we will add ε-transitions to the automaton, will start with an automaton with ε-transitions from the very beginning. As remarked in the previous section, it suffices to check whether $1 \in \pi(L(\mathcal{A}))$. For this we iterate the following operation as long as

[1] The only class of groups with a decidable membership problem for context-free sets, the author is aware of, are finitely generated virtually abelian groups.

possible: If there are transitions $(p, a, q), (r, a^{-1}, s) \in \Delta$ with $a \in X$, and state r can be reached from state q by a sequence of ε-transitions, then we add the ε-transition (p, ε, s) to Δ. The order in which we add ε-transitions is not important. Note that we only add new transitions but we do not add new states. Hence the saturation process has to terminate after at most $|Q|^2$ many steps. Let $\mathcal{B}$ be the resulting automaton with ε-transitions. Then, one can show the following:

- $\pi(L(\mathcal{A})) = \pi(L(\mathcal{B}))$ (this follows by induction on the construction of $\mathcal{B}$).
- If $w \in L(\mathcal{B})$ and w is of the form $w = uaa^{-1}v$ with $u, v \in X^*$ and $a \in X$, then also $uv \in L(\mathcal{B})$.

Hence, we have $1 \in \pi(L(\mathcal{A}))$ if and only if $1 \in \pi(L(\mathcal{B}))$ if and only if there is a word $w \in L(\mathcal{B})$ such that w can be reduced by cancellations of the form $aa^{-1} \to \varepsilon$ $(a \in X)$. But the latter condition is equivalent to $\varepsilon \in L(\mathcal{B})$. Hence, $1 \in \pi(L(\mathcal{A}))$ if and only if $\varepsilon \in L(\mathcal{B})$, and the latter means that there is a path consisting only of ε-transitions leading from the initial state q_0 to a final state. This condition can be checked by an algorithm.

It is worth mentioning that the above algorithm works in polynomial time, see [6] for a precise complexity analysis.

Next, let us consider finitely generated abelian groups. The following result was shown by Grunschlag in his thesis [19] using integer linear programming.

Theorem 4.2 *Every finitely generated abelian group has a decidable rational subset membership problem.*

Grunschlag reduces the rational subset membership problem for finitely generated abelian groups to integer linear programming, which is a classical NP-complete problem. It turns out that already the submonoid membership problem for free abelian groups $\mathbb{Z}^k$ is NP-complete if k is part of the input. To see this, we start with the NP-complete problem 1-in-3 SAT [48]. The input is a conjunction

$$\psi = \bigwedge_{i=1}^{m} C_i,$$

where every C_i is a disjunction of three literals (a literal is a boolean variable or a negated boolean variable). Let $x_1, \ldots, x_n$ be the boolean variables that appear in ψ and let $C_i = \tilde{x}_{i_1} \vee \tilde{x}_{i_2} \vee \tilde{x}_{i_3}$, where $\tilde{x}_{i_j} \in \{x_{i_j}, \neg x_{i_j}\}$. It is asked whether there exists a truth assignment for the variables $x_1, \ldots, x_n$ such that in each disjunction C_i exactly one literal becomes true. This is true if and only if the following system of linear equations in the $2n$ variables $x_1, \overline{x}_1, \ldots, x_n, \overline{x}_n$ has a solution in $\mathbb{N}$:

$$\begin{aligned} x_i + \overline{x}_i &= 1 \quad \text{for } 1 \leq i \leq n \\ \tilde{x}_{i_1} + \tilde{x}_{i_2} + \tilde{x}_{i_3} &= 1 \quad \text{for } 1 \leq i \leq m \end{aligned}$$

In the second equation, we identify the literal $\neg x_{i_j}$ with the variable $\overline{x}_{i_j}$. This system can be written as

$$\sum_{i=1}^{n} (x_i \cdot \mathbf{a}_i + \overline{x}_i \cdot \mathbf{b}_i) = \mathbf{c},$$

for $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_n, \mathbf{b}_n, \mathbf{c} \in \mathbb{Z}^{n+m}$. This system is solvable in the natural numbers if and only if $\mathbf{c}$ belongs to the submonoid generated by $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_n, \mathbf{b}_n$.

Note that in the above NP-hardness proof we have to assume that the dimension (which is $n+m$) is not fixed. In our context, it is more natural to consider the case of a fixed dimension, since in Problems 3.2–3.4 we always fix an underlying group. Using some recent results on the Parikh images of regular languages, we can show:

Theorem 4.3 *For every finitely generated abelian group the rational subset membership problem can be solved in polynomial time.*

Proof Consider a fixed finitely generated abelian group $G = \prod_{i=1}^{n} Z_i$, where every Z_i is cyclic. By Theorem 5.5 from the next section, we can assume that $Z_i \cong \mathbb{Z}$ for every $1 \leq i \leq n$. Take the generating set $X = \{x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}\}$, where x_i generates Z_i as a group. As usual, let $\pi : X^* \to G$ be the canonical morphism. Recall that the Parikh image of a language $L \subseteq X^*$ is the image of L under the canonical morphism $\Psi : X^* \to \mathbb{N}^{2n}$. Thus, if $\Psi(w) = (c_1, d_1, \ldots, c_n, d_n)$, then c_i (resp., d_i) is the number of occurrences of the symbol x_i (resp., x_i^{-1}) in the word w. It is well-known that the Parikh image $\Psi(L)$ of a regular language (and even a context-free language) is semi-linear, i.e., $\Psi(L)$ can be written as

$$\Psi(L) = \bigcup_{i=1}^{k} \{\mathbf{a}_i + \lambda_1 \mathbf{a}_{i,1} + \cdots + \lambda_{l_i} \mathbf{a}_{i,l_i} \mid \lambda_1, \ldots, \lambda_{l_i} \in \mathbb{N}\},$$

for $\mathbf{a}_i, \mathbf{a}_{i,1}, \ldots, \mathbf{a}_{i,l_i} \in \mathbb{N}^{2n}$. It has been recently shown that from a given finite automaton $\mathcal{A}$ over X one can compute a semi-linear representation of the Parikh image $\Psi(L(\mathcal{A})$ in polynomial time[2] [28]. Such a semi-linear representation consists of a list of all vectors $\mathbf{a}_i, \mathbf{a}_{i,j}$ ($1 \leq i \leq k$, $1 \leq j \leq l_i$), where the vector entries are represented as binary encoded numbers. It is crucial here that the alphabet Σ is fixed, because the running time of the algorithm from [28] is exponential in the size of the alphabet.

Let us now consider the rational subset membership problem for G. Let $\mathcal{A}$ be a finite automaton over X. We have to check whether $1 \in \pi(L(\mathcal{A}))$. First, we compute in polynomial time the Parikh image

$$\Psi(L) = \bigcup_{i=1}^{k} \{\mathbf{a}_i + \lambda_1 \mathbf{a}_{i,1} + \cdots + \lambda_{l_i} \mathbf{a}_{i,l_i} \mid \lambda_1, \ldots, \lambda_{l_i} \in \mathbb{N}\}.$$

For $\mathbf{a} = (c_1, d_1, \ldots, c_n, d_n) \in \mathbb{Z}^{2n}$ define the vector $\mathbf{a}' = (c_1 - d_1, \ldots, c_n - d_n) \in \mathbb{Z}^n$. Then, we have $1 \in \pi(L(\mathcal{A}))$ if and only if there are $1 \leq i \leq k$ such that the system

$$\lambda_1 \mathbf{a}'_{i,1} + \cdots + \lambda_{l_i} \mathbf{a}'_{i,l_i} = -\mathbf{a}'_i$$

has a solution in $\mathbb{N}$. But this is an instance of integer programming in the fixed dimension n, which can be solved in polynomial time [30, Sec. 4]. □

Let us now come to classical undecidablity results in the context of rational subsets. The first such result was shown by Mihailova [43] in 1966:

[2] In particular, all numbers $k, l_1, \ldots, l_k$ are polynomially bounded in the size of $\mathcal{A}$.

Theorem 4.4 *The direct product $F_2 \times F_2$ of two copies of the free group of rank 2 contains a fixed finitely generated subgroup with an undecidable membership problem.*

In particular, $F_2 \times F_2$ has an undecidable subgroup membership problem. Hence also the submonoid membership problem and the rational subset membership problem for $F_2 \times F_2$ are undecidable. Mihalova's result is also remarkable since $F_2 \times F_2$ is a very natural group. In contrast all known examples of finitely presented groups with an undecidable word problem are constructed from Turing machines (or other universal computation models) with undecidable acceptance problem and cannot be considered as simple or natural. Nevertheless, Mihailova's result is shown by reducing the word problem for a finitely presented group to the membership problem for a finitely generated subgroup of $F_2 \times F_2$.

A second classical undecidability result for the subgroup membership problem was shown by Rips in 1982:

Theorem 4.5 *There is a word-hyperbolic group that contains a finitely generated subgroup with an undecidable membership problem.*

So again, the subgroup membership problem, the submonoid membership problem, and the rational subset membership problem are in general undecidable for word-hyperbolic groups. The group constructed by Rips is actually a torsion-free small cancellation group satisfying the condition $C'(1/6)$. Wise [58] modified Rips' construction so that the resulting group is also residually finite.

5 Closure properties

For every group theoretic decision problem, let us call it $\mathcal{P}$, it is good to know closure properties with respect to group theoretic constructions. They allow us to construct from groups for which $\mathcal{P}$ is decidable new (and maybe more complicated) groups for which $\mathcal{P}$ is decidable. Mihailova's result (Theorem 4.4) implies that the class of groups for which the subgroup membership problem (or the submonoid membership problem, or the rational subset membership problem) is decidable is not closed under direct products: F_2 has a decidable rational subset membership problem by Benois' result (Theorem 4.1) but $F_2 \times F_2$ has an undecidable subgroup membership problem. Another important operation, which destroys the decidability of the rational subset membership problem is the wreath product; see Section 9 for more details. But fortunately, there are other important group constructions for which we can prove positive results.

Two very important constructions in combinatorial group theory are HNN-extensions and amalgamated free products. The following two results were shown in [26] (and independently in [32]) for the rational subset membership problem and in [27] for the subgroup membership problem.

Theorem 5.1 *Let $\mathcal{P}$ stand for either the rational subset membership problem or the subgroup membership problem. Assume that G is a finitely generated group for which $\mathcal{P}$ is decidable. Then $\mathcal{P}$ is decidable for every HNN-extension $\langle G, t \mid t^{-1}at = \varphi(a)\ (a \in A)\rangle$ with $A \leq G$ finite.*

Theorem 5.2 *Let $\mathcal{P}$ stand for either the rational subset membership problem or the subgroup membership problem. Assume that G_1 and G_2 are finitely generated groups for which $\mathcal{P}$ is decidable. Then $\mathcal{P}$ is decidable for every amalgamated free product $G_1 *_{A_1=A_2} G_2$ with $A_1 \leq G_1$ and $A_2 \leq G_2$ finite.*

Closure of the class of groups with a decidable subgroup membership problem under free products was already shown by Mihailova in [42].

Theorems 5.1 and 5.2 can be rephrased in terms of graphs of groups. Every fundamental group of a graph of groups with finite edge groups and vertex groups that have a decidable rational subset membership problem (resp., subgroup membership problem) has a decidable rational subset membership problem (resp., subgroup membership problem) as well.

Surprisingly, it is not known whether the decidability of the submonoid membership problem is preserved under HNN-extensions with finite associated subgroups and amalgamated free products over finite subgroups:

Problem 5.3 Assume that G is a finitely generated group with a decidable submonoid membership problem, and let $H = \langle G, t \mid t^{-1}at = \varphi(a)\ (a \in A)\rangle$ be an HNN-extension with $A \leq G$ finite. Does H have a decidable submonoid membership problem?

Assume that G_1 and G_2 are finitely generated groups with a decidable submonoid membership problem, and let $G = G_1 *_{A_1=A_2} G_2$ be an amalgamated free product with $A_1 \leq G_1$ and $A_2 \leq G_2$ finite. Does G have a decidable submonoid membership problem? Does the free product $G_1 * G_2$ have a decidable submonoid membership problem?

Actually, the author conjectures that there are specific groups where the answers to the above questions are negative. We will come back to this conjecture in Section 10 when we consider the relationship between the rational subset membership problem and the submonoid membership problem in more detail.

Let us now discuss subgroups and extensions. The following result is trivial:

Proposition 5.4 *Let $\mathcal{P}$ stand for either the rational subset membership problem, the submonoid membership problem, or the subgroup membership problem. Assume that H is a finitely generated subgroup of the finitely generated group G. If $\mathcal{P}$ is decidable for G, then $\mathcal{P}$ is decidable for H as well.*

Our last closure result concerns finite extensions and was shown by Grunschlag in his thesis [19]:

Theorem 5.5 *Let $\mathcal{P}$ stand for either the rational subset membership problem or the subgroup membership problem. Assume that G is a finite index subgroup of H. If $\mathcal{P}$ is decidable for G, then $\mathcal{P}$ is decidable for H as well. Moreover, if $\mathcal{P}$ can be solved in polynomial time for G, then the same holds for the group H.*

Let us sketch the proof. Assume that G (resp., H) is generated by the symmetric set X (resp., Y). Let $W_X(G) \subseteq X^*$ (resp., $W_Y(H) \subseteq Y^*$) be the set of all words

that evaluate to the identity of G (resp., H). There exists a rational transduction $\tau \subseteq X^* \times Y^*$ (which is just a rational subset of the monoid $X^* \times Y^*$) such that

$$W_Y(H) = \tau(W_X(G)) = \{w \in Y^* \mid \exists u \in W_X(G) : (u, w) \in \tau\},$$

see [26, Lemma 3.3]. This rational transduction is given by a fixed automaton $\mathcal{T}$ with transitions labelled by pairs from $(X \times \{\varepsilon\}) \cup (\{\varepsilon\} \times Y)$. Here, "fixed" means that we do not have to construct the automaton $\mathcal{T}$.

Take a finite automaton $\mathcal{A}$ over Y. We have to check, whether $L(\mathcal{A})$ contains a word that evaluates to the identity of H, i.e., that belongs to $W_Y(H)$. We have $L(\mathcal{A}) \cap W_Y(H) \neq \emptyset$ if and only if $L(\mathcal{A}) \cap \tau(W_X(G)) \neq \emptyset$ if and only if $\tau^{-1}(L(\mathcal{A})) \cap W_X(G) \neq \emptyset$. Finally, an automaton for $\tau^{-1}(L(\mathcal{A}))$ can be constructed in polynomial time from the automaton $\mathcal{A}$ using a product construction with the automaton $\mathcal{T}$.

For HNN-extensions and amalgamated free products, it is open whether the decidability of the submonoid membership problem is preserved by finite extensions:

Problem 5.6 Assume that G is a finite index subgroup of H and that G has a decidable submonoid membership problem. Is the submonoid membership problem for H decidable?

6 Right-angled Artin groups

Let $H = (\Gamma, E)$ be a finite simple graph. In other words, the edge relation $E \subseteq V \times V$ is irreflexive and symmetric. One associates with H the group

$$\mathbb{G}(H) = \langle \Gamma \mid ab = ba \ ((a, b) \in E)\rangle.$$

Such a group is called a *right-angled Artin group*, *graph group*, or *free partially commutative group*. Here, we use the term right-angled Artin group[3]. Right-angled Artin groups received a lot of attention in group theory during the last few years, mainly due to their rich subgroup structure [8, 12, 17].

For graphs $H_1 = (V, E)$ and H_2, we say that H_1 contains an induced H_2, if there is a subset $U \subseteq V$ such that the graph $(U, E \cap (U \times U))$ is isomorphic to H_2. In this situation, $\mathbb{G}(H_2)$ is a subgroup of $\mathbb{G}(H_1)$. With $C4$ (cycle on 4 nodes) we denote the following graph:

Note that the right-angled Artin group $\mathbb{G}(C4)$ is $F_2 \times F_2$. Hence, by Mihailova's result (Theorem 4.4), the subgroup membership problem is undecidable for every graph group $\mathbb{G}(H)$ such that H contains an induced $C4$.

On the decidability side, the following result is shown in [27]. A simplified proof[4] can be found in [34].

[3] This term comes from the fact that right-angled Artin groups are exactly the Artin groups corresponding to right-angled Coxeter groups.

[4] Actually, decidability of the subgroup membership problem is shown in [27, 34] for a much larger class of groups.

Theorem 6.1 *Let H be a finite simple graph that does not contain an induced cycle on $n \geq 4$ nodes (such a graph is called chordal). Then, the subgroup membership problem for the graph group $\mathbb{G}(H)$ is decidable.*

This result and Mihailova's result leave a gap for the decidability status of the subgroup membership problem.

Problem 6.2 For which graphs H is the subgroup membership problem for the right-angled Artin group $\mathbb{G}(H)$ decidable? More specifically, is the subgroup membership problem decidable for the right-angled Artin group $\mathbb{G}(C5)$ (where $C5$ denotes a cycle on 5 nodes)?

With $P4$ (path on 4 nodes) we denote the following graph:

The following characterization of right-angled Artin groups with a decidable rational subset membership problem (resp., submonoid membership problem) is shown in [33]:

Theorem 6.3 *Let H be a finite simple graph. Then, the following three conditions are equivalent:*

- *H does not contain an induced $P4$ or $C4$.*
- *The rational subset membership problem for $\mathbb{G}(H)$ is decidable.*
- *The submonoid membership problem for $\mathbb{G}(H)$ is decidable.*

For the undecidability statement in Theorem 6.3 one has to show that the submonoid membership problem is undecidable for $\mathbb{G}(P4)$ and $\mathbb{G}(C4)$. The latter group is covered by Mihailova's result. For $\mathbb{G}(P4)$ it is first shown in [33] that this group has an undecidable rational subset membership problem. Then, in a second step the rational subset membership problem for $\mathbb{G}(P4)$ is reduced[5] to the submonoid membership problem for $\mathbb{G}(P4)$. To prove that $\mathbb{G}(P4)$ has an undecidable rational subset membership problem, one can use a result from the theory of trace monoids. Trace monoids are the monoid counterparts of right-angled Artin groups. For a finite simple graph $H = (\Gamma, E)$ one defines the corresponding trace monoid $\mathbb{M}(H)$ as the quotient of the free monoid Γ^* by the monoid congruence generated by all pairs (ab, ba) with $(a, b) \in E$. Aalbersberg and Hoogeboom [1] proved that the following two conditions are equivalent:

- It is decidable, whether the intersection of two given rational subsets of the trace monoid $\mathbb{M}(H)$ is nonempty.
- The graph H does not contain an induced $P4$ or $C4$.

But for two rational subsets $L, K \subseteq \mathbb{M}(H)$, one has $L \cap K = \emptyset$ if and only if the set LK^{-1} (interpreted in the right-angled Artin group $\mathbb{G}(H)$) contains the identity element 1.

[5] This reduction is very similar to the reduction of the rational subset membership problem to the submonoid membership problem in case of a group with infinitely many ends. This reduction is outlined in Section 10.

The proof of the decidability statement in Theorem 6.3 uses the following characterization of graphs without induced $P4$ or $C4$, see [59]: A finite simple graph H does not contain an induced $P4$ or $C4$ if and only if H can be obtained from the graph with one node using the following two operations:

- Take the disjoint union of two graphs.
- Add a new vertex to the graph and connect it to all old nodes.

On the level of right-angled Artin groups, these two operations correspond to (i) the free product of two groups, and (ii) the direct product by $\mathbb{Z}$. Hence, one has to show that the rational subset membership problem is decidable for every group that can be produced from the trivial group 1 using the operations of free product and direct product with $\mathbb{Z}$.

The algorithm from [33] is not very efficient. To deal with the case of a free product, Parikh's theorem (stating that the Parikh image of a context-free language is semilinear) is applied, which leads to an exponential blow-up in the running time. This implies that for the *uniform* rational subset membership problem for right-angled Artin groups $\mathbb{G}(H)$, where H does not contain an induced $P4$ or $C4$ (in this problem, H is also part of the input), the proof in [33] only yields a non-elementary algorithm, i.e., an algorithm whose running time is not bounded by a tower of exponents of fixed height.

Problem 6.4 What is the computational complexity of the rational subset membership problem for a right-angled Artin group $\mathbb{G}(H)$, where H does not contain an induced $P4$ or $C4$? Is there an algorithm with elementary running time for the uniform problem, where the graph H is part of the input?

7 Nilpotent groups and polycyclic groups

The lower central series of the group G is the sequence of subgroups $G = G_1 \geq G_2 \geq G_3 \geq \cdots$ where $G_{i+1} = [G_i, G]$ (which is the subgroup of G_i generated by all commutators $g^{-1}h^{-1}gh$ for $g \in G_i$ and $h \in G$; by induction one can show that indeed $G_{i+1} \leq G_i$). The group G is nilpotent if there exists $i \geq 1$ with $G_i = 1$. A group G is polycyclic, if there exists a subnormal series $G = G_0 \rhd G_1 \rhd \cdots \rhd G_{n-1} \rhd G_n = 1$ such that every quotient G_{i-1}/G_i is cyclic. Nilpotent groups are polycyclic.

Mal'cev [39] proved that every polycyclic group G is subgroup separable, i.e., for every finitely generated subgroup $H \leq G$ and $g \in G \setminus H$ there exists a morphism $\varphi : G \to K$ to a finite group K such that $\varphi(g) \notin \varphi(H)$. Together with the finite presentability of finitely generated polycyclic groups, one gets:

Theorem 7.1 *Every finitely generated polycyclic group has a decidable subgroup membership problem.*

A more practical algorithm for the subgroup membership problem for polycyclic groups can be found in [2].

By Theorem 7.1 every finitely generated nilpotent group has a decidable subgroup membership problem. This result does not generalize to the rational subset membership problem, as Roman'kov [52] has shown:

Theorem 7.2 *There exists a number r such that the free nilpotent group of class 2 generated by r elements (this group is denoted by $N_{2,r}$) has an undecidable rational subset membership problem.*

The proof of this result in [52] uses a reduction from Hilbert's 10th problem, i.e., the question whether a Diophantine equation $P(x_1, \ldots, x_n)$ with P a polynomial with integer coefficients has an integer solution. The decidability status of the submonoid membership problem for finitely generated nilpotent groups is open:

Problem 7.3 Is there a finitely generated nilpotent group with an undecidable submonoid membership problem?

Rational subsets in nilpotent groups were also studied by Bazhenova [4]. She proved that the rational subsets of a finitely generated nilpotent group G are a Boolean algebra if and only if G is virtually abelian.

8 Metabelian groups

Recall that a group G is metabelian if the commutator subgroup $[G, G]$ is abelian. Equivalently, G is metabelian if G has an abelian normal subgroup A such that the quotient G/A is abelian too. Hall [20] has shown that one can view A as a $\mathbb{Z}[Q]$-module, which is finitely generated (as a $\mathbb{Z}[Q]$-module) if G is finitely generated. This fact allows us to apply commutative algebra to obtain decidability results for metabelian groups. In particular, in [53, 54] the following result is shown:

Theorem 8.1 *For every finitely generated metabelian group, the subgroup membership problem is decidable.*

The submonoid membership problem seems to mark the borderline between decidability and undecidability for metabelian groups.

Theorem 8.2 *The free metabelian group generated by two elements (this group is denoted by M_2 in the following) contains a fixed finitely generated submonoid with an undecidable membership problem.*

This result is shown in [36] via a reduction from the membership problem for finitely generated subsemimodules of free $(\mathbb{Z} \times \mathbb{Z})$-modules of finite rank. This latter problem is shown to be undecidable in [36] by interpreting it as a particular tiling problem of the Euclidean plane[6] that in turn is shown to be undecidable via a direct encoding of a Turing machine.

Also if one tries to generalize Theorem 8.1 to a larger classes of groups, one quickly reaches undecidability, as Umirbaev [57] has shown:

Theorem 8.3 *The free solvable group of derived length 3 and rank 2 has an undecidable subgroup membership problem.*

[6] A good introduction into tiling problems can be found in [9, Appendix A].

9 Wreath products

Let G and H be groups. Consider the direct sum

$$K = \bigoplus_{g \in G} H_g,$$

where H_g is a copy of H. We view K as the set

$$H^{(G)} = \{f \in H^G \mid f^{-1}(H \setminus \{1\}) \text{ is finite}\}$$

of all mappings from G to H with finite support together with pointwise multiplication as the group operation. The group G has a natural left action on $H^{(G)}$ given by

$$gf(a) = f(g^{-1}a)$$

where $f \in H^{(G)}$ and $g, a \in G$. The corresponding semidirect product $H^{(G)} \rtimes G$ is the wreath product $H \wr G$. In other words:

- Elements of $H \wr G$ are pairs (f, g), where $f \in H^{(G)}$ and $g \in G$.
- The multiplication in $H \wr G$ is defined as follows: Let $(f_1, g_1), (f_2, g_2) \in H \wr G$. Then $(f_1, g_1)(f_2, g_2) = (f, g_1 g_2)$, where $f(a) = f_1(a) f_2(g_1^{-1} a)$.

The following intuition might be helpful: An element $(f, g) \in H \wr G$ can be thought of as a finite multiset of elements of $H \setminus \{1\}$ that are sitting at certain elements of G (the mapping f) together with the distinguished element $g \in G$, which can be thought of as a cursor moving in G. If we want to compute the product $(f_1, g_1)(f_2, g_2)$, we do this as follows: First, we shift the finite collection of H-elements that corresponds to the mapping f_2 by g_1: If the element $h \in H \setminus \{1\}$ is sitting at $a \in G$ (i.e., $f_2(a) = h$), then we remove h from a and put it to the new location $g_1 a \in H$. This new collection corresponds to the mapping $f_2' \colon a \mapsto f_2(g_1^{-1} a)$. After this shift, we multiply the two collections of H-elements pointwise: If in $a \in G$ the elements h_1 and h_2 are sitting (i.e., $f_1(a) = h_1$ and $f_2'(a) = h_2$), then we put the product $h_1 h_2$ into the location a. Finally, the new distinguished G-element (the new cursor position) becomes $g_1 g_2$.

If H (resp., G) is generated by the set X (resp., Y) with $X \cap Y = \emptyset$, then $H \wr G$ is generated by $X \cup Y$. It is well-known and easy to see that decidability of the word problem for G and H implies decidability of the word problem for $H \wr G$. The following simple proposition is useful, see [37] for a proof:

Proposition 9.1 *Let K be a subgroup of G of finite index m and let H be a group. Then $H^m \wr K$ is isomorphic to a subgroup of index m in $H \wr G$.*

The following decidability result is shown in [37]:

Theorem 9.2 *The rational subset membership problem is decidable for every group $H \wr V$ with H finite and V virtually free.*

Note that Theorem 9.2 covers the well known lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$.

The proof of Theorem 9.2 in [37] makes use of well-quasi-order (wqo) theory. Let us briefly explain the idea for a wreath product $G = H \wr F_2$, where H is finite and F_2 is the free group generated by a and b. Given a finite automaton $\mathcal{A}$ over the alphabet

$H \cup \{a, a^{-1}, b, b^{-1}\}$, it suffices to check whether $\mathcal{A}$ accepts a word that represents the identity of G. The key ingredient is a certain language over the alphabet of triples (p, d, q), where p and q are states of the automaton $\mathcal{A}$ and $d \in \{a, a^{-1}, b, b^{-1}\}$. The idea is that such a triple may represent a path in $\mathcal{A}$ from state p to q such that the sequence of labels from $\{a, a^{-1}, b, b^{-1}\}$ along the path is a loop in the Cayley-graph of F_2 that leaves the origin in direction d and returns to the origin from direction d. The effect of such a path is the product of all transition labels; it is an element of the direct sum $K = \bigoplus_{g \in F_2} H$. A word w over the alphabet of triples is a *loop pattern* if each triple (p, d, q) in the word can be replaced by an automaton path as described above, such that the product of the effects of these paths is the identity of K. It is shown in [37] that the set of all loop patterns is a regular language. For this, it is shown that the set of loop patterns is an upward closed set of words with respect to a wqo, which is a refinement of the subsequence relation (also known as embeddability) on words (which is a wqo by Higman's Lemma). Using a saturation process one can actually compute an automaton for the set of all loop patterns. Using this, it is straightforward to check whether $\mathcal{A}$ accepts a word that represents the identity of G.

The computational complexity of the rational subset membership problem for groups $H \wr V$ with H finite and V virtually free is open. Due to the use of well quasi orders, the algorithm from [37] is not primitive recursive.

Problem 9.3 Is the rational subset membership problem for groups $H \wr V$ with H finite and V virtually free primitive recursive? In particular, is the rational subset membership problem for the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$ primitive recursive?

It should be mentioned that there exist several decision problems, for which decidability is proved using a well quasi order, and which can be shown to be not primitive recursive. An example is the membership problem for so called leftist grammars [23, 45] (these are grammars, where every production has the form $ab \to b$ or $d \to cd$).

By the following result from [37], decidability for the rational subset membership problem cannot be pushed very far beyond wreath products of the form $H \wr V$ with H finite and V virtually free:

Theorem 9.4 *There is a fixed finitely generated submonoid M of the wreath product $\mathbb{Z} \wr \mathbb{Z}$ with an undecidable membership problem.*

For the proof of Theorem 9.4 in [37], the authors encode the acceptance problem for a 2-counter machine (Minsky machine [44]) into the submonoid membership problem for $\mathbb{Z} \wr \mathbb{Z}$. One should remark that $\mathbb{Z} \wr \mathbb{Z}$ is a finitely generated metabelian group and hence has a decidable subgroup membership problem, see Theorem 8.1.

The wreath product $\mathbb{Z} \wr \mathbb{Z}$ is a subgroup of Thompson's group F (see [41]) as well as of Baumslag's [3] finitely presented metabelian group $\langle a, s, t \mid [s, t] = [a^t, a] = 1, a^s = aa^t \rangle$, see, e.g., [11]. Hence, we get:

Corollary 9.5 *Thompson's group F and Baumslag's finitely presented metabelian group both contain finitely generated submonoids with an undecidable membership problem.*

A further undecidability result for wreath products was shown in [36]:

Theorem 9.6 *For every non-trivial group H, the rational subset membership problem for $H \wr (\mathbb{Z} \times \mathbb{Z})$ is undecidable.*

The proof of this result in [36] uses an encoding of a tiling problem, which uses the grid structure of the Cayley graph of $\mathbb{Z} \times \mathbb{Z}$. It is very similar to the undecidability proof for the submonoid membership problem for free metabelian groups (Theorem 8.2). It is open, whether Theorem 9.6 can be sharpened to the submonoid membership problem:

Problem 9.7 Assume that H is a non-trivial group. Is the submonoid membership problem for $H \wr (\mathbb{Z} \times \mathbb{Z})$ undecidable?

The author conjectures that the answer to this question is positive. Another reasonable conjecture is that Theorem 9.6 can be generalized to every wreath product $H \wr G$, where H is non-trivial and G is not virtually free (note that $\mathbb{Z} \times \mathbb{Z}$ is not virtually free).

Problem 9.8 Assume that H is a non-trivial group and G is not virtually free. Is the rational subset membership problem for $H \wr G$ undecidable?

As remarked above, the author conjectures that the answer to this question is again positive. The reason is that the undecidability proof for $H \wr (\mathbb{Z} \times \mathbb{Z})$ from [36] only uses the grid-like structure of the Cayley graph of $\mathbb{Z} \times \mathbb{Z}$. In [29] it was shown that the Cayley graph of a group G has bounded tree width (a graph-theoretic measure that, roughly speaking, determines how tree-like a graph is) if and only if the group is virtually free. Hence, if G is not virtually free, then the Cayley-graph of G has unbounded tree width. By known results from graph theory, this implies that finite grids of arbitrary size appear as graph-theoretic minors in the Cayley-graph of G. There is hope to use these grids for encoding an undecidable tiling problem into the rational subset membership problem for $H \wr G$ (for H non-trivial).

Theorem 9.4 and 9.6 imply the following: For finitely generated non-trivial abelian groups G and H, the wreath product $H \wr G$ has a decidable rational subset membership problem if and only if (i) G is finite[7] or (ii) G has rank 1 and H is finite. Furthermore, for virtually free groups G and H, the rational subset membership problem is decidable for $H \wr G$ if and only if (i) G is trivial, or (ii) H is finite, or (iii) G is finite and H is virtually $\mathbb{Z}$, i.e., has $\mathbb{Z}$ as a finite index subgroup. Note that if G is finite, then Proposition 9.1 implies that $H^{|G|}$ is a finite index subgroup of $H \wr G$. Hence, if H is virtually $\mathbb{Z}$, then $H^{|G|}$ is virtually abelian and hence has a decidable rational subset membership problem. On the other hand, if H is virtually F_n for F_n a free group of rank $n > 1$ and G is nontrivial, then $H^{|G|}$ (and hence $H \wr G$) has an undecidable subgroup membership problem by Theorem 4.4.

[7] If G has size m, then by Proposition 9.1, $H^m \cong H^m \wr 1$ is isomorphic to a subgroup of index m in $H \wr G$. Since H^m is finitely generated abelian, decidability of the rational subset membership problem of $H \wr G$ follows from Theorems 4.2 and 5.5.

10 Rational subsets versus submonoids

It is a trivial obersvation that decidability of the rational subset membership problem for a group G implies decidability of the submonoid membership problem for G, and the latter implies decidability of the subgroup membership problem for G. On the other hand, we have seen groups, for which the subgroup membership problem is decidable, but the submonoid membership problem is undecidable. Examples are the free metabelian group generated by two elements (see Theorems 8.1 and 8.2) and the right-angled Artin group $\mathbb{G}(P4)$ (see Theorems 6.1 and 6.3; note that $P4$ does not contain an induced cycle, which allows us to apply Theorem 6.1). It is therefore an interesting question, whether there is a finitely generated group, for which the submonoid membership problem is decidable but the rational subset membership problem is undecidable. Unfortunately, we do not know, whether such a group exists.

Problem 10.1 Is there a finitely generated group, for which the submonoid membership problem is decidable but the rational subset membership problem is undecidable?

By the following result from [35] we know that if such a group exists, then it must have only one end. The number of ends of a finitely generated infinite group G is a geometric invariant of G that is defined as follows: Assume that G is finitely generated by the symmetric set X (the following definition is not influenced by the concrete choice of X) and consider the Cayley graph $\mathcal{G}(G, X)$. The nodes of this graph are the elements of G and there is an edge between two elements of $g, h \in G$ if and only if there is a generator $a \in X$ such that $h = ga$ in G. This graph is undirected (since X is symmetric) and connected (since X generates G). Moreover, it is vertex-transitive, which means that for all $g, h \in G$, there is a graph automorphism of $\mathcal{G}(G, X)$ that maps g to h. To define the number of ends of G, choose an arbitrary node $g \in G$ (the concrete choice of g is not important) and let $\mathcal{G}_n$ (for $n \geq 0$) be the subgraph of $\mathcal{G}(G, X)$ obtained by removing all nodes from $\mathcal{G}(G, X)$ that have distance at most n from g. Let e_n be the number of connected components of $\mathcal{G}_n$. Then the number of ends is the limit of the sequence $(e_n)_{n \geq 0}$ or ∞ if this sequence is unbounded. By the Freudenthal-Hopf Theorem, every finitely generated infinite group G has either 1, 2, or ∞ many ends, see, e.g., [41]. Here are three typical examples for each possibility:

- The number of ends of $\mathbb{Z} \times \mathbb{Z}$ is 1.
- The number of ends of $\mathbb{Z}$ is 2.
- The number of ends of the free group F_2 or rank 2 is ∞.

A group has two ends if and only if it is virtually $\mathbb{Z}$. A seminal result of Stallings [55, 56] characterizes groups with infinitely many ends: A group has infinitely many ends if and only if it is an HNN-extension with finite associated subgroups or an amalgamated product with finite amalgamated subgroups. The following result was shown in [35]:

Theorem 10.2 *Assume that G is a finitely generated group G. If G has more than one end, then the rational subset membership problem for G is decidable if and only if the submonoid membership problem for G is decidable.*

The case of group G with two ends is easy: G has $\mathbb{Z}$ as a finite index subgroup. Since the rational subset membership problem for $\mathbb{Z}$ is decidable, Theorem 5.5 implies that the rational subset membership problem (and hence also the submonoid membership problem) is decidable. So, it remains to consider a group G with infinitely many ends. By Stalling's theorem one can write G as an HNN-extension with finite associated subgroups or an amalgamated product with finite amalgamated subgroups. Let us sketch the proof of Theorem 10.2 in a simple case that nevertheless shows the main idea: Assume that $G = H * F_2$ and assume that the submonoid membership problem for G is decidable. We have to show that G has a decidable rational subset membership problem. By Theorem 5.2 (and the fact that F_2 has a decidable rational subset membership problem) it suffices to show that H has a decidable rational subset membership problem. So, let us fix a generating set X for H together with a canonical homomorphism $\pi : X^* \to H$, and let $\mathcal{A} = (Q, \Delta, q_0, F)$ be a finite automaton over X. By adding ε-transitions to Δ we can assume that F consists of a single state $q_f \neq q_0$. Since F_2 contains a copy of F_n (the free group of rank n) for any $n \geq 1$, we can assume that F_2 contains a copy of $F(Q)$, i.e., the free group generated by the states of $\mathcal{A}$. Recall that $\Delta \subseteq Q \times (X \cup \{\varepsilon\}) \times Q$ is the set of transitions. Now define a finitely generated submonoid of $G = H * F_2$ as follows. Let

$$Y = \{q^{-1}ap \mid (q, a, p) \in \Delta\} \subseteq H * F(Q) \subseteq H * F_2 = G.$$

Then, one can show that for every $w \in X^*$, we have $\pi(w) \in \pi(L(\mathcal{A}))$ if and only if $q_0^{-1}wq_f$ represents an element of the submonoid Y^* of G. The idea is that in a product of the form $(q_1^{-1}a_1p_1)(q_2^{-1}a_2p_2)\cdots(q_n^{-1}a_np_n)$ a factor of the form $p_iq_{i+1}^{-1}$ with $p_i \neq q_{i+1}$ cannot be erased. On the other hand, if $p_i = q_{i+1}$ for $1 \leq i \leq n-1$, then the word is equal to $q_1^{-1}(a_1a_2\cdots a_n)p_n$.

Problem 10.1 is related to Problem 5.3: Assume that the class of finitely generated groups with a decidable submonoid membership problem is closed under free product (whether this is true was asked in Problem 5.3). Let G be an arbitrary finitely generated group with a decidable submonoid membership problem. Hence, by our assumption, also the free product $G * F_2$ has a decidable submonoid membership problem. But this group has infinitely many ends. So, by Theorem 10.2, $G * F_2$ has a decidable rational subset membership problem. But then, also the finitely generated subgroup G has a decidable rational subset membership problem.

The author conjectures that one can construct a finitely generated group with a decidable submonoid membership problem and an undecidable rational subset membership problem. This leads to the conjecture that the class of groups with a decidable submonoid membership problem is not closed under free products.

11 Further results on the submonoid membership problem

Let us briefly mention some further results on the submonoid membership problem. In [46] the bounded submonoid membership problem for a finitely generated group G was introduced:

Decision problem 11.1 (Bounded submonoid membership problem)

- INPUT: Elements $g, g_1, \ldots, g_n \in G$ and a unary encoded number k.

- QUESTION: Can g be written as a product $g = g_{i_1} g_{i_2} \cdots g_{i_l}$ with $l \leq k$ and $1 \leq i_1, \ldots, i_l \leq n$.

It was shown in [46] that the bounded submonoid membership problem can be solved in polynomial time for finitely generated virtually nilpotent groups and word hyperbolic groups.

In [22] it was shown that the word problem for a one-relator inverse monoid $\mathsf{Inv}\langle X \mid r = 1\rangle$ is decidable if and only if the submonoid of the one-relator group $\mathsf{Gp}\langle X \mid r = 1\rangle$ that is generated by all prefixes of r has a decidable membership problem. The latter problem is also called the *prefix monoid membership problem* for the one-relator group $\mathsf{Gp}\langle X \mid r = 1\rangle$. Motivated by this result, the submonoid membership problem was further studied in [40], where a general technique based on distortion functions for solving submonoid membership problems is introduced. Using this technique, the authors show that the prefix membership problem is decidable for Baumslag-Solitar groups, surface groups of genus at least two (for which decidability was already shown in [22]), and certain one-relator groups given by Adian type presentations.

12 The rational subset membership problem for monoids and semigroups

We defined the notion of a rational subset for all monoids. Hence, it makes sense to study the rational subset membership problem for finitely generated monoids (and, by replacing the monoid closure by the semigroup closure, even for finitely generated semigroups). Kambites and Render proved several interesting results in this context. They showed that the rational subset membership problem is decidable for the following classes of finitely generated monoids:

- Polycyclic and bicyclic monoids [49],
- Finitely generated Rees matrix semigroups (with or without zero) over a semigroup with decidable rational subset membership problem [50],
- Monoids that satisfy the small overlap condition $C(4)$ (which is inspired by small cancellation theory for groups) [25].

References

[1] I.J. Aalbersberg & H.J. Hoogeboom, Characterizations of the decidability of some problems for regular trace languages, *Math. Systems Theory* **22** (1989), 1–19.

[2] J. Avenhaus & D. Wißmann, Using rewriting techniques to solve the generalized word problem in polycyclic groups, *Proc. ISSAC 1989*, 322–337, ACM Press, 1989.

[3] G. Baumslag, A finitely presented metabelian group with a free abelian derived group of infinite rank, *Proc. Amer. Math. Soc.* **35** (1972), 61–62.

[4] G. Bazhenova, Rational sets in finitely generated nilpotent groups, *Algebra Log.* **39** (2000), 379–394.

[5] M. Benois, Parties rationnelles du groupe libre, *C. R. Acad. Sci. Paris Sér. A–B* **269** (1969), 1188–1190.

[6] M. Benois & J. Sakarovitch, On the complexity of some extended word problems defined by cancellation rules, *Inform. Process. Lett.* **23** (1986), 281–287.

[7] J. Berstel & J. Sakarovitch, Recent results in the theory of rational sets, *Proc. MFCS 1986*, 15–28, Lecture Notes in Comput. Sci. **233**, Springer, 1986.

[8] M. Bestvina & N. Brady, Morse theory and finiteness properties of groups, *Invent. Math.* **129** (1997), 445–470.
[9] E. Börger, E. Grädel & Y. Gurevich, *The Classical Decision Problem*, Springer, 2001.
[10] P. Buckheister & G. Zetzsche, Semilinearity and context-freeness of languages accepted by valence automata, *Proc. MFCS 2013*, 231–242, Lecture Notes in Comput. Sci. **8087**, Springer, 2013.
[11] S. Cleary, Distortion of wreath products in some finitely-presented groups, *Pacific J. Math.* **228** (2006), 53–61.
[12] J. Crisp & B. Wiest, Embeddings of graph braid and surface groups in right-angled Artin groups and braid groups, *Algebr. Geom. Topol.* **4** (2004), 439–472.
[13] M. Dehn, Über unendliche diskontinuierliche Gruppen, *Math. Ann.* **71** (1911), 116–144.
[14] V. Diekert & A. Muscholl, Solvability of equations in free partially commutative groups is decidable, *Internat. J. Algebra Comput.* **16** (2006), 1047–1069.
[15] S. Eilenberg & M.P. Schützenberger, Rational sets in commutative monoids, *J. Algebra* **13** (1969), 173–191.
[16] H. Fernau & R. Stiebe, Sequential grammars and automata with valences, *Theor. Comput. Sci.* **276** (2002), 377–405.
[17] R. Ghrist & V. Peterson, The geometry and topology of reconfiguration, *Adv. in Appl. Math.* **38** (2007), 302–323.
[18] R.H. Gilman, Formal languages and infinite groups, *Geometric and computational perspectives on infinite groups (Minneapolis, MN & New Brunswick, NJ, 1994)*, 27–51, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **25**, Amer. Math. Soc., 1996.
[19] Z. Grunschlag, *Algorithms in Geometric Group Theory*, PhD thesis, University of California at Berkley, 1999.
[20] P. Hall, Finiteness conditions for soluble groups, *Proc. London Math. Soc. (3)* **4** (1954), 419–436.
[21] J.E. Hopcroft & J.D. Ullman, *Introduction to Automata Theory, Languages and Computation*, Addison–Wesley, 1979.
[22] S.V. Ivanov, S.W. Margolis & J.C. Meakin, On one-relator inverse monoids and one-relator groups, *J. Pure Appl. Algebra* **159** (2001), 83–111.
[23] T. Jurdzinski, Leftist grammars are non-primitive recursive, *Proc. ICALP 2008, Part II*, 51–62, Lecture Notes in Comput. Sci. **5126**, Springer, 2008.
[24] M. Kambites, Formal languages and groups as memory, *Comm. Algebra* **37** (2009), 193–208.
[25] M. Kambites, Small overlap monoids, II: Automatic structures and normal forms, *J. Algebra* **321** (2009), 2302–2316.
[26] M. Kambites, P.V. Silva & B. Steinberg, On the rational subset problem for groups, *J. Algebra* **309** (2007), 622–639.
[27] I. Kapovich, R. Weidmann & A. Myasnikov, Foldings, graphs of groups and the membership problem, *Internat. J. Algebra Comput.* **15** (2005), 95–128.
[28] E. Kopczynski & A.W. To, Parikh images of grammars: complexity and applications, *Proc. LICS 2010*, 80–89, IEEE Computer Soc., 2010.
[29] D. Kuske & M. Lohrey, Logical aspects of Cayley-graphs: the group case, *Ann. Pure Appl. Logic* **131** (2005), 263–286.
[30] H. Lenstra, Integer programming with a fixed number of variables, *Math. Oper. Res.* **8** (1983), 538–548.
[31] M. Lohrey & G. Sénizergues, Theories of HNN-extensions and amalgamated products, *Proc. ICALP 2006*, 681–692, Lecture Notes Comput. Sci. **4052**, Springer, 2006.
[32] M. Lohrey & G. Sénizergues, Rational subsets in HNN-extensions and amalgamated products, *Internat. J. Algebra Comput.* **18** (2008), 111–163.
[33] M. Lohrey & B. Steinberg, The submonoid and rational subset membership problems for graph groups, *J. Algebra* **320** (2008), 728–755.

[34] M. Lohrey & B. Steinberg, An automata theoretic approach to the generalized word problem in graphs of groups, *Proc. Amer. Math. Soc.* **138** (2010), 445–453.
[35] M. Lohrey & B. Steinberg, Submonoids and rational subsets of groups with infinitely many ends, *J. Algebra* **324** (2010), 970–983.
[36] M. Lohrey & B. Steinberg, Tilings and submonoids of metabelian groups, *Theory Comput. Syst.* **48** (2011), 411–427.
[37] M. Lohrey, B. Steinberg & G. Zetzsche, Rational subsets and submonoids of wreath products, *Proc. ICALP 2013, Part II*, 361–372, Lecture Notes Comput. Sci. **7966**, Springer, 2013.
[38] R.C. Lyndon & P.E. Schupp, *Combinatorial Group Theory*, Springer, 1977.
[39] A.I. Mal'cev, On homomorphisms onto finite groups, *Amer. Math. Soc. Transl. (2)* **119** (1983), 67–79. Translation from *Ivanov. Gos. Ped. Inst. Ucen. Zap.* **18** (1958) 49–60.
[40] S.W. Margolis, J. Meakin & Z. Šuniḱ, Distortion functions and the membership problem for submonoids of groups and monoids, *Geometric methods in group theory*, 109–129, Contemp. Math. **372**, Amer. Math. Soc., 2005.
[41] J. Meier, *Groups, Graphs and Trees: an Introduction to the Geometry of Infinite Groups*, London Math. Soc. Student Texts **73**, CUP, 2008.
[42] K.A. Mihaĭlova, The occurrence problem for free products of groups, *Doklady Akademii Nauk SSSR* **127** (1959), 746–748.
[43] K.A. Mihaĭlova, The occurrence problem for direct products of groups, *Math. Sb. (N.S.)* **70 (112)** (1966), 241–251.
[44] M.L. Minsky, *Computation: Finite and Infinite Machines*, Prentice-Hall, 1967.
[45] R. Motwani, R. Panigrahy, V.A. Saraswat & S. Venkatasubramanian, On the decidability of accessibility problems, *Proc. STOC 2000*, 306–315. ACM Press, 2000.
[46] A. Myasnikov, A. Nikolaev & A. Ushakov, Knapsack problems in groups, arXiv.org, 2013. http://arxiv.org/abs/1302.5671.
[47] J. Nielsen, Om regning med ikke kommutative faktoren og dens anvendelse i gruppeteorien, *Matematisk Tidsskrift, B.* (1921), 77–94 (Danish).
[48] C.H. Papadimitriou, *Computational Complexity*, Addison Wesley, 1994.
[49] E. Render & M. Kambites, Rational subsets of polycyclic monoids and valence automata, *Inform. and Comput.* **207** (2009), 1329–1339.
[50] E. Render & M. Kambites, Semigroup automata with rational initial and terminal sets, *Theor. Comput. Sci.* **411** (2010), 1004–1012.
[51] H. Rogers, *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, 1968.
[52] V. Roman'kov, On the occurrence problem for rational subsets of a group, *International Conference on Combinatorial and Computation Methods in Mathematics*, 76–81, 1999.
[53] N.S. Romanovskiĭ, Some algorithmic problems for solvable groups, *Algebra Log.* **13** (1974), 26–34.
[54] N.S. Romanovskiĭ, The occurrence problem for extensions of abelian groups by nilpotent groups, *Sibirsk. Mat. Zh.* **21** (1980), 170–174.
[55] J.R. Stallings, On torsion-free groups with infinitely many ends, *Ann. of Math. (2)* **88** (1968), 312–334.
[56] J.R. Stallings, *Group Theory and Three-Dimensional Manifolds*, Yale Mathematical Monographs **4**, Yale University Press, 1971.
[57] U.U. Umirbaev, The occurrence problem for free solvable groups, *Algebra i Logika* **34** (1995), 211–232, 243.
[58] D.T. Wise, A residually finite version of Rips's construction, *Bull. London Math. Soc.* **35** (2003), 23–29.
[59] E.S. Wolk, A note on the "The comparability graph of a tree", *Proc. Amer. Math. Soc.* **16** (1965), 17–20.
[60] G. Zetzsche, On the capabilities of grammars, automata, and transducers controlled by

monoids, *Proc. ICALP 2008, Part II*, 222–233, Lecture Notes in Comput. Sci. **6756**, Springer, 2011.

[61] G. Zetzsche, Silent transitions in automata with storage, in *Proc. ICALP 2013, Part II*, 434–445, Lecture Notes in Comput. Sci. **7966**, Springer, 2013.

A SURVEY ON MILNOR LAWS

OLGA MACEDOŃSKA

Institute of Mathematics, Silesian University of Technology, Gliwice, Poland
Email: Olga.Macedonska@polsl.pl

Abstract

The group laws which are not satisfied in product-varieties $\mathfrak{A}_p\mathfrak{A}$ for any prime p will be called, after F. Point, the Milnor laws. Let G be a 2-generator relatively free group defined by a nontrivial law $w \equiv 1$. We show that this law is the Milnor law if and only if G'/G'' is finitely generated. Moreover, the properties $G' = G''$ and G' is finitely or infinitely generated, allow us to split the Milnor laws for three disjoint classes. We describe their properties.

1 Introduction

We denote by $F = \langle x, y\rangle$ the free group of rank 2 and by F_∞ the free group on $X = \{x_1, x_2, \dots\}$. A word $w = w(x_1, x_2, \dots, x_n)$ is called a law for a group G if $w(g_1, g_2, \dots, g_n) = e$ for all $g_1, g_2, \dots, g_n$ in G. The law w can be written as $w \equiv 1$. We denote $[x, y] = x^{-1}y^{-1}xy = x^{-1}x^y$, $[x,_0\, y] = x$, and $[x,_n\, y] = [[x,_{n-1}\, y], y]$.

A variety of groups is the class of all groups satisfying every law in a given set of laws (see [20] Chapter 1). Each variety is defined by a verbal subgroup $V \subseteq F_\infty$ and consists of all groups G satisfying $V(G) = \{e\}$.

By $\mathfrak{A}_p$ we denote the variety of all abelian groups of a prime exponent p, by $\mathfrak{A}$ – the variety of all abelian groups. $\mathfrak{N}_c$ denotes the variety of nilpotent groups of class $\leq c$ and $\mathfrak{B}_e$ – the variety of all groups of exponent dividing e.

In the following section we describe different types of laws, such as positive laws, pseudo-abelian laws, $\mathfrak{R}$-laws, Milnor laws. The corresponding varieties have similar names. The aim of the paper is to show that the commutator subgroup G' of a 2-generator free group G in the variety defined by a law $w \equiv 1$ is responsible for the crucial properties of the law.

We prove that the number of generators in G' and G'/G'' allow to split all the laws into four disjoint classes, where for the classes (I), (II), (III), G'/G'' is finitely generated and for (II), (III), (IV), $G' \neq G''$.

We write f.g. and inf.g. for 'finitely generated' and 'infinitely generated' respectively. We show the connection of G' and the following law properties:

(I) $G' = G''$ – abelian and pseudo-abelian laws
(II) $G' \neq G''$, G' f.g. – restraining laws ($\mathfrak{R}$-laws)
(III) $G' \neq G''$, G' inf.g., G'/G'' f.g. – there are no examples known
(IV) $G' \neq G''$, G'/G'' inf.g. – non-Milnor laws.

2 Different types of laws and varieties

We recall definitions and facts concerning different types of laws and varieties. We use the same name for a law and the variety, it defines.

Positive laws and varieties

Definition 2.1 A law is called positive if it implies a nontrivial law of the form $u(x_1, x_2, \ldots, x_n) \equiv v(x_1, x_2, \ldots, x_n)$ where u, v are positive words (not involving the inverses of the x_i's)

The law is called balanced if the exponent sum of each x_i is the same in u and v. Note that each positive law implies a two-variable positive law $u(x, y) \equiv v(x, y)$ if we substitute x_i by xy^i. This law can be assumed balanced because each law $u \equiv v$ implies the law $uv \equiv vu$. So we have the following.

Proposition 2.2 *Let G be a 2-generator free group in a variety $\mathfrak{V}$. Every group in $\mathfrak{V}$ satisfies a positive law if and only if G satisfies a binary balanced positive law.*

In 1953 A. I. Maltsev [15] and independently in 1963 B. H. Neumann and T. Taylor [19] proved that nilpotency can be defined by a positive law. It follows that groups which are nilpotent-by-(finite exponent), in particular nilpotent-by-finite groups, satisfy positive laws.

Pseudo-abelian laws and varieties

Definition 2.3 A non-abelian law is called pseudo-abelian if every metabelian group satisfying this law is abelian. A variety defined by such a law is called a pseudo-abelian variety.

The problem of existence of such a variety was formulated as a Problem 5 in the book of Hanna Neumann [20] and solved by A. Yu. Ol'shanskii in [21].

It is clear that a relatively free group G in a pseudo-abelian variety has $G' = G''$, which is if and only if G satisfies a law of the form

$$(*): \ [x, y] \equiv u, \quad \text{where} \quad u := u(x, y) \in F''.$$

Proposition 2.4 *Let $\mathfrak{V}$ be a variety of groups and let G be the free group of rank 2 in $\mathfrak{V}$. We have $G' = G''$ if and only if every finite group in $\mathfrak{V}$ is abelian.*

Proof Assume that there is a finite non-abelian group satisfying the law $(*)$ then there is such a group H of the smallest order, all whose proper subgroups are abelian. By [16], the group H must be metabelian. Then H is abelian. The contradiction proves that every finite group satisfying $(*)$ is abelian.

Conversely, let every finite group satisfying $(*)$ be abelian and H be a non-abelian metabelian group satisfying $(*)$. By [18], every 2-generator non-abelian metabelian group has a finite non-abelian quotient. This contradiction proves that every metabelian group satisfying $(*)$ is abelian. □

Answering a question of H. Neumann (Problem 5 of [20]), A. Yu. Ol'shanskii [21] proved existence of pseudo-abelian varieties by constructing a family of pseudo-abelian laws of the form

$$[x,y] \equiv v^n\,[y,x]^{\varepsilon_1} v^{n+1}\,[y,x]^{\varepsilon_2} v^{n+2} \cdots [y,x]^{\varepsilon_{h-1}} v^{n+h-1} \tag{1}$$

where $v := [\,[y^d,x^d]^d,\,[x^d,y^{-d}]^d]$, $h \equiv 1 \pmod{10}$, $\varepsilon_{10k+1} = \varepsilon_{10k+2} = \varepsilon_{10k+3} = \varepsilon_{10k+5} = \varepsilon_{10k+6} = 1$, $\varepsilon_{10k+4} = \varepsilon_{10k+7} = \varepsilon_{10k+8} = \varepsilon_{10k+9} = \varepsilon_{10k+10} = -1$, $k = 0,1,\ldots,(h-1)/10$ and d,n,h are sufficiently large natural numbers chosen with respect to the restrictions that are given in Chapter 7 of [22]. Note that $\varepsilon_1 + \cdots + \varepsilon_{h-1} = 0$ and the right side of (1) is in F''. The varieties defined by the law (1) are of exponent zero [11]. The pseudo-abelian varieties of finite exponent are constructed in [9].

Proposition 2.5 *A law $[x,y] \equiv [x,_n y]$, $n > 1$, is either abelian or pseudo-abelian.*

Proof If we substitute $[y,x]$ for y, we get $[x,[y,x]] \equiv [x,_n [y,x]]$, which implies the law $[y,x,x] \equiv u(x,y)$, where $u(x,y) \in F''$. It follows that a metabelian group satisfying the law $[x,y] \equiv [x,_n y]$, also satisfies $[x,y,y] \equiv 1$. Since $[x,_n y] = [[x,y,y],_{n-2} y]$, every metabelian group satisfying the initial law is abelian, which finishes the proof. □

It was shown by N. Gupta [6] that for $n \leq 3$ the law $[x,y] \equiv [x,_n y]$ is abelian.

Question 1 Is the law $[x,y] \equiv [x,_n y]$ pseudo-abelian for $n > 3$?

Varieties $\mathfrak{A}_p\mathfrak{A}$ for a prime p

The variety $\mathfrak{A}_p\mathfrak{A}$ is metabelian and every group G in it has $(G')^p = \{e\}$. The variety $\mathfrak{A}_p\mathfrak{A}$ is generated by the restricted wreath product $W := C_p \wr C$, where $C_p = \langle a\rangle_p$, $C = \langle b\rangle_\infty$ denote the cyclic group of order p and the infinite cyclic group respectively (see e.g. [20], Corollary 22.44). V. V. Belyaev, N. F. Sesekin [2] proved that the group $C_p \wr C$ contains a free subsemigroup of rank 2, hence the varieties $\mathfrak{A}_p\mathfrak{A}$ do not satisfy positive laws.

Proposition 2.6 *The commutator subgroup W' of the group $W := C_p \wr C$ is infinitely generated.*

Proof The group W contains elements $[a,b^i] = a^{-1}a^{b^i}$ for all $i \in \mathbb{Z}$, hence W' has an infinite support and cannot be finitely generated. □

Just not *p.l.*-varieties

Every variety which does not satisfy a positive law contains (by Zorn Lemma) a minimal subvariety without positive laws, which is called a *just not p.l.*-variety. We can see that the varieties of the form $\mathfrak{A}_p\mathfrak{A}$ are *just not p.l.*-varieties. In 1971 J. R. J. Groves [5, Theorem C(*ii*)] proved, in particular, that each soluble variety either lies in a variety of the form $\mathfrak{N}_c\mathfrak{B}_e$ or contains a subvariety $\mathfrak{A}_q\mathfrak{A}$ for some prime q. Since a proper subvariety in $\mathfrak{A}_p\mathfrak{A}$ cannot contain any of $\mathfrak{A}_q\mathfrak{A}$, we conclude that each proper

subvariety in $\mathfrak{A}_p\mathfrak{A}$ lies in some $\mathfrak{N}_c\mathfrak{B}_e$, hence satisfies a positive law. So the varieties $\mathfrak{A}_p\mathfrak{A}$ for any prime p are *just not p.l.*-varieties.

The problem *whether the varieties of the form $\mathfrak{A}_p\mathfrak{A}$ are the only just not p.l.-varieties* posed in [26, 19.2], was solved in [10, Corollary]. It was shown that there exist continuum of pseudo-abelian *just not p.l.*-varieties. The question arises

Question 2 Must every just not *p.l.*-variety be either pseudo-abelian or of the form $\mathfrak{A}_p\mathfrak{A}$ for some prime p?

$\mathfrak{R}$-laws and varieties

The property that *for all $g, h \in G$ the subgroup $\langle\, h^{-i}gh^i,\ i \in \mathbb{Z}\,\rangle$ is finitely generated* is called the Milnor property by F. Point in [23] because this property first was used by J. Milnor [17, Lemma 3]. Later it attracted attention of many authors, e.g., [25], [8], [3], [13], [14]. The groups satisfying this property are called restrained groups [8]. We call a law providing this property an $\mathfrak{R}$-law (or a restraining law) [12].

Definition 2.7 A law $w \equiv 1$ is called an $\mathfrak{R}$-law if for all elements g, h in a group satisfying this law, the subgroup $\langle\, h^{-i}gh^i,\ i \in \mathbb{Z}\,\rangle$ is finitely generated.

We need the following property of $\mathfrak{R}$-laws, which can be found in [7].

Lemma 2.8 *Let G be an n-generator group satisfying an $\mathfrak{R}$-law. If G/H is cyclic then H is finitely generated.*

Proof We can assume that G/H is generated by a coset gH and $\langle g\rangle \cap H = \{e\}$. Since G is an n-generator group, there exist $h_1, \dots, h_n \in H$ such that $G = \langle g, h_1, \dots, h_n\rangle$. Let $T := \langle h_1, \dots, h_n\rangle$. Then its normal closure T^G is generated by subgroups $g^{-i}Tg^i$, $i \in \mathbb{Z}$, hence by subgroups $\langle\, g^{-i}h_kg^i,\ i \in \mathbb{Z}\,\rangle$, $k = 1, 2, \dots, n$, which are finitely generated by assumption. So T^G is finitely generated. Moreover, $\langle g\rangle T^G = G$ and $T^G \subseteq H$. Then by means of the modular law: $H = G \cap H = \langle g\rangle T^G \cap H = (\langle g\rangle \cap H)T^G = T^G$. So H is finitely generated. □

Proposition 2.9
1. *Every finitely generated group G satisfying an $\mathfrak{R}$-law has its commutator subgroup G' finitely generated.*
2. *If a 2-generator relatively free group G defined by a law $w \equiv 1$ has G' finitely generated, then $w \equiv 1$ is the $\mathfrak{R}$-law.*

Proof 1. Let G satisfy an $\mathfrak{R}$-law. Since G/G' is abelian and finitely generated, there exists a finite normal series with, say, m cyclic factors: $G = N_0 \triangleright N_1 \triangleright \cdots \triangleright N_m = G'$. If N_i is finitely generated then since N_i/N_{i+1} is cyclic, we have by Lemma 2.8 that N_{i+1} is finitely generated. By repeating this step we get that G' is finitely generated.

2. Let G be a relatively free group defined by a law $w \equiv 1$, with free generators a, b and let G' be finitely generated. Then the normal closure of a is equal to $\langle\, b^{-i}ab^i,\ i \in \mathbb{Z}\,\rangle = \langle a\rangle\big[\langle a\rangle, \langle b\rangle\big] = \langle a\rangle G'$, hence is finitely generated.

Since for all elements g, h in any group satisfying the law $w \equiv 1$, the subgroup $\langle\, h^{-i}gh^i,\ i \in \mathbb{Z}\,\rangle$ is an image of $\langle\, b^{-i}ab^i,\ i \in \mathbb{Z}\,\rangle$, we conclude that the law $w \equiv 1$ is an $\mathfrak{R}$-law. □

Lemma 2.10 *A law $w \equiv 1$ is an $\mathfrak{R}$-law if and only if for some natural n it implies a law of the form*

$$x^{y^n} \equiv v(x,y) \qquad \text{where} \qquad v(x,y) \in \langle x, x^y, \ldots, x^{y^{n-2}}, x^{y^{n-1}} \rangle \tag{2}$$

Proof Let G be a relatively free group of rank 2, defined by the law $w \equiv 1$. Let G be generated by elements a, b. Assume that G satisfies an $\mathfrak{R}$-law. By Definition 2.7 the subgroup $\langle\, b^{-i}ab^i,\ i \in \mathbb{Z} \,\rangle$ is finitely generated. Then it is generated by some n consequent elements in the row $\ldots, a^{b^{-3}}, a^{b^{-2}}, a^{b^{-1}}, a, a^b, a^{b^2}, a^{b^3}, \ldots$. Since this subgroup is invariant to conjugation by b, it is generated by $a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}}$ and then

$$a^{b^n} \in \langle a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}} \rangle, \tag{3}$$

which implies the law (2) because a and b are the free generators in G.

Conversely, the law (2) implies (3) which, conjugated by b, gives

$$a^{b^{n+1}} \in \langle a^b,\ a^{b^2}, \ldots, a^{b^{(n-1)}}, a^{b^n} \rangle$$

and in view of (3) can be written as

$$a^{b^{n+1}} \in \langle a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}} \rangle$$

Repeated conjugation gives by induction $a^{b^i} \in \langle a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}} \rangle$, $i \geq 0$. By (3) we obtain

$$a^{b^{-n}} \in \langle a, a^{b^{-1}}, \ldots, a^{b^{-(n-1)}} \rangle. \tag{4}$$

Conjugating by b^{n-1}, it follows that $a^{b^{-1}}$ is in $\langle a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}} \rangle$, which implies that $\langle\, b^{-i}ab^i,\ i \in \mathbb{Z} \,\rangle = \langle a, a^b, \ldots, a^{b^{n-2}}, a^{b^{n-1}} \rangle$ is finitely generated. □

Example 2.11 Every Engel law $[x,_n y] \equiv 1$ is an $\mathfrak{R}$-law.

Proof Let v_i denote any word in $\langle x, x^y, \ldots, x^{y^i} \rangle$. In view of Lemma 2.10 it suffices to show that the law $[x,_n y] \equiv 1$ is equivalent to a law $v_{n-1}x^{y^n} \equiv 1$.

For $n = 1$, $[x, y] = x^{-1}x^y = v_0 x^y$, and if assume $[x,_{n-1} y] = v_{n-2}x^{y^{n-1}}$, then

$$[x,_n y] = (v_{n-2}x^{y^{n-1}})^{-1}(v_{n-2}x^{y^{n-1}})^y = \underbrace{(x^{-y^{n-1}}v_{n-2}^{-1}v_{n-2}^y)}_{v_{n-1}}x^{y^n} = v_{n-1}x^{y^n}.$$

Hence the law $[x,_n y] \equiv 1$ implies (2), which finishes the proof. □

Example 2.12 Every positive law is an $\mathfrak{R}$-law.

Proof Each positive law implies a balanced positive law $u(x,y) \equiv v(x,y)$ where we can assume that u has xy as initial segment, and the first letter in v is y. Then

$$u = xy^{r_1}x^{k_1}y^{r_2}x^{k_2}y^{r_3}\cdots = x\cdot(x^{k_1})^{y^{-a_1}}(x^{k_2})^{y^{-a_2}}\ldots(x^{k_m})^{y^{-a_m}}\cdot y^{-\sum r_i},$$

$$a_k = \sum_{i=1}^{k} r_i,$$

$$v = y^{s_1}x^{t_1}y^{s_2}x^{t_2}y^{s_3}\cdots = (x^{t_1})^{y^{-b_1}}(x^{t_2})^{y^{-b_2}}\ldots(x^{t_q})^{y^{-b_q}}\cdot y^{-\sum s_i}, \quad b_k = \sum_{i=1}^{k} s_i.$$

Since the law $u \equiv v$ is balanced, $\sum r_i = \sum s_i =: n$, and the law can be written as

$$x \equiv w(x, y), \quad w \in \langle x^{y^{-1}}, x^{y^{-2}}, \ldots, x^{y^{-n}} \rangle.$$

If conjugate this by y^n we obtain the law (2) as required. □

The common properties of groups satisfying positive laws and Engel laws are studied in [12] and [13]. It is shown in [1] that finitely generated, locally graded groups satisfying an $\mathfrak{R}$-law is virtually nilpotent.

3 Milnor laws and varieties

The name of Milnor laws is due to F. Point [23] who introduced so called Milnor identities, which by G. Endimioni [4] are exactly those laws not satisfied in any variety of the form $\mathfrak{A}_p\mathfrak{A}$ for a prime p (see [24, Proposition 1.1]). So we suggest the following

Definition 3.1 A law $w \equiv 1$ is called the Milnor law if it is not satisfied in any variety of the form $\mathfrak{A}_p\mathfrak{A}$, with p a prime number.

A Milnor variety is the one which does not contain any of $\mathfrak{A}_p\mathfrak{A}$ as a subvariety.

By another words the Milnor law is a law which is not satisfied in any group $C_p \wr C$, with a prime p. By Proposition 2.6, the group $W := C_p \wr C$ has infinitely generated W', hence by Proposition 2.9, we conclude that W does not satisfy an $\mathfrak{R}$-law. Hence the $\mathfrak{R}$-laws are Milnor laws. By Examples 1 and 2 we have that positive laws and Engel laws are Milnor laws. Each pseudo-abelian law is a Milnor law because by definition it cannot be satisfied in non-abelian metabelian groups.

An algorithm which allows to check whether $w(x, y) \equiv 1$ is the Milnor law is given in [27].

Proposition 3.2 *Let G be a 2-generator free group in the variety defined by the law $w \equiv 1$. Then G'/G'' is finitely generated if and only if $w \equiv 1$ is the Milnor law.*

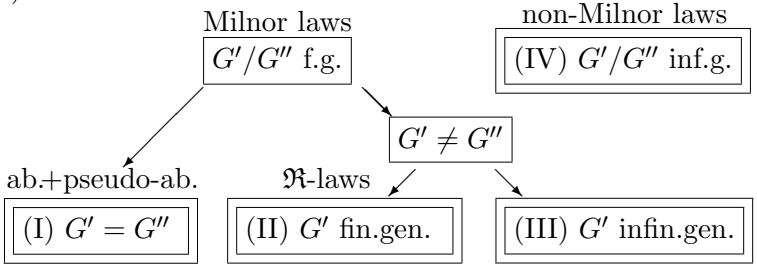
Proof If $w \equiv 1$ is the Milnor law then, by definition, it is not satisfied in any variety $\mathfrak{A}_p\mathfrak{A}$. Hence the metabelian variety generated by the group G/G'' does not contain any of $\mathfrak{A}_p\mathfrak{A}$ as a subvariety. By J. R. J. Groves [5, Theorem C(ii))], the group G/G'' is nilpotent-by-(finite exponent). Hence by result of A. I. Maltsev [15], G/G'' satisfies a positive law which is the $\mathfrak{R}$-law (Example 1). Then the commutator $(G/G'')' = G'/G''$ is finitely generated (see Proposition 2.9), as required.

Conversely, if $w \equiv 1$ is not the Milnor law, then for some p, $\mathrm{var}(G) \supseteq \mathfrak{A}_p\mathfrak{A}$ and hence also $\mathrm{var}(G/G'') \supseteq \mathfrak{A}_p\mathfrak{A}$. It follows that G/G'' has a quotient $C_p \wr C$ which has the commutator subgroup infinitely generated. Hence $(G/G'')' = (G'/G'')$ is not finitely generated, which finishes the proof. □

Let G be a 2-generator free group in the variety defined by the law $w \equiv 1$. The properties of G' and G'' allow us to classify the laws and the varieties.

Theorem 3.3 (Main Theorem) *Every law $w \equiv 1$ belongs to one of four disjoint classes* (I)–(IV)*, defined by properties of the commutator subgroup of the free* 2*-generator group in the variety defined by this law.*

The first three classes consist of the Milnor laws. All the laws in the class (II) *and some in* (I) *are the $\mathfrak{R}$-laws.*



By Proposition 3.2 we can split all laws for Milnor and non-Milnor laws accordingly to finite or infinite number of generators in G'/G''. It is clear that the classes on the picture are disjoint. We describe some details.

3.1 (I) ($G' = G''$)

Corollary 3.4 *By Proposition* 2.4 *the class* (I) *contains all abelian and all pseudo-abelian Milnor laws.*

Question 3 Each law of the form $[x, y] \equiv u(x, y)$, where $u(x, y) \in F''$ is in the class (I). Which words $u(x, y)$ define abelian and which pseudo-abelian varieties?

3.2 (II) ($G' \neq G''$, G' is finitely generated)

Corollary 3.5 *By Propositions* 2.4 *and* 2.9 *the class* (II) *contains all $\mathfrak{R}$-laws, which are not pseudo-abelian.*

Question 4 Are there $\mathfrak{R}$-laws in the class (II) different from the Engel laws and positive laws?

3.3 (III) ($G' \neq G''$, G' inf. gen., G'/G'' finit. gen.)

Corollary 3.6 *By Propositions* 2.4, 2.9 *and* 3.2 *the class* (III) *contains all Milnor laws which are not $\mathfrak{R}$-laws, and are not pseudo-abelian.*

Question 5 Are there examples of laws in the class (III)?

We can prove the following

Proposition 3.7 *Let $[x, y] \equiv u(x, y)$ with $u(x, y) \in F''$ be a pseudo-abelian law which is not an $\mathfrak{R}$-law. Then the law*

$$[x, y, y] \equiv [u(y, x),\ y] \tag{5}$$

is in the class (III)*.*

Proof Let G be the 2-generator relatively free group defined by the law (5). The law (5) is a Milnor law because each metabelian group satisfying (5) is a 2-Engel group, while none of $C_p \wr C$ is. So by Proposition 3.2, G'/G'' is finitely generated. The law (5) is not a $\mathfrak{R}$-law, because it is the consequence of the law which, by assumption, is not the $\mathfrak{R}$-law. So in view of Proposition 2.9, G' is infinitely generated. The law (5) is satisfied in the quaternion group Q_8, which is not abelian, hence by Proposition 2.4, $G' \neq G''$. Hence all conditions of the class (III) are satisfied which finishes the proof. □

Question 6 Are the pseudo-abelian laws of the form (1) $\mathfrak{R}$-laws?

References

[1] B. Bajorska, O. Macedońska and W. Tomaszewski, A defining property of virtually nilpotent groups, *Publ. Math. Debrecen* 81/3-4 (2012), 415–420.

[2] V. V. Belyaev and N. F. Sesekin, Free subsemigroups in solvable groups. (Russian) *Ural. Gos. Univ. Mat. Zap.* **12**(3) (1981), 13–18.

[3] R. G. Burns, O. Macedońska and Y. Medvedev, Groups satisfying semigroup laws, and nilpotent-by-Burnside varieties, *J. Algebra* **195** (1997), 510–525.

[4] G. Endimioni, On the locally finite p-groups in certain varieties of groups, *Quart. J. Math. Oxford Ser. (2)* **48** (1997), 169–178.

[5] J. R. J. Groves, Varieties of soluble groups and a dichotomy of P. Hall, *Bull. Austral. Math. Soc.* **5** (1971), 391–410.

[6] N. D. Gupta, Some group-laws equivalent to the commutative law. *Arch. Math. (Basel)* **17** (1966), 97–102.

[7] Y. K. Kim and A. H. Rhemtulla, Weak maximality condition and polycyclic groups, *Proc. Amer. Math. Soc.* **123** (1995), 711–714.

[8] Y. K. Kim and A. H. Rhemtulla, On locally graded groups, *Proceedings of the Third International Conference on Group Theory, Pusan, Korea 1994*, (Springer-Verlag, 1995), 189–197.

[9] P. A. Kozhevnikov, On nonfinitely based varieties of groups of large prime exponent, *Comm. Algebra* **40** (2012), 2628–2644.

[10] P. Kozhevnikov and O. Macedońska, On varieties of groups without positive laws, *Comm. Algebra* **30** (2002), 4331–4334.

[11] O. Macedońska and A. Storozhev, Varieties of t-groups, *Comm. Algebra* **25** (1997), 1589–1593.

[12] O. Macedońska, What do the Engel laws and positive laws have in common, *Fundam. Prikl. Mat.* **14**, (2008) 175–183.

[13] O. Macedońska and W. Tomaszewski, On Engel and positive laws, *Groups St Andrews 2009, Vol. 2* (C. M. Campbell et al., eds.), London Math. Soc. Lecture Note Ser. **388** (CUP, Cambridge 2011), 461–472.

[14] O. Macedońska, W. Tomaszewski, Group laws $[x, y^{-1}] = u(x, y)$ and varietal properties, *Comm. Algebra* **40** (2012), 4661–4667.

[15] A. I. Maltsev, Nilpotent semigroups, *Ivanov. Gos. Ped. Inst. Uc. Zap.* **4** (1953), 107–111 (Russian).

[16] G. A. Miller and H. C. Moreno, Non-abelian groups in which every subgroup is abelian, *Trans. Amer. Math. Soc.* **4** (1903), 398–404.

[17] J. Milnor, Growth of finitely generated solvable groups, *J. Diff. Geom.* **2** (1968), 447–449.

[18] B. H. Neumann, Ascending derived series. *Compositio Math.* **13** (1956), 47–64.

[19] B. H. Neumann and T. Taylor, Subsemigroups of nilpotent groups, *Proc. Roy. Soc. (Series A)* **274** (1963), 1–4.

[20] H. Neumann, *Varieties of groups*, (Springer-Verlag, 1967).

[21] A. Yu. Ol'shanskii, Varieties in which all finite groups are abelian, *Mat. Sb.* **126**(168), (1985), 59–82 (in Russian).
[22] A. Yu. Ol'shanskii, *Geometry of defining relations in groups*, Mathematics and its Applications (Soviet Series) **70** (Kluwer Academic Publishers, 1991).
[23] F. Point, Milnor identities, *Comm. Algebra* **24** (1996), 3725–3744.
[24] F. Point, Milnor property in finitely generated soluble groups, *Comm. Algebra* **31** (2003), 1475–1484.
[25] S. Rosset, A property of groups of non-exponential growth, *Proc. Amer. Math. Soc.* **54** (1976), 24–26.
[26] L. N. Shevrin and M. V. Volkov, Identities of semigroups, *Izv. Vyssh. Uchebn. Zaved. Mat.* **11** (1985), 3–47 (in Russian); English translation: *Soviet Math. (Iz. VUZ)* **29**(11) (1985), 1–64.
[27] W. Tomaszewski, How to recognize the Milnor laws?, manuscript.

CAPABLE p-GROUPS

ARTURO MAGIDIN* and ROBERT FITZGERALD MORSE†

*Mathematics Department, University of Louisiana at Lafayette, PO Box 41010, Lafayette, LA 70504-1010, USA
Email: magidin@member.ams.org

†Department of Electrical Engineering and Computer Science, University of Evansville, Evansville, IN 47722, USA
Email: rfmorse@evansville.edu

Abstract

A group G is said to be *capable* if it is isomorphic to a central quotient of some group K; that is, $G \cong K/Z(K)$. We survey three different approaches to the question of which groups are and which are not capable, with particular emphasis on p-groups. We discuss the historical connection with Philip Hall's program to classify p-groups up to isoclinism, and how the different approaches interconnect.

1 Introduction

A common exercise in a first course in group theory is to prove that if G is a group, and $G/Z(G)$ is cyclic, then G is abelian; in other words, no group can have an inner automorphism group that is nontrivial cyclic.

This may raise two questions in the mind of the reader. First, what groups can and cannot occur as an inner automorphism group? And second, does it matter?

The purpose of this survey is two-fold: we will review some of the recent advances in answering the first question, and put them into the context of the ideas and tools that have come to be used in studying it. And we will try to put those advances in context, and thereby answer the second question. We are particularly interested in the recent progress in determining which p-groups occur as inner automorphism groups, but we will discuss the problem in more generality when it does not take us too far afield.

Today, a group that is the central quotient of another group is said to be "capable," a term coined by J.K. Senior and Marshal Hall, Jr. [32]. Thus, a group G is *capable* if there exists a group K such that $G \cong K/Z(K)$; and is *not capable* or *incapable* otherwise. The exercise mentioned at the beginning shows that a nontrivial cyclic group is never capable.

Many papers that discuss capability invoke the great group theorist Philip Hall, who made the following oft-quoted observation in his paper on the classification of finite p-groups [34] (we change the name of the groups to match our nomenclature above):

> *The question of what conditions a group G must fulfil in order that it may be the central quotient group of another group K, $G \cong K/Z(K)$, is an interesting one. But while it is easy to write down a number of necessary conditions, it is not so easy to be sure that they are sufficient.*

However, the exact reason why this comment was made in a paper about classifying p-groups seems to be less well known. We want to take the opportunity afforded by this survey to recount that connection and make it explicit. Since the connection traces back to the Schur multiplier and the homological approach, and so connects the three strands we discuss below, we believe it makes sense to include it in this survey to provide a more complete picture of both the results that are known, and the reason why we might care about them.

Many different techniques have been brought to bear on the problem of determining which groups are capable. Rather than give a purely chronological presentation, we organize our presentation along three of the major 'flavors', as we see them.

The first part of the paper deals with what we are calling the classical approach, where we deal with the groups more or less directly. This approach includes the early success of Baer, who described precisely which groups that are direct sums of cyclic groups are capable, and recent generalizations; the techniques are also used to find structural conditions that a group must satisfy in order to be capable.

The second part of the paper deals with an approach that we have called the "modern approach", even though it harkens back to the first investigations into this problem which arose in the work of Schur, Speiser, and Brauer, among others, when studying extensions of groups. The modern period began with the introduction of the *epicenter of a group* by Beyl, Felgner, and Schmid [13] (the name is due to Burns and Ellis [20]), which is a characteristic central subgroup that measures the obstruction to the capability of G. Using this approach, the first major additions to Baer's characterization were obtained: the characterization of precisely which extraspecial p-groups are capable, and precisely which metacyclic groups are capable.

The third approach is the most recent one, and we explore it in the third part of the paper. This approach, which we call the "homological approach", brings together several homological functors (in particular the nonabelian tensor and nonabelian exterior square of a group) and connects to both the motivations for the classical approach of Philip Hall through the Schur multiplier, and to the epicenter of Beyl, Felgner, and Schmid. This final approach has brought new results which we will discuss below.

Within each approach we will attempt to give a more-or-less chronological overview; we note, however, that the three approaches overlap in time.

In the fourth and final part, we briefly mention some related concepts and generalizations of capability, as well as some of what we think are the more tractable open questions about capability of p-groups lying just beyond our current knowledge.

Part I: The classical approach

2 History and beginnings

2.1 Baer: capability of abelian groups

Perhaps the first investigation that dealt directly and specifically with what would come to be known as "capable groups" occurs in the work of Reinholt Baer. In a 1934 paper [5] in which he considers the automorphisms of a group extension, Baer determined those finitely generated abelian groups that can occur as a central quotient. Four years later, Baer published a pair of articles [6, 7]; the first one

considered the general construction of "groups with abelian central quotient group" (i.e., what we would now call nilpotent groups of class 2). The second paper considers the following related question: given two groups A and B, what are the necessary and sufficient conditions for the existence of a group G such that $A \cong Z(G)$ and $B \cong G/Z(G)$; and what are the necessary and sufficient conditions for such a group to be unique up to isomorphism? Baer answers the first question (existence) provided that B is a direct sum (finite or infinite) of cyclic groups; and solves the uniqueness problem under the further assumption that B is a finitely generated abelian group.

In the latter paper, Baer does not suggest any other consideration that might have led him to ask these questions, but rather says only that the existence and uniqueness problem "seem to be most elementary." However, it seems fair to guess that he was led to these questions by his previous work on constructions of nilpotent groups of class two and on determining the automorphism group of extensions.

Baer takes the time to state explicitly the corollary that describes which abelian groups that are direct sums of cyclic groups can be realized as a central quotient. To state the result, we need a bit of notation: given an abelian group G that is a direct sum of cyclic groups, $r(G,0)$ denotes the number of direct summands that are infinite in any decomposition of G into indecomposable direct summands; and for a prime p and positive integer i, $r(G,p^i)$ denotes the number of direct summands of order p^i in such a decomposition. These invariants can be computed directly from the structure of G (that is, without explicitly giving the decomposition) in well-known ways. We use G_{tor} to denote the torsion subgroup of an abelian group.

Baer's result can then be stated as follows:

Theorem 2.1 ([7], Corollary on p. 389) *A group G that is a direct sum of cyclic groups is isomorphic to the central quotient group of a suitable group K, $G \cong K/Z(K)$, if and only if*

(a) If $r(G,0) = 1$, then the orders of the elements in G_{tor} are not bounded;

(b) If G is a torsion group, and $r(G,p^i) = 1$, then G contains elements of order p^{i+1}.

In the case of a *finitely* generated abelian group, this becomes the oft-quoted result.

Corollary 2.2 *If G is a finitely generated abelian group, and we write*

$$G \cong C_{a_1} \oplus C_{a_2} \oplus \cdots \oplus C_{a_n}, \qquad a_1 \neq 1, \quad a_i | a_{i+1}, \quad i = 1, \ldots, n-1,$$

where C_m is cyclic of order m if $m > 0$ and C_0 is infinite cyclic, then G is isomorphic to the central quotient of a suitable group if, and only if, $n > 1$ and $a_{n-1} = a_n$.

Baer establishes the sufficiency constructively, in the sense that given a group G that satisfies the given conditions Baer uses the material in [6] to construct a witness K for which $G \cong K/Z(K)$. Baer's conditions would later be generalized in several directions: Philip Hall proved that the necessary condition on the orders of the invariant factors extends to regular p-groups, while one of us (AM) proved that the characterization extends from direct sums to nilpotent products in the small class case; see the discussion in Section 3.1 below.

Baer's papers do not seem to have made much of an impact, at least at the time of their publication and in the years that followed. The 1934 paper [5] is cited only five times in the next twenty years in the MathSciNet database, and most of its citations came after 2000. The other two papers fared even worse, with [6] cited a total of eight times, and [7] cited for the first time in 2002.

In fact, Baer's characterization of the finitely generated abelian groups that are capable was re-proven in [32], with no indication that the authors were aware that the result had already been proven over twenty years earlier.

2.2 P. Hall, J.K. Senior, and M. Hall, Jr.: *p*-groups, isoclinism and the groups of order 2^n, $n \leq 6$.

The next appearance of the concept bears directly on our own interest (p-groups), in the work of Philip Hall. In a 1940 paper [34], Philip Hall sets forth a general scheme which he hoped would help in the classification of p-groups. The paper was submitted to the German *J. für die Reine und Angewandte Mathematik* on July 1st, 1939, two months before the beginning of World War II, and the paper was published in 1940. The onset of the war delayed further development.

This was not Philip Hall's first foray into the problem of classifying p-groups; he had made seminal contributions eight years earlier in [33], where he introduced many of the invariants that are still used today. The main goal of the earlier paper, according to Hall, was to present a general theory for p-groups, in contrast to previous efforts that had been dominated by considering special classes of p-groups, such as those with "large" abelian subgroups, those that occur as Sylow subgroups of important insoluble groups like the symmetric and modular linear groups, etc. Philip Hall divides p-groups into *regular* and *irregular*, with regular groups being those in which the operations of commutation interacts well with what Hall terms the "order-power structure" of the group (the characteristic series of subgroups $\Omega_i = \langle\, x \in G \mid x^{p^i} = 1 \,\rangle$ and $\mho_i = \langle\, x^{p^i} \mid x \in G \,\rangle$). By restricting himself to regular groups, Philip Hall was able to show that "most of the classical theory of abelian groups ... is valid for the more general class of regular groups." In particular, Philip Hall proves that a finite regular p-group has invariants $\omega, \mu_1, \ldots, \mu_\omega$ such that $|G| = p^{\mu_1+\cdots+\mu_\omega}$, and such that G has elements $x_1, \ldots, x_\omega$ with x_i of order p^{μ_i}, with the property that any element of G can be expressed uniquely in the form $x_1^{a_1} \cdots x_\omega^{a_\omega}$ with $0 \leq a_i < p^{\mu_i}$. The quantities $\mu_1, \ldots, \mu_\omega$ are called the "type invariants" of the group. When the group is abelian, the type invariants correspond precisely to the invariant factors.

Parenthetically (top of page 33) Philip Hall remarks that he "hope[s], in a later paper, to deal with irregular groups on similar principles." Hall's hope was never realized, and perhaps the difficulties he encountered eventually led to [34].

In [34], Philip Hall changes strategies. He notes that it is not so difficult to construct all possible groups of order p^n once we know those of order p^{n-1}, by realizing them as extensions by a cyclic group of order p; but rather, that the difficulty lies in recognizing those that are isomorphic. Hall proposes instead to consider a coarser classification scheme, which he terms *isoclinism*. Capable groups make a remarkable appearance in these considerations, as we will see below. Philip Hall describes two ideas that guide his definition (bottom of p. 132 in [34]):

> *First, two groups G and K are considered to be the same in the abstract sense if they can be placed in the relation of isomorphism to one another: $G \cong K$. Clearly then, if we replace the relation of isomorphism by some weaker equivalence relation, $G \sim K$, we shall obtain a classification of all groups into mutually exclusive classes. The second idea is that the Abelian groups, at least those of finite order, do not need to be classified, since they may be regarded as completely known. Thus we shall choose the relation $G \sim K$ in such a way that the statement $G \sim 1$ means the same as: G is Abelian. For this reason, we call the equivalence relation $\sim$, on which the system of classification is based,* **isoclinism,** *and read the relation $G \sim K$ as saying that G and K are* **isoclinic** *groups. (The word isoclinic might perhaps be translated: gleich schief.)*

Explicitly, given two groups G and K, we say that G is isoclinic to K if and only if we have two isomorphisms, $\theta\colon G/Z(G) \to K/Z(K)$ and $\psi\colon [G,G] \to [K,K]$, which are compatible in the sense that $\psi([g_1,g_2]) = [\theta(g_1),\theta(g_2)]$. The attentive reader will no doubt note that this equation does not literally make sense, since θ is a map defined on the cosets of $Z(G)$ and with images that are cosets of $Z(K)$, whereas the image of ψ is supposed to be an element of K; however, it is easy to verify that in any group H, if $h_1Z(H) = h_1'Z(H)$, $h_2Z(H) = h_2'Z(H)$, then $[h_1,h_2] = [h_1',h_2']$; thus, we can view $[\theta(g_1),\theta(g_2)]$ as the commutation of any element of $\theta(g_1Z(G))$ with any element of $\theta(g_2Z(G))$, since the result will be independent of the choice of coset representative. Intuitively, we might say that G is isoclinic to K if G and K fail to be abelian in the same manner.

The concept of isoclinism fails to simplify the isomorphism problem for many classes of groups; for example, two simple groups are isomorphic if and only if they are isoclinic. But Philip Hall believed it held great promise for nonabelian p-groups. For one thing, in this setting both $[G,G]$ and $Z(G)$ must be proper nontrivial subgroups. For another, many of the standard invariants of p-groups remain invariants under isoclinism, notably the terms of the lower central series (and not just the commutator subgroup), as well as the class of the group.

The idea of using isoclinism to study p-groups seems to have arisen in large part during extended discussions between Philip Hall and James K. Senior about the classification of the groups of order 64; the background is related in the introduction to [32] and we refer the reader there. They believed they were within months of sending out the classification in the summer of 1939, but circumstances conspired against them. After many delays, Hall withdrew from the project permanently in 1959, and Marshall Hall, Jr. joined Senior to complete the effort (with the approval of Philip Hall), leading to the publication of [32].

Let us discuss in a bit more detail the ideas at play in the classification of p-groups up to isoclinism, and how capable groups enter into the picture; the connection of capable groups to the classification of p-groups is often mentioned, but seldom explained, so it seems worthwhile to expound on it here.

First we begin by recounting some observations made by Philip Hall. If H is a subgroup of G, then H is isoclinic to $HZ(G)$, and so in particular G is isoclinic to any of its *cocentral* subgroups (subgroups H such that $HZ(G) = G$); conversely, if $G/Z(G)$ is finite, then G is isoclinic to a subgroup H of itself if and only if $HZ(G) = G$.

Dually, if H is a normal subgroup of G that is not contained in $[G,G]$, then G/H is isoclinic to $G/(H\cap[G,G])$; so G/H is isoclinic to G if $H\cap[G,G]=\{1\}$, and the converse holds when $[G,G]$ is finite.

Philip Hall identifies an important class of groups, which he calls the *stem groups*: these are the groups G in which $Z(G)\subseteq[G,G]$. An important observation that Philip Hall makes is contained in the following theorem, which says that every equivalence class under isoclinism contains a stem group:

Theorem 2.3 ([34], discussion on p. 135) *Every group is isoclinic to a group G in which we have $Z(G)$ contained in $[G,G]$.*

To give the reader a feel for the kind of arguments employed by Hall (and more generally, those that go into investigations of this sort), we sketch the proof:

Sketch. Given an arbitrary group K, pick a generating set S; let A be the free abelian group generated by S, and let H be the subgroup of $K\times A$ generated by the pairs (s,s) with $s\in S$. It is then easy to verify that K is isoclinic to $K\times A$, and the latter to H. Moreover, $H^{\mathrm{ab}}\cong A$, and

$$\frac{Z(H)}{Z(H)\cap[H,H]}\cong\frac{Z(H)[H,H]}{[H,H]};$$

the latter corresponds to a subgroup of A, so $Z(H)/(Z(H)\cap[H,H])$ is free abelian (possibly trivial), hence projective. Therefore, $Z(H)\cong(Z(H)\cap[H,H])\oplus L$ for some complement L. Then $L\cap[H,H]=\{1\}$, and H/L is isoclinic to H. Finally, $Z(H/L)=Z(H)/L\subseteq[H/L,H/L]=L[H,H]/L$. Thus, $G=H/L$ has the desired properties. □

Philip Hall refers to the stem groups (those G with $Z(G)\subseteq[G,G]$) in a given isoclinism family[1] as the *stem* of the family. Two stem groups in the same family are necessarily of the same order, and if K is any group, then the order of the stem groups in the isoclinism family of K is given by $[K:Z(K)]\big|Z(K)\cap[K,K]\big|$. For finite groups, the stem groups of an isoclinism family are *precisely* all the groups in the family that are of minimal order (within the family). Much of the effort in describing an isoclinism family goes into determining the stem groups of the family.

In the case of isoclinism families of p-groups, the stem groups will themselves be p-groups. If the stem groups of a given family have order p^ρ, then ρ is called the *rank* of the family. If we determine all stem groups of a given order, we will determine all isoclinism families of that rank.

But how can we determine the stem groups of a given rank? It is here that capable groups make an important appearance. Philip Hall sketches the ideas in [34], and the details are fleshed out by M. Hall and Senior in [32].

The question Philip Hall asks is: suppose that G and $\widehat{G}$ are groups with isomorphic central quotients, $G/Z(G)\cong\widehat{G}/Z(\widehat{G})\cong H$. What relations, if any, hold between the isoclinism family of G and that of $\widehat{G}$? What Philip Hall proves is that there is a most general stem group with the same central quotient as G (and $\widehat{G}$), and such that each

[1] We use "family" to describe an equivalence class under isoclinism, because the term "class" may be confused with the nilpotency class of a group.

of G and $\widehat{G}$ are quotients of this group. What follows is the argument as presented in [32].

Let $h_1, \ldots, h_r$ be generators of H, and let F be the free group on $x_1, \ldots, x_r$. Let K be the kernel of the map $F \to H$ induced by $x_i \mapsto h_i$. Let $N_0 = [K, F]$ and $N_1 = K \cap [F, F]$; then K/N_0 is abelian, and $K/N_1 \cong [F, F]K/[F, F]$ is a subgroup of the free abelian group $F/[F, F]$ and so is itself free abelian, and hence K/N_0 is the direct product of N_1/N_0 and a free abelian group of the form L/N_0 for a suitable subgroup L of F. The group N_1/N_0 is none other than the Schur multiplier of H, and depends only on H and not on the choice of presentation.

Recall now that G is a group with $G/Z(G) \cong H$. Let $g_1, \ldots, g_r$ be elements of G that map to $h_1, \ldots, h_r$ under the isomorphism, and let $\overline{G} = \langle g_1, \ldots, g_r \rangle$. Then $G = \overline{G}Z(G)$, hence $G \sim \overline{G}$, so $\overline{G}$ is also a witness to the capability of H. Let M be the kernel of the map $F \to \overline{G}$ induced by $x_i \longmapsto g_i$. Then $M \subseteq K$, and K/M is the center of F/M, hence $N_0 = [K, F] \leq M$, and K/N_0 is the center of F/N_0.

Let $N = M \cap N_1$, so that $N_0 \leq N \leq N_1$. Since $M \leq K$, then $M \cap [F, F] = M \cap N_1 = N$, so the subgroup M/N of F/N intersects $[F, F]/N$ trivially. Therefore, $F/M \sim F/N$, and so $G \sim F/N$, and $K/N = Z(F/N)$.

The conclusion of the above is that if G is a group such that $G/Z(G) \cong H$, then the isoclinism family to which G belongs has at least one representative of the form F/N, where $N_0 \leq N \leq N_1$ and K/N is the center of F/N.

By selecting $N = N_0$ we obtain a group that belongs to a family $\mathcal{M}(H)$ that is maximal, in the sense that any group with central quotient isomorphic to H is isoclinic to a suitable quotient of a group from $\mathcal{M}(H)$. The family $\mathcal{M}(H)$ is called the *maximal family associated to the given H*. In particular, we can obtain the stem groups of every family $\mathcal{M}$ that has central quotients isomorphic to H by considering F/N_0 and suitable quotients thereof.

The question that remains is whether we have a way of recognizing that two quotients F/N and F/M are isoclinic. We have:

Theorem 2.4 ([32], Theorem 2.2) *The outer automorphisms of H are represented in a natural way by automorphisms of the Schur multiplier of H, N_1/N_0. The subgroups N/N_0 such that $N_0 \leq N \leq N_1$ and $K/N = Z(F/N)$ are permuted among themselves by this representation; and $F/N \sim F/M$ if and only if N and M belong to the same orbit of the action.*

Say we want to find all isoclinism families of rank n; if we know all capable groups of order p^k, $k < n$, then we can proceed by considering each such capable group H, constructing the corresponding maximal family $\mathcal{M}(H)$, and determining the suitable quotients F/N that lead to stem groups of order p^n. Then we use Theorem 2.4 to ensure we have not listed any isoclinism family more than once.

Philip Hall shows the usefulness of these ideas by obtaining a classification up to isoclinism of the p-groups of order at most p^5, with $p > 3$ (the restriction simplifies calculations by ensuring all groups are regular; some minor but potentially obfuscating difficulties arise when dealing with $p = 3$, while $p = 2$ was being dealt with in the work with Senior). Philip Hall does not need the approach through capable groups and maximal families to deal with the families of smallest rank, since we already know the structure of p-groups of small order: the family of rank 0 corresponds to

abelian groups; there are no families of rank 1 or 2 (since the stem groups would be abelian); the two nonabelian groups of order p^3 are isoclinic and yield a single family; and there is a single family of rank 4, corresponding to a stem group that is of order p^4 and maximal class (Philip Hall characterizes the groups in this isoclinism class as those that have an abelian subgroup of order p, commutator subgroup of order p^2, and third term of the lower central series of order p).

But in order to deal with the family of rank 5, Philip Hall uses the approach outlined above; he does not go into all the details (in particular, he does not derive the necessary and sufficient condition for $F/N \sim F/M$ to hold) because he does not need them for the application in question and presumably because he expected them to appear in the planned work with Senior. Philip Hall determines that the groups of order at most p^5, $p > 3$, fall into ten families: one each of ranks zero, three, and four; and seven of rank five.

These ideas were also used to great effect in [32], where all 267 groups of order 64 are described by generators and relations, along with all groups of orders 2, 4, 8, 16, and 32. A lot of information is provided for each group; e.g., the nilpotency class, the isoclinism family to which it belongs, the number of elements of a given order, the order of the automorphism groups, and the lattice of normal subgroups.

The ideas presented by Philip Hall in [34] play a central role in the exposition given in [32], though in the latter they are further refined and expanded. The groups of order 2^n, $n \leq 6$, fall into 27 isoclinism families: one each of ranks zero, three, and four; five of rank five; and nineteen of rank six.

Philip Hall's isoclinism program was later used to classify the groups of order p^6 for odd prime p; they fall into 43 isoclinism families [40]. It was also used to classify the groups of order 2^7, which fall into 115 isoclinism families [41].

2.3 Beyond Hall

As far as we are aware, there was no sustained effort to extend the work of Philip Hall, Marshall Hall Jr., and Senior beyond that mentioned above. Although the groups of order p^7, $p > 2$, and the groups of orders 2^n with $n = 8, 9$ have been described [56, 57, 11], and those of order 2^{10} have been counted [11], the methods used do not follow the outline of Philip Hall's isoclinism program, so there was no need to determine capable groups and maximal families to obtain these results.

This may help explain why work on capable groups often mentions Philip Hall's dictum that determining the capable p-groups is interesting and important for their classification, but more recent work that focuses on understanding p-groups does not seem to touch on the problem at all.

Still, results on capable groups that proceed by working directly with the groups (and do not involve the homological machinery we will discuss in the next two sections) continue to appear, e.g., [36, 38, 39, 47, 49]. We discuss some of these results in the next section.

3 Further results

We call the approach above "classical": though the Schur multiplier shows up in the considerations of Philip Hall, the actual determination of which groups are capable

or, more often, necessary conditions for a group to be capable, take a very hands-on approach.

In this section we discuss some further developments that proceed along those same lines: by working directly with the groups, without homological constructions or other high-power results.

3.1 Orders of elements and beyond Baer.

Baer proves his results in [7] constructively; the sufficiency of the conditions is established by explicitly constructing the required groups, and the necessity by directly analyzing the consequences that the required relations would imply in a group whose central quotient is the given group. Some remarks made by Philip Hall in [34] suggest that Hall was unaware of Baer's results, and the same seems to be true about Senior and M. Hall: Corollary 2.2 is proven *de nuovo* in [32], without any reference to Baer. The proof is different from Baer's, though again it is very hands on. It is based on the lemma below; we give its proof and a sketch of how it was used to establish Corollary 2.2, again to give the reader a feel for the type of arguments that are typical of the classical approach.

Lemma 3.1 ([32], Lemma 3.1) *Let G be generated by $x_1, \ldots, x_r$, and suppose that there is an element $u \neq 1$ of G such that $u \in \langle x_i \rangle$ for each i; then G is not capable.*

Proof Suppose K is such that $K/N \cong G$, with $N \subseteq Z(K)$. If $k \in K$ is such that k maps to u, and $k_1, \ldots, k_r \in K$ map to $x_1, \ldots, x_r$, respectively, then there exist integers $m_1, \ldots, m_r$ such that $k \equiv k_i^{m_i} \pmod{Z(K)}$. Therefore, k commutes with each k_i; for example, there exists $z \in Z(K)$ such that $k = k_1^{m_1} z$, hence $k_1 k = k_1^{m_1+1} z = z k_1 k_1^{m_1} = k k_1$. Since $K = \langle Z(K), k_1, \ldots, k_r \rangle$ it follows that $k \in Z(K)$, hence $N \neq Z(K)$, since $kN \neq N$. □

From here one can prove Corollary 2.2: let $G = C_{r_1} \times \cdots \times C_{r_k}$, where C_{r_j} is cyclic of order r_j generated by x_j, $r_1 \neq 1$, and $r_i | r_{i+1}$. If $r_{k-1} < r_k$, let $y_i = x_i x_k$ for $i < k$ and $y_k = x_k$; then H is generated by $y_1, \ldots, y_k$, $u = y_k^{r_{k-1}} \neq 1$, and $u = y_i^{r_{k-1}}$ for each i; by the previous lemma, it follows that G is not capable. Conversely, if G satisfies the conditions of Corollary 2.2, one can define a group of class 2 (which will lie in the maximal family associated to G) generated by $y_1, \ldots, y_r$ and subject to the relations $[y_i, y_j]^{r_i} = 1$ with $i < j$, together with the relations that make commutators central. The center of this group is then easily seen to be generated by the commutators and by $y_k^{r_{k-1}}$, so that $K/Z(K) \cong C_{r_1} \oplus \cdots \oplus C_{r_{k-1}} \oplus C_{r_{k-1}}$ which is isomorphic to G when $r_{k-1} = r_k$. See also Theorem 3.4 below.

The same ideas can be used to establish necessary conditions on the orders of elements of a generating set of a capable group. Immediately after noting that necessary conditions are not difficult to come by, but sufficient conditions are harder, Philip Hall illustrates his remark by mentioning the following generalization of the necessary half of Baer's Theorem:

Proposition 3.2 ([34], bottom of p. 137) *If G is a regular p-group, and G is capable, then its two largest type invariants are the same.*

This can be shown using a similar argument as that used above, augmented with some commutator calculus for regular groups.

Philip Hall notes that this condition is sufficient for regular p-groups of order less than p^5 with p odd, which suffices for his application in [34]. However, the condition is not sufficient in general for regular groups, and in fact already fails for groups of order p^6; that is, just after the orders considered by Philip Hall (perhaps it was this failure that led Hall to consider only groups of order at most p^5). For example, the extraspecial group of order p^5 and exponent p, $E_5 = \langle x, y\rangle \circ \langle z, w\rangle$ is not capable, where $\langle x, y\rangle \cong \langle z, w\rangle$ are nonabelian groups of order p^3 and exponent p (p an odd prime), and $\circ$ represents the central product (the quotient of their direct product obtained by identifying their isomorphic centers). To see that E_5 is not capable (even though it is regular and its four type invariants are equal), suppose that H is a group and $N \subseteq Z(H)$ is such that $H/N \cong E_5$; we show that $N \neq Z(H)$. If h_x, h_y, h_z, h_w map onto x, y, z, w, respectively, in the isomorphism, then $[h_x, h_z]$, $[h_x, h_w]$, $[h_y, h_z]$ and $[h_y, x_w]$ are all central in H. Since H is of class 3, the Hall-Witt identity yields that in H, $[h_x, h_y, h_z] = 1$ and so $[h_x, h_y]$ commutes with h_z; similar calculations show that $[h_x, h_y]$ also commutes with h_w, and that $[h_z, h_w]$ commutes with both h_x and h_y. But since $[h_x, h_y][h_w, h_z]$ is also central, we also obtain $1 = \big[[h_x, h_y][h_w, h_z], h_x\big] = [h_x, h_y, h_x][h_w, h_z, h_x] = [h_x, h_y, h_x]$, so $[h_x, h_y]$ also commutes with h_x. Symmetrically, it commutes with h_y, and so we conclude that $[h_x, h_y] \in Z(H) \setminus N$, so E_5 is not capable.

The necessary condition in terms of type invariants does not make sense for irregular groups (which do not have type invariants); however, one of us (AM) proved that there is a generalization for arbitrary p-groups, whose proof only requires commutator calculus.

Proposition 3.3 ([47], Lemma 3.12) *Let G be a p-group of class c, with $c \geq 1$ and p a prime. Furthermore, let $\{x_1, \ldots, x_r\}$ be a generating set for G with x_i of order p^{a_i}, where $a_1 \leq a_2 \leq \cdots \leq a_r$. If G is capable, then $r > 1$ and*

$$a_r \leq a_{r-1} + \left\lfloor \frac{c-1}{p-1} \right\rfloor,$$

where $\lfloor x \rfloor$ is the floor of x, i.e., largest integer less than or equal to x.

If $c < p$, often called the "small class case", the group will necessarily be regular, and the theorem yields Philip Hall's condition on the type invariants. When $c = 1$, the condition readily yields the necessity clause of Baer's theorem for finite abelian groups.

The proof requires only some careful calculations with commutators: if H is a p-group of class $c+1$, y and z are elements of H, and a is a positive integer with the property that $[z, y^{p^i}, z] = [z, y^{p^i}, y] = 1$ for all $i \geq a$, then $[z^{p^N}, y] = [z, y]^{p^N} = [z, y^{p^N}]$ for all $N \geq a + \lfloor (c-1)/(p-1) \rfloor$ (see Theorem 3.9 in [47]). From this it follows that if H is a p-group of class $c+1$, and the elements $y_1, \ldots, y_r$ have images that generate $H/Z(H)$ and have order $p^{a_1}, \ldots, p^{a_r}$, respectively, then $y_r^{p^\ell} \in Z(H)$, where $\ell = a_{r-1} + \lfloor (c-1)/(p-1) \rfloor$. The necessary condition then follows immediately.

The inequality is best possible: for $p = 2$, the dihedral groups provide examples where we have equality. For odd primes, a modification of a construction of East-

erfield [22] provides the necessary examples; see [46]. However, the condition is not sufficient in general, as witnessed again by the extraspecial group E_5.

How far can we generalize Baer's Theorem for finitely generated abelian groups? One may attempt to generalize Baer's Theorem by noting that the direct sum of abelian groups is their coproduct in the category of abelian groups. Any variety of groups (in the sense of universal algebra: a class of groups closed under subgroups, homomorphic images, and arbitrary direct products) has a coproduct; in the case of nilpotent groups of class at most c, the coproduct is the "c-nilpotent product", originally introduced by Golovin [31] in a slightly more general setting. At least for values of c that are no larger than p, we do get the generalization of Baer's Theorem. Recall that if $A_1, \ldots, A_k$ are nilpotent groups of class at most c, then their c-nilpotent product $A_1 \amalg^{\mathfrak{N}_c} \cdots \amalg^{\mathfrak{N}_c} A_k$ is defined to be

$$A_1 \amalg^{\mathfrak{N}_c} \cdots \amalg^{\mathfrak{N}_c} A_k = \frac{A_1 * \cdots * A_k}{(A_1 * \cdots * A_k)_{c+1}},$$

where $*$ denotes the free product, and B_{c+1} is the $(c+1)$st term of the lower central series of a group B. For instance, the explicit group that witnesses the capability of a direct sum of cyclic groups with $a_{r-1} = a_r$ mentioned above is the 2-nilpotent product of the same cyclic groups. We have:

Theorem 3.4 ([47], Theorem 4.4; [49], Theorem 3.11) *Let p be a prime, and let $C_1, \ldots, C_r$ be cyclic groups generated by $x_1, \ldots, x_r$, respectively; assume that the order of x_i is p^{a_i}, and that $a_1 \leq \cdots \leq a_r$. If $c \leq p$ and G is the c-nilpotent product of the C_i, $C_1 \amalg^{\mathfrak{N}_c} \cdots \amalg^{\mathfrak{N}_c} C_r$, then G is capable if and only if $r > 1$ and $a_r \leq a_{r-1} + \lfloor (c-1)/(p-1) \rfloor$.*

Necessity follows from Proposition 3.3; sufficiency can be established constructively: the given nilpotent product is the central quotient of the nilpotent product "one class up", that is, of $C_1 \amalg^{\mathfrak{N}_{c+1}} \cdots \amalg^{\mathfrak{N}_{c+1}} C_r$. The proof in [47] uses a normal form that was found by R.R. Struik [65, 66], together the same commutator calculus that is used to establish Proposition 3.3.

We do not know whether the restriction on c can be dropped, though one of us (AM) conjectures that this is indeed the case. The generalization does have one weakness that Baer's Theorem does not: whereas in the abelian case ($c = 1$) all finite p-groups can be expressed as a coproduct of cyclic groups, the same is not true for $c > 1$. Of course, the universal property of the coproduct guarantees that any finite p-group of class c is a quotient of a c-nilpotent product of cyclic groups, so some progress can be made.

3.2 Index of the center

Another consequence of capability that can be established through elementary methods is the following result from Isaacs:

Theorem 3.5 ([39]) *There exists a function $t(n)$ defined on the natural numbers such that if G is finite and capable, then $|G : Z(G)| \leq t(|G'|)$.*

Isaacs does not provide an estimate for the function $t(n)$. Podoski and Szegedy take up the matter in [58], providing an upper bound and eliminate the need that G be finite.

Theorem 3.6 ([58]) *If G is capable and $|G'| = n$, then $|G : Z(G)| \leq n^{c \log_2 n}$ with $c = 2$.*

They conjecture that the bound on the central index might be improved to

$$|G : Z(G)| \leq n^{\frac{1}{2} \log_2 n + c_2}$$

for some constant c_2.

A classical result of Schur states for any group G, if $|G : Z(G)|$ is finite then $|G'|$ is finite. The converse of this result is not true in general. However Philip Hall shows that if $|G'|$ is finite then $|G : Z_2(G)|$ is finite, where $Z_2(G)$ is the second center of G. Since for any group G the quotient $G/Z(G)$ is capable, we can apply Theorem 3.6 whenever $|(G/Z(G))'|$ is finite and obtain a bound for $|G : Z_2(G)|$. Suppose $|(G/Z(G))'| = n$. Then by Theorem 3.6

$$|G/Z(G) : Z(G/Z(G))| = |G : Z_2(G)| \leq n^{c \log_2 n}.$$

Observing $|(G/Z(G))'| = |G' : G' \cap Z(G)|$ we obtain the following result.

Theorem 3.7 ([58]) *Let G be an arbitrary group. If $|G' : G' \cap Z(G)| = n$ then $|G : Z_2(G)| \leq n^{c \log_2 n}$ with $c = 2$.*

In a second paper [59] Podoski and Szegedy look to bound $|G : Z(G)|$ of a finite capable group by the size of certain generating sets. This can lead to improvements to Theorem 3.6. Denote by $\mathrm{d}(G)$ the minimum number of generators of G, and by $\mathrm{rk}(G)$ the rank of G, which is the minimal number such that each subgroup of G is generated by $\mathrm{rk}(G)$ elements. It is clear $\mathrm{d}(G) \leq \mathrm{rk}(G)$. Since any subgroup of a finite group G can be generated by at most $\log_2(|G|)$ elements this is an upper bound on $\mathrm{rk}(G)$.

Theorem 3.8 ([59]) *If G is a finite capable group and $\mathrm{rk}(G') = r$, then*

$$|G : Z(G)| \leq |G'|^{4r}.$$

Placing conditions on the structure of $Z(G)$ or $[G, G]$ provides further leverage to improve the bounds on the index of $Z(G)$ in G. All centerless groups are capable and in this case we can improve Theorem 3.8.

Theorem 3.9 ([59]) *If G is a finite group with $Z(G) = 1$ and $d = \mathrm{d}(G')$, then $|G| \leq |G'|^{d+1}$.*

If G' is cyclic, Theorem 3.8 can be further improved:

Theorem 3.10 ([58]) *If G is a finite capable group whose commutator subgroup G' is cyclic, then $|G : Z(G)| \leq |G'|^2$.*

For certain capable p-groups of nilpotency class two it is possible to obtain exact bounds on the index of the center. For example, by considering the bilinear map

$$(G/Z(G)) \times (G/Z(G)) \to [G,G].$$

Heineken proved the following:

Theorem 3.11 ([36], Prop. 3) *If G is a finite capable group with $[G,G] \subseteq Z(G)$ and $C_p \times C_p \cong [G,G]$, then $p^2 < |G/Z(G)| < p^6$.*

For groups of nilpotency class two and exponent p, a much better result is possible:

Theorem 3.12 (Heineken and Nikolova [38]) *Assume that G is a finite group of exponent p, $Z(G) = [G,G]$, and that G is capable. If $Z(G)$ is of rank k, then $G/Z(G)$ is of rank at most $2k + \binom{k}{2}$.*

It is worth noting that the condition $Z(G) = [G,G]$ is not an obstruction to the applicability of this result: if G is a p-group of class exactly 2 and exponent p, then it is easy to show that G can be expressed as $G = K \times A$, where A is a elementary abelian p group of rank ≥ 0 and K is a group such that $Z(K) = [K,K]$. It is straightforward to show that G is capable if and only if K is capable. So the problem of determining the capability of any group of class two and exponent p can be reduced to one for a group whose center is equal to its commutator subgroup.

3.3 A sufficient condition

As we have repeated a number of times, it seems to be difficult to obtain sufficient conditions for capability. The results summarized above hint at this: there are very few conditions guaranteeing capability.

One class of groups that seems like a good candidate for a complete characterization is the class of p-groups of class two and exponent p. Combining an argument of Isaacs (Lemma 2.1 of [39]) and one of Heineken and Nikolova (mentioned *en passant* in the proof of Theorem 1 in [38]), one can show that if G is a finite capable group, then we can always find a finite group H that acts as witness for the capability of G (that is, a finite H such that $G \cong H/Z(H)$) and moreover, if G is generated by elements $g_1, \ldots, g_n$, then one may choose such an H to be generated by elements $h_1, \ldots, h_n$, with h_i mapping to g_i under the isomorphism $G \cong H/Z(H)$, and with h_i of the same order as g_i (see Theorem 3.1 in [51]).

This means that if G is a capable p-group of class 2 and odd exponent p, then there necessarily exists a witness H to that capability which is a finite p-group of class 3 and generated by elements of exponent p. These groups have a straightforward commutator structure (which is an abelian group of exponent p with basis given by the basic commutators on the generators), and so there is a lot of structure with which to perform computations. One can even write down a "canonical witness" for the capability of G in the following sense: if we let N be the kernel of the map from the relatively free group of class 2, exponent p, and rank r, onto G, then we can view N as a subgroup of the commutator subgroup of F, the 3-nilpotent product of r cyclic groups of order p. Then G is capable if and only if $F/[N,F]$ is a witness to the capability of G.

By relating the size of N and that of $[N,F]$, one can test to see whether there can exist an M, $N \subsetneq M$, with $[N,F] = [M,F]$; the existence of such an M is equivalent to the incapability of G. This gives a nice counterpart to Theorem 3.12:

Theorem 3.13 ([51], Theorem 5.28) *Let G be a p-group of class two and exponent p, where p is an odd prime, and assume that $Z(G) = [G,G]$. Let $\mathrm{rk}(G^{\mathrm{ab}}) = n$ and $\mathrm{rk}([G,G]) = m$. Define a function f on positive integers by $f(n) = \binom{r}{3} + \binom{s}{2}$, if n is a positive integer and $n = \binom{r}{2} + s$, $0 \leq s < r$. If $f\left(\binom{n}{2} - m + 1\right) < n$, then G is capable.*

In essence, Theorem 3.13 says that if the commutator subgroup of G is "big enough", then G will necessarily be capable; on the other hand, Theorem 3.12 can be thought of as saying that if G is capable, then the commutator subgroup of G cannot be "too small." The two results, together with some classification work of Brahana [16], suffice to characterize the capable groups of class at most 2 and exponent p with $G/Z(G)$ of rank at most 5.

Corollary 3.14 ([51], Theorem 6.1) *Let p be an odd prime, and let G be a p-group of class at most 2 and exponent p, and assume that $G/Z(G)$ is of rank at most 5. Then G is one and only one of:*

(i) Cyclic and nontrivial;

(ii) A nontrivial central product AB with $[A,A] \cap [B,B] \cong C_p$; or

(iii) Capable.

Unfortunately, as n grows the gap between the necessary condition of Theorem 3.12 and the sufficient condition of Theorem 3.13 also grows. However, the two results suggest that a full characterization for the class of groups of class 2 and prime exponent may be within reach using current ideas and techniques.

Part II: The modern approach

4 The Schur multiplier, extensions, and the epicenter

The approach that we dub "modern" originated in the work of Beyl, Felgner, and Schmid [13], and the introduction of the "precise center" or "epicenter" $Z^*(G)$ of a group G. This led to the determination of exactly which metacyclic groups and which extraspecial p-groups are capable, the first classes to be completely dealt with since Baer's result on finitely generated abelian groups. In this section, we discuss these developments, starting with the definition of the epicenter and its connection with the Ganea map.

The subject of capable groups seems to have lain more or less dormant for several years after the appearance of [32]. The next major development came in 1979, with a paper of Beyl, Felgner, and Schmid [13]. In a sense, this paper harkens back to ideas of Schur [61] and Speisel [64] on representation theory mentioned by Philip Hall in [34].

Beyl, Felgner, and Schmid introduced a central subgroup of a given group G, which they denote by $Z^*(G)$; this subgroup was sometimes called the "precise center" of G,

but would later be christened *the epicenter of* G by Burns and Ellis [20], the name by which it is generally known today.[2]

The first definition of $Z^*(G)$ comes from considering group extensions. We say that an extension of G is a pair, (E,φ), with $\varphi\colon E\to G$ a surjective group morphism. If (F,π) is another extension of G, then a homomorphism $f\colon F\to E$ is said to be a homomorphism *over* G if $\varphi f=\pi$. And we say that (E,φ) is a *central extension* of G if $\ker(\varphi)\subseteq Z(E)$. Beyl, Felgner, and Schmid define the epicenter as: $Z^*(G)=\bigcap\{\varphi(Z(E))\mid (E,\varphi)\text{ is a central extension of }G\}$. The connection with capability is the following:

Theorem 4.1 ([13], Corollary 2.2) *The epicenter $Z^*(G)$ is the intersection of all normal subgroups N of G such that G/N is capable. In particular, $Z^*(G)$ is the smallest subgroup of $Z(G)$ such that $G/Z^*(G)$ is capable.*

Corollary 4.2 ([13], Corollary 2.3) *G is capable if and only if $Z^*(G)=\{1\}$.*

Thus, we can view $Z^*(G)$ as the obstruction to the capability of G, or as a measure of how far G is from being capable. This also connects the study of capability with the study of *unicentrality*, introduced by Evens [28]. Evens was interested in computing the Schur multiplier of a semidirect product $A\rtimes K$, with A abelian, in terms of the Schur multipliers of A and of K. Evens defines a group G to be *unicentral* if and only if for every central extension (E,φ) of G, $\varphi(Z(E))=Z(G)$. Thus, we see that, in terms of the epicenter, a group G is unicentral if and only if $Z^*(G)=Z(G)$; that is, unicentral groups are at the opposite end of the spectrum from capable groups.

The epicenter had been looked at elsewhere; for example, it occurs in the work of Read [28].

As defined above, the epicenter seems difficult to compute; inspired by the result that G is unicentral if and only if the canonical map $M(G)\to M(G/Z(G))$ is one-to-one (where $M(K)$ is the Schur multiplier of K), Beyl, Felgner, and Schmid looked for an alternate description of the epicenter. Let $1\to N\to G\to Q\to 1$ be a central extension. Ganea [30] associates to this extension the exact sequence

$$N\otimes G/G'\to M(G)\to M(Q)\to N\to G/G'\to Q/Q'\to 1.$$

The map $N\otimes G/G'\to M(G)$ is called the Ganea map associated with N. The following theorem connects the epicenter with the Ganea map.

Theorem 4.3 ([13], Thm 5.1) *Let γ_G be the Ganea map associated to $Z(G)$. Then $Z^*(G)$ is the left kernel of γ_G.*

This characterization provides a more computational way to find the epicenter. With this in hand, in [13] a number of interesting results regarding capable groups are proven: that reduced and subdirect products of capable groups are capable ([13], Prop. 6.1); that if $H\triangleleft G$ has finite index, both H and G/H are capable, and the transfer map $G\to H/[H,H]$ is onto, then G is capable ([13], Cor. 6.5); a simple proof of Baer's Theorem for finitely generated abelian groups is given, and extended

[2] According to Ellis, they came up with the name when, while they were working on these ideas in the mid-90s, there was an earthquake with epicenter in the Irish Sea.

to show that a torsion-free abelian group (whether or not it is a direct sum of cyclic groups) is capable if and only if its rank is not 1 ([13], Prop. 7.5).

Beyl, Felgner, and Schmid also characterize exactly which extraspecial p-groups and which metacyclic groups are capable. Recall that a p-group G is *extraspecial* if and only if $Z(G) = [G,G] = \Phi(G) \cong C_p$, where $\Phi(G)$ is the Frattini subgroup of G. Extraspecial p-groups arise as iterated central products of nonabelian groups of order p^3. We have:

Theorem 4.4 ([13], Prop. 8.1 and Cor. 8.2) *If G is the central product of nilpotent groups G_i with $[G_i,G_i] = Z(G_i)$ $(i \in I)$, and $|I| \geq 2$, then G is unicentral. In particular, an extraspecial p-group is capable if and only if it is either dihedral of order 8, or of order p^3 and exponent $p > 2$.*

As for metacyclic groups, if G has a cyclic normal subgroup of order m with quotient cyclic of order n, we can present G as

$$G(m,n,r,s) = \langle\, x, y \mid x^m = 1,\ y^{-1}xy = x^r,\ y^n = x^s \,\rangle,$$

where r and s are positive integers such that $\gcd(m, 1 + r + \cdots + r^{n-1}) \equiv 0 \pmod{s}$, and $r^n \equiv 1 \pmod{m}$. With that notation, the following result is proven:

Theorem 4.5 ([13], Prop. 9.2 and Cor. 9.3) *Let $G = G(m,n,r,s)$, and let $\overline{n}$ be the smallest positive divisor of n such that $1 + r + \cdots + r^{\overline{n}-1} \equiv 0 \pmod{s}$. Then $Z^*(G)$ is the cyclic group of order $nm/\overline{n}s$ generated by $y^{\overline{n}}$. In particular, G is capable if and only if $s = m$ and $n = \overline{n}$.*

The capable groups in other families would be characterized later by combining the epicenter and the approach through homological algebra that we will describe in Part III.

5 Results using the epicenter: subgroup structure

The epicenter has proven useful in studying the normal subgroup structure of capable groups. For example, Shahriari considered in [62] the situation in which we have a group G that has a known incapable subgroup Q and whether one can establish that certain nontrivial subgroups of the epicenter $Z^*(Q)$ of Q must also be subgroups of the epicenter $Z^*(G)$ of G, thus proving that G will also not be capable. Shahriari proves:

Proposition 5.1 ([62], Prop. 3.2) *Let G be a finite group. If $G = QC_Q(G)$ for some $Q \leq G$, and $1 \neq M \subseteq Z^*(Q) \cap [Q,Q]$, then $M \subseteq Z^*(G)$. In particular, G is not capable.*

The idea of the proof is to show that if (E,φ) is a central extension of G, then the inverse image of M under φ will be contained in $Z(E)$, which implies that M is contained in the epicenter of G.

Shahriari also relaxes the condition that $G = QC_G(Q)$ in a number of ways (see in particular Proposition 3.3 in [62]), leading to:

Theorem 5.2 ([62], Theorems 4.2, 4.3, and 5.2) *Let G be a finite group.*

(i) If $Q_{2^n} \triangleleft G$, where Q_{2^n} is the generalized quaternion group of order 2^n, then $Z(Q_{2^n}) \subseteq Z^*(G)$. In particular, G is not capable.

(ii) If $S_{2^n} \triangleleft G$, where S_{2^n} is the semidihedral group of order 2^n, $n > 3$, then $Z(S_{2^n}) \subseteq Z^*(G)$. In particular, G is not capable.

(iii) If G is nilpotent and $E \triangleleft G$, where E is the extraspecial p-group of order p^3 and exponent p^2, $p > 2$, then $Z(E) \subseteq Z^*(G)$. In particular, G is not capable.

On the other hand, trying to restrict the subgroup structure itself seems to be a more difficult matter. By using the "coproduct with amalgamation" (a construction that is to the nilpotent product as the free product with amalgamation is to the free product of groups), and a description of the epicenter more suitable for computations that was developed by Ellis [25], one of us (AM) proved;

Proposition 5.3 ([50], Theorems 3.1 and 3.2) *Let G be any nontrivial group of class at most two and odd prime exponent. Then there exist groups G_1 and G_2 such that:*

(i) Both G_1 and G_2 have a subgroup isomorphic to G.

(ii) Both G_1 and G_2 are of class two and prime exponent.

(iii) Neither G_1 nor G_2 can be decomposed as a nontrivial direct or central product.

(iv) G_1 is capable and G_2 is not capable.

Part III: The homological approach

6 The nonabelian tensor product

In the 1980's Brown and Loday [18, 19] defined the *nonabelian tensor product* of two groups G and H whenever the groups act compatibly on each other and by conjugation on themselves. Brown and Loday's motivation for developing the nonabelian tensor product of two groups was topological: for instance it provides an algebraic description for certain homotopy groups and allows for a characterization of low dimensional homology groups. For further details on this subject, see [17]. The construction had forerunners in the work of Miller [54] and Dennis [21]. In this section, we first survey the construction and the work that preceded it.

The connection of the epicenter and the Ganea map to capability described in Section 4 provides a means for determining the capability of a group by connecting it to the study of the nonabelian tensor square of a group; this gives a new description of the epicenter which makes it easier to compute (for example, using a computer algebra system). We will describe this connection and discuss some of the results obtained via this description after defining the nonabelian tensor square.

6.1 The nonabelian tensor square

A special case of the nonabelian tensor product occurs when the two groups are equal and act compatibily on each other by conjugation. The resulting object is called the *nonabelian tensor square*.

Explicitly, if G is a group, then the nonabelian tensor square $G \otimes G$ of G has generating set consisting of formal elements labeled $g \otimes h$ for $g, h \in G$, with relations

$$gg' \otimes h = ({}^g g' \otimes {}^g h)(g \otimes h) \quad \text{and} \quad g \otimes hh' = (g \otimes h)({}^h g \otimes {}^h h'). \tag{1}$$

In the nonabelian tensor square there are two important central subgroups: $\nabla(G) = \langle\, x \otimes x \mid x \in G \,\rangle$, and $\Delta(G) = \langle\, (x \otimes y)(y \otimes x) \mid x, y \in G \,\rangle$. We define the nonabelian exterior square $G \wedge G$ as the quotient $(G \otimes G)/\nabla(G)$ and the nonabelian symmetric square $G \tilde{\otimes} G$ as the quotient $(G \otimes G)/\Delta(G)$. See [17, 18, 19].

Although Brown and Loday's interest arose from topological considerations, the nonabelian tensor square is connected to prior work that had its roots in algebraic K-theory [21] and in the work of Clair Miller [54].

Miller was interested in determining a group theoretic interpretation of the second homology group with integer coefficients. For a group G, Miller defines the free group $\langle G, G\rangle$ whose generators are all pairs (x, y) for all $x, y \in G$. This group has a natural epimorphism $\chi\colon \langle G, G\rangle \to [G, G]$. The kernel of this epimorphism, which we will denote as $N(G)$, is then a measure of all the commutator identities in G. Miller then defines the normal subgroup $B(G)$ to be the normal closure in $\langle G, G\rangle$ of certain words in $\langle G, G\rangle$ whose image under χ are commutator identities satisfied in the free group. Such identities are referred to as universal commutator relations. The generators Miller chooses to generate $B(G)$ are

$$(x, x), \tag{2}$$

$$(x, y)(y, x), \tag{3}$$

$$(xy, z)^{-1}({}^x y, {}^x z)(x, z), \quad ({}^x y, {}^x z)^{-1}(x, [y, z])(y, z) \tag{4}$$

for all x, y, z in G where ${}^u v = uvu^{-1}$. Identity (2) corresponds to the fact that $[x, x]$ is trivial for all x; (3) to the identity $[y, x] = [x, y]^{-1}$; and the identities in (4) to the product identity $[xy, z] = {}^x[y, z][x, z]$ when we use the convention that $[a, b] = aba^{-1}b^{-1}$ (a trivial modification is needed if one wants to follow the alternative convention $[a, b] = a^{-1}b^{-1}ab$), together with ${}^x[y, z] = [{}^x y, {}^x z]$ and ${}^u v = [u, v]u$.

We can think of the formal symbol (x, y) in Miller's construction as corresponding to the formal symbol $x \otimes y$ in Brown and Loday's definition of the nonabelian tensor square. With this in mind, note that the first relation in (1) is the same as the first relator in (4). In [17] it is shown that G acts on $G \otimes G$ so that ${}^h(g \otimes g') = {}^h g \otimes {}^h g'$. This action along with the relations of the nonabelian tensor square lead to the identity $g' \otimes [g, h] = ({}^{g'} g \otimes {}^{g'} h)(g \otimes h)^{-1}$ which is equivalent to the second relator of (4). In the other direction, the relators of (4) imply the relations (1). Hence $G \otimes G$ and $\langle G, G\rangle / B_1(G)$ are the same group where $B_1(G)$ is the normal closure of the subgroup generated by the elements given in (4).

Miller then shows that the quotient $M(G) = N(G)/B(G)$ is isomorphic to $H_2(G, \mathbb{Z})$, the second homology group of G with integer coefficients. The notation is appropriate as $M(G)$ is also the Schur multiplier of G. Miller proves that $M(G)$ is trivial for the free group and hence any universal commutator relation is a consequence of the ones chosen to generate $B(G)$. Therefore, $M(G)$ can be interpreted as a measure of the extent to which relations among commutators in G fail to be consequences of the universal commutator relations.

Another strand of the story arises in the work of Keith Dennis, whose 1976 preprint [21] is one of the most cited preprints in recent memory, being cited at least 17 times since 1999 and more than 20 times prior to that. Dennis was interested in defining a new sequence of homology functors that he labels $\widetilde{H}_i(G,M)$ that should satisfy certain axioms related to K-theory; here, G is a group and M is a G-module. These axioms hold for $\widetilde{H}_i(G,M) = H_i(G,M)$ when $i = 0,1$ but there is an obstruction to the equality $\widetilde{H}_2(G,M) = H_2(G,M)$. Dennis follows Miller's construction and forms a quotient of $\langle G,G\rangle$ with the subgroup $B_0(G)$ which is the normal closure of the words (3) and (4) in $\langle G,G\rangle$. He then defines $\widetilde{H}_2(G,\mathbb{Z}) = N(G)/B_0(G)$ and shows this functor satisfies the required axioms.

The extension $\langle G,G\rangle/B(G)$ of $M(G)$ by $[G,G]$ of Miller is called the nonabelian exterior square of G, and is readily seen to be isomorphic to the group $G \wedge G$ defined above following Brown and Loday's construction. Likewise, the extension $\langle G,G\rangle/B_0(G)$ of $\widetilde{H}_2(G,\mathbb{Z})$ by $[G,G]$ of Dennis is called the nonabelian symmetric square, and corresponds to the group $G\tilde{\otimes}G$ of the same name described above. These isomorphisms establish the connection between the works of Miller, Dennis, and of Brown and Loday.

We also mention that the functor $\widetilde{H}_2(G,\mathbb{Z})$ has topological significance: it is $\pi_4 S^2 K(G,1)$, the fourth homotopy group of the double suspension of the Eilenberg-Mac Lane space $K(G,1)$ with fundamental group G [19].

In 1988, R. Brown visited L.-C. Kappe at the State University of New York at Binghamton and gave a talk about the group theoretic aspects of the nonabelian tensor product and square as outlined in what is now considered a seminal paper on the subject [17]. L.-C. Kappe was taken with this topic, since commutator calculus is one of her specialties, and she and her students began working on problems related to the nonabelian tensor product in earnest. Five of her Ph.D. students from 1992 to 2010 wrote dissertations on topics related to the nonabelian tensor product, and about 12 research papers and one expository paper on the subject were published by her and/or her students. The theme of several of these papers was to describe the isomorphism types of a class of groups, and use this description to compute their nonabelian tensor squares. Among the classes they considered were the 2-generator p-groups of class 2 [2], infinite metacyclic groups [12], and the infinite 2-generator groups of class 2 [42].

6.2 The nonabelian exterior square and the epicenter

With this body of knowledge in place, two papers by Graham Ellis [26, 25] are of interest. In these papers Ellis defines two central subgroups in a group G:

$$Z^{\otimes}(G) = \{x \mid x \otimes y = 1_{\otimes} \text{ for all } y \text{ in } G\}$$
$$Z^{\wedge}(G) = \{x \mid x \wedge y = 1_{\wedge} \text{ for all } y \text{ in } G\}$$

called the tensor center and exterior center, respectively. Ellis proves the following:

Theorem 6.1 ([26], Prop. 16) *Let G be a group. Then $Z^{\wedge}(G) = Z^*(G)$.*

With this result in hand and the established computations for the nonabelian tensor square for various classes of groups in place, several authors then computed the

nonabelian exterior square and the exterior center to identify those groups that are capable in these classes. See for example [3, 12, 42].

6.3 The group $\tau(G)$ and the epicenter

The computations of the epicenter using the nonabelian exterior square were still somewhat impeded by difficulties in computing the latter, in the sense of finding some way to recognize the group $G \wedge G$ in a more familiar guise (for example, as a direct sum of cyclic groups when it is abelian). Computing the nonabelian exterior square required first computing the nonabelian tensor square. Hand methods to compute the nonabelian tensor square typically used crossed pairings (see [17]) which became computationally difficult as the nonabelian tensor square become more complex and nonabelian. Examples of this complexity can be seen in [4] and [15] where the nonabelian tensor squares are computed for the free 2-Engel groups of finite rank using crossed pairings.

Given a finite group G, initial computer methods for determining the nonabelian tensor square used the defining generators and relations from (1) to obtain a finite presentation for $G \otimes G$. This presentation has $|G|^2$ generators and $2|G|^3$ relations. Tietze transformations were then applied to the presentation to simplify it so its structure might be determined. The nonabelian tensor squares of the nonabelian groups up to order 30 were computed this way with some optimizations in [17].

A better way of approaching the problem was developed independently by Ellis and Leonard [27] and by Rocco [60]. In each case, they define the group $\nu(G) = (G*G^\varphi)/J$ where G^φ is an isomorphic copy of G via the mapping $\varphi : g \mapsto g^\varphi$, $G * G^\varphi$ is the free product, and J is the normal closure in $G * G^\varphi$ of the subgroup generated by the set of words

$$\{{}^z[g,h^\varphi][({}^zh)^\varphi,{}^zg],\ {}^{z^\varphi}[g,h^\varphi][({}^zh)^\varphi,{}^zg] \mid \text{ for all } z,g,h \in G\}. \tag{5}$$

The groups G and G^φ isomorphically embed into $\nu(G)$, and their images intersect trivially. We abuse notation and label these isomorphic subgroups of $\nu(G)$ as G and G^φ. The main feature of $\nu(G)$ is given in the following theorem:

Theorem 6.2 ([27] Claim 6; [60] Prop. 2.6) *Let $\nu(G)$, G, and G^φ be as above. The normal subgroup $[G,G^\varphi]$ of $\nu(G)$ is isomorphic to $G \otimes G$ via the natural mapping $[g,h^\varphi] \mapsto g \otimes h$.*

We obtain two major advantages by using this construction. The first is the ability to extend computer calculations by having a smaller presentation for $G \otimes G$. This was pursued by Ellis and Leonard in [27]. They show if $\mathcal{G}$ is a generating set for G and $\mathcal{D}$ is the union of a transversal of $Z(G)$ and a generating set of $Z(G)$ then the set of words generating J can be limited to

$$\{{}^z[g,h^\varphi][({}^zh)^\varphi,{}^zg],\ {}^{z^\varphi}[g,h^\varphi][({}^zh)^\varphi,{}^zg] \mid \text{ for all } g,h \in \mathcal{G} \text{ and } z \in \mathcal{D}\}.$$

Hence, if G is a finite group with presentation $\langle \mathcal{G} \mid \mathcal{R} \rangle$, then we can find a presentation for $\nu(G)$ with $2|\mathcal{G}|$ generators and $2|\mathcal{R}||\mathcal{G}|^2|\mathcal{D}|$ relations. This is a significant reduction in the size of the presentation to work with over that of definition (1). Ellis and Leonard were able to compute the nonabelian tensor squares for significantly larger

groups than those found in [17], such as the Burnside group of rank 2 and exponent 4 which has order 2^{12}.

The second advantage of working with $\nu(G)$ is that tensor calculations become commutator calculations within $[G, G^\varphi]$. Certainly all the usual commutator identities hold, but there are further identities (derived from the relations in $\nu(G)$) that also hold in $[G, G^\varphi]$. The investigation of this expanded commutator calculus was initiated by Rocco [60], and extended by Blyth and Morse [14]. The list below gives a flavor of these identities. For all $g_1, g_2, g_3, g_4 \in G$:

$$[g_1, [g_2, g_3]^\varphi] = [g_2, g_3, g_1^\varphi]^{-1} \tag{6}$$

$$[g_1, g_2, g_3^\varphi] = [g_1^\varphi, g_2^\varphi, g_3] \tag{7}$$

$$[[g_1, g_2^\varphi], [g_3, g_4^\varphi]] = [[g_1, g_2], [g_3, g_4]^\varphi]. \tag{8}$$

If we take the union of the set containing the words $[g, g^\varphi]$ and $[g, h^\varphi][h, g^\varphi]$ for all $g, h \in G$ with the set (5), and let J' be the normal closure in $G * G^\varphi$ of the subgroup generated by this new larger set of words, then we obtain a new group $\tau(G) \cong (G * G^\varphi)/J'$. This group is defined in [14] and has analogous properties to $\nu(G)$ that apply to the nonabelian exterior square. That is, both G and G^φ embed in $\tau(G)$, their images intersect trivially, and the corresponding commutator subgroup $[G, G^\varphi]$ is isomorphic to $G \wedge G$ via the natural mapping $[g, h^\varphi] \mapsto g \wedge h$. This construction leads to the following characterization of the epicenter.

Theorem 6.3 ([14], Theorem 19) *Let G be a group. Then $G \cap Z(\tau(G)) \cong Z^*(G)$ where we identify G with its isomorphic image inside $\tau(G)$.*

Now suppose that G is a group with generating set $\mathcal{G}$, and let $z \in Z(G)$ be an element that we want to test for membership in the epicenter of G. Applying the theorem we can show z is in the epicenter of G by demonstrating $[z, h^\varphi] = 1_{\tau(G)}$ for all $h \in \mathcal{G}$. This characterization of the epicenter combined with the commutator calculus for $[G, G^\varphi]$ makes determining whether a nontrivial central element is in the epicenter independent from computing the nonabelian exterior square, and so, in fact, means that we do not need to recognize a "more tractable" description of $G \wedge G$ in order to describe the epicenter. These ideas have been used to determine the capability of the 2-generator p-groups of class 2 [52] and the special p-groups of rank 2 [37].

7 Capable p-groups of nilpotency class two

As we mentioned above, the nonabelian exterior square and related constructions have been used to identify the capable groups within various classes of groups. These classes include the 2-generator p-groups of class 2 [52], the infinite metacyclic groups [12], and the 2-generator nontorsion groups of class 2 [43]. These groups in part were singled out because if G is metacyclic or nilpotent of class 2, then the nonabelian tensor and nonabelian exterior square of G are abelian, which makes them easier to compute and describe.

General results identifying capable p-groups are rare even for p-groups of class 2. The following is from ongoing work of H. Heineken, L.-C. Kappe, and R. F. Morse:

Proposition 7.1 ([37]) *Let G be a p-group of nilpotency class 2. If $\mho_k(G)$ is non-trivial and cyclic for some $k \in \mathbb{N}$, then G is not capable, provided that the exponent of G' divides p^k, if p is odd, and the exponent of G' divides p^{k-1}, if $p = 2$.*

In addition to the abelian, extraspecial, and metacyclic p-groups, the only class of p-groups where we have a complete characterization of the capable groups is that of the 2-generator p-groups of class 2. We describe some of this work in the next subsection, and conclude with a few more results obtained by homological methods for p-groups of class 2 and exponent p.

7.1 Two generator p-groups of class 2

The classification of the 2-generator p-groups of class 2 was initiated by Trebenko [67]. This classification, unfortunately, contained several errors and was incomplete. Bacon and L.-C. Kappe made corrections to this work for odd primes in [2], and used their description to compute the nonabelian tensor squares of these groups. The $p = 2$ case was taken up by L.-C. Kappe, Visscher, and Sarmin [43], again presenting a classification and using it to compute their nonabelian tensor squares.

Bacon and L.-C. Kappe [3] also used their computations of the nonabelian tensor squares for odd p to compute the nonabelian symmetric and exterior squares of these groups. The paper noted some errors that had remained in the classification from their earlier paper [2], and also some errors in computing the tensor squares. They attempted to corrected these errors, and with the new computations found the exterior centers, thus identifying most of the capable groups within this class. Classical methods were later used to complete the determination, and deal with the case of $p = 2$ [47, 48].

Unfortunately, in 2008 it was discovered that the classification for $p = 2$ found in [43] was incomplete, and further review unearthed a similar problem for the odd prime case analyzed in [2] and [3], as well as further issues with the classification for $p = 2$; however, it should be noted that these issues turn out not to affect the determination of capable groups, as the missing families consisted only of incapable groups.

Faced with these issues, a new classification for these groups using a different approach was done in [1]. This new classification lists the groups in terms of five parameters, and identifies precisely when two 5-tuples of parameters correspond to isomorphic groups. We will not go into the details, directing the interested reader to [1], and describe only how the five parameters yield a presentation for the group:

Theorem 7.2 *Let p be a prime and $n > 2$ a positive integer. Every 2-generator p-group of order p^n and class 2 corresponds to an ordered 5-tuple of integers, $(\alpha, \beta, \gamma; \rho, \sigma)$, such that: $\alpha \geq \beta \geq \gamma \geq 1$, $\alpha + \beta + \gamma = n$, $0 \leq \rho \leq \gamma$, and $0 \leq \sigma \leq \gamma$, with $(\alpha, \beta, \gamma; \rho, \sigma)$ corresponding to the group presented by*

$$G = \langle\, a, b \mid [a,b]^{p^\gamma} = [a,b,a] = [a,b,b] = 1, a^{p^\alpha} = [a,b]^{p^\rho}, b^{p^\beta} = [a,b]^{p^\sigma} \,\rangle. \qquad (9)$$

The same group may be represented by different 5-tuples (this is handled in [1], which provides a distinguished 5-tuple for each isomorphism class); for example, if $\alpha =$

$\beta > \gamma$, then the groups corresponding to $(\alpha, \beta, \gamma; \rho, \sigma)$ and to $(\alpha, \beta, \gamma; \min(\rho, \sigma), \gamma)$ are isomorphic.

In [52], this description is used to compute the nonabelian tensor and exterior squares of these groups, describing them in terms of the parameters. This then makes a determination of the capable (and likewise the unicentral) 2-generator p-groups of class 2 easy to describe using the exterior center and Theorem 6.3. For odd p, the result is:

Theorem 7.3 ([52], Theorem 63) *Let G be any 2-generator p-groups of class 2 with presentation (9) and $p > 2$. Then the group G is capable if it meets one of the following conditions:*

(i) $\rho \leq \sigma$, $\alpha = \beta$, *and* $\gamma = \rho$;

(ii) $\rho > \sigma$, $(\alpha - \beta) = (\rho - \sigma)$, *and* $\rho = \gamma$.

A similar result holds for $p = 2$.

This process is a good illustration of what is perhaps the most successful way in which the nonabelian tensor square has been used to date to determine capable groups: describe a class of groups via some parameterization, use the parameterization to compute the exterior center, and then identify the capable groups by constraints on the parameters. The end result is then theorems that echo Baer's classic Corollary 2.2, as well as the characterization of the capable metacyclic groups in Theorem 4.5.

7.2 Groups of class 2 and exponent p

The p-groups of class 2 and exponent p are an elusive class of p-groups when trying to characterize those which are capable. In Section 3 we saw conditions obtained via classical methods that either require or disallow the groups to be capable. In many ways, it seems that we are tantalizingly close to a full characterization, yet the full characterization has not yet been realized.

Using homological methods, there are two more results on p-groups of class 2 and exponent p. We include them here to show some of the ideas that are used to find nontrivial elements in the epicenter, and give the reader some idea of how the homological methods come into play.

Proposition 7.4 ([25], Prop. 9) *Let G be a finitely generated group of nilpotency class two and of prime exponent. Let $\{x_1, \ldots, x_k\}$ be a subset of G corresponding to a basis of the vector space $G/Z(G)$, and suppose that those non-trivial commutators of the form $[x_i, x_j]$ with $1 \leq i < j \leq k$ are distinct and constitute a basis for the vector space $[G, G]$. Then G is capable.*

The next result is similar; though it has not been included in any published paper that we are aware of:

Proposition 7.5 *Let G be a p-group of class 2 and exponent p with generating set $g_1, \ldots, g_n$ and $Z(G) = G'$. If $\bigcap [C_G(g_i), C_G(g_i)] \neq 1$, then G is not capable.*

Proposition 7.5 is proved with the aid the following lemma already found in [60].

Lemma 7.6 *Let G be a group and $x, y, z \in G$. If $[x, z] = 1$ and $[y, z] = 1$ in G then $[z, [x, y]^{\varphi}] = 1_{\tau(G)}$.*

Proof [Proposition 7.5] Suppose $\bigcap [C_G(g_i), C_G(g_i)] \neq 1$. Let c be in the intersection. Then by Lemma 7.6, $[c, g_i^{\varphi}] = 1_{\tau(G)}$ for $i = 1, \ldots, n$. Hence c is in the epicenter and G is not capable by Theorem 6.3. □

The converse of Proposition 7.5 does not hold. To see this, let A be the group of class 2 and exponent p generated by x, y, and z, with $[x, z] = 1$; and let B be the nonabelian group of order p^3 and exponent p, generated by w and t. Let G be the central product of A and B identifying $[x, y]$ with $[w, t]$. If we let $M = \{x, y, z, w, t\}$, then we have $[x, y] \in \bigcap_{m \in M} [C_G(m), C_G(m)]$. However, we can perturb the generator x and then take a new generating set $N = \{xw, y, z, w, t\}$; with this set, $C_G(xw) = \langle G', x, w, z\rangle$ is abelian, we have $\bigcap_{n \in N} [C_G(n), C_G(n)] = 1$, proving that the converse of Proposition 7.5 does not hold.

Nonetheless, we (AM and RFM) conjecture that for G of class two, exponent p, and such that $Z(G) = G'$, G is capable if and only if for *any* choice of minimal generating set $\{g_1, \ldots, g_n\}$, we have $\bigcap [C_G(g_i), C_G(g_i)] = 1$.

Part IV: Generalizations and future directions

8 Related concepts

There are many directions in which the notions in the previous sections may be generalized; because of space limitations, we only briefly sample a few of them with no claim of completeness.

Perhaps the simplest extension is to replace the center of a group with another term of the upper central series, $Z(G) = Z_1(G) \leq Z_2(G) \leq \cdots$. We say that a group G is *c-capable* if there exists a group K such that $G \cong K/Z_c(K)$ [20]. Clearly, if G is c-capable, then it is also d-capable for all $d \leq c$. Does the converse hold? For finitely generated abelian groups, Burns and Ellis proved it does:

Theorem 8.1 ([20], Theorem 1.3) *Fix $c \geq 1$. A finitely generated abelian group is c-capable if and only if it is capable.*

However, the converse does not hold in general: Burns and Ellis exhibit a group of order 2^{11} that is nilpotent of class 2, is 1-capable (that is, capable), but is not 2-capable.

For each c, there is an analogue of the epicenter, the *c-epicenter* $Z_c^*(G)$, which measures the obstruction to being c-capable, and analogues of the Schur multiplier.

In fact, these analogues fit into the much richer picture of Baer invariants and isologisms. In [35], a short follow-up to [34], Philip Hall introduced the notion of *marginal subgroups*. We can replace the commutator word $w(x, y) = [x, y]$ with an arbitrary group-theoretic word w (or a set of words), and the center with a subgroup, called the *marginal subgroup* $w^*(G)$ associated to w. The corresponding relation is called *isologism*, and Hall remarks that many, but not all, of the theorems of isoclinism carry over to the general theory. See [35] for the details.

In this setting, the role of the Schur multiplier is played by the Baer invariants; these were first introduced in the context of associative algebras by Frölich [29], who named them after Baer's work in [8, 9, 10]. A thorough exploration of these ideas, together with a short historical summary and many references, can be found in [44]. The corresponding analogues of capability and the epicenter have been studied by Moghaddam and Kayvanfar in several papers, beginning with [55].

Finally, in [24], Ellis points to the work of Shahriari mentioned above, and introduces the notion of *"relative capability"*: a group N is said to be "relatively capable" if there exists a capable group G containing N as a normal subgroup. Ellis then extends the notions of capability, Schur multipliers, and central series to *pairs of groups*, by which he means a pair (G, N) where G is a group and $N \triangleleft G$. When $N = G$, the notion reduces to the usual one for a single group. See [24] for relevant definitions and details. The notion was then extended to Baer invariants by Moghaddam and others.

9 Open questions

The homological approach has opened many doors in the study of capable groups. In particular, it is now relatively easy to determine whether a given finitely presented group is capable or not; this is particularly simple to do for polycyclic groups, and the algorithms developed to study that class apply, of course, to any particular p-group. The questions of interest are therefore no longer of the form "is this particular G capable?" as that question can be answered in a straightforward way for the majority of specific given groups G. Instead, we look for characterizations of the capable groups among some suitable class of groups, ideally along the lines of Baer's theorem, Corollary 2.2; this is a kind of gold standard for the type of theorems we want: it covers a fairly large class of groups, and it characterizes capability in terms of standard invariants of the group that are easy to compute.

The list of classes for which such a characterization has been obtained is still surprisingly short: (i) abelian groups that are direct sums of cyclic groups (Baer [7]), and in particular finitely generated abelian groups. (ii) Torsionfree abelian groups, even if they are not direct sums of cyclic groups (Beyl, Felgner, and Schmid, [13]). (iii) Extraspecial p-groups (Beyl, Felgner, and Schmid [13]). (iv) Metacyclic groups (Beyl, Felgner, and Schmid [13]). And (v) 2-generated p-groups of class two (AM and RFM [52]).

Not only is the list short, we may also note that with the exception of the very first (and oldest) result, the classes are somewhat restricted.

As we mentioned earlier, progress in the last few years has been along the following lines: if a class of p-groups (or more generally, polycyclic groups) can be parameterized in some way, then we can attempt to determine the capability of the groups in that class via those parameters: this is done by computing the nonabelian tensor square, and then the epicenter in terms of the nonabelian tensor square. The parameterization usually facilitates computational exploration (for example, using GAP), and a description of the epicenter in terms of the parameters. This was the situation for 2-generated p-groups of class two in [52], and other attempts (e.g., [3]). Our ability to obtain further results following this line depends, then, on our ability to

provide suitable descriptions of the families of groups in question. Once a suitable description can be found, it is possible to use computer packages such as GAP to accumulate computational evidence to help guide the desired characterization, as the authors of this survey did in [52].

Among the classes that seem within reach, we mention four:

- The *semiextraspecial p-groups.* A p-group G is semiextraspecial if for every maximal subgroup N of $Z(G)$, G/N is extraspecial. Moretó has shown that if a semiextraspecial p-group is capable, then its order must be of the form p^{3n} and the group must be of exponent p; if the condition is sufficient as well, as Moretó conjectures, this would give a nice generalization of the description of the capable extraspecial p-groups.
- The p-groups of class 2 with commutator subgroup isomorphic to $C_p \times C_p$; it was shown by Heineken [36] that any capable group satisfying this condition must have $p^2 < |G/Z(G)| < p^6$; as mentioned above, one of us (RFM) has been working with Heineken and L.-C. Kappe on determining up to isomorphism of all the capable groups in this class [37].
- The 2-generator p-group with cyclic derived subgroup and p odd. These groups were described by Miech [53], and again more recently but following a different scheme by Song [63]. The descriptions should allow a determination of the nonabelian tensor square, and from there the epicenter, in terms of the same parameters that describe the groups.
- The p-groups of class 2 and prime exponent, discussed above. Intuitively, it seems that a group in this class will be capable if and only if it is "non-abelian enough", but making this precise is proving challenging.

In each of these cases, ideally we would prefer a full description of the epicenter, which would yield both the capable and the unicentral groups in the class.

Finally, we briefly mentioned a line of inquiry suggested to us by Primož Moravec. Recall that if G is a finite p-group of order p^n and nilpotency class c, then the coclass of G is defined as $n - c$. Great strides have been made in understanding p-groups via coclass; see for example [45]. Associated with every prime p and positive integer r, we have a directed graph $\mathcal{G}(p, r)$; the vertices correspond to isomorphism types of p-groups of coclass r, and we have a directed edge $G \to H$ if $G \cong H/H_{\mathrm{cl}(H)}$, where $\mathrm{cl}(H)$ is the nilpotency class of H and H_i is the ith term of the lower central series of H. The graphs are trees, and they are usually pictured with nodes in "levels", with the nodes in level n corresponding to groups of order p^n. For more on the graphs, see for example [23]. One of the aims of coclass theory is to understand these graphs. Moravec noted in a personal communication that in the case of coclass 1, the capable groups are precisely the groups that are not leaves in the graphs. It would be interesting to see whether the capable groups are placed in some predictable pattern on the graph $\mathcal{G}(p, r)$ for arbitrary p and r.

References

[1] A. Ahmad, A. Magidin & R.F. Morse, Two generator p-groups of nilpotency class 2 and their conjugacy classes, *Publ. Math. Debrecen* **81** (2012), 145–166.

[2] M.R. Bacon & L.-C. Kappe, The nonabelian tensor square of a 2-generator p-group of class 2, *Arch. Math. (Basel)* **61** (1993), 508–516.

[3] M.R. Bacon & L.-C. Kappe, On capable p-groups of nilpotency class two, *Illinois J. Math.* **47** (2003), 49–62.
[4] M.R. Bacon, L.-C. Kappe & R.F. Morse, On the nonabelian tensor square of a 2-Engel group, *Arch. Math. (Basel)* **69** (1997), 353–364.
[5] R. Baer, Erweiterung von Gruppen und ihren Isomorphismen, *Math. Z.* **38** (1934), 375–416.
[6] R. Baer, Groups with abelian central quotient group, *Trans. Amer. Math. Soc.* **44** (1938), 357–386.
[7] R. Baer, Groups with preassigned central and central quotient group, *Trans. Amer. Math. Soc.* **44** (1938), 387–412.
[8] R. Baer, Representations of groups as quotient groups. I, *Trans. Amer. Math. Soc.* **58** (1945), 295–347.
[9] R. Baer, Representations of groups as quotient groups. II. Minimal central chains of a group, *Trans. Amer. Math. Soc.* **58** (1945), 348–389.
[10] R. Baer, Representations of groups as quotient groups. III. Invariants of classes of related representations, *Trans. Amer. Math. Soc.* **58** (1945), 390–419.
[11] H.U. Besche, B. Eick & E.A. O'Brien, A millennium project: constructing small groups, *Internat. J. Algebra Comput.* **12** (2002), 623–644.
[12] J.R. Beuerle & L.-C. Kappe, Infinite metacyclic groups and their non-abelian tensor squares, *Proc. Edinburgh Math. Soc. (2)* **43** (2000), 651–662.
[13] F.R. Beyl, U. Felgner & P. Schmid, On groups occurring as center factor groups, *J. Algebra* **61** (1979), 161–177.
[14] R.D. Blyth & R.F. Morse, Computing the nonabelian tensor squares of polycyclic groups, *J. Algebra* **321** (2009), 2139–2148.
[15] R.D. Blyth, R.F. Morse & J.L. Redden, On computing the non-abelian tensor squares of the free 2-Engel groups, *Proc. Edinb. Math. Soc. (2)* **47** (2004), 305–323.
[16] H.R. Brahana, Finite metabelian groups and the lines of a projective four-space, *Amer. J. Math.* **73** (1951), 539–555.
[17] R. Brown, D.L. Johnson & E.F. Robertson, Some computations of nonabelian tensor products of groups, *J. Algebra* **111** (1987), 177–202.
[18] R. Brown & J.-L. Loday, Excision homotopique en basse dimension, *C. R. Acad. Sci. Paris Sér. I Math.* **298** (1984), 353–356.
[19] R. Brown & J.-L. Loday, Van Kampen theorems for diagrams of spaces, *Topology* **26** (1987), 311–335.
[20] J. Burns & G. Ellis, On the nilpotent multipliers of a group, *Math. Z.* **226** (1997), 405–428.
[21] R.K. Dennis, In search of new "homology" functors having a close relationship to k-theory, unpublished preprint, 1976.
[22] T.E. Easterfield, The orders of products and commutators in prime-power groups, *Proc. Cambridge Philos. Soc.* **36** (1940), 14–26.
[23] B. Eick & T. Rossmann, Periodicities for graphs of p-groups beyond coclass, *Computational group theory and the theory of groups, II*, 11–23, Contemp. Math. 511, Amer. Math. Soc., 2010.
[24] G. Ellis, Capability, homology, and central series of a pair of groups, *J. Algebra* **179** (1996), 31–46.
[25] G. Ellis, On the capability of groups, *Proc. Edinburgh Math. Soc. (2)* **41** (1998), 487–495.
[26] G. Ellis, Tensor products and q-crossed modules, *J. London Math. Soc. (2)* **51** (1995), 243–258.
[27] G. Ellis & F. Leonard, Computing Schur multipliers and tensor products of finite groups, *Proc. Roy. Irish Acad. Sect. A* **95** (1995), 137–147.
[28] L. Evens, Terminal p-groups, *Illinois J. Math.* **12** (1968), 682–699.
[29] A. Fröhlich, Baer-invariants of algebras, *Trans. Amer. Math. Soc.* **109** (1963), 221–244.

[30] T. Ganea, Homologie et extensions centrales de groupes, *C. R. Acad. Sci. Paris Sér. A–B* **266** (1968), A556–A558.
[31] O.N. Golovin, Nilpotent products of groups, *Amer. Math. Soc. Transl. (2)* **2** (1956), 89–115.
[32] M. Hall, Jr. & J.K. Senior, *The Groups of Order* 2^n *(*$n \leq 6$*)*, Macmillan, 1964.
[33] P. Hall, A contribution to the theory of groups of prime-power order, *Proc. London Math. Soc.* **S2-36** (1932), 29–95.
[34] P. Hall, The classification of prime-power groups, *J. Reine Angew. Math.* **182** (1940), 130–141.
[35] P. Hall, Verbal and marginal subgroups, *J. Reine Angew. Math.* **182** (1940), 156–157.
[36] H. Heineken, Nilpotent groups of class two that can appear as central quotient groups, *Rend. Sem. Mat. Univ. Padova* **84** (1990), 241–248 (1991).
[37] H. Heineken, L.-C. Kappe & R.F. Morse, The classification of special p-groups of rank two that appear as central quotient groups, in preparation.
[38] H. Heineken & D. Nikolova, Class two nilpotent capable groups, *Bull. Austral. Math. Soc.* **54** (1996), 347–352.
[39] I.M. Isaacs, Derived subgroups and centers of capable groups, *Proc. Amer. Math. Soc.* **129** (2001), 2853–2859.
[40] R. James, The groups of order p^6 (p an odd prime), *Math. Comp.* **34** (1980), 613–637.
[41] R. James, M.F. Newman & E.A. O'Brien, The groups of order 128, *J. Algebra* **129** (1990), 136–158.
[42] L.-C. Kappe, N.M. Mohd Ali & N.H. Sarmin, On the capability of finitely generated non-torsion groups of nilpotency class 2, *Glasg. Math. J.* **53** (2011), 411–417.
[43] L.-C. Kappe, M.P. Visscher & N.H. Sarmin, Two-generator two-groups of class two and their nonabelian tensor squares, *Glasg. Math. J.* **41** (1999), 417–430.
[44] C.R. Leedham-Green & S. McKay, Baer-invariants, isologism, varietal laws and homology, *Acta Math.* **137** (1976), 99–150.
[45] C.R. Leedham-Green & S. McKay, *The Structure of Groups of Prime Power Order*, London Math. Soc. Monographs, OUP, 2002.
[46] A. Magidin, On the orders of generators of capable p-groups, *Bull. Austral. Math. Soc.* **70** (2004), 391–395.
[47] A. Magidin, Capability of nilpotent products of cyclic groups, *J. Group Theory* **8** (2005), 431–452.
[48] A. Magidin, Capable 2-generator 2-groups of class two, *Comm. Algebra* **34** (2006), 2183–2193.
[49] A. Magidin, Capability of nilpotent products of cyclic groups. II, *J. Group Theory* **10** (2007), 441–451.
[50] A. Magidin, Embedding groups of class two and prime exponent in capable and noncapable groups, *Bull. Aust. Math. Soc.* **79** (2009), 303–308.
[51] A. Magidin, On the capability of finite groups of class two and prime exponent, *Publ. Math. Debrecen*, to appear, 2014; see also expanded version in arXiv:0708.2391.
[52] A. Magidin & R.F. Morse, Certain homological functors of 2-generator p-groups of class 2, *Computational group theory and the theory of groups, II*, 127–166, Contemp. Math. 511, Amer. Math. Soc., 2010.
[53] R.J. Miech, On p-groups with a cyclic commutator subgroup, *J. Austral. Math. Soc.* **20** (1975), 178–198.
[54] C. Miller, The second homology group of a group; relations among commutators, *Proc. Amer. Math. Soc.* **3** (1952), 588–595.
[55] M.R.R. Moghaddam & S. Kayvanfar, A new notion derived from varieties of groups, *Algebra Colloq.* **4** (1997), 1–11.
[56] E.A. O'Brien, The groups of order 256, *J. Algebra* **143** (1991), 219–235.
[57] E.A. O'Brien & M.R. Vaughan-Lee, The groups with order p^7 for odd prime p, *J. Alge-*

bra **292** (2005), 243–258.
[58] K. Podoski & B. Szegedy, Bounds for the index of the centre in capable groups, *Proc. Amer. Math. Soc.* **133** (2005), 3441–3445.
[59] K. Podoski & B. Szegedy, On finite groups whose derived subgroup has bounded rank, *Israel J. Math.* **178** (2010), 51–60.
[60] N.R. Rocco, On a construction related to the nonabelian tensor square of a group, *Bol. Soc. Brasil. Mat. (N.S.)* **22** (1991), 63–79.
[61] J. Schur, Über die Darstellung der endlichen Gruppen durch gebrochen lineare Substitutionen, *J. Reine Angew. Math.* **127** (1904), 20–50.
[62] S. Shahriari, On normal subgroups of capable groups, *Arch. Math. (Basel)* **48** (1987), 193–198.
[63] Q. Song, Finite two-generator p-groups with cyclic derived group, *Comm. Algebra* **41** (2013), 1499–1513.
[64] A. Speiser, Gruppendeterminante und Körperdiskriminante, *Math. Ann.* **4** (1916), 546–562.
[65] R.R. Struik, On nilpotent products of cyclic groups, *Canad. J. Math.* **12** (1960), 447–462.
[66] R.R. Struik, On nilpotent products of cyclic groups. II, *Canad. J. Math.* **13** (1961), 557–568.
[67] D.Ya. Trebenko, Nilpotent groups of class two with two generators, *Current analysis and its applications (Russian)*, 201–208, 228, "Naukova Dunka", 1989.

ON THE NORMAL STRUCTURE OF A FINITE GROUP WITH RESTRICTIONS ON THE MAXIMAL SUBGROUPS

N. V. MASLOVA* and D. O. REVIN†

*N.N. Krasovskii Institute of Mathematics and Mechanics of Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russia

Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia

†S.L. Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Novosibirsk State University, Novosibirsk, Russia

Email: butterson@mail.ru, revin@math.nsc.ru

1 Introduction

Our terminology and notation are mostly standard (see, for example, [1, 2]). We use the term "group" to mean "finite group."

Let π be a set of primes. Denote by π' the set of primes not in π. Given a natural n, we denote by $\pi(n)$ the set of prime divisors of n. A natural number n with $\pi(n) \subseteq \pi$ is called a *π-number*, and a group G such that $\pi(G) \subseteq \pi$ is called a *π-group*. For a group G, the set $\pi(G) = \pi(|G|)$ is the *prime spectrum* of G. A subgroup H of a group G is called a *π-Hall subgroup* if $\pi(H) \subseteq \pi$ and $\pi(|G : H|) \subseteq \pi'$. Thus, if π consists of a single prime p then a π-Hall subgroup is exactly a Sylow p-subgroup. A *Hall subgroup* is a π-Hall subgroup for some set π of primes. A group G is *prime spectrum minimal* if $\pi(H) \neq \pi(G)$ for every proper subgroup H of G.

We say that G is *a group with Hall maximal subgroups* if every maximal subgroup of G is a Hall subgroup. It is easy to see that every group with Hall maximal subgroups is prime spectrum minimal.

A group G is *a group with complemented maximal subgroups* if for every maximal subgroup M of G, there exists a subgroup H such that $MH = G$ and $M \cap H = 1$.

The study of groups with Hall maximal subgroups was started in 2006 by Levchuk and Likharev [3] and Tyutyanov [4], who established that a nonabelian simple group with complemented maximal subgroups is isomorphic to one of the groups $PSL_2(7) \cong PSL_3(2)$, $PSL_2(11)$ or $PSL_5(2)$. In all these groups, every maximal subgroup is a Hall subgroup. In 2008, Tikhonenko and Tyutyanov [5] showed that the nonabelian simple groups with Hall maximal subgroups are exhausted up to isomorphism by the groups $PSL_2(7)$, $PSL_2(11)$, and $PSL_5(2)$.

The research was supported by RFBR (projects 13-01-00469 and 13-01-00505), by the Program of the Joint Investigations of the Ural Branch of the Russian Academy of Sciences with Siberian Branch of the Russian Academy of Sciences (project 12--1-10018) and with Belorussian National Academy of Sciences (project 12-C-1-1009), by the grant of the President of Russian Federation for young scientists (project MK-3395.2012.1), by the Dmitry Zimin Foundation "Dynasty" and by the Program of the State support of leading universities of the Russia (agreement no. 02.A03.21.0006 of 27.08.2013).

In 2008, Monakhov [6] studied the normal structure and other properties of a solvable group with Hall maximal subgroups. Also in [6], he formulated the following problem.

Problem 1. *What are the nonabelian composition factors of a nonsolvable group whose all maximal subgroups are Hall?*

In 2010, Problem 1 was written by Monakhov into the "Kourovka Notebook" [7] as Problem 17.92.

We have solved Problem 1, determined the normal structure of groups with Hall maximal subgroups and investigated the nonabelian composition factors of prime spectrum minimal groups. In this paper we give a survey of these results.

2 Nonabelian composition factors of a group with Hall maximal subgroups

Despite the fact that simple groups with Hall maximal subgroups are known, nonabelian composition factors of a nonsolvable group with Hall maximal subgroups need not be *a priori* groups with Hall maximal subgroups.

Although Hall subgroups in simple groups and groups close to them were studied by various authors and are at present completely described (see, for example, the surveys [8, 9]), the study of Hall maximal subgroups in an arbitrary group cannot be reduced only to the study of the Hall maximal subgroups of its composition factors. For example, each subgroup $P \in Syl_2(Aut(A_6))$ is maximal in $Aut(A_6)$ but $P \cap A_6 < S_4$ and S_4 is a maximal subgroup in A_6 which is not Hall. But the maximal subgroups of a simple group give some information about whether or not the group is isomorphic to a composition factor of some group with Hall maximal subgroups.

The following lemma gives an approach to solving Problem 1.

Lemma 2.1 *Let S be a nonabelian simple group having a subgroup X such that*

(1) *the conjugacy class $X^S = \{X^s \mid s \in S\}$ is invariant under $Aut(S)$;*

(2) *$|Z|$ and $|S : Z|$ are not coprime for every subgroup Z such that $X \leq Z < S$.*

Then there is no a group G with Hall maximal subgroups such that S is isomorphic to a composition factor of G.

The first author has obtained a full description of nonabelian composition factors for nonsolvable groups with Hall maximal subgroups. Thus, Problem 17.92 in the "Kourovka Notebook" was solved. The solution of Problem 1 is given by the following theorem [10, Theorem 1].

Theorem 2.2 *The nonabelian composition factors of a group with Hall maximal subgroups are exhausted by the groups $PSL_2(7)$, $PSL_2(11)$, and $PSL_5(2)$.*

In the proof of Theorem 2.2 we use the approach from Lemma 2.1, the classification of maximal subgroups of odd index in almost simple groups [11, 12, 13, 14, 15] and the description of π-Hall subgroups of Chevalley groups whose characteristic belongs to π [16].

3 Normal structure of a group with Hall maximal subgroups

The following problem is closely related to Problem 1.

Problem 2. *What are the chief factors of a nonsolvable group with Hall maximal subgroups*?

Monakhov in [6, Corollary 1] proved the following theorem.

Theorem 3.1 *Let G be a solvable group. The following conditions are equivalent:*

(1) *G is a group with Hall maximal subgroups*;

(2) *every maximal subgroup of G is complemented by a Sylow subgroup of G*;

(3) *every chief factor of G is isomorphic to a Sylow subgroup of G*;

(4) *every normal subgroup of G is a Hall subgroup in G.*

In other words, a solvable group G is a group with Hall maximal subgroups if and only if in G there exists a normal series

$$G = G_0 > G_1 > \ldots > G_n = 1$$

such that every factor G_i/G_{i+1} is an elementary abelian p-group isomorphic to a Sylow subgroup of G for some prime divisor p of $|G|$ and G/G_i acts irreducibly on G_i/G_{i+1}.

Evidently, the normal structure of an arbitrary group with Hall maximal subgroups can be significantly different from the normal structure of a solvable group with Hall maximal subgroups. Problem 2 was considered by the authors. Using Theorem 2.2 we proved the following theorem [17, Theorem 1].

Theorem 3.2 *Let G be a group. Then G is a group with Hall maximal subgroups if and only if in G there exists a normal series*

$$G = G_0 > G_1 > \ldots > G_n = 1$$

such that

(1) *for $i \geq 1$, the factor G_i/G_{i+1} is an elementary abelian p-group isomorphic to a Sylow subgroup of G for some prime divisor p of $|G|$ and G/G_i acts irreducibly on G_i/G_{i+1}*;

(2) *for the factor $\bar{G} = G_0/G_1$, one of the following conditions holds:*

   (i) $\bar{G} \cong Z_p$, where p is a prime;

   (ii) $\bar{G}$ is isomorphic to either $PSL_2(11)$ or $PSL_5(2)$;

   (iii) $\bar{G}/\Phi(\bar{G}) \cong PSL_2(7)$ and $\Phi(\bar{G})$ is a 3-group.

In particular, a group G with Hall maximal subgroups contains at most one nonabelian composition factor, and the solvable radical $S(G)$ of G possesses a Sylow normal chain. Furthermore, G acts irreducibly on the factors of this chain and the factor group $G/S(G)$ is either trivial or isomorphic to one of the following groups: $PSL_2(7)$, $PSL_2(11)$ or $PSL_5(2)$.

In 2013, using Theorems 2.2 and 3.2, Vedernikov [18, Theorem 2] described the nonabelian compositional factors of a group in which every maximal subgroup is solvable or Hall.

4 Groups with complemented maximal subgroups

Since every solvable or simple group with Hall maximal subgroups is a group with complemented maximal subgroups, the following conjecture was formulated in [5].

Conjecture 1. *Every group with Hall maximal subgroups is a group with complemented maximal subgroups.*

Using Theorem 3.2, we have confirmed this conjecture. We have proved the following theorem [17, Theorem 2].

Theorem 4.1 *If G is a group with Hall maximal subgroups then G is a group with complemented maximal subgroups.*

Note that the converse to Theorem 4.1 is false even for a solvable group. For example, in an elementary abelian p-group, all maximal subgroups are complemented but are not Hall subgroups. In the groups $PGL_2(7)$ and $Z_3 \times PSL_2(7)$, all maximal subgroups are complemented but these groups possess maximal subgroups which are not Hall subgroups. It is interesting to study the normal structure of a group with complemented maximal subgroups. Because of simple groups $PSL_2(7)$, $PSL_2(11)$ and $PSL_5(2)$ in Theorem 2.2 are the only nonabelian simple groups with complemented maximal subgroups, the following open problem is interesting.

Problem 3. *What are nonabelian composition factors of a group with complemented maximal subgroups*?

We have written Problem 3 into the "Kourovka Notebook" as Problem 18.68.

5 Generation of a group with Hall maximal subgroups by a pair of conjugate elements

In 2010, Shumyatsky has written into the "Kourovka Notebook" [7, Problem 17.125] the following conjecture.

Conjecture 2. *In any group G, there is a pair a, b of conjugate elements such that $\pi(G) = \pi(\langle a, b\rangle)$.*

Note that $\pi(G) = \pi(\langle x, y\rangle)$ for every group G and some $x, y \in G$ [20].

It is easy to prove that Shumyatsky's conjecture is equivalent to the following conjecture.

Conjecture 3. *Every prime spectrum minimal group is generated by a pair of conjugate elements.*

Moreover, a minimal counterexample to either of conjectures 2 or 3 will also be a minimal counterexample to the other one (see [19, Lemma 1]).

Because every group with Hall maximal subgroups is prime spectrum minimal and every solvable prime spectrum minimal group is a group with Hall maximal subgroups (see [19, Lemma 2]), it is interesting to study the following question.

Question. *Is every group with Hall maximal subgroups generated by a pair of conjugate elements*?

Using Theorem 3.2, we have proved the following theorem [19, Theorem].

Theorem 5.1 *Every group with Hall maximal subgroups is generated by a pair of conjugate elements.*

Moreover, we have given an algorithm that constructs explicitly for every group with Hall maximal subgroups a pair of conjugated elements generating this group [19, P. 204].

Thus, we have obtained a partial confirmation of Conjecture 2 and a partial solution of Problem 17.125 from the "Kourovka Notebook". Note that we succeeded to prove Theorem 5.1 because the normal structure of a group with Hall maximal subgroups is known. Thus, the following open problem is of interest.

Problem 4. *Let G be a simple group. Is G isomorphic to a nonabelian composition factor of a prime spectrum minimal group*?

6 Nonabelian composition factors of a prime spectrum minimal group

Note that the class of prime spectrum minimal groups is more general than the class of groups with Hall maximal subgroups. Indeed, by using [1, 24], it is possible to show that the following groups are prime spectrum minimal: $Aut(PSL_2(32))$ and $Sz(2^{p_1})\times Sz(2^{p_2})\times\ldots\times Sz(2^{p_n})$ where $p_1, p_2, \ldots, p_n$ are pairwise different odd primes. Thus, a prime spectrum minimal group can be almost simple, but not simple, and the number of nonabelian composition factors of a prime spectrum minimal group is unbounded, while every nonsolvable group G with Hall maximal subgroups has exactly one nonabelian composition factor and the quotient $G/S(G)$ is simple.

Simple prime spectrum minimal groups were described by Liebeck, Praeger and Saxl in [21, Corollary 5, Table 10.7].

Theorem 6.1 *If G is a simple group then G is prime spectrum minimal except of the following cases:*

(i) $G \cong A_n$ *where n is not a prime*;

(ii) $G \cong PSp_{2m}(q)$ *where m and q are even*;

(iii) $G \cong PSp_4(q)$ *where q is odd*;

(iv) $G \cong P\Omega_{2m+1}(q)$ *where m is even and q is odd*;

(v) $G \cong P\Omega^+_{2m}(q)$ *where m is even*;

(vi) G *is isomorphic to one of the following simple groups*: $PSL_6(2)$, $PSU_3(3)$, $PSU_3(5)$, $PSU_4(2)$, $PSU_4(3)$, $PSU_5(2)$, $PSU_6(2)$, $PSp_6(2)$, $G_2(3)$, ${}^2F_4(2)'$, M_{11}, M_{12}, M_{24}, HS, McL, Co_2, Co_3.

As follows from Theorem 3.2, every nonabelian composition factor of a group with Hall maximal subgroups is a group with Hall maximal subgroups. In the case of prime spectrum minimal groups the same statement is false. V. I. Trofimov has constructed

the prime spectrum minimal group G such that $G \cong (McL)^{104} \leftthreetimes SL_2(103)$, and the unique involution in $SL_2(103)$ induces a nontrivial outer authomorphism on every component of the socle of G. We have proved the following theorem [22, Theorem 2].

Theorem 6.2 *The McLaughlin group McL is not prime spectrum minimal, but there exists a prime spectrum minimal group G containing McL as a composition factor. In every such group G there exists a normal series $G \geq Y > Z \geq 1$ such that*

a) *groups G/Y and Z are not containing McL as a composition factor;*

b) *the factor Y/Z is a chief factor of G and is isomorphic to a direct product of simple groups every of which is isomorphic to the group McL;*

c) *the factor G/Y is nonsolvable, its order is not divisible by 7 and by 11 and its nonabelian composition factors are isomorphic to groups from the following list: $PSL_2(q)$, $PSL_3(q)$, $PSL_4(q)$, $PSL_5(q)$, $PSU_3(q)$, $PSU_4(q)$, $PSU_5(q)$, $PSp_4(2^m)$, ${}^2B_2(2^{2m+1})$, ${}^2F_4(2^{2m+1})$, J_3.*

Note that our proof of the existence of a prime spectrum minimal group having a composition factor isomorphic to the McLaughlin group is based on Trofimov's idea (see above) and on an embedding of $SL_2(103)$ into the the permutation wreath product $Z_2 \wr PSL_2(103)$ associated with the natural projective action of $PSL_2(103)$ (see [23]).

Within studying Problem 4 we have proved the following theorem [22, Theorem 1].

Theorem 6.3 *The following simple groups are not isomorphic to composition factors of prime spectrum minimal groups:*

(1) *sporadic groups M_{11}, M_{12}, M_{24}, HS, Co_3, Co_2 and the Tits group ${}^2F_4(2)'$;*

(2) *A_n where n is not a prime;*

(3) *$PSp_4(q)$ where q is odd;*

(4) *$PSp_{2m}(q)$ where $m \geq 4$ and q are even;*

(5) *$P\Omega_{2m+1}(q)$ where $m \geq 4$ is even and q is odd;*

(6) *simple groups $PSU_3(3)$, $PSU_4(2)$, $PSU_5(2)$, $PSp_6(2)$, $PSL_6(2)$, $G_2(3)$.*

A good approach to prove Theorem 6.3 gives the following lemma.

Lemma 6.4 *Let S be a nonabelian simple group having a subgroup X such that*

(1) *the conjugacy class $X^S = \{X^s \mid s \in S\}$ is invariant under $Aut(S)$;*

(2) *$\pi(X) = \pi(S)$.*

Then there is no a prime spectrum minimal group G such that S is isomorphic to a composition factor of G.

However, there are some simple groups S such that $\pi(X) = \pi(S)$ for some proper subgroup $X < S$ and the conjugacy class $X^S = \{X^s \mid s \in S\}$ is not invariant under $Aut(S)$ for every such subgroup X. For such a group S the solution of Problem 4 can be positive (for example, for $S = McL$) as well as negative (for example, for $S = A_6$). In every such "unregular" case the solution of Problem 4 demands individual approach.

Remark. For simple groups $P\Omega^+_{4k}(q)$, $PSp_4(2^w)$, $PSU_3(5)$, $PSU_4(3)$ and $PSU_6(2)$ Problem 4 is open.

7 Prime spectrum critical groups

A group G is said to be *prime spectrum critical* if for every subgroups K and L such that $K \trianglelefteq L \leq G$ the equality $\pi(L/K) = \pi(G)$ implies $L = G$ and $K = 1$.

It was proved in [19, Proposition 3] that a minimal counterexample to Conjecture 2 is a prime spectrum critical group. The set of all nonabelian composition factors of all prime spectrum minimal groups coincides with the set of all nonabelian composition factors of all prime spectrum critical groups in view of the following criterion.

Proposition 7.1 *Let G be a prime spectrum minimal group. Then G is prime spectrum critical if and only if its Fitting subgroup $F(G)$ is a Hall subgroup of G.*

It's easy to see, every solvable or simple prime spectrum minimal group is prime spectrum critical.

Acknowledgement The authors wish to thank Professors A. S. Kondratiev, E. I. Khukhro, V. D. Mazurov, V. I. Trofimov and A. V. Zavarnitsine for valuable remarks which contributed to our results.

References

[1] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups*, Clarendon Press, Oxford, 1985.

[2] P. B. Kleidman and M. W. Liebeck, *The subgroup structure of finite classical groups*, Cambridge Univ. Press, Cambridge, 1990.

[3] V. M. Levchuk, A. G. Likharev, *Finite simple groups with complemented maximal subgroups*, Siberian Math. J., **47:4** (2006), 659–681.

[4] V. N. Tyutyanov, *Finite groups with complemented subgroups*, Izv. F. Skorina Gomel Univ., **36:3** (2006), 178-183 (In Russian).

[5] T. V. Tikhonenko, V. N. Tyutyanov, *Finite groups with maximal Hall subgroups*, Izv. F. Skorina Gomel Univ., **50:5** (2008), 198-206 (In Russian).

[6] V. S. Monakhov, *Finite π-solvable groups whose maximal subgroups have the Hall property*, Math. Notes, **84:3** (2008), 363–366.

[7] *Kourovka Notebook, Unsolved Problems of Group Theory.* 18 ed., Institute of Mathematics Siberian Division RAS, Novosibirsk, 2014. arXiv:1401.0296 [math.GR]

[8] E. P. Vdovin, D. O. Revin, *Theorems of Sylow type*, Russian Mathematical Surveys, **66:5** (2011), 829–870.

[9] D. O. Revin, E. P. Vdovin, *Generalizations of the Sylow theorem*, in C. M. Campbell, M. R. Quick, E. F. Robertson, C. M. Roney-Dougal, G. C. Smith (eds), Groups St Andrews 2009 in Bath, LMS Lecture Note Series, **388:2** (2011), Cambridge Univ. Press, Cambridge, 488–519.

[10] N. V. Maslova, *Nonabelian composition factors of a finite group whose all maximal subgroups are Hall*, Siberian Math. J., **53:5** (2012), 853–861.

[11] W. M. Kantor, *Primitive permutation groups of odd degree, and an application to finite projective planes*, J. Algebra **106** (1987), 15–45.

[12] M. W. Liebeck and J. Saxl, *The primitive permutation groups of odd degree*, J. London Math. Soc. (2) **31** (1985), 250–264.

[13] N. V. Maslova, *Classification of maximal subgroups of odd index in finite simple classical groups*, Proceedings of the Steklov Institute of Mathematics, Suppl. **3** (2009), S164–S183.
[14] N. V. Maslova, *Classification of maximal subgroups of odd index in finite groups with alternatng socle*, Trudy IMM UrO RAN, **16:3** (2010), 182–184 (in Russian).
[15] N. V. Maslova, *Maximal subgroups of odd index in finite groups with simple classical socle*, in C. M. Campbell, M. R. Quick, E. F. Robertson, C. M. Roney-Dougal, G. C. Smith (eds), Groups St Andrews 2009 in Bath, LMS Lecture Note Series, **388:2** (2011), Cambridge Univ. Press, Cambridge, 473–478.
[16] D. O. Revin, *Hall π-subgroups of finite Chevalley groups whose characteristic belongs to π*, Siberian Adv. Math., **9:2** (1999), 25–71.
[17] N. V. Maslova, D. O. Revin, *Finite groups whose maximal subgroups have the hall property*, Siberian Adv. in Math., **23:3** (2013), 196–209.
[18] V. A. Vedernikov, *Finite groups with Hall nonsolvable maximal subgroups*, Trudy IMM UrO RAN, **19:3** (2013), 71–82 (in Russian).
[19] N. V. Maslova, D. O. Revin, *Generation of a finite group with Hall maximal subgroups by a pair of conjugate elements*, Trudy IMM UrO RAN, **19:3** (2013), 199–206 (in Russian).
[20] A. Lucchini, M. Morigi, P. Shumyatsky *Boundedly generated subgroups of finite groups*, Forum Math. **24:4** (2012), 875–887.
[21] M. W. Liebeck, C. E. Praeger, J. Saxl, *Transitive subgroups of primitive permutation groups*, J. Algebra, **234:2** (2000), 291–361.
[22] N. V. Maslova, D. O. Revin, *On the nonabelian composition factors of a finite prime spectrum minimal group*, Trudy IMM UrO RAN, **19:4** (2013), 155–166 (in Russian).
[23] A. V. Zavarnitsine, *Subextensions for a permutation $PSL_2(q)$-module*, Siberian Electronic Mathematical Reports, **10** (2013), 551–557.
[24] M. Suzuki, *On a class of doubly transitive groups*, Ann. Math., **75:1** (1962), 105–145.

CERTAIN MONOMIAL CHARACTERS AND THEIR NORMAL CONSTITUENTS

GABRIEL NAVARRO* and CAROLINA VALLEJO†

Departament d'Àlgebra, Universitat de València, 46100 Burjassot, Spain
*Email: gabriel.navarro@uv.es
†Email: carolina.vallejo@uv.es

Abstract

Suppose that G is a finite p-solvable group such that $\mathbf{N}_G(P)/P$ has odd order, where $P \in \mathrm{Syl}_p(G)$. If χ is an irreducible complex character with degree not divisible by p and field of values contained in a cyclotomic field $\mathbb{Q}_{p^a}$, then every subnormal constituent of χ is monomial. Also, the number of such irreducible characters is the number of $\mathbf{N}_G(P)$-orbits on P/P'.

1 Introduction

There are few results guaranteeing that a single irreducible complex character $\chi \in \mathrm{Irr}(G)$ of a finite group G is monomial. Recall that $\chi \in \mathrm{Irr}(G)$ is *monomial* if there is $\lambda \in \mathrm{Irr}(U)$ linear such that $\lambda^G = \chi$. It is known that every irreducible character of a supersolvable group is monomial, for instance, but this result depends more on the structure of the group rather than on the properties of the characters themselves. An exception is a theorem by R. Gow of 1975 ([3]): an odd degree real valued irreducible character of a solvable group is monomial. Recently, we gave in [8] an extension of this theorem which also dealt with the degree and the field of values of the character. (Yet another similar monomiality criterium was given in [9]: if the field of values $\mathbb{Q}(\chi)$ of χ is contained in the cyclotomic field $\mathbb{Q}_n$ and $(\chi(1), 2n) = 1$, then χ is monomial whenever G is solvable.) In this note, we apply non-trivial Isaacs π-theory of solvable groups to give a shorter proof of the above result at the same time that we gain some new information about the subnormal constituents of the characters, among other things. It does not seem easy at all to prove these new facts without using this deep theory.

Recall that for every solvable group and any set of primes π, M. Isaacs defined a canonical subset $B_\pi(G)$ of $\mathrm{Irr}(G)$ with remarkable properties ([4]). Since, by definition, every $\chi \in B_\pi(G)$ is induced from a character of π-degree, it is clear that B_π-characters of π'-degree are monomial.

Theorem 1.1 *Let p be a prime, let G be a p-solvable finite group, and let $P \in \mathrm{Syl}_p(G)$. Let $\chi \in \mathrm{Irr}(G)$ be such that p does not divide $\chi(1)$ and such that $\mathbb{Q}(\chi) \subseteq \mathbb{Q}_{p^a}$ for some $a \geq 0$. If $|\mathbf{N}_G(P)/P|$ is odd, then $\chi \in B_p(G)$. In particular, if $N \lhd\lhd G$ and θ is an irreducible constituent of χ_N, then θ is monomial.*

We obtain the following consequence, in which a global invariant of a finite group (that can be calculated in its character table) is computed locally.

Corollary 1.2 *Let p be a prime, let G be a p-solvable finite group, and let $P \in \mathrm{Syl}_p(G)$. Assume that $\mathbf{N}_G(P)/P$ has odd order. Then the number of irreducible characters χ of G such that $\chi(1)$ is not divisible by p and $\mathbb{Q}(\chi) \subseteq \mathbb{Q}_{|G|_p}$ is the number of orbits of the natural action of $\mathbf{N}_G(P)$ on P/P'.*

2 Proofs

The notation for characters is that of [6]. The π-special characters were defined in [2], while B_π-characters were defined in [4].

Lemma 2.1 *Suppose that G is a finite p-solvable group. Let $P \in \mathrm{Syl}_p(G)$, and assume that $\mathbf{N}_G(P)/P$ has odd order. If $\alpha \in \mathrm{Irr}(G)$ is p'-special and real, then α is the trivial character.*

Proof We argue by induction on $|G|$. Let $K = \mathbf{O}_p(G)$. If $K > 1$, then $K \subseteq \ker\alpha$ by Corollary (4.2) of [2], and we apply induction in G/K. Otherwise, let $K = \mathbf{O}_{p'}(G)$. Since α has p'-degree, then the set Ω of irreducible constituents of α_K has a p'-number of elements, using Clifford's Theorem. Also, by Clifford's Theorem, we have that G acts transitively on Ω. By elementary group theory, it follows that P fixes a point in Ω, and that two points fixed by P are $\mathbf{N}_G(P)$-conjugate. Let $\theta \in \Omega$. Since α is real, then $\bar\theta$ is also under α, and therefore there is $g \in \mathbf{N}_G(P)$ such that $\bar\theta = \theta^g$. Now g^2 fixes θ, and since $\mathbf{N}_G(P)/P$ has odd order, we see that $\bar\theta = \theta$. Since $\mathbf{N}_G(P)/P$ has odd order, we have that $\mathbf{C}_K(P)$ has odd order. Let $\theta^* \in \mathrm{Irr}(\mathbf{C}_K(P))$ be the P-Glauberman correspondent of θ. Since the Glauberman correspondence commutes with Galois automorphisms, we have that θ^* is a real irreducible character of a group of odd order. By Burnside's theorem, $\theta^* = 1$ and $\theta = 1$ by the uniqueness of the Glauberman correspondence. Thus $K \subseteq \ker\theta$, and we apply induction. □

We are now ready to prove our main results.

Proof [Proof of Theorem 1.1 and Corollary 1.2] By Theorem (3.6) of [5], there exists a subgroup $P \subseteq W \subseteq G$ and a p-special linear character $\lambda \in \mathrm{Irr}(W)$, such that: $\psi = \lambda^G \in \mathrm{Irr}(G)$ is a B_p-character, and (W, λ) is a nucleus of ψ. Also, there is a p'-special character $\alpha \in \mathrm{Irr}(W)$ such that $\chi = (\lambda\alpha)^G$. By Theorem 4.2 of [7], the pair $(W, \lambda\alpha)$ is unique up to G-conjugacy. Now, let $\sigma \in \mathrm{Gal}(\mathbb{Q}_{|G|}/\mathbb{Q}_{|G|_p})$ be the unique Galois automorphism that complex conjugates the p'-roots of unity and fixes p-power roots of unity. Since χ and λ are fixed by σ, then we deduce that there is $g \in G$ such that $(W^g, \lambda^g\alpha^g) = (W, \lambda\alpha^\sigma)$ by uniqueness. Hence $\alpha^g = \alpha^\sigma$ by Proposition (7.1) of [2]. Since $P, P^g \subseteq W$, then $P^{gw} = P$ for some $w \in W$, and we may assume that $g \in \mathbf{N}_G(P)$. Also, $\alpha^{g^2} = \alpha$, and therefore since $\mathbf{N}_G(P)/P$ has odd order, we see that $\alpha^\sigma = \alpha$. Now, let H be a p-complement of W. Then

$$\bar\alpha_H = \overline{\alpha_H} = (\alpha^\sigma)_H = \alpha_H$$

and we deduce that $\bar\alpha = \alpha$, by using Proposition (6.1) of [2]. Since $\mathbf{N}_W(P)/P$ has odd order, by Lemma (2.1), we have that $\alpha = 1$. Thus $\chi = \psi \in B_p(G)$ and χ is monomial. Now, to finish the proof of the theorem, use that subnormal constituents of B_p-characters are B_p-characters (Corollary (7.5) of [4]).

Recall that B_π-characters have their values in $\mathbb{Q}_{|G|_\pi}$ by their definition and part (a) of Proposition 6.3 of [2]. We have proved that the irreducible characters of G of p'-degree and field of values contained in $\mathbb{Q}_{|G|_p}$ are exactly the B_p-characters of p'-degree of G. Suppose now that $\gamma \in B_p(\mathbf{N}_G(P))$ has p'-degree. Let $\lambda \in \mathrm{Irr}(P)$ be linear under γ and let $\hat{\lambda} \in \mathrm{Irr}(T)$ be the canonical extension of λ to T (Corollary (8.16) of [6]). By Gallagher's Corollary (6.17) of [6] and the Clifford correspondence, let $\alpha \in \mathrm{Irr}(T/P)$ be such that $(\alpha\hat{\lambda})^{\mathbf{N}_G(P)} = \delta$. Now the argument in the first paragraph of this proof (with T instead of W) shows that $\alpha = 1$. Hence the number of B_p-characters in $\mathbf{N}_G(P)$ of p'-degree is the number of orbits of the action of $\mathbf{N}_G(P)$ on $\mathrm{Irr}(P/P')$. Now Corollary 1.2 follows from Theorem 2.2 and Corollary 2.3 of [1]. □

References

[1] P. Centella and G. Navarro, Correspondences between constituents of projective characters, *Arch. Math.* **90** (2008), 289–294.

[2] D. Gajendragadkar, A characteristic class of characters of finite π-separable groups, *J. Algebra* **59** (1979), 237–259.

[3] R. Gow, Real-valued characters of solvable groups, *Bull. London Math. Soc.* **7** (1975), 132.

[4] M. Isaacs, Characters of π-separable groups, *J. Algebra* **86** (1984), 98–128.

[5] M. Isaacs and G. Navarro, Characters of p'-degree of p-solvable groups, *J. Algebra* **246** (2001), 394–413.

[6] M. Isaacs, *Character Theory of Finite Groups* (AMS Chelsea Publishing, Providence, RI, 2006).

[7] G. Navarro, New properties of the π-special characters, *J. Algebra* **187** (1997), 201–213.

[8] G. Navarro and C. Vallejo, Certain monomial characters, *Archiv. Math.* **99** (2012), 407–411.

[9] C. Vallejo, A criterium for monomiality, *Red. Sem. Mat. Uni. Padova*, to appear.

RECOGNITION OF FINITE QUASI-SIMPLE GROUPS BY THE DEGREES OF THEIR IRREDUCIBLE REPRESENTATIONS

HUNG NGOC NGUYEN* and HUNG P. TONG-VIET†

*Department of Mathematics, The University of Akron, Akron, Ohio 44325, United States
Email: hungnguyen@uakron.edu

†School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Scottsville 3209, South Africa
Email: Tongviet@ukzn.ac.za

Abstract

We present some recent advances in the study of the problem of recognizing finite groups by the degrees of their irreducible complex representations. We especially focus on simple groups and more generally quasi-simple groups.

1 Introduction

Representation theory of finite groups was originally developed to analyze groups in terms of linear transformations or matrices. A representation of *degree* n (where n is a positive integer) over a field $\mathbb{F}$ of a group is a way to represent elements in the group by $n \times n$ invertible matrices with entries in $\mathbb{F}$ in such a way that the rule of group operation corresponds to matrix multiplication. Degree certainly is the most important piece of information in a representation, and therefore the degrees of irreducible representations are a key tool to study the structure of finite groups.

This is an expository paper in which we survey some recent advances on the problem of recognizing finite groups by the degrees of their (complex) representations, especially for simple groups and more generally quasi-simple groups. For a finite group G, we denote the set of degrees of irreducible representations of G by $\mathrm{cd}(G)$ and call it the *degree set* of G. The multiplicity of each degree is the number of irreducible representations of that degree, and when these numbers are taken into account, we will similarly have the *degree multiset* of G, denoted by $\mathrm{cd}^*(G)$.

A fundamental question in group representation theory is whether one can recover a group or some of its properties from the degrees of its irreducible representations. In the late 1980s, I. M. Isaacs [19] proved that if $\mathrm{cd}^*(G) = \mathrm{cd}^*(H)$ and p is a prime, then G has a normal p-complement if and only if H has a normal p-complement, and therefore the *nilpotency* of a group is determined by its degree multiset. Later, T. Hawkes [14] provided a counterexample showing that the same assertion does not hold for *super-solvability*. It is still unknown whether the *solvability* of a finite group is determined by its degree multiset, see [30, Problem 11.8].

In his famous list of problems in representation theory of finite groups [9], R. Brauer asked: *when do non-isomorphic groups have isomorphic complex group algebras*? (see [9, Problem 2]). The complex group algebra of a finite group G, denoted by $\mathbb{C}G$, is isomorphic to a direct sum of matrix algebras over $\mathbb{C}$ whose dimensions are exactly

the degrees of irreducible representations of G. Therefore, Brauer's question leads to the following:

Problem 1.1 Given a finite group G, determine all finite groups (up to isomorphism) having the same degree multiset as G.

A complete solution to this problem seems out of reach based on the present knowledge of group representation theory, but one may hope to obtain a partial solution.

Problem 1.1 is easy for abelian groups but difficult for solvable groups in general. This is due to the fact that the connection between a solvable group and its degree multiset is rather loose, in the sense that there are often several non-isomorphic groups having the same degree multiset as a given solvable group. In contrast to solvable groups, simple groups or groups 'close' to simple seem to have a stronger connection with their representation degrees.

In Section 2, we sketch some main ideas in the solution of Problem 1.1 for simple groups, mainly due to the second author. In Section 3, we discuss the problem for other groups close to simple such as *quasi-simple groups* and *almost simple groups*. Especially, we will present a method to approach the conjecture that every quasi-simple group is determined uniquely up to isomorphism by its degree multiset.

We have seen a tight connection between a quasi-simple group and its multiset of irreducible representation degrees. In the late 1990s, B. Huppert proposed that the connection should be tighter, at least for non-abelian simple groups. In fact, he conjectured in [16] that if G is a finite group and S is a finite non-abelian simple group such that the degree sets of G and S are equal, then G is isomorphic to the direct product of S and an abelian group. To give some evidence, Huppert himself verified the conjecture on a case-by-case basis for many simple groups, including the Suzuki groups, many of the sporadic simple groups, and a few of the simple groups of Lie type. Recently, there has been substantial progress on verifying Huppert's conjecture by the authors and their collaborators, especially for various families of simple groups of Lie type of small rank. This is discussed in Section 4.

Recent success on Problem 1.1 for quasi-simple groups suggests that Huppert's conjecture might be extended from non-abelian simple groups to quasi-simple groups. The following conjecture has been recently proposed in [15] and will be discussed in Section 5.

Conjecture 1.2 *Let G be a finite group and H a finite quasi-simple group. If $\mathrm{cd}(G) = \mathrm{cd}(H)$, then $G \cong H \circ A$, a central product of H with an abelian group A. In other words, every finite quasi-simple group is determined up to an abelian central product factor by its degree set.*

In the last section, we report some recent results concerning the recognition of non-abelian simple groups by using the multiplicity of irreducible representation degrees. Notice that if the complex group algebra $\mathbb{C}G$ of some finite group G is given, then we know both $\mathrm{cd}(G)$ and the multiplicity pattern $\mathrm{mp}(G)$ (defined in Section 6). It is shown in [52] that several families of non-abelian simple groups are uniquely determined by the multiplicity patterns and we conjecture that every non-abelian simple group is uniquely determined by its multiplicity pattern.

Perhaps the best way to describe complex representations (and indeed modular representations as well) is by characters, as a complex representation of a finite group is determined (up to equivalence) by its character. The character afforded by a group representation is a function on the group which associates to each group element the trace of the corresponding matrix and therefore it carries the essential information about the representation in a more condensed form. The degree of a character is exactly the degree of the representation affording it. This explains why we have used the notations $\mathrm{cd}(G)$ and $\mathrm{cd}^*(G)$ (c.d. stands for character degree) for the set and multiset of irreducible representation degrees. Throughout the paper, we will go back and forth between representations and characters, depending on which one is more convenient.

2 Simple groups and their degree multisets

If G is any finite abelian group of order n, then $\mathbb{C}G$ is isomorphic to a direct sum of n copies of $\mathbb{C}$ so that the degree multisets of any two abelian groups having the same order are equal. For p-groups or more generally nilpotent groups, the probability that two groups have equal degree multisets is also fairly 'high'. For instance, Huppert pointed out in [16] that among 2328 groups of order 2^7, there are only 30 different degree multisets. In [13], Dade even managed to construct two non-isomorphic metabelian groups G and H with isomorphic group algebras $\mathbb{F}G$ and $\mathbb{F}H$ over an arbitrary field $\mathbb{F}$.

We now turn our attention to finite non-abelian simple groups. As mentioned above, simple groups seem to have a stronger connection with their representation degrees. In [49, 50, 51], the second author has succeeded in proving that, if G is a finite group and S is a finite non-abelian simple group such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$, then $G \cong S$. This substantially improves a classical result of W. Kimmerle in [20] where it was proved that $G \cong S$ if $\mathbb{F}G \cong \mathbb{F}S$ for every field $\mathbb{F}$.

Obverse that knowing $\mathrm{cd}^*(G)$ is equivalent to knowing $\mathbb{C}G$ and $\mathrm{cd}^*(G)$ is just the first column of the ordinary character table of G. There are several papers in the literature devoted to characterizing the non-abelian simple groups by their character tables, see [25, 41, 42, 43] for instance. Upon the completion of the classification, it is easy to see that all non-abelian simple groups are uniquely determined by their character tables. In fact, as the normality of subgroups of a group can be detected from the character table, if finite groups G and S have the same character table, where S is non-abelian simple, then G is also non-abelian simple, and furthermore $|G| = |S|$. Now by applying Artin's Theorem [21, Theorem 5.1], we have

$$\{G, S\} = \{\mathrm{PSL}_4(2), \mathrm{PSL}_3(4)\}$$

or

$$\{G, S\} = \{\mathrm{PSp}_{2n}(q), \Omega_{2n+1}(q)\},$$

where $n \geq 3$ and q is an odd prime power. For these exceptions, we can easily check that they have distinct character tables. Thus, we have proved that if a finite group G and a finite non-abelian simple group S have the same character table, then $G \cong S$.

Indeed, the ordinary character table reflects much more of the group structure. By W. Kimmerle [22], the character table of a finite group determines the chief series and chief factors of the group, and therefore its composition factors as well.

The main result in this section, stated below, gives a new characterization of finite non-abelian simple groups by using the first column of their ordinary character tables or equivalently, by their complex group algebras.

Theorem 2.1 ([49, 50, 51]) *If G is a finite group and S is a finite non-abelian simple group such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$, then $G \cong S$. In other words, every non-abelian simple group is determined uniquely up to isomorphism by its degree multiset.*

The result above is a weaker version of Huppert's conjecture proposed by B. Huppert in the late 1990s, which will be considered in Section 4. Indeed, if H is any non-abelian simple group and G is a finite group such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$, then $\mathrm{cd}(G) = \mathrm{cd}(S)$ and $|G| = |S|$. In particular, if Huppert conjecture holds for the simple group S, then since $\mathrm{cd}(G) = \mathrm{cd}(S)$, we deduce that $G \cong S \times A$, for some abelian group A. By comparing the orders, we deduce that $G \cong S$ as wanted.

As expected, we have made use of the classification of finite simple groups in the proof of Theorem 2.1. We know that every non-abelian simple group is a sporadic simple groups, an alternating group of degree at least 5, or a finite simple group of Lie type. The latter class can be divided into the simple classical groups and the simple exceptional groups of Lie type. We first record some easy properties of groups G and S under the hypothesis that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$, where S is non-abelian simple. In the following lemma, the notation $\pi(G)$ stands for the set of prime divisors of the order of G.

Lemma 2.2 *Assume that G and S satisfy the hypothesis of Theorem 2.1. Then*

(1) $|G| = |S|$,

(2) $\mathrm{cd}(G) = \mathrm{cd}(S)$,

(3) *G is perfect, i.e., $G = G'$,*

(4) *If N is a maximal normal subgroup of G, then G/N is a non-abelian simple group and $\mathrm{cd}(G/N) \subseteq \mathrm{cd}(S)$. Furthermore, the ith-smallest nontrivial degree of G/N is greater than or equal that of S and $\pi(G/N) \subseteq \pi(S)$.*

For simple groups which are not classical, we managed to prove the following result, which is the main part of the proof of Theorem 2.1 for these groups.

Proposition 2.3 *Let S be a sporadic simple group, the Tits group, an alternating group of degree at least 7, or a finite simple exceptional group of Lie type and let T be any non-abelian simple group. If $\mathrm{cd}(T) \subseteq \mathrm{cd}(S)$, then $T \cong S$.*

For each of the sporadic simple groups, the Tits group and the alternating groups of degree at least 5, and each possibility of the non-abelian simple group T, we compare several smallest nontrivial degrees of S and T using results of F. Lübeck [26] and R. Rasala [44] and also the classification of prime power character degrees of quasisimple groups by G. Malle and A. E. Zalesskii [29] to eliminate all but the simple groups which are isomorphic to the given simple group S. For the exceptional simple

groups of Lie type, apart from the previous results, we also made use of the explicit list of character degrees of these simple groups by F. Lübeck [27].

Let G be a finite group and let S be any non-abelian simple group which appears in Proposition 2.3 such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$. By Lemma 2.2(4), $\mathrm{cd}(G/N) \subseteq \mathrm{cd}(S)$ and G/N is non-abelian simple. Proposition 2.3 yields that $G/N \cong S$ and thus $|G/N| = |S|$. By Lemma 2.2(1), we obtain that $|G| = |S|$ and so $|N| = 1$, which implies that $G \cong S$. This gives a proof of Theorem 2.1 for the simple group S.

For finite simple classical groups, we can prove Proposition 2.3 for these groups by a similar method. However the proof is quite long and complicated. Instead, we have used a different approach. Assume now that S is a finite simple classical group in characteristic p and G is a finite group such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(S)$. With the same method as above, we can deduce that G/N is a finite simple group of Lie type in the same characteristic p.

If N is trivial, then G is a finite simple group of Lie type in characteristic p, $\mathrm{cd}(G) = \mathrm{cd}(S)$ and $|G| = |S|$. Using Artin's Theorem mentioned earlier, we have $\{G, S\} = \{\mathrm{PSp}_{2n}(q), \Omega_{2n+1}(q)\}$, where $n \geq 3$ and q is an odd prime power or $\{G, S\} = \{\mathrm{PSL}_4(2), \mathrm{PSL}_3(4)\}$. The latter case can be eliminated easily. For the former case, using the existence of the Weil characters of $\mathrm{PSp}_{2n}(q)$ with odd prime power q and the minimal characters of $\Omega_{2n+1}(q)$ in [45, 46], one sees that these two groups have different degree sets and thus $G \cong S$ in this case.

Assume that N is nontrivial. As the Steinberg character St_S of S of degree $|S|_p$ is the only irreducible character of S of nontrivial p-power degree and $\mathrm{St}_{G/N} \in \mathrm{Irr}(G/N)$ also has nontrivial p-power degree which is $|G/N|_p$ we deduce that $|G/N|_p = |G|_p$. If N is non-solvable, then N possesses a nontrivial irreducible character φ which is extendible to $\varphi_0 \in \mathrm{Irr}(G)$ (see [8, Lemma 5]) and then by Gallagher's Theorem [18, Corollary 6.17], we have $\varphi_0\mathrm{St}_{G/N} \in \mathrm{Irr}(G)$ and hence $\varphi_0(1)|G/N|_p = \varphi(1)|S|_p \in \mathrm{cd}(S)$, which is impossible as St_S is the only irreducible character of S of p-defect zero (see [12, Theorem 4]). Recall that an irreducible character $\chi \in \mathrm{Irr}(G)$ is of r-defect zero for some prime r if $\chi(1)_r = |G|_r$. Therefore, N must be a nontrivial solvable group. In order to arrive at a contradiction, we need the following result.

Lemma 2.4 *Let S be a finite simple group of Lie type and let G be a finite group. If $\mathrm{cd}(G) = \mathrm{cd}(S)$ and $|G| = |S|$, then the Fitting subgroup $F(G)$ of G is trivial.*

The proof of this lemma is an easy application of the existence of blocks of defect zero of finite simple groups of Lie type [59] and Ito's Theorem [18, Theorem 6.15].

Return to our problem, we see that $F(G)$ is nontrivial as N is a nontrivial solvable subgroup of G and furthermore $\mathrm{cd}(G) = \mathrm{cd}(S)$ and $|G| = |S|$. Now Lemma 2.4 will provide a contradiction.

Problem 1.1 is now done for finite simple groups. The next natural groups to be considered are perhaps the characteristically simple groups.

Question 2.5 Let G be a finite group and H a direct product of copies of a non-abelian simple group such that $\mathrm{cd}^*(G) = \mathrm{cd}^*(H)$. Can we conclude that $G \cong H$?.

3 Quasi-simple groups and their degree multisets

We have seen in Section 2 that every simple group is determined uniquely up to isomorphism by its degree multiset. What other groups have the same property? It was proved that all the symmetric groups are determined by their degree multisets by M. Nagl in [33] and independently by the second author in [48], improving an old result of H. Nagao [32] that if G is a finite group whose character table agrees, up to a permutation of its rows and columns, with the character table of the symmetric group S_n, then $G \cong \mathsf{S}_n$. This suggests that finite groups 'close' to simple might be determined by their degree multisets. We have proposed in [38]:

Conjecture 3.1 *Let G be a finite group and H a finite quasi-simple group. If $\mathrm{cd}^*(G) = \mathrm{cd}^*(H)$, then $G \cong H$. In other words, every finite quasi-simple group is determined uniquely up to isomorphism by its degree multiset.*

Let us recall that a finite group H is said to be *quasi-simple* if H is perfect and $H/\mathbf{Z}(H)$ is non-abelian simple, in which case we also say that H is a *perfect central cover* or simply a *cover* of $H/\mathbf{Z}(H)$.

Conjecture 3.1 indeed has been predicted earlier by the first author in [37], where he proved that every quasi-simple *classical group* H is uniquely determined up to isomorphism by its degree multiset, except possibly when $H/\mathbf{Z}(H)$ is isomorphic to $\mathrm{PSL}_3(4)$ or $\mathrm{PSU}_4(3)$. The main ideas can be summarized as follows.

As H is perfect, it has a unique linear character. Therefore G has a unique linear character as well and hence G is perfect. It follows that, if M is a maximal normal subgroup of G then G/M is non-abelian simple. The *first step* in the proof is quite similar to what we have done for simple groups; that is, to show that

$$G/M \text{ is isomorphic to } S := H/Z(H).$$

This basically eliminates the involvement of all non-abelian simple groups other than S in the structure of G. Let $\mathrm{Schur}(S)$ denote the Schur cover (or the covering group) of S. We have

$$\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(G) = \mathrm{cd}^*(H) \subseteq \mathrm{cd}^*(\mathrm{Schur}(S)),$$

where the last containment comes from the fact that every cover of S is a quotient of the Schur cover of S. This condition together with some others if necessary can be used to force two non-abelian simple groups G/M and $H/\mathbf{Z}(H)$ to be isomorphic. On the way to the proof of Conjecture 3.1 for quasi-simple classical groups, we in fact prove the following:

Proposition 3.2 *Let S be a simple classical group. Let G be a perfect group such that $|S| \mid |G| \mid |\mathrm{Schur}(S)|$ and $\mathrm{cd}(S) \subseteq \mathrm{cd}(G) \subseteq \mathrm{cd}(\mathrm{Schur}(S))$. If M is a maximal normal subgroup of G, then $G/M \cong S$.*

The proof of this proposition depends heavily on the results on prime power representation degrees, due to G. Malle and A. E. Zalesskii [29], and relatively small character degrees of quasi-simple classical groups, due to P. H. Tiep and A. E. Zalesskii [45] and H. N. Nguyen [35].

The *second step* is to show that $G \cong H$. Since $|G| = |H|$ and $G/M \cong H/Z(H)$, we deduce that $|M| = |Z(H)|$. It follows that, if H is simple then M is trivial and we have immediately that $G \cong H$. However, since we are working with quasi-simple groups, the problem becomes more difficult; especially for quasi-simple groups with complicated centers such as the covers $\mathrm{PSL}_3(4)$ or $\mathrm{PSU}_4(3)$ whose the Schur multipliers are very exceptional. We have done this in a series of lemmas. We reproduce the proofs of some of them.

The first two lemmas follow from the classification of finite simple groups.

Lemma 3.3 *Let S be a non-abelian simple group. Let A be an abelian group such that $|A| \leq |\mathrm{Mult}(S)|$. Then $|S| > |\mathrm{Aut}(A)|$.*

Lemma 3.4 *Let S be simple group of Lie type. Then no proper multiple of $\mathrm{St}_S(1)$ is a degree of $\mathrm{Schur}(S)$.*

For each nonnegative integer i, let $M^{(i)}$ denote the ith derived subgroup of M.

Lemma 3.5 *Let S be a non-abelian simple group. Let G be a perfect group and $M \lhd G$ such that $G/M \cong S$ and $|M| \leq |\mathrm{Mult}(S)|$. Then, for every nonnegative integer i, the quotient $G/M^{(i)}$ is isomorphic to a quotient of $\mathrm{Schur}(S)$.*

Proof We prove by induction that $G/M^{(i)}$ is isomorphic to a quotient of $\mathrm{Schur}(S)$ for every i. The induction base $i = 0$ is exactly the hypothesis. Assuming that $G/M^{(i)} \cong \mathrm{Schur}(S)/Z_i$ for some normal subgroup Z_i of $\mathrm{Schur}(S)$, we will show that $G/M^{(i+1)}$ is also a quotient of $\mathrm{Schur}(S)$.

As $M^{(i)}/M^{(i+1)}$ is abelian and normal in $G/M^{(i+1)}$, we have

$$\frac{M^{(i)}}{M^{(i+1)}} \leq \mathbf{C}_{G/M^{(i+1)}}\left(\frac{M^{(i)}}{M^{(i+1)}}\right) \trianglelefteq \frac{G}{M^{(i+1)}}.$$

We first consider the case $\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)}) = G/M^{(i+1)}$. Then $M^{(i)}/M^{(i+1)}$ is central in $G/M^{(i+1)}$. As G is perfect, $G/M^{(i+1)}$ is a stem extension of $G/M^{(i)} \cong \mathrm{Schur}(S)/Z_i$. As $\mathrm{Schur}(S)/Z_i$ is a quasi-simple group whose quotient by the center is S, we deduce that $G/M^{(i+1)}$ is a quotient of the Schur cover of $\mathrm{Schur}(S)/Z_i$. Therefore, $G/M^{(i+1)}$ is a quotient of $\mathrm{Schur}(S)$, as wanted.

The lemma is completely proved if we can show that $\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)})$ cannot be a *proper* normal subgroup of $G/M^{(i+1)}$. Assume so, then it follows by the induction hypothesis that

$$\frac{\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)})}{M^{(i)}/M^{(i+1)}} \lhd \frac{G/M^{(i+1)}}{M^{(i)}/M^{(i+1)}} \cong \frac{G}{M^{(i)}} = \frac{\mathrm{Schur}(S)}{Z_i}.$$

Therefore,

$$\left|\frac{\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)})}{M^{(i)}/M^{(i+1)}}\right| \leq \left|\frac{\mathrm{Mult}(S)}{Z_i}\right| = \left|\frac{M}{M^{(i)}}\right|$$

and hence

$$|\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)})| \leq |M/M^{(i+1)}|.$$

Thus

$$\Big|\frac{G/M^{(i+1)}}{\mathbf{C}_{G/M^{(i+1)}}(M^{(i)}/M^{(i+1)})}\Big| \geq |G/M| = |S|.$$

Since the quotient group on the left side can be embedded in $\mathrm{Aut}(M^{(i)}/M^{(i+1)})$ and $M^{(i)}/M^{(i+1)}$ is abelian of order less than or equal to $|M|$, this last inequality leads to a contradiction by Lemma 3.3. □

Since the Schur multipliers of the alternating groups and sporadic simple groups are 'small', the proof for these groups are easier and therefore, let us just focus on simple groups of Lie type.

Lemma 3.6 *Let S be a simple group of Lie type. Let G be a perfect group and $M \lhd G$ such that $G/M \cong S$, $|M| \leq |\mathrm{Mult}(S)|$, and $\mathrm{cd}(G) \subseteq \mathrm{cd}(\mathrm{Schur}(S))$. Then G is isomorphic to a quotient of $\mathrm{Schur}(S)$.*

Proof By Lemma 3.5, we are done if M is solvable. So it remains to consider the case when M is nonsolvable. If M is nonsolvable, there is an integer i such that

$$M^{(i)} = M^{(i+1)} > 1.$$

Let $N \leq M^{(i)}$ be a normal subgroup of G so that $M^{(i)}/N \cong T^k$ for some non-abelian simple group T. By [31, Lemma 4.2], T has a non-principal irreducible character φ that extends to $\mathrm{Aut}(T)$. Now [8, Lemma 5] implies that φ^k extends to G/N. Therefore, by Gallagher's lemma, $\varphi^k\chi \in \mathrm{Irr}(G/N)$ for every $\chi \in \mathrm{Irr}(G/M^{(i)})$. In particular,

$$\varphi(1)^k\chi(1) \in \mathrm{cd}(G/N) \subseteq \mathrm{cd}(G) \subseteq \mathrm{cd}(\mathrm{Schur}(S)).$$

Taking χ to be the Steinberg character of S. By Lemma 3.5, S is a quotient of $G/M^{(i)}$ and hence χ can be considered as a character of of $G/M^{(i)}$. We now get a contradiction since $\varphi(1)^k\chi(1)$, which is larger than $\chi(1) = \mathrm{St}_S(1)$, can not be degree of $\mathrm{Schur}(S)$ by Lemma 3.4. □

Lemma 3.7 *Let S be a simple group of Lie type different from $\mathrm{PSL}_3(4)$, $\mathrm{PSU}_4(3)$, and $\mathrm{P\Omega}^+_{2n}(q)$ with n even and q odd. Let G be a perfect group and $M \lhd G$ such that $G/M \cong S$, $|M| \leq |\mathrm{Mult}(S)|$, and $\mathrm{cd}(G) \subseteq \mathrm{cd}(\mathrm{Schur}(S))$. Then G is uniquely determined (up to isomorphism) by S and the order of G.*

Remark 3.8 The exceptions in the lemma are true exceptions. For instance, let $S = \mathrm{PSL}_3(4)$. Then $\mathrm{Mult}(S) = \mathbb{Z}_4 \times \mathbb{Z}_4 \times \mathbb{Z}_3$. Let Z_1 and Z_2 be subgroups of $\mathrm{Mult}(S)$ isomorphic respectively to $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$. The non-isomorphic groups $\mathrm{Schur}(S)/Z_1$ and $\mathrm{Schur}(S)/Z_2$ (see [10] where these groups are denoted by $12_1.S$ and $12_2.S$) both satisfies the hypothesis of the lemma.

Proof First we consider the case $S = \mathrm{P\Omega}^+_8(2)$ or $\mathrm{Suz}(8)$. Then $\mathrm{Mult}(S) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. By Lemma 3.6, S is isomorphic to a quotient of $\mathrm{Schur}(S)$ so that we can assume $G \cong \mathrm{Schur}(S)/Z$ with $Z \leq \mathrm{Mult}(S)$ (note that Z cannot be $\mathrm{Schur}(S)$). If $|M| = 1$ or 4 then $Z = \mathrm{Mult}(S)$ or 1, respectively, and so we are done. Thus it remains to consider $|M| = 2$. We then have $|Z| = 2$ and hence Z is generated by an involution of $\mathrm{Mult}(S)$.

However, as these three involutions are permuted by an outer automorphism of S of degree 3, the quotient groups of the form $\mathrm{Schur}(S)/\langle t\rangle$ for any involution $t \in \mathrm{Mult}(S)$ are isomorphic and we are done again. Next, we assume that $S = \mathrm{PSU}_6(2)$ or ${}^2\mathrm{E}_6(2)$. Though the Schur multipliers of these groups are more complicated, these cases in fact can be argued similarly as above.

If S is none of the groups already considered and also $S \neq \mathrm{PSL}_3(4)$ and $\mathrm{PSU}_4(3)$, then $\mathrm{Mult}(S)$ indeed is cyclic. Again, S is isomorphic to a quotient of $\mathrm{Schur}(S)$ and we can assume

$$G \cong \mathrm{Schur}(S)/Z,$$

where $Z \leq \mathrm{Mult}(S)$. As $G/M \cong S$, we then deduce that $|Z| = |\mathrm{Mult}(S)|/|M| = |\mathrm{Schur}(S)|/|G|$. Since the cyclic group $\mathrm{Mult}(S)$ has a unique subgroup of order $|\mathrm{Schur}(S)|/|G|$, Z is uniquely determined by S and $|G|$ and the lemma follows. □

Following the method outlined above, we have shown in [38] that:

Theorem 3.9 *Every finite quasi-simple group except possibly the Schur double covers of the alternating groups is uniquely determined up to isomorphism by its degree multiset.*

For now, we are unable to establish the first step for the Schur double covers of the alternating groups. Specifically, we do not yet know how to eliminate the case where $H = \mathrm{Schur}(\mathsf{A}_n)$ and G/M is a simple group of Lie type in even characteristic.

There is an ongoing work on this in [6] where the authors also study Problem 1.1 for the Schur double covers of symmetric groups. The symmetric group S_n has two isomorphism classes of Schur double covers, denoted by $\hat{\mathsf{S}}_n^-$ and $\hat{\mathsf{S}}_n^+$. It is well known that the group algebras $\mathbb{C}\hat{\mathsf{S}}_n^+$ and $\mathbb{C}\hat{\mathsf{S}}_n^-$ are canonically isomorphic and therefore $\hat{\mathsf{S}}_n^+$ and $\hat{\mathsf{S}}_n^-$ are not uniquely determined by their degree multisets. Nevertheless, it is anticipated in [6] that

$$\text{if } \mathrm{cd}^*(G) = \mathrm{cd}(\hat{\mathsf{S}}_n^+) = \mathrm{cd}^*(\hat{\mathsf{S}}_n^-) \text{ then } G \cong \hat{\mathsf{S}}_n^+ \text{ or } \hat{\mathsf{S}}_n^-.$$

The proof of this, as expected, depends heavily on the representation theory of the symmetric and alternating groups, their Schur double covers, and quasi-simple groups in general. We particularly needs the results on relatively small degrees and prime power degrees of the regular and spin representations of the alternating and symmetric groups. These results are due to various authors, including A. Balog, C. Bessenrodt, J. B. Olsson, and K. Ono [4], C. Bessenrodt and J. B. Olsson [7], and A. Kleshchev and P. H. Tiep [23, 24].

A group is said to be *almost simple* if it contains a non-abelian simple group and is contained within the automorphism group of that simple group. We end this section by a question.

Question 3.10 What are almost simple groups that are uniquely determined (up to isomorphism) by their degree multisets?

4 Character degrees of simple groups and Huppert's conjecture

In Sections 2 and 3, we have seen that the degree multiset of a finite group encodes a lot of structural information of the group. In this section and the next one, we will drop the multiplicities of the character degrees and focus on the degrees only. The multiplicities will be considered in Section 6. We will see that the degree set also provide some information on the group structure.

In general, the character degree set of G does not completely determine the structure of G. As abelian groups have only the trivial character degree, it is easy to see that $\mathrm{cd}(G) = \mathrm{cd}(G \times A)$ for any abelian group A. There are many other examples without an abelian direct factor. For instance, the non-isomorphic groups D_8 and Q_8 not only have the same set of character degrees, but also share the same character table. The degree set also cannot be used to distinguish between solvable and nilpotent groups, as the groups Q_8 and S_3 show. Recently, Gabriel Navarro constructed a finite perfect group of order 37500 and a finite solvable group, both have the same character degree set. Thus, the degree set cannot distinguish solvable and non-solvable groups either. However, it remains open whether the complex group algebras determine the solvability of the groups or not.

Huppert conjectured in the late 1990s that the non-abelian simple groups are essentially determined by the set of their character degrees. More explicitly, he proposed in [16] the following.

Conjecture 4.1 (Huppert's Conjecture) *Let S be any non-abelian simple group and let G be a group such that $\mathrm{cd}(G) = \mathrm{cd}(S)$. Then $G \cong S \times A$, where A is abelian.*

As the character degrees of $S \times A$ are the products of the character degrees of S and those of A, this result is the best possible. The hypothesis that S is a non-abelian simple group is critical. There cannot be a corresponding result for solvable groups. For example, if we consider the solvable group Q_8, then $\mathrm{cd}(\mathrm{Q}_8) = \mathrm{cd}(\mathrm{S}_3)$ but $\mathrm{Q}_8 \ncong \mathrm{S}_3 \times A$ for any abelian group A.

Huppert in [16, 17] and his unpublished preprints verified the conjecture on a case-by-case basis for many non-abelian simple groups, including the Suzuki groups, many of the sporadic simple groups, and a few of the simple groups of Lie type. Except for the Suzuki groups and the family of simple linear groups $\mathrm{PSL}_2(q)$, Huppert proved the conjecture only for specific simple groups of Lie type of small, fixed rank. He indeed provided a pattern to approach the conjecture.

(1) Show that $G' = G''$.

(2) Suppose that G'/M is a chief factor of G. Show that $G'/M \cong S$.

(3) Show that any linear character $\theta \in \mathrm{Irr}(M)$ is stable under G', which implies $[M, G'] = M'$.

(4) Show that M is trivial.

(5) Show that $G \cong G' \times \mathbf{C}_G(G')$.

This pattern and variants thereof have been successfully used to make significant progress on the verification of the conjecture for a number of families of simple groups, notably the sporadic simple groups, alternating groups of small degree, and simple groups of Lie type of small rank. It is hoped that more general techniques can be

developed to aid in the verification of the conjecture for simple groups of Lie type of higher rank and alternating groups of higher degree.

T. P. Wakefield verified Huppert's conjecture for most of the simple groups of Lie type of rank two as part of his dissertation research. These results appear in [60, 61, 62]. The conjecture was verified for $\mathrm{G}_2(5)$ by S. H. Alavi and A. Daneshkah in [1] and established more generally for $\mathrm{G}_2(q)$ by Wakefield and the second author in [57]. The conjecture for simple group $\mathrm{F}_4(2)$ was verified in [58]. The remaining sporadic simple groups also have been shown to satisfy the conjecture in [2, 3, 55]. In [47, 56], the result is established for the Steinberg triality and the Ree groups. For the alternating groups, it has been confirmed up to the degree 13 in [39]. We summarize all these results in the following theorem.

Theorem 4.2 *Let G be a finite group and S a sporadic simple group, an alternating group of degree at most* 13*, a simple group of Lie type of rank at most* 2 *or* $\mathrm{F}_4(2)$*. If* $\mathrm{cd}(G) = \mathrm{cd}(S)$*, then $G \cong S \times A$, where A is an abelian group.*

Let us now describe in more details some techniques as well as difficulties in accomplishing the steps in Huppert's method.

1. The first step is to show that if S is a non-abelian simple group and G is a group such that $\mathrm{cd}(G) = \mathrm{cd}(S)$, then $G' = G''$. This step can be done by using the techniques in [16] and is described in details in [60]. Suppose that $G'' < G'$, and take K to be maximal subject to K being normal in G and G/K being a non-abelian solvable group. This means G'/K is the unique minimal normal subgroup of G/K. The structure of finite groups with this property has been described explicitly in [18]. In particular, it then can be deduced that G/K is either a p-group or a Frobenius group whose kernel is an elementary abelian p-group for some prime p. With this special structure of G/K, we can identify a degree of G that is not a degree of S to get a contradiction.

2. The second step is to show that if G'/M is a chief factor of G then it is isomorphic to the simple group S. By Step 1, the quotient G'/M is a non-abelian chief factor of G and therefore it is the direct product of k copies of a non-abelian simple group T:

$$G'/M \cong \underbrace{T \times T \times \cdots \times T}_{k \text{ times}}.$$

To prove Step 2, we must show that $k = 1$ and $T \cong S$. Huppert's proofs of this step for some non-abelian simple groups mentioned above relies upon either very specific properties of the groups (for the Suzuki groups $\mathrm{Suz}(q)$ and the linear groups $\mathrm{PSL}_2(q)$) or the fact that the orders of the simple groups under consideration are divisible by very few primes.

The approach by Wakefield and the authors for Step 2 in [39, 40], described as follows, has proved to be effective for more complicated simple groups, especially finite groups of Lie type of higher rank. If α is an irreducible character of S that extends to $\mathrm{Aut}(S)$, then, by tensor induction, it can be shown that

$$\underbrace{\alpha \times \alpha \times \cdots \times \alpha}_{k \text{ times}} \text{ is in } \mathrm{Irr}(G'/M) \text{ and extends to } G/M.$$

It in particular follows that $\alpha(1)^k \in \mathrm{cd}(G)$ and therefore

$$\alpha(1)^k \in \mathrm{cd}(S).$$

If T is a simple group with "small" outer automorphism group such as an alternating or sporadic group, this gives many different characters α so that $\alpha(1)^k$ is a degree in $\mathrm{cd}(S)$. It is unlikely that $\mathrm{cd}(S)$ contains many degrees that are powers whose exponents are divisible by k, so it should be possible to show that $k = 1$.

The other possibility is when T is a simple group of Lie type. We know that the Steinberg character St of T extends to $\mathrm{Aut}(T)$. This yields $\mathrm{St}(1)^k$ as a degree in $\mathrm{cd}(S)$. Recall that $\mathrm{St}(1)$ is a power of p, where the prime p is the defining characteristic for S. It is known that prime powers as degrees of quasi-simple groups are relatively rare [29]. Other than a known finite list, the only possibilities are the Steinberg characters of groups of Lie type. Unless S is one of the exceptions, this forces $\mathrm{St}(1) \in \mathrm{cd}(T)$ to be the degree of the Steinberg character of S. From this, we can obtain a bound on the characteristic and size of the underlying field of T. We then find a degree of T which divide degrees of S of relatively large degree and this will allow us to establish a contradiction if $T \ncong S$.

3. In Step 3, we have to prove a technical result that every linear character of M is invariant in G'. Let $\theta \in \mathrm{Irr}(M)$ be such a character and let $I = I_{G'}(\theta)$ be the stabilizer of θ under the action of G' on $\mathrm{Irr}(M)$. If θ is *not* invariant in G', the inertia group I would be a proper subgroup of G'. This means I would be contained in a maximal subgroup, say U, of G'. Suppose that

$$\theta^I = \sum_i \phi_i, \text{ where } \phi_i \in \mathrm{Irr}(I).$$

Then $\phi_i^{G'} \in \mathrm{Irr}(G')$ and hence $\phi_i(1)|G' : I| \in \mathrm{cd}(G')$ by Clifford theory. It follows that

$$\phi_i(1)|G' : U||U : I| \text{ divides some degree of } G.$$

In particular, the index of U/M in $G'/M (\cong S$, by Step 2) divides a degree of G.

We have seen that, in Step 3, knowledge of maximal subgroups of finite simple groups plays an important role. In particular, we need to consider the maximal subgroups of S that have indices dividing degrees in $\mathrm{cd}(S)$. The idea is to show that if U/M is a maximal subgroup of G'/M, then U is not the stabilizer of any character in $\mathrm{Irr}(M)$. Fortunately, it seems rare that S has maximal subgroups whose indices divide degrees in $\mathrm{cd}(S)$.

4. The result obtained in Step 3 implies in particular that $[M, G'] = M'$ and $|M : M'|$ divides the order of $\mathrm{Mult}(G'/M) = \mathrm{Mult}(S)$. We recall that $\mathrm{Mult}(G)$ denote the Schur multiplier of G. In order to prove that M is trivial or equivalently $G' \cong S$, we have to analyze the differences between character degrees of various central extensions of the simple group S. In other words, the knowledge on the degrees of the irreducible projective representations of S is crucial in this step. This is expected to be complicated when the Schur multiplier of S is large.

5. The proof of Step 5 requires an understanding of the action of the automorphism group of a quasi-simple group on its irreducible representations. Let $C := \mathbf{C}_G(G')$ be the centralizer of G' in G. Then G/C is embedded into the automorphism group of

$G' \cong S$. One sees that $G/G'C$ will correspond to a subgroup of the outer automorphism group of S. The goal is to show that any outer automorphism of G' will cause fusion among the irreducible characters in $\mathrm{Irr}(G')$. This will produce a character degree in $\mathrm{cd}(G)$ that is not a degree in $\mathrm{cd}(S)$. In other words, we aim to show that

$$G = G'C.$$

Once this is proved, G will be a direct product of G' and C. It follows that $C \cong G/G'$ and hence C is abelian, as wanted.

Step 5 can be difficult when the outer automorphism group $\mathrm{Out}(S)$ of S is complicated. For instance, if S is simple of Lie type, the structure of $\mathrm{Out}(S)$ is

$$d \cdot f \cdot g,$$

where d is the group of *diagonal automorphisms*, f is the (cyclic) group of *field automorphisms* (generated by a Frobenius automorphism), and g is the group of *graph automorphisms* (coming from automorphisms of the Dynkin diagram). To understand the action of $\mathrm{Out}(S)$ on the set of irreducible characters of H, one has to study the action of these kinds of automorphisms individually. This topic might be of independent interest and we hope that Huppert's conjecture will motivate more study on the behavior of irreducible representations of a simple group under the action of its outer automorphisms.

With Wakefield, we have recently succeeded in applying these arguments to establish the conjecture for the simple linear and unitary groups in dimension 4. The result for the linear groups appears in [40] and its proof requires modifications in the five steps outlined above.

Theorem 4.3 ([40]) *Let $q \geq 13$ be a prime power and let G be a finite group such that $\mathrm{cd}(G) = \mathrm{cd}(\mathrm{PSL}_4(q))$. Then G is isomorphic to the direct product of $\mathrm{PSL}_4(q)$ and an abelian group.*

With further modifications of Huppert's method, the second author has recently completed the verification of Huppert conjecture for the remaining simple exceptional groups of Lie type in [54].

5 Extending Huppert's conjecture to quasi-simple groups

Recent success on Problem 1.1 for quasi-simple groups suggests that Huppert's conjecture might be extended from non-abelian simple groups to quasi-simple groups. In an ongoing work with Majozi and Wakefield [15], we put forward the following, which is Conjecture 1.2 in the Introduction.

Conjecture 5.1 *Let G be a finite group and H a finite quasi-simple group. If $\mathrm{cd}(G) = \mathrm{cd}(H)$, then $G \cong H \circ A$, a central product of H with an abelian group A. In other words, every finite quasi-simple group is determined up to an abelian central product factor by its degree set.*

As mentioned in Section 4, Huppert [16] outlined a pattern consisting of five steps to study his conjecture. This pattern and variants thereof have been successfully used

in verifying the conjecture for several non-abelian simple groups. Drawing upon his method, the following pattern is proposed in [15] to approach Conjecture 5.1.

(1) Show that $G' = G''$.
(2) Suppose that G'/M is a chief factor of G. Show that $G'/M \cong H/\mathbf{Z}(H)$.
(3) Show that G' is isomorphic to a perfect central cover of $H/\mathbf{Z}(H)$.
(4) Show that $G = G' \circ \mathbf{C}_G(G')$. It follows in particular that $\mathrm{cd}(G) = \mathrm{cd}(G')$.
(5) Show that covers of $H/\mathbf{Z}(H)$ have distinct sets of character degrees. Steps 3 and 4 then imply that $G' \cong H$ and hence G is isomorphic to the central product of H and the abelian group $\mathbf{C}_G(G')$.

It is worth pointing out that Steps $1, 2$ and 4 here correspond respectively to Steps $1, 2$ and 5 in Huppert's method for non-abelian simple groups, while Steps 3 and 5 are fundamentally different.

With extra work, the proof of Huppert conjecture for a certain non-abelian simple group can be extended to obtain the proof of Conjecture 5.1 for the perfect central cover covers of that simple group, although it is not always obvious. In [15], by following the pattern outlined above, the authors have established Conjecture 5.1 for all quasi-simple linear groups in dimensions 2 and 3.

6 Characterizing non-abelian simple groups by their multiplicity patterns

In this last section, we shift our focus on the multiplicities of character degrees of finite groups. Before we can present some results and open conjectures, we need some notation. For a finite group G, we write $\mathrm{cd}(G) = \{d_0, d_1, \cdots d_t\}$ with $d_0 = 1 < d_1 < \cdots < d_t$. For each positive integer d, the multiplicity of d in G, denoted by $m_G(d)$, is the number of irreducible characters of G of degree d, that is

$$m_G(d) = |\{\chi \in \mathrm{Irr}(G) : \chi(1) = d\}|.$$

The *multiplicity pattern* $\mathrm{mp}(G)$ of G is defined to be the vector

$$(m_G(d_0), m_G(d_1), \cdots, m_G(d_t)).$$

Clearly, the first coordinate of $\mathrm{mp}(G)$ is the number of linear characters of G, which is $|G : G'|$, and for $i \geq 1$, the $(i+1)$th-coordinate of $\mathrm{mp}(G)$ is the multiplicity of the ith-smallest nontrivial character degree of G. Similarly, we can define $\mathrm{mp}_1(G)$ to be the vector $(m_G(d_1), m_G(d_2), \cdots, m_G(d_t))$. If the complex group algebra $\mathbb{C}G$ of G is given, then both $\mathrm{mp}(G)$ and $\mathrm{mp}_1(G)$ are known. Also, if we know $\mathrm{mp}(G)$, then $k(G)$, the number of conjugacy classes of G, can be computed by taking the sum of all entries of $\mathrm{mp}(G)$.

It seems that the multiplicities of character degrees of a finite group G also have strong influence on the structure of the group. For example, it was proved by A. Moretó and D. Craven in [11, 31] that the order of a finite group G is bounded above by a function of the maximum multiplicity of character degrees of G. A conjugacy class analogue of this was studied by the first author in [36]. In [5], Y. Berkovich

and L. Kazarin proved that the nonsolvable groups in which only two nonlinear irreducible characters have equal degrees are exactly $\mathrm{PSL}_2(5)$ and $\mathrm{PSL}_2(7)$. This result has recently been generalised in [53], where it was showed that two simple groups above are also the only nonsolvable groups which have a unique nontrivial multiplicity of nontrivial character degrees. In [52], we proposed the following problem which is much stronger than Problem 1.1.

Problem 6.1 Given a finite group G, determine all finite groups (up to isomorphism) having the same multiplicity pattern $\mathrm{mp}(G)$ as that of G.

We can also ask the following question: What does $\mathrm{mp}(G)$ (or $\mathrm{mp}_1(G)$) know about G? Notice that we use the vector $\mathrm{mp}(G)$ rather than the set of multiplicities as there are many non-isomorphic groups having the same set of multiplicity of character degrees, for instance, D_8, Q_8, S_3, A_5, $\mathrm{PSL}_2(7)$, J_2, M_{22} and B all have the same set $\{1, 2\}$ of multiplicities. As the quaternion group Q_8 and the dihedral group D_8 have the same character table, we deduce that $\mathrm{mp}(\mathrm{Q}_8) = \mathrm{mp}(\mathrm{D}_8)$, which is equal to $(4, 2)$. Hence, solvable groups are not uniquely determined by the multiplicity patterns. Using [10], we see that the multiplicities of the non-abelian simple groups in the Atlas are all distinct.

Conjecture 6.2 Let H be a non-abelian simple group. If G is a finite group such that $\mathrm{mp}(G) = \mathrm{mp}(H)$, then $G \cong H$.

This conjecture, if true, would be a generalisation of Theorem 2.1. We do not know whether $G \cong H$ or not if $\mathrm{mp}_1(G) = \mathrm{mp}_1(H)$, with H a non-abelian simple group. Notice that $\mathrm{mp}_1(\mathrm{S}_6 \cdot 2) = \mathrm{mp}_1(\mathrm{S}_6)$. In support of Conjecture 6.2, we have proved:

Theorem 6.3 *Let G be a finite group and let H be a non-abelian simple group with at most seven distinct character degrees. If $\mathrm{mp}(G) = \mathrm{mp}(H)$, then $G \cong H$.*

The non-abelian simple groups with at most seven distinct character degrees have been classified by G. Malle and A. Moretó [28].

Lemma 6.4 ([28, Theorem C]) *Let S be a non-abelian simple group. If $|\mathrm{cd}(S)| \leq 7$, then one of the following cases holds.*

(1) $|\mathrm{cd}(S)| = 4$ *and* $S \cong \mathrm{PSL}_2(q)$, $q = 2^f \geq 4$.
(2) $|\mathrm{cd}(S)| = 5$ *and* $S \cong \mathrm{PSL}_2(q)$, $q = p^f > 5$ *and* $p > 2$.
(3) $|\mathrm{cd}(S)| = 6$ *and* $S \cong {}^2\mathrm{B}_2(\mathrm{q}^2)$, $q^2 = 2^{2m+1}, m \geq 1$ *or* $\mathrm{PSL}_3(4)$.
(4) $|\mathrm{cd}(S)| = 7$ *and* $S \cong \mathrm{PSL}_3(3), \mathrm{A}_7, \mathrm{M}_{11}$ *or* J_1.

In the next lemma, we obtain some consequences under the assumption that $\mathrm{mp}(G) = \mathrm{mp}(H)$, where H is a finite perfect group and G is a finite group.

Lemma 6.5 *Let G be a group and let H be a nontrivial perfect group. Assume that $\mathrm{mp}(G) = \mathrm{mp}(H)$. Then the following hold.*

(1) $|\mathrm{cd}(G)| = |\mathrm{cd}(H)|$ *and* $k(G) = k(H)$.
(2) *G is perfect, i.e., $G' = G$.*

(3) *If M is a maximal normal subgroup of G, then G/M is a non-abelian simple group, $k(G/M) \leq k(H)$ and $|\mathrm{cd}(G/M)| \leq |\mathrm{cd}(H)|$. Moreover, if $d \in \mathrm{cd}(G/M)$, then $m_G(d) \geq m_{G/M}(d)$.*

The proof of Theorem 6.3 is based on the previous two lemmas together with the explicit lists of the multiplicity patterns of those simple groups in Lemma 6.4. Indeed, if the finite groups G and H satisfy the hypotheses of Theorem 6.3, then by Lemma 6.5, G is perfect, $k(G) = k(H), |\mathrm{cd}(G)| = |\mathrm{cd}(H)|$ and G/M is a non-abelian simple group with $|\mathrm{cd}(G/M)| \leq |\mathrm{cd}(G)|$, where M is a maximal normal subgroup of G. It follows that G/M is one of the simple groups in Lemma 6.4 as $|\mathrm{cd}(G/M)| \leq |\mathrm{cd}(G)| \leq 7$. Now if $|\mathrm{cd}(G/M)| = |\mathrm{cd}(G)|$, then $\mathrm{cd}(G) = \mathrm{cd}(G/M)$ and so by applying Theorem 4.2 and the fact that G is perfect, we obtain that $G \cong G/M$. Hence, $M = 1$ and G is non-abelian simple with $|\mathrm{cd}(G)| = |\mathrm{cd}(H)|$. Using Lemma 6.4 again and the fact that G and H have the same number of conjugacy classes, we deduce that $G \cong H$ as required. Assume that $|\mathrm{cd}(G/M)| < |\mathrm{cd}(G)|$. For each possible G/M, using Clifford Theory and the theory of character triple isomorphisms in [18], we eliminate these cases by showing that $\mathrm{mp}(G) \neq \mathrm{mp}(H)$.

For quasi-simple groups, we obtain the following.

Theorem 6.6 *Let G be a finite group and let $q \geq 4$ be a prime power. If $\mathrm{mp}(G) = \mathrm{mp}(\mathrm{SL}_2(q))$, then $G \cong \mathrm{SL}_2(q)$.*

As with complex group algebras, we predict that all quasi-simple or symmetric groups are uniquely determined up to isomorphism by their multiplicity patterns.

Acknowledgement

We are grateful to the referee for several helpful suggestions that have improved the exposition of the paper.

References

[1] S.H. Alavi & A. Daneshkhah, A new characterization of the simple group $\mathrm{G}_2(5)$, *Int. J. Algebra* **2** (2008), 327–337.

[2] S.H. Alavi, A. Daneshkah, H.P. Tong-Viet & T.P. Wakefield, Huppert's Conjecture and Fi_{23}, *Rend. Semin. Mat. Univ. Padova* **126** (2011), 201–211.

[3] S.H. Alavi, A. Daneshkah, H.P. Tong-Viet & T.P. Wakefield, On Huppert's Conjecture for the Conway and Fischer families of sporadic simple groups, *J. Aust. Math. Soc.* **94** (2013), 289–303.

[4] A. Balog, C. Bessenrodt, J.B. Olsson & K. Ono, Prime power degree representations of the symmetric and alternating groups, *J. London Math. Soc.* **64** (2001), 344–356.

[5] Y. Berkovich & L. Kazarin, Finite nonsolvable groups in which only two nonlinear irreducible characters have equal degrees, *J. Algebra* **184** (1996), 538–560.

[6] C. Bessenrodt, H.N. Nguyen, J.B. Olsson & H.P. Tong-Viet, Complex group algebras of the Schur double covers of the symmetric and alternating groups. *Algebra Number Theory* **9** (2015), 601–628.

[7] C. Bessenrodt & J.B. Olsson, Prime power degree representations of the double covers of the symmetric and alternating groups, *J. London Math. Soc.* **66** (2002), 313–324.

[8] M. Bianchi, D. Chillag, M.L. Lewis & E. Pacifici, Character degree graphs that are complete graphs, *Proc. Amer. Math. Soc.* **135** (2007), 671–676.

[9] R. Brauer, *Representations of finite groups*, Lectures on modern math., Vol.I, 1963.

[10] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker & R.A. Wilson, *Atlas of Finite Groups*, Clarendon Press, Oxford, 1985.
[11] D. Craven, Symmetric group character degrees and hook numbers, *Proc. Lond. Math. Soc.* (3) **96** (2008), no. 1, 26–50.
[12] C.W. Curtis, The Steinberg character of a finite group with a (B, N)-pair, *J. Algebra* **4** (1966), 433–441.
[13] E.C. Dade, Deux groupes finis distincts ayant la même algèbre de groupe sur tout corps, *Math. Z.* **119** (1971), 345–348.
[14] T. Hawkes, On groups having isomorphic group algebras, *J. Algebra* **167** (1994), 557–577.
[15] H.N. Hung, P.R. Majozi, H.P. Tong-Viet & T.P. Wakefield, Extending Huppert's conjecture from non-abelian simple groups to quasi-simple groups, submitted, 2015.
[16] B. Huppert, Some simple groups which are determined by the set of their character degrees I, *Illinois J. Math.* **44** (2000), 828–842.
[17] B. Huppert, Some simple groups which are determined by the set of their character degrees II, *Rend. Sem. Mat. Univ. Padova* **115** (2006), 1–13.
[18] I.M. Isaacs, *Character Theory of Finite Groups*, Academic Press, San Diego, 1976.
[19] I.M. Isaacs, Recovering information about a group from its complex group algebra, *Arch. Math.* **47** (1986) 293–295.
[20] W. Kimmerle, Group rings of finite simple groups, Around group rings, Resenhas **5** (2002), 261–278.
[21] W. Kimmerle, R. Lyons, R. Sandling & D.N. Teague, Composition factors from the group ring and Artin's theorem on orders of simple groups, *Proc. London Math. Soc* **60** (1990), 89–122.
[22] W. Kimmerle, Beiträge zur ganzzahligen Darstellungstheorie endlicher Gruppen. *Bayreuth. Math. Schr. No.* **36** (1991), 139 pp.
[23] A. Kleshchev & P.H. Tiep, On restrictions of modular spin representations of symmetric and alternating groups, *Trans. Amer. Math. Soc.* **356** (2004), 1971–1999.
[24] A. Kleshchev & P.H. Tiep, Small dimension projective representations of symmetric and alternating groups, *Algebra Number Theory* **6** (2012), 1773–1816.
[25] P.J. Lambert, Characterizing groups by their character tables I, *Quart. J. Math. Oxford* **23** (1972), 427–433.
[26] F. Lübeck, Smallest degrees of representations of exceptional groups of Lie type, *Comm. Algebra* **29** (2001), 2147–2169.
[27] F. Lübeck, Character degrees and their multiplicities for some groups of Lie type, `http://www.math.rwth-aachen.de/~Frank.Luebeck/chev/DegMult/index.html`.
[28] G. Malle & A. Moretó, Nonsolvable groups with few character degrees, *J. Algebra* **294** (2005), 117–126.
[29] G. Malle & A.E. Zalesskii, Prime power degree representations of quasi-simple groups, *Arch. Math.* **77** (2001), 461–468.
[30] V.D. Mazurov & E.I. Khukhro, *Unsolved problems in group theory, the Kourovka notebook*, 17th edition, Novosibirsk, 2010.
[31] A. Moretó, Complex group algebras of finite groups: Brauer's problem 1, *Adv. Math.* **208** (2007), 236–248.
[32] H. Nagao, On the groups with the same table of characters as symmetric groups, *J. Inst. Polytech. Osaka City Univ. Ser. A.* **8** (1957), 1–8.
[33] M. Nagl, Charakterisierung der Symmetrischen Gruppen durch ihre komplexe Gruppenalgebra, Stuttgarter Mathematische Berichte, http://www.mathematik.uni-stuttgart.de/preprints/downloads/2011/2011-007.pdf.
[34] G. Navarro, The set of character degrees of a finite group does not determine its solvability, *Proc. Amer. Math. Soc.* **143** (2015), 989–990.
[35] H.N. Nguyen, Low-dimensional complex characters of the symplectic and orthogonal groups, *Comm. Algebra* **38** (2010), 1157–1197.

[36] H.N. Nguyen, Multiplicities of conjugacy class sizes of finite groups, *J. Algebra* **341** (2011), 250–255.
[37] H.N. Nguyen, Quasisimple classical groups and their complex group algebras, *Israel J. Math.* **195** (2013), 973–998.
[38] H.N. Nguyen & H.P. Tong-Viet, Characterizing finite quasisimple groups by their complex group algebras, *Algebr. Represent. Theory* **17** (2014), 305–320.
[39] H.N. Nguyen, H.P. Tong-Viet & T.P. Wakefield, On Huppert's conjecture for alternating groups of low degrees, *Algebra Colloq.* **22** (2015), 293–308.
[40] H.N. Nguyen, H.P. Tong-Viet & T.P. Wakefield, Projective special linear groups $\mathrm{PSL}_4(q)$ are determined by the set of their character degrees, *J. Algebra Appl.* **11** (2012), 1250108.
[41] T. Oyama, On the groups with the same table of characters as alternating groups, *Osaka J. Math.* **1** (1964), 91–101.
[42] H. Pahlings, Characterization of groups by their character tables. I, *Comm. Algebra* **4** (1976), 111–153.
[43] H. Pahlings, Characterization of groups by their character tables. II. *Comm. Algebra* **4** (1976), 155–178.
[44] R. Rasala, On the minimal degrees of characters of S_n, *J. Algebra* **45** (1977), 132–181.
[45] P.H. Tiep & A.E. Zalesskii, Minimal characters of the finite classical groups, *Comm. Algebra* **24** (1996), 2093–2167.
[46] P.H. Tiep & A.E. Zalesskii, Some characterizations of the Weil representations of the symplectic and unitary groups, *J. Algebra* **192** (1997),130–165.
[47] H.P. Tong-Viet, The simple Ree groups ${}^2\mathrm{F}_4(2^{2m+1})$ are determined by the set of their character degrees, *J. Algebra* **339** (2011), 357–369.
[48] H.P. Tong-Viet, Symmetric groups are determined by their character degrees, *J. Algebra* **334** (2011), 275–284.
[49] H.P. Tong-Viet, Alternating and Sporadic simple groups are determined by their character degrees, *Algebr. Represent. Theory* **15** (2012), 379–389
[50] H.P. Tong-Viet, Simple exceptional groups of Lie type are determined by their character degrees, *Monatsh. Math.* **166** (2012), 559–577.
[51] H.P. Tong-Viet, Simple classical groups are determined by their character degrees, *J. Algebra* **357** (2012), 61–68.
[52] H.P. Tong-Viet, Characterization of some simple groups by the multiplicity pattern, *Monatsh. Math.* **172** (2013), 189–206.
[53] H.P. Tong-Viet, Finite nonsolvable groups with many distinct character degrees, *Pacific J. Math.* **268** (2014), 477–492.
[54] H.P. Tong-Viet, Character degree sets of simple exceptional groups of Lie type, preprint.
[55] H.P. Tong-Viet & T.P. Wakefield, On Huppert's Conjecture for the Monster and Baby Monster, *Monatsh. Math.* **167** (2012), 589-600.
[56] H.P. Tong-Viet & T.P. Wakefield, Verifying Huppert's Conjecture for ${}^3\mathrm{D}_4(q)$ for $q > 2$, *Algebr. Represent. Theory* **16** (2013), 471–490.
[57] H.P. Tong-Viet & T.P. Wakefield, Verifying Huppert's Conjecture for $\mathrm{G}_2(q)$, *J. Pure Appl. Algebra* **216** (2012), 2720–2729.
[58] H.P. Tong-Viet & T.P. Wakefield, On Huppert's conjecture for $\mathrm{F}_4(2)$, *Int. J. Group Theory* **1** (2012), 1–9.
[59] W. Willems, Blocks of defect zero in finite simple groups of Lie type, *J. Algebra* **113** (1988), 511–522.
[60] T.P. Wakefield, Verifying Huppert's Conjecture for $\mathrm{PSL}_3(q)$ and $\mathrm{PSU}_3(q^2)$, *Comm. Algebra* **37** (2009), 2887–2906.
[61] T.P. Wakefield, Verifying Huppert's Conjecture for ${}^2\mathrm{G}_2(q^2)$, *Algebr. Represent. Theory* **14** (2011), 609–623.
[62] T.P. Wakefield, Verifying Huppert's Conjecture for $\mathrm{PSp}_4(q)$ when $q > 7$, *Algebr. Represent. Theory* **15** (2012), 427–448.

GENERALIZED BAUMSLAG-SOLITAR GROUPS: A SURVEY OF RECENT PROGRESS

DEREK J. S. ROBINSON

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
Email: dsrobins@illinois.edu

1 Introduction

A *generalized Baumslag-Solitar group* is the fundamental group of a graph of groups with infinite cyclic vertex and edge groups. These groups have been the subject of numerous investigations over the last twenty five years. Here we give an account of some of what has been discovered, particularly relating to homology, the determination of the centre and the maximum cyclic normal subgroup, and the relation to 3-manifold groups. Generally proofs are omitted.

We begin by recalling that the *Baumslag-Solitar groups* are the groups with a presentation of the form

$$BS(m,n) = \langle\, t, x \mid (x^m)^t = x^n \,\rangle,$$

where $m, n \in \mathbb{Z}^* = \mathbb{Z}\backslash\{0\}$. A similar type of group is

$$K(m,n) = \langle\, x, y \mid x^m = y^n \,\rangle,$$

where $m, n \in \mathbb{Z}^*$. When m and n are relatively prime, $K(m,n)$ is a torus knot group.

Aside from being 1-relator groups with simple presentations, what these groups have in common is that they are the fundamental groups of certain simple graphs of groups.

Generalized Baumslag-Solitar graphs and groups

Let Γ be a finite connected graph, loops and multiple edges being allowed. Let $V(\Gamma)$ and $E(\Gamma)$ denote the respective sets of vertices and edges of Γ. For each $e \in E(\Gamma)$ label the endpoints e^- and e^+, so that $e = \langle e^-, e^+ \rangle$. Infinite cyclic groups $\langle g_x \rangle$ and $\langle u_e \rangle$ are assigned to each vertex x and edge e and injective homomorphisms $\langle u_e \rangle \to \langle g_{e^+} \rangle$ and $\langle u_e \rangle \to \langle g_{e^-} \rangle$ are defined by

$$u_e \mapsto g_{e^+}^{\omega^+(e)} \text{ and } \; u_e \mapsto g_{e^-}^{\omega^-(e)},$$

where $\omega^+(e),\ \omega^-(e) \in \mathbb{Z}^*$. This means that we have a *weight function*

$$\omega : E(\Gamma) \to \mathbb{Z}^* \times \mathbb{Z}^*,$$

where $\omega(e) = (\omega^-(e), \omega^+(e))$ is defined up to $\pm$. The weighted graph (Γ, ω) is called a *generalized Baumslag-Solitar graph* or *GBS-graph*.

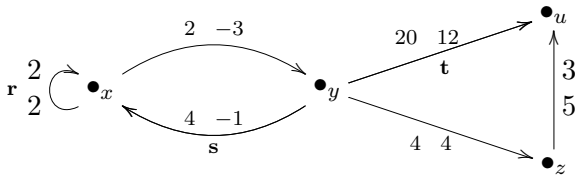
The *generalized Baumslag-Solitar group* (or GBS-group) determined by the GBS-graph (Γ, ω) is its fundamental group $G = \pi_1(\Gamma, \omega)$. If T is a maximal subtree of Γ,

then G has a presentation with generators g_x, $(x \in V(\Gamma))$, and t_e, $(e \in E(\Gamma)\backslash E(T))$, and relations

$$\begin{cases} g_{e^+}^{\omega^+(e)} = g_{e^-}^{\omega^-(e)}, & \text{if } e \in E(T), \\ (g_{e^+}^{\omega^+(e)})^{t_e} = g_{e^-}^{\omega^-(e)}, & \text{if } e \in E(\Gamma)\backslash E(T). \end{cases}$$

The t_e are the *stable letters*. Thus $G = K(m,n)$ if Γ is a single edge e and $G = BS(m,n)$ if Γ is a loop e, where $m = \omega^+(e), n = \omega^-(e)$.

Example 1.1 Consider the GBS-graph



Let the maximal subtree T be the path x, y, z, u. The stable letters are r, s, t and the GBS-group has a presentation in generators $r, s, t, g_x, g_y, g_z, g_u$ with relations

$$g_x^2 = g_y^{-3},\ g_y^4 = g_z^4,\ g_z^5 = g_u^3,$$
$$(g_x^2)^r = g_x^2,\ (g_x^4)^s = g_y^{-1},\ (g_u^{12})^t = g_y^{20}.$$

It is worthwhile reflecting on the concepts introduced so far. We began with a GBS-graph (Γ, ω), an arithmetic-combinatorial object, and we constructed an algebraic object, its fundamental group $\pi_1(\Gamma, \omega)$. In fact there is also a topological structure associated with the GBS-graph (Γ, ω), namely a 2-dimensional simplicial complex $C(\Gamma, \omega)$ whose fundamental group is isomorphic with $\pi_1(\Gamma, \omega)$. Details of the construction may be found in [5: 7]. Thus we have a combinatorial object, an algebraic object and a topological object linked by the isomorphism

$$\pi_1(\Gamma, \omega) \simeq \pi_1(C(\Gamma, \omega)).$$

It is the interplay between combinatorics, algebra and topology that makes this area a fertile one for research.

Some properties of GBS-groups

Let $\pi_1(\Gamma, \omega)$ be a GBS-graph and put $G = \pi_1(\Gamma, \omega)$. Then G has the following properties.

(i) *Up to isomorphism the group G is independent of the choice of maximal subtree.* This is a standard fact from the theory of graphs of groups – see [3] or [18].

The next two results are also well known.

(ii) *G is finitely presented and torsion-free.*

(iii) *If Γ is a tree, then G is residually finite and hence is hopfian.*

The next result, which is due to P. Kropholler [13], shows clearly the central position of GBS-groups in combinatorial group theory.

(iv) *The non-cyclic GBS-groups are exactly the finitely generated groups of cohomological dimension* 2 *which have a commensurable infinite cyclic subgroup.*

(v) *If H is a finitely generated subgroup of a GBS-group G, then either H is a GBS-group or else it is free. Hence G is coherent.*

Proof We have $cd(H) \leq cd(G) \leq 2$. If $cd(H) = 1$, then H is free by the Stallings-Swan Theorem. Otherwise $cd(H) = 2$. If H contains a commensurable element, it is a GBS-group by (iv). If H has no commensurable elements, it intersects every conjugate of a vertex group trivially, so it it is free. □

Another interesting result of Kropholler [13] is:

(vi) *The second derived subgroup of a GBS-group is free.*

This shows that there is a Tits Alternative for GBS-groups: either a GBS-group is soluble or it has a free subgroup of rank 2.

(vii) *The GBS-graphs with soluble fundamental groups have been classified in* [6].

The automorphism groups of GBS-groups have been studied by G. Levitt [15], where the following is established.

(viii) *If G is a GBS-group, then either* $\mathrm{Out}(G)$ *has a free subgroup of rank* 2 *or it is virtually nilpotent of class at most* 2.

Open problems

We mention some open problems about GBS-groups.

(i) Find necessary and sufficient conditions for a GBS-group to be residually finite or hopfian. This has been done for Baumslag-Solitar groups – see [1] and [16].

(ii) Is the Isomorphism Problem for GBS-groups soluble? In other words, given two GBS-graphs, is there an algorithm which can decide if their fundamental groups are isomorphic? The answer is positive for graphs for which the outer automorphism group of the fundamental group has no free subgroups of rank 2: this is due to Levitt [15]. For further results see [2] and [10].

(iii) Can the structure of the abelianization of a GBS-group G be determined directly from the underlying GBS-graph? Or is there at least a method for determining whether G_{ab} is torsion-free: note that the torsion-free rank is given by Theorem 2.3 below. Of course, the structure of G_{ab} can be determined algebraically from a presentation for G_{ab}.

2 The weight of a path in a GBS-graph

The concept of the weight of a path in a maximal subtree of a GBS-graph is a useful one. Let (Γ, ω) be a GBS-graph with a maximal subtree T and let $e = \langle x, y \rangle$ be a non-tree edge where $x \neq y$. Then there is a unique path in T from x to y, say

$$x = x_0, x_1, \ldots, x_n = y.$$

By reading along the path, we detect a relation

$$g_x^{p_1(e)} = g_y^{p_2(e)}$$

in $G = \pi_1(\Gamma, \omega)$, where $p_1(e)$ and $p_2(e)$ are the products of the left and right weight values of the edges in the tree path $[x, y]$. The elementary result that follows establishes the existence of a minimal relation of this type.

Lemma 2.1 *Let (Γ, ω) be a GBS-graph with a maximal subtree T and let $\alpha = [x, y]$ be a path in T. Then there exist $a, b \in \mathbb{Z}^*$ such that $g_x^a = g_y^b$ in $\pi_1(\Gamma, \omega)$, and if $g_x^m = g_y^n$, then $(m, n) = (a, b)q$ for some $q \in \mathbb{Z}^*$.*

Proof As already remarked, there exist $a, b \in \mathbb{Z}^*$ such that $g_x^a = g_y^b$. Assume that the pair (a, b) has been chosen with $|a|$ minimal. Suppose that $g_x^m = g_y^n$ and write $m = aq + r$ where $q, r \in \mathbb{Z},\ 0 \leq r < |a|$. Then $g_y^n = g_x^m = g_x^{aq+r} = g_y^{bq} g_x^r$, so that $g_x^r = g_y^{n-bq}$. From the minimality of $|a|$ we deduce that $r = 0$ and $m = aq$. Then $g_x^{aq} = g_y^n$ and $g_y^{bq} = g_y^n$, so that $n = bq$, as required. □

We call the integer pair (a, b) in Lemma 2.1 the *weight of the path* α in T and denote it by $\omega_T(\alpha)$ or

$$\omega_T(x, y) = (\omega_T^{(1)}(x, y), \omega_T^{(2)}(x, y)).$$

By convention if $x = y$, so that $[x, y]$ is a loop, then $\omega_T([x, y]) = (1, 1)$. Keep in mind that the weight is unique only up to $\pm$.

Computing the weight of a path

Let (Γ, ω) be a GBS-graph with a maximal subtree T. Let α be the path $x = x_0, x_1, \ldots, x_n = y$ in T and write $\omega(\langle x_i, x_{i+1}\rangle) = (u_i^{(1)},\ u_i^{(2)}),\ i = 0, 1, \ldots, n-1$. Define a pair of non-zero integers $(\ell_i, m_i),\ 0 \leq i \leq n$, recursively by $\ell_0 = 1 = m_0$ and

$$\ell_{i+1} = \frac{\ell_i u_i^{(1)}}{\gcd(m_i, u_i^{(1)})}, \qquad m_{i+1} = \frac{m_i u_i^{(2)}}{\gcd(m_i, u_i^{(1)})}.$$

With this notation we state a result which allows the weight of a path to be calculated efficiently.

Lemma 2.2 $\omega_T(x, y) = (\ell_n, m_n)$.

Tree and skew tree dependence

Let (Γ, ω) be a GBS-graph with a maximal subtree T. A non-tree edge $e = \langle x, y\rangle$ with $x \neq y$ is called *T-dependent* or *skew T-dependent* if

$$\frac{\omega^-(e)}{\omega^+(e)} = \frac{\omega_T^{(1)}(e)}{\omega_T^{(2)}(e)} \quad \text{or} \quad -\frac{\omega_T^{(1)}(e)}{\omega_T^{(2)}(e)}$$

respectively. If e is a loop, then e is said to be T-dependent or skew T-dependent if and only if $\omega^-(e) = \omega^+(e)$ or $-\omega^+(e)$ respectively.

If every non-tree edge of a GBS-graph is T-dependent, the graph is called *tree dependent*. If every non-tree edge is T-dependent or skew T-dependent, with at least one edge of the latter type, the GBS-graph is called *skew tree dependent*. It is a fact that these properties are independent of the choice of maximal subtree T – see Corollary 1 below.

Tree dependence is relevant to the computation of homology in low dimensions. Recall that the homology of a GBS-group vanishes in dimensions 3 and higher. The following is what is known about the integral homology in dimensions 1 and 2.

Theorem 2.3 (Levitt [15], Robinson [17]) *Let $G = \pi_1(\Gamma, \omega)$ be a GBS-group. Then the torsion-free rank of $H_1(G) = G_{ab}$ is*

$$r_0(G) = |E(\Gamma)| - |V(\Gamma)| + 1 + \epsilon(\Gamma, \omega)$$

where $\epsilon(\Gamma, \omega) = 1$ if (Γ, ω) is tree dependent and otherwise $\epsilon(\Gamma, \omega) = 0$.

Theorem 2.4 (Robinson [17]) *Let $G = \pi_1(\Gamma, \omega)$ be a GBS-group. Then the Schur multiplier $H_2(G)$ is free abelian of rank $r_0(G) - 1$.*

For example, consider the GBS-group G arising from the GBS-graph in Example 1.1, where the maximal subtree chosen is the path x, y, z, u. The non-tree edges with the exception of $\langle y, x\rangle$ are T-dependent. Therefore (Γ, ω) is not tree-dependent, $\epsilon(\Gamma, \omega) = 0$ and $r_0(G) = |E(\Gamma)| - |V(\Gamma)| + 1 = 3$. Thus $M(G) \simeq \mathbb{Z} \oplus \mathbb{Z}$.

The Δ-function

Let G be a group with a commensurable element x of infinite order. If $g \in G$, then $\langle x\rangle \cap \langle x\rangle^g \neq 1$ and $(x^n)^g = x^m$ for some $m, n \in \mathbb{Z}^*$. Define $\Delta_x(g) = m/n$. A simple calculation reveals that

$$\Delta_x : G \mapsto \mathbb{Q}^* = \mathbb{Q}\backslash\{0\}$$

is a well defined homomorphism. This useful function was introduced by Kropholler in [12] and is sometimes referred to as the modular homomorphism.

If $y \in G$ is commensurable and $\langle x\rangle \cap \langle y\rangle \neq 1$, it is easily seen that $\Delta_x = \Delta_y$. If this holds for all commensurable elements y, the function Δ_x does not depends on x and we denote it by Δ_G.

The Δ-function of a GBS-group

A GBS-graph (Γ, ω), or the corresponding group $G = \pi_1(\Gamma, \omega)$, is called *elementary* if $G \simeq \mathbb{Z}$ or $BS(1, \pm 1)$. If G is non-elementary, then each commensurable element of G is elliptic in its action on the Bass-Serre tree – see [5: 3.1] – and hence is conjugate to a power of some vertex generator. Therefore the Δ-function does not depend on any particular element. Suppose that T and $\bar{T}$ are two maximal subtrees in (Γ, ω) leading to GBS-groups G and $\bar{G}$. There is an isomorphism $\psi : G \to \bar{G}$ and clearly $\Delta_G = \Delta_{\bar{G}}\psi$. Hence $\mathrm{Im}(\Delta_G) = \mathrm{Im}(\Delta_{\bar{G}})$ is independent of the maximal subtree chosen, a fact that will be important in the sequel.

The Δ-function of a GBS-group is easily calculated from the graph by using the next result.

Lemma 2.5 *Let (Γ, ω) be a non-elementary GBS-graph with a maximal subtree T and let $G = \pi_1(\Gamma, \omega)$. Then:*

(i) $\Delta_G(g_x) = 1$ *for all* $x \in V(\Gamma)$*;*

(ii) If $e \in E(\Gamma)\backslash E(T)$, $\omega(e) = (a, b)$ *and* $\omega_T(e) = (m, n)$, *then* $\Delta_G(t_e) = an/bm$.

The Δ-function provides criteria for tree dependence and skew tree dependence:

Corollary 2.6 *Let (Γ, ω) be a non-elementary GBS-graph with a maximal subtree T and let $G = \pi_1(\Gamma, \omega)$. If $e \in E(\Gamma)\backslash E(T)$, then:*

(i) e is T-dependent if and only if $\Delta_G(t_e) = 1$. *Hence (Γ, ω) is tree dependent if and only if* $\mathrm{Im}(\Delta_G) = \{1\}$.

(ii) e is skew T-dependent if and only if $\Delta_G(t_e) = -1$. *Hence (Γ, ω) is skew tree dependent if and only if* $\mathrm{Im}(\Delta_G) = \{\pm 1\}$.

A consequence of the corollary is that the properties of tree dependence and skew tree dependence do not depend on the choice of maximal subtree. If $\mathrm{Im}(\Delta_G) \subseteq \{\pm 1\}$, the graph (Γ, ω), or $G = \pi_1(\Gamma, \omega)$, is called *unimodular*.

3 The cyclic radical

A non-elementary GBS-group contains a unique maximum cyclic normal subgroup.

Lemma 3.1 *Let (Γ, ω) be a non-elementary GBS-graph and set $G = \pi_1(\Gamma, \omega)$. Then:*

(i) G has a unique maximum cyclic normal subgroup $C(G)$ and $Z(G) \leq C(G)$.

(ii) Exactly one of the following is true: $1 = C(G)$, $1 = Z(G) < C(G)$ *and* $1 < Z(G) = C(G)$.

Proof Suppose that $\{ C_i \mid i \in I \}$ is an infinite ascending chain of cyclic normal subgroups of G. Each C_i is commensurable and hence is elliptic since G is not elementary. Therefore each C_i lies in a vertex subgroup. But then infinitely many of the C_i lie in some $\langle g_v \rangle$, which is impossible. Therefore G has a maximal cyclic normal subgroup, say C.

Next let D be any non-trivial cyclic normal subgroup of G. We show that $D \leq C$. Since CD is nilpotent, no element of C can induce inversion in D and therefore CD is abelian. Also C and D are commensurable and hence are contained in vertex subgroups. It follows that $C \cap D \neq 1$ and CD/D is finite, which shows that CD is cyclic. Hence $D \leq C$, so we may define $C(G)$ to be C. Clearly $Z(G) \leq C(G)$. Finally, if $Z(G) \neq 1$, then elements of G centralize a non-trivial subgroup of $C(G)$, so they centralize $C(G)$. □

We will refer to the subgroup $C(G)$ in Lemma 3.1 as the *cyclic radical* of G. The Δ-function can also be used to give criteria for the center or cyclic radical of a GBS-group to be non-trivial.

Proposition 3.2 *Let (Γ, ω) be a non-elementary GBS-graph and put $G = \pi_1(\Gamma, \omega)$. Then the following are equivalent.*

(i) (Γ, ω) *is tree dependent or skew dependent;*

(ii) $\mathrm{Im}(\Delta_G) = \{1\}$, *respectively* $\mathrm{Im}(\Delta_G) = \{-1, 1\}$;

(iii) $Z(G) \neq 1$, *respectively* $1 = Z(G) < C(G)$.

The the equivalence of conditions (ii) and (iii) in Proposition 3.2 was proved by Levitt [15: 2.5, 2.6]. We remark that as a consequence of Proposition 3.2, if G is a unimodular GBS-group, then $G/C(G)$ is the fundamental group of a graph of finite cyclic groups. In consequence $G/C(G)$ is virtually free and from this it follows readily that G *is residually finite and hence is hopfian.*

Locating the cyclic radical and centre of a GBS-group

Let (Γ, ω) be a GBS-graph with $G = \pi_1(\Gamma, \omega)$. By Lemma 3.1 it is sufficient to show how to find $C(G)$. In this task we may assume the graph is non-elementary and we can also assume (Γ, ω) is unimodular, since otherwise $C(G) = 1$.

First some useful terminology: in a GBS-graph the *distal weight* of a leaf in a maximal subtree is the weight occurring at the vertex of degree 1. In finding the cyclic radical there is no loss in assuming there are no leaves with distal weight ± 1, since the vertex generator corresponding to the vertex of degree 1 can be deleted. Under these circumstances we can make an initial determination of the location of the cyclic radical.

Lemma 3.3 *Let (Γ, ω) be a non-elementary GBS-graph with a maximal subtree T which has no leaves of distal weight ± 1. If $G = \pi_1(\Gamma, \omega)$, then*

$$C(G) \leq J = \bigcap_{v \in V(\Gamma)} \langle g_v \rangle.$$

For any $x, v \in V(\Gamma)$ we have $\langle g_x \rangle \cap \langle g_v \rangle = \langle g_v^{\omega_T^{(1)}(v,x)} \rangle$ by Lemma 2.1. Hence $J = \langle {g_v}^{h_v} \rangle$ where

$$h_v = \mathrm{lcm}\{\omega_T^{(1)}(v, x) \mid x \in V(\Gamma)\} = \omega_T^{tot}(v),$$

which is called the *total weight* of v in T: this is just the smallest positive power of g_v which belongs to to every vertex subgroup.

There is a more economic expression for the total weight. Let $y_1, y_2, \ldots, y_k$ be the vertices of degree 1 in T. Then

$$\omega_T^{tot}(v) = \mathrm{lcm}\{\omega_T^{(1)}(v, y_i) \mid i = 1, 2, \ldots, k\}.$$

This is true since by Lemma 2.2 the weight of a path from v in T is divisible by the weight of the subpath from v to any previous vertex on the path.

Corollary 3.4 *Let (T, ω) be a non-elementary GBS-tree with no distal weights ± 1 and let $G = \pi(T, \omega)$. Then for any vertex v of Γ*

$$Z(G) = \langle g_v^{\omega_T^{tot}(v)} \rangle.$$

Of course $Z(G) = C(G)$ in this corollary since $Z(G) \neq 1$. Turning to the case of a general GBS-group, we deduce from Lemma 3.3 the following result.

Corollary 3.5 *Let (Γ, ω) be a non-elementary GBS-graph with a maximal subtree T. Assume that no leaf of T has distal weight ± 1 and let $G = \pi_1(\Gamma, \omega)$. If $I = \bigcap_{v \in V(\Gamma)} \langle g_v \rangle$, then*

$$C(G) = \bigcap_{e \in E(\Gamma) \setminus E(T)} I_{\langle t_e \rangle},$$

where $I_{\langle t_e \rangle}$ is the $\langle t_e \rangle$-core of I, that is, the largest $\langle t_e \rangle$-invariant subgroup of I.

It remains to show how to compute the cores in Corollary 3.5. The next lemma provides the crucial step.

Lemma 3.6 *Let (Γ, ω) be a GBS-graph with a maximal subtree T and let $e = \langle v, u \rangle$ be a non-tree edge which is T-dependent or skew T-dependent. If $\omega(e) = (m, n)$ and $\omega_T(v, u) = (a, b)$, then*

$$I_{\langle t_e \rangle} = \langle g_v^{\operatorname{lcm}(a,m)} \rangle \cap I.$$

By combining the Lemmas 3.3 and 3.6 we arrive at our principal result about the cyclic radical of a non-elementary GBS-group.

Theorem 3.7 *Let (Γ, ω) be a non-elementary, unimodular GBS-graph with $G = \pi_1(\Gamma, \omega)$. Assume that T is a maximal subtree with no leaves of distal weight ± 1. Let v be any fixed vertex of Γ and let the non-tree edges of Γ be $e_i = \langle x_i, y_i \rangle \mid i = 1, 2, \ldots, k$. Write $\omega(e_i) = (m_i, n_i)$, $\omega_T(x_i, y_i) = (a_i, b_i)$, $\omega_T(v, x_i) = (c_i, d_i)$, and put $\ell_i = \operatorname{lcm}(a_i, m_i)$. Then*

$$C(G) = \langle g_v^{\omega^{tot}_{(\Gamma,\omega)}(v)} \rangle$$

where

$$\omega^{tot}_{(\Gamma,\omega)}(v) = \operatorname{lcm} \left\{ \frac{c_i \ell_i}{\gcd(\ell_i, d_i)},\ \omega^{tot}_T(v) \middle| i = 1, 2, \ldots, k \right\}.$$

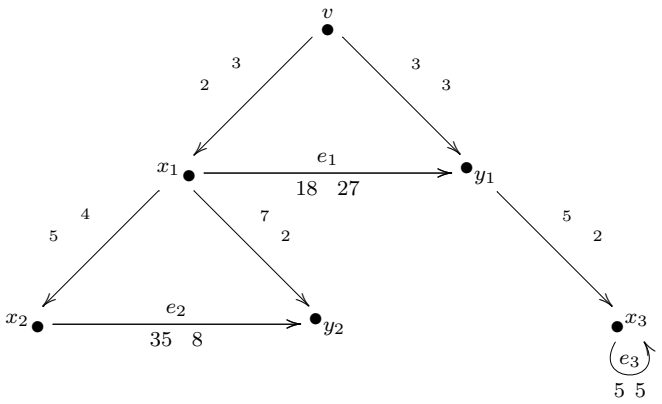
The positive integer

$$\omega^{tot}_{(\Gamma,\omega)}(v) = \omega^{tot}(v)$$

is called the *total weight* of the vertex v: it is the positive power of g_v which generates the cyclic radical in a non-elementary, unimodular GBS-group. Theorem 3.7 provides a method for computing total weights: for small graphs the computations can be done by hand, while larger examples may require machine computation, (and a program to achieve this has in fact been implemented).

Example 3.8 Consider the GBS-graph that follows below. Here the edges which do not belong to the maximal subtree T are $e_i = \langle x_i, y_i \rangle$, $i = 1, 2$, and the loop $e_3 = \langle x_3, x_3 = y_3 \rangle$. The vertices of degree 1 in T are x_2, y_2, x_3. By inspection of the graph we see that the edges e_i are T-dependent, so (Γ, ω) is tree dependent and $C(G) = Z(G) \neq 1$. Suppose that we want to express $C(G)$ in terms of g_v: simply read off the data required from the GBS-graph. Firstly

$$\omega^{tot}_T(v) = \operatorname{lcm} \left(\omega^{(1)}_T(v, x_2),\ \omega^{(1)}_T(v, y_2),\ \ \omega^{(1)}_T(v, x_3) \right) = \operatorname{lcm}(6, 21, 15),$$



which equals 210.

Next $\omega(e_1) = (m_1, n_1) = (18, 27)$, $\omega(e_2) = (m_2, n_2) = (35, 8)$, $\omega(e_3) = (m_3, n_3) = (5, 5)$, $\omega_T(x_1, y_1) = (a_1, b_1) = (2, 3)$, $\omega_T(x_2, y_2) = (a_2, b_2) = (35, 8)$, $\omega_T(x_3, y_3) = (a_3, b_3) = (1, 1)$, $\omega_T(v, x_1) = (c_1, d_1) = (3, 2)$, $\omega_T(v, x_2) = (c_2, d_2) = (6, 5)$, $\omega_T(v, x_3) = (c_3, d_3) = (15, 2)$. Hence $\ell_1 = 18$, $\ell_2 = 35$, $\ell_3 = 5$ and $\omega^{tot}(v) = \text{lcm}(27, 42, 75, 210) = 9450$. Therefore

$$C(G) = Z(G) = \langle g_v^{9450} \rangle.$$

Now suppose we change the weights on the edge e_2 to $(-35, 8)$. The GBS-graph then becomes skew tree dependent and Theorem 3.7 yields $C(G) = \langle g_v^{9450} \rangle$. Conjugation in this subgroup by t_{e_2} induces inversion since $\Delta_G(t_{e_2}) = -1$. Of course $Z(G) = 1$.

4 GBS-groups and 3-manifold groups

As the final topic in this survey, we consider the relation between GBS-groups and 3-manifold groups, i.e., the fundamental groups of compact 3-manifolds. We begin with some interesting examples due to W. Heil [11].

(i) $K(m, n) = \langle x, y \mid x^m = y^n \rangle$ *is a 3-manifold group.* Note that the underlying GBS-graph is: $\bullet_x \xrightarrow{m \quad n} \bullet_y$

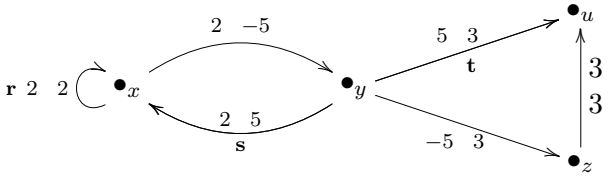
(ii) *The group* $\langle x_1, x_2, x_3 \mid {x_1}^m = x_2^n,\ {x_2}^m = x_3^n \rangle$ *is a 3-manifold group if and only if* $|m| = 1$ *or* $|n| = 1$ *or* $|m| = |n|$. Here the GBS-graph is: $\bullet_{x_1} \xrightarrow{m \quad n} \bullet_{x_2} \xrightarrow{m \quad n} \bullet_{x_3}$

(iii) $B(m, n)$ *is a 3-manifold group if and only if* $|m| = |n|$. Of course the GBS-graph is a loop in this case.

These examples suggest the problem of finding necessary and sufficient conditions on a GBS-graph (Γ, ω) for $\pi_1(\Gamma, \omega)$ to be the fundamental group of a compact 3-manifold. To solve this problem we introduce two special types of GBS-graphs. A GBS-graph (Γ, ω) is called *locally weight constant* if at every vertex v all weights are equal to a constant c_v and *locally* $\pm$ *weight constant* if all weights at v equal $\pm c_v$ for some constant c_v. We begin with a simple observation.

Lemma 4.1 *A GBS-graph is tree dependent if it is locally weight constant, and it is tree or skew tree dependent, i.e., unimodular, if it is locally ± weight constant.*

Example 4.2 The GBS-graph shown is locally ± weight constant, but not locally weight constant.



The following result furnishes a complete description of the non-elementary GBS-groups that are 3-manifold groups. (The elementary GBS-groups are easily seen to be 3-manifold groups).

Theorem 4.3 (Delgado, Robinson and Timm [8]) *Let (Γ, ω) be a non-elementary GBS-graph. Then the following properties are equivalent.*

(i) $\pi_1(\Gamma, \omega)$ *is a 3-manifold group.*

(ii) $\pi_1(\Gamma, \omega)$ *is an orientable 3-manifold group.*

(iii) (Γ, ω) *is locally ± weight constant.*

This result explains Heil's examples: for example, $B(m, n)$ is a 3-manifold group if and only if $|m| = |n|$.

3-Manifold GBS-group covers

Let (Γ, ω) be a non-elementary GBS-graph. If $\pi_1(\Gamma, \omega)$ is not a 3-manifold group, it might still be a quotient of a GBS-group which is a 3-manifold group. With this possibility in mind, we define a 3-*manifold GBS-group cover* of $\pi_1(\Gamma, \omega)$ to be a surjective homomorphism

$$\varphi : \pi_1(\Gamma, \tau) \to \pi_1(\Gamma, \omega)$$

where (Γ, τ) is a GBS-graph such that $\pi_1(\Gamma, \tau)$ is a 3-manifold group, and φ is a *pinch map*, i.e., it is a composite of *pinches*. Here a pinch is a map arising from division of the weights on a fixed edge of Γ by a common factor. (For a detailed account of pinch maps and other "geometric homomorphisms" between GBS-groups, see [5]).

The next theorem tells us exactly which GBS-groups possess 3-manifold GBS-group covers.

Theorem 4.4 (Delgado, Robinson and Timm [8]) *Let (Γ, ω) be a non-elementary GBS-graph. Then the following properties are equivalent:*

(i) $\pi_1(\Gamma, \omega)$ *has a 3-manifold GBS-group cover.*

(ii) $\pi_1(\Gamma, \omega)$ *has an orientable 3-manifold GBS-group cover.*

(iii) $\pi_1(\Gamma, \omega)$ *is unimodular, i.e.,* (Γ, ω) *is tree dependent or skew tree dependent.*

Thus in Example 4.2 the GBS-groups are unimodular and so have 3-manifold GBS-group covers, but they are not 3-manifold groups.

The total ± weight cover of a GBS-group

While we do not present a complete proof of Theorem 4.4, we will explain how the 3-manifold GBS-group covers in the theorem are constructed. Suppose that (Γ, ω) is a non-elementary GBS-graph such that $\pi_1(\Gamma, \omega)$ unimodular and let T be a maximal subtree with no distal weights ± 1: we can also assume that the edges of T have been labelled so that all weights are positive. Define a new GBS-graph $\pi_1(\Gamma, \tau)$ in the following manner.

(i) Case: (Γ, ω) *is tree dependent.*
Define the new weight function τ by

$$\tau(e) = (\omega^{tot}(e^-), \omega^{tot}(e^+)), \quad e \in E(\Gamma).$$

The GBS-graph (Γ, τ) is called the *total weight cover* of (Γ, ω). Notice that (Γ, τ) is locally weight constant, so $\pi_1(\Gamma, \tau)$ is a compact (orientable) 3-manifold group by Theorem 4.3. The identity map on Γ and a suitable sequence of pinches give rise to a surjective homomorphism $\varphi : \pi_1(\Gamma, \tau) \to \pi_1(\Gamma, \omega)$, which is a 3-manifold GBS-group cover of $\pi_1(\Gamma, \omega)$.

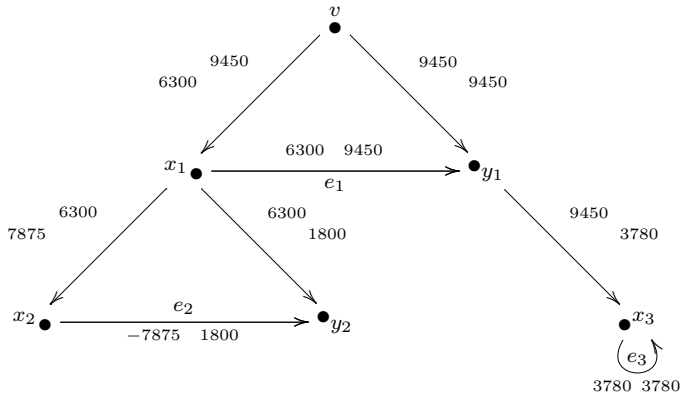
(ii) Case: (Γ, ω) *is skew tree dependent.*
In this situation we partition the set of non-tree edges of Γ into two subsets, $E(\Gamma) \backslash E(T) = P \cup N$ where P is the set of edges with positive weights and N is the set of remaining edges. Define the weight function τ by

$$\tau(e) = \begin{cases} (\omega^{tot}(e^-), \omega^{tot}(e^+)), & e \in E(T) \cup P, \\ (-\omega^{tot}(e^-), \omega^{tot}(e^+)), & e \in N. \end{cases}$$

Thus (Γ, τ) is a locally ± weight constant GBS-graph, which is called the *total* ± *weight cover* of (Γ, ω). From Theorem 4.3 we see that $\pi_1(\Gamma, \tau)$ is a 3-manifold group, so that once again we have a 3-manifold GBS-group cover $\varphi : \pi_1(\Gamma, \tau) \to \pi_1(\Gamma, \omega)$ defined by the identity map on Γ and a sequence of suitable pinches.

Finally, we remark that the 3-manifold GBS-group covers constructed above are unique with respect to a minimality property, in the sense that all other covers factor through them. Also the pinch maps involved in these minimal 3-manifold GBS-group covers can be computed using the algorithm of Theorem 3.7. Details of the proofs can be found in [7] and [8].

Example 4.5 Consider the GBS-graph in Example 3.8 in the skew dependent case. To find the the 3-manifold GBS cover, compute the total weights of all the vertices, using the formulas in Theorem 3.7. This yields the locally ± weight constant GBS-graph below, which determines the canonical GBS-covering group of the original GBS-group.



References

[1] G. Baumslag & D. Solitar, Some two-generator one-relator non-Hopfian groups, *Bull. Amer. Math. Soc.* **68** (1962), 199–201.

[2] M. Clay & M. Forester, On the isomorphism problem for generalized Baumslag-Solitar groups, *Algebra Geom. Topol.* **8** (2008), 2289–2322.

[3] D.E. Cohen, *Combinatorial Group Theory: a Topological Approach*, Cambridge, 1989.

[4] M. Culler & J.W. Morgan, Group actions on $\mathbb{R}$-trees, *Proc. London Math. Soc.* **55** (1987), 571–604.

[5] A.L. Delgado, D.J.S. Robinson & M. Timm, Generalized Baumslag-Solitar groups and geometric homomorphisms, *J. Pure Appl. Algebra* **215** (2011), 398–410.

[6] A.L. Delgado, D.J.S. Robinson & M. Timm, Generalized Baumslag-Solitar graphs with soluble fundamental groups, *Algebra Colloquium* **21** (2014), 53–58.

[7] A.L. Delgado, D.J.S. Robinson & M. Timm, Cyclic normal subgroups in generalized Baumslag-Solitar groups, preprint.

[8] A.L. Delgado, D.J.S. Robinson & M. Timm, 3-Manifold groups and generalized Baumslag-Solitar groups, preprint.

[9] M. Forester, On uniqueness of JSJ decompositions of finitely generated groups, *Comment. Math. Helv.* **78** (2003), 740–751.

[10] M. Forester, Splittings of generalized Baumslag-Solitar groups, *Geom. Dedicata* **121** (2006), 43–59.

[11] W.H. Heil, Some finitely presented non-3-manifold groups, *Proc. Amer. Math. Soc.* **53** (1975), 497–500.

[12] P.H. Kropholler, A note on centrality in 3-manifold groups, *Math. Proc. Camb. Phil. Soc.* **107** (1990), 261–266.

[13] P.H. Kropholler, Baumslag-Solitar groups and some other groups of cohomological dimension two, *Comment. Math. Helv. J.* **65** (1990), 547–558.

[14] G. Levitt, Characterizing rigid simplicial actions on trees, *Geometric Methods in Group Theory*, 27–33, Contemp. Math. 372, 2005.

[15] G. Levitt, On the automorphism group of generalized Baumslag-Solitar groups, *Geom. Topol.* **11** (2007), 473–515.

[16] D.I. Moldavanskiĭ, On the intersection of subgroups of finite index in the Baumslag-Solitar groups, *Mat. Zametki* **87** (2010), 92–100 = *Math. Notes* **87** (2010), 88–95.

[17] D.J.S. Robinson, The Schur multiplier of a generalized Baumslag-Solitar group, *Rend. Sem. Mat. Univ. Padova* **125** (2011), 207–215.

[18] J.-P. Serre. *Trees*, Springer, Berlin, 1980.

[19] J. Taback & P. Wong, Twisted conjugacy and quasi-isometry invariance for generalized solvable Baumslag-Solitar groups, *J. London Math. Soc. (2)* **75** (2007), 705–717.

ZETA FUNCTIONS OF GROUPS AND RINGS – RECENT DEVELOPMENTS

CHRISTOPHER VOLL

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany
Email: C.Voll.98@cantab.net

Abstract

I survey some recent developments in the theory of zeta functions associated to infinite groups and rings, specifically zeta functions enumerating subgroups and subrings of finite index or finite-dimensional complex representations.

1 About these notes

Over the last few decades, zeta functions have become important tools in various areas of asymptotic group and ring theory. With the first papers on zeta functions of groups published barely 25 years ago, the subject area is still comparatively young. Recent developments have led to a wealth of results and given rise to new perspectives on central questions in the field. The aim of these notes is to introduce the nonspecialist reader informally to some of these developments.

I concentrate on two types of zeta functions: firstly, zeta functions in *subgroup and subring growth* of infinite groups and rings, enumerating finite-index subobjects. Secondly, representation zeta functions in *representation growth* of infinite groups, enumerating finite-dimensional irreducible complex representations. I focus on common features of these zeta functions, such as Euler factorizations, local functional equations, and their behaviour under base extension.

Subgroup growth of groups is a relatively mature subject area, and the existing literature reflects this: zeta functions of groups feature in the authoritative 2003 monograph [39] on "Subgroup Growth", are the subject of the Groups St Andrews 2001 survey [16] and the report [18] to the ICM 2006. The book [21] contains, in particular, a substantial list of explicit examples. Some more recent developments are surveyed in [32, Chapter 3].

On the other hand, few papers on representation zeta functions of infinite groups are older than ten years. Some of the lecture notes in [32] touch on the subject. The recent survey [31] on representation growth of groups complements the current set of notes.

In this text I use, more or less as blackboxes, the theory of p-adic integration and the Kirillov orbit method. The former provides a powerful toolbox for the treatment of a number of group-theoretic counting problems. The latter is a general method to parametrize the irreducible complex representations of certain groups in terms of co-adjoint orbits. Rather than explain in detail how these tools are employed I will refer to specific references at appropriate places in the text. I all but ignore the rich subject of zeta functions enumerating representations or conjugacy classes of finite groups of Lie type; see, for instance, [34].

These notes grew out of a survey talk I gave at the conference Groups St Andrews 2013 in St Andrews. I kept the informal flavour of the talk, preferring instructive examples and sample theorems over the greatest generality of the presented results. As a consequence, the text is not a systematic treatment of the subject, but rather the result of a subjective choice.

2 Zeta functions in asymptotic group and ring theory

We consider counting problems of the following general form. Let Γ be a – usually infinite – algebraic object, such as a group or a ring, and assume that, for each $n \in \mathbb{N}$, we are given integers $d_n(\Gamma) \in \mathbb{N}_0$, encoding some algebraic information about Γ. Often this data will have a profinite flavour, in the sense that, for every n, there exists a finite quotient Γ_n of Γ such that $d_n(\Gamma)$ can be computed from Γ_n. In any case, we encode the sequence $(d_n(\Gamma))$ in a generating function.

Definition 2.1 The *zeta function of* $(\Gamma, (d_n(\Gamma)))$ is the Dirichlet generating series

$$\zeta_{(d_n(\Gamma))}(s) = \sum_{n=1}^{\infty} d_n(\Gamma) n^{-s}, \tag{2.1}$$

where s is a complex variable. If $(d_n(\Gamma))$ is understood from the context, we simply write $\zeta_\Gamma(s)$ for $\zeta_{(d_n(\Gamma))}(s)$.

In the counting problems we consider Dirichlet series often turn out to be preferable over other generating functions, in particular if the arithmetic function $n \mapsto d_n(\Gamma)$ satisfies some of the following properties.

(A) *Polynomial growth*, i.e., the coefficients $d_n(\Gamma)$ – or, equivalently, their partial sums – have polynomial growth: $D_n(\Gamma) := \sum_{\nu \leq n} d_\nu(\Gamma) = O(n^a)$ for some $a \in \mathbb{R}$.

(B) *Multiplicativity* in the sense of elementary number theory: if $n = \prod_i p_i^{e_i}$ is the prime factorization of n, then $d_n(\Gamma) = \prod_i d_{p_i^{e_i}}(\Gamma)$.

Indeed, polynomial growth implies that $\zeta_{(d_n(\Gamma))}(s)$ converges absolutely on some complex half-plane. If $d_n(\Gamma) \neq 0$ for infinitely many n, then the *abscissa of convergence* of $\zeta_{(d_n(\Gamma))}(s)$ is equal to

$$\alpha((d_n(\Gamma))) := \limsup_{n \to \infty} \frac{\log \sum_{\nu \leq n} D_n(\Gamma)}{\log n}.$$

Thus $\alpha((d_n(\Gamma)))$ gives the precise degree of polynomial growth of the partial sums $D_n(\Gamma)$ as n tends to infinity. If the sequence $(d_n(\Gamma))$ is understood from the context, we sometimes write $\alpha(\Gamma)$ for $\alpha((d_n(\Gamma))$.

Multiplicativity implies that – at least formally – the series (2.1) satisfies an *Euler factorization*, indexed by the prime numbers:

$$\zeta_{(d_n(\Gamma))}(s) = \prod_{p \text{ prime}} \zeta_{(d_n(\Gamma)),p}(s), \tag{2.2}$$

where, for a prime p, the function

$$\zeta_{(d_n(\Gamma)),p}(s) = \zeta_{(d_{p^i}(\Gamma))}(s) = \sum_{i=0}^{\infty} d_{p^i}(\Gamma)p^{-is}$$

is called the *local factor of* $\zeta_{(d_n(\Gamma))}(s)$ *at the prime* p. We will later consider other Euler factorizations, indexed by places of a number field rather than rational prime numbers, which reflect multiplicativity features of the underlying counting problem which are subtler than the multiplicativity of $n \mapsto d_n(\Gamma)$. In any case, there are often *rationality results* which establish that the Euler factors are rational functions, rendering them — at least in principle — amenable to computation. In practice, the study of many (global) zeta functions of the form (2.1) proceeds via a uniform description of local factors in Euler factorizations like (2.2).

Key questions regarding zeta functions of groups and rings concern the following:

1. Analytic properties regarding, e.g., the abscissa of convergence, analytic continuation, natural boundaries, location and multiplicities of zeros and poles, residue formulae, special values, etc.,
2. arithmetic properties of the local factors, e.g., rationality; if so, structure of numerators and denominators, special symmetries (functional equations), etc.,
3. the variation of these properties as Γ varies within natural families of groups.

In the sequel we survey some key results and techniques in the study of zeta functions in the context of subgroup and subring growth (Section 3) and of representation growth (Section 4).

3 Subgroup and subring growth

3.1 Subgroup growth of finitely generated nilpotent groups

A finitely generated group Γ has only finitely many subgroups of each finite index n. We set, for $n \in \mathbb{N}$,

$$a_n(\Gamma) := \#\{H \leq \Gamma \mid |\Gamma : H| = n\}.$$

If $s_n(\Gamma) := \sum_{\nu \leq n} a_\nu(\Gamma) = O(n^a)$ for some $a \in \mathbb{R}$, then Γ is said to be of *polynomial subgroup growth* (*PSG*). Finitely generated, residually finite groups of PSG have been characterized as the virtually solvable groups of finite rank; see [37]. This class of groups includes the torsion-free, finitely generated nilpotent (or $\mathfrak{T}$-)groups. Let Γ be a $\mathfrak{T}$-group. Then the sequence $(a_n(\Gamma))$ is multiplicative. This follows from the facts that every finite index subgroup H of Γ contains a normal such subgroup, and that a finite nilpotent group is isomorphic to the direct product of its Sylow p-subgroups. In [27], Grunewald, Segal, and Smith pioneered the use of zeta functions in the theory of subgroup growth of $\mathfrak{T}$-groups. They studied the *subgroup zeta function*

$$\zeta_\Gamma(s) := \zeta_{(a_n(\Gamma))}(s) = \sum_{n=1}^{\infty} a_n(\Gamma)n^{-s}$$

of Γ via the Euler factorization

$$\zeta_\Gamma(s) = \prod_{p \text{ prime}} \zeta_{\Gamma,p}(s), \tag{3.1}$$

where, for each prime p, the local factor at p is defined via $\zeta_{\Gamma,p}(s) = \sum_{i=0}^{\infty} a_{p^i}(\Gamma)p^{-is}$. One of the main result of [27] is the following fundamental theorem.

Theorem 3.1 ([27, Theorem 1]) *For all primes p, the function $\zeta_{\Gamma,p}(s)$ is rational in p^{-s}, i.e., there exist polynomials $P_p, Q_p \in \mathbb{Q}[Y]$ such that*

$$\zeta_{\Gamma,p}(s) = P_p(p^{-s})/Q_p(p^{-s}).$$

The degrees of P_p and Q_p in Y are bounded.

The following is by now a classical example.

Example 3.2 ([27, Proposition 8.1]) Let

$$\mathbf{H}(\mathbb{Z}) = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ & 1 & \mathbb{Z} \\ & & 1 \end{pmatrix} \tag{3.2}$$

be the integral Heisenberg group. Then

$$\zeta_{\mathbf{H}(\mathbb{Z})}(s) = \zeta(s)\zeta(s-1)\zeta(2s-2)\zeta(2s-3)\zeta(3s-3)^{-1}, \tag{3.3}$$

where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \text{ prime}}(1-p^{-s})^{-1}$ is the Riemann zeta function.

It is of great interest to understand how the rational functions giving the local zeta functions in Euler factorizations like (3.1) vary with the prime p. It is known that the denominator polynomials $Q_p(Y)$ can be chosen to be of the form $\prod_{i\in I}(1-p^{a_i-b_is})$, for a finite index set I and nonnegative integers a_i, b_i, all depending only on Γ. Computing these integers, or even just a reasonably small set of candidates, however, remains a difficult problem. The numerator polynomials' variation with the prime p is even more mysterious. It follows from fundamental work of du Sautoy and Grunewald that there are finitely many varieties $V_1, \ldots, V_N$ defined over $\mathbb{Q}$, and rational functions $W_1(X,Y), \ldots, W_N(X,Y) \in \mathbb{Q}(X,Y)$ such that, for almost all primes p,

$$\zeta_{\Gamma,p}(s) = \sum_{i=1}^{N} |\overline{V_i}(\mathbb{F}_p)|W_i(p,p^{-s}), \tag{3.4}$$

where $\overline{V_i}$ denotes the reduction of V_i modulo p; cf. [17]. One may construct $\mathfrak{T}$-groups where the numbers $|\overline{V_i}(\mathbb{F}_p)|$ are not all polynomials in p; cf., for instance, [15]. Recent results determine the degree in Y of the rational functions $P_p/Q_p \in \mathbb{Q}(Y)$ in Theorem 3.1 for almost all primes p; cf. Corollary 3.9.

Variations of the sequence $(a_n(\Gamma))$ include the normal subgroup sequence $(a_n^{\triangleleft}(\Gamma))$, where

$$a_n^{\triangleleft}(\Gamma) := \#\{H \triangleleft \Gamma \mid |\Gamma : H| = n\}.$$

It gives rise to the *normal (subgroup) zeta function*

$$\zeta_{\Gamma}^{\triangleleft}(s) := \zeta_{(a_n^{\triangleleft}(\Gamma))}(s) = \sum_{n=1}^{\infty} a_n^{\triangleleft}(\Gamma)n^{-s}$$

of Γ. It also has an Euler factorization whose factors are rational in p^{-s} and, in principle, given by formulae akin to (3.4). The normal zeta function of the integral Heisenberg group (cf. (3.2)), for example, is equal to

$$\zeta^{\triangleleft}_{\mathbf{H}(\mathbb{Z})}(s) = \zeta(s)\zeta(s-1)\zeta(3s-2) = \prod_{p \text{ prime}} \frac{1}{(1-p^{-s})(1-p^{1-s})(1-p^{2-3s})};$$

cf. [27, Section 8].

It is interesting to ask how subgroup zeta functions of $\mathcal{T}$-groups, or their variations, vary under base extension. Given a number field K with ring of integers $\mathcal{O}$ one may consider, for instance, the $\mathcal{T}$-group $\mathbf{H}(\mathcal{O})$ of upper-unitriangular 3×3-matrices over $\mathcal{O}$. Then

$$\zeta^{\triangleleft}_{\mathbf{H}(\mathcal{O})}(s) = \prod_{p \text{ prime}} \zeta^{\triangleleft}_{\mathbf{H}(\mathcal{O}),p}(s). \tag{3.5}$$

The following result extends parts of [27, Theorem 2] and makes it more precise.

Theorem 3.3 ([44]) *For every $r \in \mathbb{N}$ and every finite family $\mathbf{f} = (f_1, \ldots, f_r) \in \mathbb{N}^r$, there exist* explicitly given *rational functions $W_{\mathbf{f}}(X, Y) \in \mathbb{Q}(X, Y)$, such that the following hold.*

1. *If p is a prime which is unramified in K and decomposes in K as $p\mathcal{O} = \prod_{i=1}^{r} \mathfrak{P}_i$ for prime ideals $\mathfrak{P}_i$ of $\mathcal{O}$ with inertia degrees $f_i = \log_p |\mathcal{O} : \mathfrak{P}_i|$ for $i = 1, \ldots, r$, then*
$$\zeta^{\triangleleft}_{\mathbf{H}(\mathcal{O}),p}(s) = W_{\mathbf{f}}(p, p^{-s}).$$
2. *Setting $d = |K : \mathbb{Q}| = \sum_{i=1}^{r} f_i$, we have*
$$W_{\mathbf{f}}(X^{-1}, Y^{-1}) = (-1)^n X^{\binom{3d}{2}} Y^{5d} W_{\mathbf{f}}(X, Y). \tag{3.6}$$

The proof of Theorem 3.3 is essentially combinatorial. In the case that p splits completely, i.e., $\mathbf{f} = (1, \ldots, 1)$, it proceeds by organizing the infinite sums defining the local zeta functions as sums indexed by pairs of partitions (λ, μ), each of at most n parts, where λ dominates μ. We further partition the infinite set of such pairs into $C_n = \frac{1}{n+1}\binom{2n}{n}$ (the n-th Catalan number) parts, indexed by the Dyck words of length $2n$, determined by the "overlap" between λ and μ. This subdivision by Dyck words is suggested by a simple lemma, attributed to Birkhoff, that determines the numbers of subgroups of type μ in a finite abelian p-group of type λ. For each fixed Dyck word, we express the corresponding partial sum of the local zeta function in terms of natural generalizations of combinatorially defined generating functions, first studied by Igusa (cf. [49, Theorem 4]) and Stanley [47]. Remarkably, a functional equation of the form (3.6) is already satisfied by each of the C_n partial sums. If p does not split completely, the strategy above still works after some moderate modification.

The functional equation (3.6) reflects the Gorenstein property of certain face rings. That such a functional equation holds for *almost all* primes p follows from [50, Theorem B]; that it holds in fact for all unramified primes is additional information. Note that $3d = h(\mathbf{H}(\mathcal{O}))$ and $5d = h(\mathbf{H}(\mathcal{O})) + h(\mathbf{H}(\mathcal{O})/Z(\mathbf{H}(\mathcal{O})))$, the sums of the Hirsch lengths of the nontrivial quotients by the terms of the upper central series of $\mathbf{H}(\mathcal{O})$. Here, given a $\mathcal{T}$-group G, we write $h(G)$ for the Hirsch length of G, i.e., the number of infinite cyclic factors in a decomposition series of G.

Formulae for the Euler factors in (3.5) indexed by primes which are nonsplit (but possibly ramified) in K are given in [45].

3.2 Subring growth of additively finitely generated rings

By a *ring* we shall always mean a finitely generated, torsion-free abelian group, together with a bi-additive multiplication – not necessarily associative, commutative, or unital. Examples of such rings include $\mathbb{Z}^d$ (e.g., with null-multiplication or with componentwise multiplication), the rings of integers in number fields, and Lie rings, that is rings with a multiplication (or "Lie bracket") which is alternating and satisfies the Jacobi identity. Examples of Lie rings include "semi-simple" matrix rings such as $\mathfrak{sl}_N(\mathbb{Z})$ and the Heisenberg Lie ring

$$\mathfrak{h}(\mathbb{Z}) = \begin{pmatrix} 0 & \mathbb{Z} & \mathbb{Z} \\ & 0 & \mathbb{Z} \\ & & 0 \end{pmatrix},$$

with Lie bracket induced from $\mathfrak{gl}_3(\mathbb{Z})$.

The subring sequence of a ring Λ is $(a_n(\Lambda))$, where

$$a_n(\Lambda) := \#\{H \leq \Gamma \mid |\Gamma : H| = n\}.$$

It is encoded in the *subring zeta function* of Λ, that is the Dirichlet generating series

$$\zeta_\Lambda(s) = \zeta_{(a_n(\Lambda))}(s) = \sum_{n=1}^{\infty} a_n(\Lambda) n^{-s}.$$

In contrast to the case of subgroup growth, polynomial growth requires no assumption on the multiplicative structure: indeed, the null-multiplication on $\mathbb{Z}^d$ yields a trivial polynomial upper bound on $s_n(\Lambda) := \sum_{\nu \leq n} a_\nu(\Lambda)$. Also, multiplicativity of the subring growth function $n \mapsto a_n(\Lambda)$ follows from the Chinese Reminder Theorem. Consequently, the subring zeta function of Λ satisfies the following Euler factorization:

$$\zeta_\Lambda(s) = \prod_{p \text{ prime}} \zeta_{\Lambda,p}(s).$$

Many of the structural results for local zeta functions of $\mathfrak{T}$-groups have analogues in the setting of zeta functions of rings. One example is the following.

Theorem 3.4 ([27, Theorem 3.5]) *For all primes p, the function $\zeta_{\Lambda,p}(s)$ is rational in p^{-s}, i.e., there exist polynomials $P_p, Q_p \in \mathbb{Q}[Y]$ such that*

$$\zeta_{\Lambda,p}(s) = P_p(p^{-s})/Q_p(p^{-s}).$$

The degrees of P_p and Q_p in Y are bounded.

As in the context of subgroup growth of $\mathfrak{T}$-groups, one also considers variations such as ideal growth of rings. The *ideal zeta function* of a ring enumerates its ideals of finite additive index. These zeta functions, too, enjoy Euler factorizations indexed by the rational primes. A rationality result analogous to Theorem 3.4 holds for the

local factors. From this perspective one recovers, for example, the classical Dedekind zeta function of a number field, enumerating ideals of finite index in the number field's ring of integers.

In fact, the study of subgroup zeta functions of $\mathcal{T}$-groups as outlined in Section 3.1 may – to a large extent – be reduced to the study of subring zeta functions of nilpotent Lie rings. Indeed, a key tool in the analysis of [27] is a linearization technique: the Mal'cev correspondence associates to each $\mathcal{T}$-group Γ a nilpotent Lie ring $\Lambda(\Gamma)$, that is a Lie ring whose additive group is isomorphic to $\mathbb{Z}^d$, where $d = h(\Gamma)$ is the Hirsch length of Γ, which is nilpotent with respect to the Lie bracket; see [27, Section 4] for details on the Mal'cev correspondence and its consequences for zeta functions of $\mathcal{T}$-groups. One of these consequences is the fact that, for almost all primes p,

$$\zeta_{\Gamma,p}(s) = \zeta_{\Lambda(\Gamma),p}(s); \tag{3.7}$$

cf. [27, Theorem 4.1]. In nilpotency class at most 2 this equality holds for all primes p. The formula (3.3), for instance, coincides with the subring zeta function of the Heisenberg Lie ring $\mathfrak{h}(\mathbb{Z}) = \Lambda(\mathbf{H}(\mathbb{Z}))$.

Maybe it is due to connections to subgroup growth like the ones just sketched that the study of subring growth has long focussed on *Lie* rings. The following example does not arise in this context.

Example 3.5 Let $\mathcal{O}$ be the ring of integers in a number field K and, for $n \in \mathbb{N}$, let $b_n(\mathcal{O})$ denote the number of subrings of $\mathcal{O}$ of index n, containing $1 \in \mathcal{O}$. The resulting zeta function $\zeta_{(b_n(\mathcal{O}))}(s)$ may be called the *order zeta function* $\eta_K(s)$ of K. The function η_K has an Euler factorization indexed by the rational primes, though – in contrast to the Dedekind zeta function $\zeta_K(s)$ – not generally by the prime ideals of $\mathcal{O}$. Clearly $\eta_{\mathbb{Q}} = 1$. If $d = |K : \mathbb{Q}| = 2$, then $\eta_K(s) = \zeta(s)$, the Riemann zeta function. For $d = 3$ it is known that

$$\eta_K(s) = \frac{\zeta_K(s)}{\zeta_K(2s)}\zeta(2s)\zeta(3s-1);$$

see [12]. For $d = 4$, Nakagawa computes in [41] the Euler factors $\eta_{K,p}(s)$, where p ranges over the primes with arbitrary but fixed decomposition behaviour in K. Remarkably, the resulting formulae are rational functions in p and p^{-s} though not, in general, expressible in terms of translates of local Dedekind zeta functions. It is interesting to establish whether this uniformity on sets of primes with equal decomposition behaviour is a general phenomenon. Of particular interest is the case of primes which split totally in K, i.e., primes p such that $p\mathcal{O} = \mathfrak{P}_1 \cdots \mathfrak{P}_d$, where $\mathfrak{P}_1, \ldots, \mathfrak{P}_d$ are pairwise distinct prime ideals of $\mathcal{O}$ with trivial residue field extension. For such primes, it is not hard to see that

$$\eta_{K,p}(s) = \zeta_{\mathbb{Z}^{d-1},p}(s),$$

where we consider $\mathbb{Z}^{d-1}$ as a ring with componentwise multiplication.

Theorem 3.6 ([41] and [35, Proposition 6.3]) *Consider $\mathbb{Z}^3$ as ring with componentwise multiplication. Then $\zeta_{\mathbb{Z}^3}(s) = \prod_{p \text{ prime}} \zeta_{\mathbb{Z}_p^3,p}(s)$, where, setting $t = p^{-s}$, we*

have

$$\zeta_{\mathbb{Z}^3,p}(s) = \Big(1 + 4t + 2t^2 + (4p-3)t^3 + (5p-1)t^4 + (p^2-5p)t^5 + (3p^2-4p)t^6 \\ - 2p^2t^7 - 4p^2t^8 - p^2t^9\Big)\Big/(1-t)^2(1-p^2t^4)(1-p^3t^6). \quad (3.8)$$

The evidence available for $d \leq 3$ suggests a positive answer to the following question.

Question 3.7 Do there exist rational functions $W_d(X,Y) \in \mathbb{Q}(X,Y)$, for $d \in \mathbb{N}$, such that, for all primes p,

$$\zeta_{\mathbb{Z}_p^d,p}(s) = W_d(p, p^{-s})?$$

Local zeta functions such as the ones given in (3.8) exhibit a curious palindromic symmetry under inversion of p. This is no coincidence, as the following result shows.

Theorem 3.8 ([50, Theorem A]) *Let Λ be a ring with $(\Lambda,+) \cong \mathbb{Z}^d$. Then, for almost all primes p,*

$$\zeta_{\Lambda,p}(s)|_{p\to p^{-1}} = (-1)^d p^{\binom{d}{2}-ds}\zeta_{\Lambda,p}(s). \quad (3.9)$$

Corollary 3.9 *For almost all primes p, $\deg_{p^{-s}}(\zeta_{\Lambda,p}(s)) = -d$.*

Via the Mal'cev correspondence, Theorem 3.8 yields an analogous statement for almost all of the local factors $\zeta_{\Gamma,p}(s)$ of the subgroup zeta function $\zeta_\Gamma(s)$ of a $\mathfrak{T}$-group Γ; cf. (3.7). There are analogous results giving functional equations akin to (3.9) for ideal zeta functions of $\mathfrak{T}$-groups – or equivalently, again by the Mal'cev correspondence, nilpotent Lie rings of finite additive rank – of nilpotency class at most 2. There are, however, examples of $\mathfrak{T}$-groups of nilpotency class 3 whose local normal subgroup zeta functions do not satisfy functional equations like (3.9); cf. [21, Theorem 1.1].

Other variants of subgroup zeta functions of $\mathfrak{T}$-groups which have been studied include those encoding the numbers of finite-index subgroups whose profinite completion is isomorphic to the one of the ambient group. These *pro-isomorphic zeta functions* also enjoy Euler product decompositions, indexed by the rational primes, whose factors are rational functions. It is an interesting open problem to characterise the $\mathfrak{T}$-groups for which these local factors satisfy functional equations comparable to (3.9). For positive results in this direction see [20, 10]. An example of a $\mathfrak{T}$-group (of nilpotency class 4 and Hirsch length 25) whose pro-isomorphic zeta function's local factors do not satisfy such functional equations was recently given in [11].

3.3 Taking the limit $p \to 1$: reduced and topological zeta functions of groups and rings

Numerous mathematical concepts, theorems, and identities allow natural q-analogues. Featuring an additional parameter q, often interpreted as a prime power, these analogues return the original object upon setting $q = 1$. Examples include the Gaussian q-binomial coefficients, generalizing classical binomial coefficients and Heine's basic hypergeometric series, generalizing ordinary hypergeometric series.

An idea that only recently took hold in the theory of zeta functions of groups and rings is to interpret local such zeta functions as "p-analogues" of certain limit objects as $p \to 1$ and to investigate the limit objects with tools from combinatorics or commutative algebra.

3.3.1 Reduced zeta functions

One way to make this idea rigorous leads, for instance, to the concept of the *reduced zeta function* $\zeta_{\Lambda,\mathrm{red}}(t)$ of a ring Λ. Informally, this rational function in a variable t over the rationals is obtained by setting $p = 1$ in the coefficients of the p-adic subring zeta function of Λ, considered as a series in $t = p^{-s}$; formally, it arises by specializing the coefficients of the motivic zeta function associated to Λ via the Euler characteristic; cf. [19, 22]. Under some very restrictive conditions on Λ, the reduced zeta function $\zeta_{\Lambda,\mathrm{red}}(t)$ is known to enumerate the integral points of a rational polyhedral cone. In the language of commutative algebra this means that $\zeta_{\Lambda,\mathrm{red}}(t)$ is the Hilbert series of an affine monoid algebra attached to a Diophantine system of linear inequalities. For general rings a somewhat more multifarious picture seems to emerge, as the following example indicates.

Example 3.10 Consider $\Lambda = \mathbb{Z}^3$ viewed as a ring with componentwise multiplication. Heuristically, setting $p = 1$ in (3.8) we obtain

$$\zeta_{\mathbb{Z}^3,\mathrm{red}}(t) = \frac{1 + 5t + 6t^2 + 3t^3 + 6t^4 + 5t^5 + t^6}{(1-t)(1-t^2)(1-t^6)}.$$

Intriguingly, this rational function is not the generating function of a polyhedral cone, but does exhibit some tell-tale signs of the Hilbert series of a graded Cohen-Macaulay (even Gorenstein) algebra of dimension 3.

3.3.2 Topological zeta functions

Topological zeta functions offer another way to define a limit as $p \to 1$ of families of p-adic zeta functions. They were first introduced in the realm of Igusa's p-adic zeta function as singularity invariants of hypersurfaces [13]. Informally, the topological zeta function is the leading term of the expansion of the p-adic zeta function in $p-1$. Formally, it may be obtained by specialising the motivic zeta function; cf. [14]. Whereas the latter lives in the power series ring over a certain completion of a localization of a Grothendieck ring of algebraic varieties, the topological zeta function is just a rational function in one variable s, say, over the rationals. The topological zeta function $\zeta_{\Lambda,\mathrm{top}}(s)$ of a ring was introduced in [19].

Example 3.11 The topological zeta function of $\mathbb{Z}^3$ (cf. Example 3.10) is

$$\zeta_{\mathbb{Z}^3,\mathrm{top}}(s) = \frac{9s-1}{s^2(2s-1)^2}.$$

In [43] Rossmann develops an effective method for computing topological zeta functions associated to groups, rings, and modules. It is built upon explicit convex-geometric formulae for a class of p-adic integrals under suitable non-degeneracy conditions with respect to associated Newton polytopes. This method yields examples of

explicit formulae for topological zeta functions of objects whose p-adic zeta functions are well out of computational reach. For a number of intriguing conjectures about arithmetic properties of topological zeta functions see [43, Section 8]. Rossmann implemented his algorithm in Sage; together with a sequel to [43] it will be publicly available shortly.

4 Representation growth

Let Γ be a group. Consider, for $n \in \mathbb{N}$, the set $\mathrm{Irr}_n(\Gamma)$ of n-dimensional irreducible complex representations of Γ up to isomorphism. If Γ has additional structure, we restrict our attention to representations respecting this structure. For instance, if Γ is a topological group, we only consider continuous representations. The group Γ is called (*representation*) *rigid* if $r_n(\Gamma) := \#\,\mathrm{Irr}_n(\Gamma)$ is finite for all n. In this case, the Dirichlet generating series

$$\zeta_\Gamma^{\mathrm{irr}}(s) := \zeta_{(r_n(\Gamma))}(s) = \sum_{n=1}^{\infty} r_n(\Gamma) n^{-s}$$

is called the *representation zeta function* of Γ. We discuss several classes of groups whose representation zeta functions (or natural variants thereof) have recently attracted attention. These are

1. finitely generated nilpotent groups,
2. arithmetic groups in characteristic 0,
3. algebraic groups,
4. compact p-adic analytic groups,
5. iterated wreath products and branch groups.

Throughout, let K be a number field with ring of integers $\mathcal{O} = \mathcal{O}_K$. We write

$$\zeta_K(s) = \sum_{I \lhd \mathcal{O}} |\mathcal{O} : I|^{-s} = \prod_{\mathfrak{p} \in \mathrm{Spec}\, \mathcal{O}} (1 - |\mathcal{O}/\mathfrak{p}|^{-s})^{-1}$$

for the Dedekind zeta function of K. Note that $\zeta_{\mathbb{Q}}(s) = \zeta(s)$. By representations we will always mean complex representations.

4.1 Finitely generated nilpotent groups

Let Γ be a $\mathcal{T}$-group. Unless Γ is trivial, the sets $\mathrm{Irr}_n(\Gamma)$ are not all finite. Indeed, a nontrivial $\mathcal{T}$-group surjects onto the infinite cyclic group and thus has infinitely many one-dimensional representations. We therefore consider finite-dimensional representations up to twists by one-dimensional representations. More precisely, two representations $\rho_1, \rho_2 \in \mathrm{Irr}_n(\Gamma)$ are said to be *twist-equivalent* if there exists $\chi \in \mathrm{Irr}_1(\Gamma)$ such that ρ_1 is isomorphic to $\rho_2 \otimes \chi$. The numbers $\widetilde{r}_n(\Gamma)$ of isomorphism classes of irreducible, complex n-dimensional representations of Γ are all finite; cf. [36, Theorem 6.6]. We define the *representation zeta function* of Γ to be the Dirichlet generating series

$$\widetilde{\zeta_\Gamma^{\mathrm{irr}}}(s) := \zeta_{(\widetilde{r}_n(\Gamma))}(s) = \sum_{n=1}^{\infty} \widetilde{r}_n(\Gamma) n^{-s}.$$

The coefficients $\widetilde{r}_n(\Gamma)$ grow polynomially, so $\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma}(s)$ converges on some complex right-half plane. The precise abscissa of convergence of $\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma}(s)$ is an interesting invariant of Γ.

The function $n \mapsto \widetilde{r}_n(\Gamma)$ is multiplicative, which yields the Euler factorization

$$\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma}(s) = \prod_{p \text{ prime}} \zeta^{\widetilde{\mathrm{irr}}}_{\Gamma,p}(s), \tag{4.1}$$

where, for a prime p, the local factor $\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma,p}(s) = \sum_{i=0}^{\infty} \widetilde{r}_{p^i}(\Gamma)(p^{-s})^i$ enumerates twist-isoclasses of representations of Γ of p-power dimension.

Theorem 4.1 ([28, Theorem 8.5]) *For all primes p, the function $\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma,p}(s)$ is rational in p^{-s}, i.e., there exist polynomials $P_p, Q_p \in \mathbb{Q}[Y]$ such that*

$$\zeta^{\widetilde{\mathrm{irr}}}_{\Gamma,p}(s) = P_p(p^{-s})/Q_p(p^{-s}).$$

The degrees of P_p and Q_p in Y are bounded.

The proof uses model-theoretic results on definable equivalence classes. We illustrate this important rationality result with a simple but instructive example.

Example 4.2 Consider the integral Heisenberg group $\mathbf{H}(\mathbb{Z})$; cf. (3.2). Then

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathbb{Z})}(s) = \sum_{n=1}^{\infty} \varphi(n) n^{-s} = \frac{\zeta(s-1)}{\zeta(s)} = \prod_{p \text{ prime}} \frac{1-p^{-s}}{1-p^{1-s}}, \tag{4.2}$$

where φ is the Euler totient function; cf. [42].

It turns out that the formula in (4.2) behaves uniformly under some base extensions, as we shall now explain. Consider, for example, the Heisenberg group $\mathbf{H}(\mathcal{O})$ over $\mathcal{O}$, i.e., the group of upper-unitriangular 3×3-matrices over $\mathcal{O}$. Then

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathcal{O})}(s) = \frac{\zeta_K(s-1)}{\zeta_K(s)} = \prod_{\mathfrak{p} \in \operatorname{Spec} \mathcal{O}} \frac{1-|\mathcal{O}/\mathfrak{p}|^{-s}}{1-|\mathcal{O}/\mathfrak{p}|^{1-s}}. \tag{4.3}$$

For quadratic number fields this was proven by Ezzat in [24]. The general case follows from [48, Theorem B].

Each factor $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathcal{O}),\mathfrak{p}}(s) := (1-|\mathcal{O}/\mathfrak{p}|^{-s})/(1-|\mathcal{O}/\mathfrak{p}|^{1-s})$ of the Euler factorization (4.3) is interpretable as a representation zeta function associated to a pro-p group. Indeed, for $\mathfrak{p} \in \operatorname{Spec} \mathcal{O}$, we denote by $\mathcal{O}_{\mathfrak{p}}$ the completion of $\mathcal{O}$ at $\mathfrak{p}$. Then $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathcal{O}),\mathfrak{p}}(s)$ is equal to the zeta function $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathcal{O}_{\mathfrak{p}})}(s)$ of the pro-p group

$$\mathbf{H}(\mathcal{O}_{\mathfrak{p}}) = \begin{pmatrix} 1 & \mathcal{O}_{\mathfrak{p}} & \mathcal{O}_{\mathfrak{p}} \\ & 1 & \mathcal{O}_{\mathfrak{p}} \\ & & 1 \end{pmatrix},$$

enumerating *continuous* irreducible representations of $\mathbf{H}(\mathcal{O}_{\mathfrak{p}})$ up to twists by *continuous* one-dimensional representations. We note the following features of $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{H}(\mathcal{O})}(s)$.

1. Whilst the Euler factorization (4.2) illustrates the general factorization (4.1), the factorization (4.3) is finer than (4.1). In fact, for each rational prime p,
$$\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{H}(\mathcal{O}),p}}(s) = \prod_{\mathfrak{p}|p\mathcal{O}} \widetilde{\zeta^{\mathrm{irr}}_{\mathbf{H}(\mathcal{O}_\mathfrak{p})}}(s).$$
2. The factors of the "fine" Euler factorization (4.3) are indexed by the nonzero prime ideals $\mathfrak{p}$ of $\mathcal{O}$, and are each given by a rational functions in q^{-s}, where $q = |\mathcal{O}/\mathfrak{p}|$ denotes the residue field cardinality.
3. Each factor of the Euler factorization (4.3) satisfies the functional equation
$$\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{H}(\mathcal{O}),\mathfrak{p}}}(s)\Big|_{q\to q^{-1}} = \frac{1-q^{-s}}{1-q^{1-s}}\Big|_{q\to q^{-1}} = \frac{1-q^{s}}{1-q^{-1+s}} = q\,\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{H}(\mathcal{O}),\mathfrak{p}}}(s).$$

As we shall see, all of these points are special cases of general phenomena.

We consider in the sequel families of groups obtained from Lie lattices. Let, more precisely, Λ be an $\mathcal{O}$-Lie lattice, i.e., a free and finitely generated $\mathcal{O}$-module, together with an antisymmetric, bi-additive form $[\,,\,]\colon \Lambda\times\Lambda\to\Lambda$, called 'Lie bracket', which satisfies the Jacobi identity. Assume further that Λ is nilpotent with respect to $[\,,\,]$ of class c, and let Λ' denote the derived Lie lattice $[\Lambda,\Lambda]$. If Λ satisfies $\Lambda'\subseteq c!\Lambda$, then it gives rise to a unipotent group scheme $\mathbf{G}_\Lambda$ over $\mathcal{O}$, via the Hausdorff series as we shall now explain. The Hausdorff series $F(X,Y)$ is a formal power series in two noncommuting variables X and Y, with rational coefficients. The Hausdorff formula gives an expression for this series in terms of Lie terms:
$$F(X,Y) = X+Y+\tfrac{1}{2}[X,Y]+\tfrac{1}{12}\big([[X,Y],Y]-[[X,Y],X]\big)+\ldots, \qquad (4.4)$$
where $[A,B] := AB-BA$. See, e.g., [32, Chapter I, Section 7.4] for further details on the Hausdorff series.

For an $\mathcal{O}$-algebra R, let $\Lambda(R) := \Lambda\otimes_\mathcal{O} R$. The assumption that $\Lambda'\subseteq c!\Lambda$ allows one to define on the set $\Lambda(R)$ a group structure $*$ by setting, for $x,y\in\Lambda(R)$,
$$x*y := F(x,y), \quad x^{-1} = -x.$$
Note that, by the nilpotency of Λ, the Hausdorff formula (4.4) yields an expression for $x*y$ as a linear combination of Lie terms in x and y. In this way one obtains a unipotent group scheme $\mathbf{G}_\Lambda$ over $\mathcal{O}$, representing the functor $R\mapsto(\Lambda(R),*)$. In nilpotency class $c=2$, one may define the group scheme $\mathbf{G}_\Lambda$ directly and avoiding the condition $\Lambda'\subseteq c!\Lambda$; cf. [48, Section 2.4].

By taking rational points of $\mathbf{G}_\Lambda$ we obtain a multitude of groups, all originating from the same global Lie lattice Λ. The group $\mathbf{G}_\Lambda(\mathcal{O})$ of $\mathcal{O}$-rational points, for instance, is a $\mathcal{T}$-group of Hirsch length $\mathrm{rk}_\mathbb{Z}(\mathcal{O})\,\mathrm{rk}_\mathcal{O}(\Lambda)$. By considering the $\mathcal{O}_\mathfrak{p}$-rational points of $\mathbf{G}_\Lambda$ for a nonzero prime ideal $\mathfrak{p}$ of $\mathcal{O}$, we obtain the nilpotent pro-p group $\mathbf{G}_\Lambda(\mathcal{O}_\mathfrak{p})$. It is remarkable that many features of the representation growth of groups of the form $\mathbf{G}_\Lambda(\mathcal{O}_\mathfrak{p})$ only depend on the lattice Λ, and not on the local ring $\mathcal{O}_\mathfrak{p}$.

Remark 4.3 We comment on connections between the above construction and the Mal'cev correspondence between $\mathcal{T}$-groups and nilpotent Lie rings. Starting from a

$\mathcal{T}$-group Γ, there exists a $\mathbb{Q}$-Lie algebra $\mathcal{L}_\Gamma(\mathbb{Q})$ and an injective mapping $\log : \Gamma \to \mathcal{L}_\Gamma(\mathbb{Q})$, such that $\log(\Gamma)$ spans $\mathcal{L}_\Gamma(\mathbb{Q})$ over $\mathbb{Q}$. Whilst $\log(\Gamma)$ needs not, in general, be a Lie lattice inside $\mathcal{L}_\Gamma(\mathbb{Q})$, there always exists a subgroup H of Γ of finite index with this property, satisfying $\log(H)' \subseteq c! \log(H)$, where c is the nilpotency class of Γ. Setting $\Lambda = \log(H)$, we recover H as the group of $\mathbb{Z}$-rational points of $\mathbf{G}_\Lambda$.

Let now Λ be again a nilpotent $\mathcal{O}$-Lie lattice of class c, and suppose that $\Lambda' \subseteq c!\Lambda$. Denote by $\mathbf{G}_\Lambda$ the associated unipotent group scheme. For every finite extension L of K, with ring of integers $\mathcal{O}_L$, we obtain a $\mathcal{T}$-group $\mathbf{G}_\Lambda(\mathcal{O}_L)$ and, for every nonzero prime ideal $\mathfrak{P} \in \operatorname{Spec} \mathcal{O}_L$, a pro-$p$ group $\mathbf{G}_\Lambda(\mathcal{O}_{L,\mathfrak{P}})$.

Theorem 4.4 ([48]) *For every finite extension L of K, with ring of integers $\mathcal{O}_L$,*

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathcal{O}_L)}(s) = \prod_{\mathfrak{P} \in \operatorname{Spec} \mathcal{O}_L} \zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathcal{O}_{L,\mathfrak{P}})}(s), \tag{4.5}$$

where, for each prime ideal $\mathfrak{P} \in \operatorname{Spec} \mathcal{O}_L$, the factor $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathcal{O}_{L,\mathfrak{P}})}(s)$ enumerates the continuous finite-dimensional irreducible representations of $\mathbf{G}_\Lambda(\mathcal{O}_{L,\mathfrak{P}})$ up to twisting by continuous one-dimensional representations. Moreover, the following hold.

1. *For each rational prime p,*

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathcal{O}_L),p}(s) = \prod_{\mathfrak{P} \mid p\mathcal{O}} \zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathcal{O}_{L,\mathfrak{P}})}(s).$$

2. *There exists a finite subset $S \subset \operatorname{Spec} \mathcal{O}$, an integer $t \in \mathbb{N}$, and a rational function $R(X_1, \ldots, X_t, Y) \in \mathbb{Q}(X_1, \ldots, X_t, Y)$ such that, for every prime ideal $\mathfrak{p} \notin S$, the following holds. There exist algebraic integers $\lambda_1, \ldots, \lambda_t$, depending on $\mathfrak{p}$, such that, for all finite extensions $\mathfrak{O}$ of $\mathfrak{o} = \mathcal{O}_\mathfrak{p}$,*

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathfrak{O})}(s) = R(\lambda_1^f, \ldots, \lambda_t^f, q^{-fs}), \tag{4.6}$$

   where $q = |\mathcal{O} : \mathfrak{p}|$ and $|\mathcal{O}_L : \mathfrak{P}| = q^f$.

3. *Setting $d = \dim_K(\Lambda' \otimes_\mathcal{O} K)$, the following functional equation holds:*

$$\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathfrak{O})}(s)\Big|_{\substack{q \to q^{-1} \\ \lambda_i \to \lambda_i^{-1}}} = q^{fd} \zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathfrak{O})}(s). \tag{4.7}$$

As a corollary, we obtain that $\zeta^{\widetilde{\mathrm{irr}}}_{\mathbf{G}_\Lambda(\mathfrak{O})}(s)$ is rational in q^{-fs}. In particular, the dimensions of the continuous representations of the pro-p group $\mathbf{G}_\Lambda(\mathfrak{O})$ are all powers of q^f.

Example 4.2 illustrates Theorem 4.4. Indeed, the Heisenberg group scheme $\mathbf{H}$ is defined over $K = \mathbb{Q}$. We have $d = 1$ and, in (2), we may take $S = \varnothing$, $t = 1$, $R(X, Y) = (1 - Y)/(1 - XY)$ and $\lambda_1 = p$.

We say a few words about the proof of Theorem 4.4, referring to [48] for all details. The Euler factorization (4.5) and the statement (1) follow easily from strong approximation for unipotent groups. The key tool to enumerate the representation zeta functions of pro-p groups like $\mathbf{G}_\Lambda(\mathfrak{O})$ is the Kirillov orbit method. Wherever this method

is applicable, it parametrizes the irreducible representations of a group in terms of the co-adjoint orbits in the Pontryagin dual of a corresponding Lie algebra. In the case at hand, it reduces the problem of enumerating twist-isoclasses of continuous finite-dimensional irreducible representations of groups of the form $\mathbf{G}_\Lambda(\mathfrak{O})$ to that of enumerating certain orbits in the duals of the derived $\mathfrak{O}$-Lie lattices $\Lambda(\mathfrak{O})' = (\Lambda \otimes_{\mathcal{O}} \mathfrak{O})'$. By translating the latter into the problem of evaluating p-adic integrals, one reduces the problem further to the problem of enumerating p-adic points on certain algebraic varieties, which only depend on Λ. In this way, one can show that there exist finitely many smooth projective varieties V_i defined over $\mathcal{O}$, and rational functions $W_i(X,Y) \in \mathbb{Q}(X,Y)$, $i = 1, \ldots, N$, such that, if $\mathfrak{p}$ avoids a finite set $S \subset \operatorname{Spec} \mathcal{O}$,

$$\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{G}_\Lambda(\mathfrak{O})}}(s) = \sum_{i=1}^{N} |\overline{V_i}(\mathbb{F}_{q^f})| W_i(q^f, q^{-fs}),$$

where $\overline{V_i}$ denotes reduction modulo $\mathfrak{p}$. By the Weil conjectures there exist, for each $i \in \{1, \ldots, N\}$, algebraic integers λ_{ij}, $j = \{0, \ldots, 2 \dim V_i\}$, such that

$$|\overline{V_i}(\mathbb{F}_{q^f})| = \sum_{j=0}^{2 \dim V_i} (-1)^j \lambda_{ij}^f$$

and

$$\sum_{j=0}^{2 \dim V_i} (-1)^j \lambda_{ij}^{-f} = q^{f \dim W_i} \sum_{j=0}^{2 \dim V_i} (-1)^j \lambda_{ij}^f.$$

This remarkable symmetry is behind the functional equations for the Hasse-Weil zeta functions of the varieties $\overline{V_i}$ and also functional equations such as (3.9). The rational functions W_i come from the enumeration of rational points of rational polyhedral cones.

Question 4.5 Let $\mathbf{G}_\Lambda$ and $\mathcal{O}_L$ be as specified above. Is the abscissa $\alpha^{\mathrm{irr}}(\mathbf{G}_\Lambda(\mathcal{O}_L))$ of $\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{G}_\Lambda(\mathcal{O}_L)}}(s)$ always a rational number? Is it independent of L?

In general, the algebraic varieties V_i are obtained from resolutions of singularities of certain – in general highly singular – varieties, and are difficult to compute explicitly. We give some of the relatively few explicit examples of representation zeta functions of $\mathcal{T}$-groups we have at the moment.

Example 4.6 Let $d \in \mathbb{N}_{>1}$ and $\mathfrak{f}_{d,2}$ the free nilpotent Lie ring on d generators of nilpotency class 2, of additive rank $d + \binom{d}{2} = \binom{d+1}{2}$. We write $\mathbf{F}_{d,2}$ for the unipotent group scheme $\mathbf{G}_{\mathfrak{f}_{d,2}}$ associated to this $\mathbb{Z}$-Lie lattice. For $d = 2$ we obtain $\mathbf{F}_{2,2} = \mathbf{H}$, the Heisenberg group scheme. We also recover the free class-2-nilpotent group on d generators as $\mathbf{F}_{d,2}(\mathbb{Z})$. We write $d = 2\lfloor d/2 \rfloor + \varepsilon$ for $\varepsilon \in \{0, 1\}$. The following generalizes (4.3).

Theorem 4.7 ([48, Theorem B]) *Let $\mathcal{O}$ be the ring of integers of a number field K. Then*

$$\widetilde{\zeta^{\mathrm{irr}}_{\mathbf{F}_{d,2}(\mathcal{O})}}(s) = \prod_{i=0}^{\lfloor d/2 \rfloor} \frac{\zeta_K(s - 2(\lfloor d/2 \rfloor + i + \varepsilon) + 1)}{\zeta_K(s - 2i)}.$$

E. Avraham has computed the local factors of the representation zeta function of the groups $\mathbf{F}_{2,3}(\mathcal{O}[\frac{1}{6}])$; see [7]. For further explicit examples of representation zeta functions of $\mathcal{T}$-groups see [23, 46].

4.2 Arithmetic lattices in semisimple groups

Let S be a finite set of places of a number field K, including all archimedean ones, and let $\mathcal{O}_S$ denote the S-integers of K. Let further $\mathbf{G}$ be an affine group scheme over $\mathcal{O}_S$ whose generic fibre is connected, simply-connected semi-simple algebraic group defined over K, together with a fixed embedding $\mathbf{G} \hookrightarrow \mathrm{GL}_N$ for some $N \in \mathbb{N}$. Let $\Gamma = \mathbf{G}(\mathcal{O}_S)$. Then Γ has polynomial representation growth if and only if Γ has the weak Congruence Subgroup Property, i.e., the congruence kernel, that is the kernel of the natural surjection

$$\widehat{\mathbf{G}(\mathcal{O}_S)} \to \mathbf{G}(\widehat{\mathcal{O}_S}) \cong \prod_{\mathfrak{p} \in (\operatorname{Spec} \mathcal{O}) \setminus S} \mathbf{G}(\mathcal{O}_{\mathfrak{p}}), \tag{4.8}$$

is finite; cf. [38]. Here $\mathbf{G}(\widehat{\mathcal{O}_S})$ denotes the congruence completion of $\mathbf{G}(\mathcal{O}_S)$. For simplicity we assume in the sequel that Γ actually has the strong Congruence Subgroup Property, i.e., that the congruence kernel is trivial, so that the surjection (4.8) is an isomorphism. A prototypical example of such a group is the group $\mathrm{SL}_N(\mathbb{Z})$ for $N \geq 3$.

On the level of representation zeta functions, the triviality of the congruence kernel is reflected by an Euler factorization, similar to but different from those previously discussed, be it in the context of subgroup and subring growth or of representation growth of $\mathcal{T}$-groups. The Euler factorization features two types of factors: the *archimedean factors* are equal to $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathbb{C})}(s)$, the so-called *Witten zeta function*, that is the Dirichlet generating series enumerating the rational finite-dimensional irreducible complex representations of the algebraic group $\mathbf{G}(\mathbb{C})$. The *non-archimedean factors*, on the other hand, are the representation zeta functions $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_{\mathfrak{p}})}(s)$, where $\mathfrak{p} \notin S$. These Dirichlet generating series enumerate the continuous finite-dimensional irreducible complex representations of the p-adic analytic groups $\mathbf{G}(\mathcal{O}_{\mathfrak{p}})$.

Proposition 4.8 ([33, Proposition 4.6]) *The following Euler factorization holds:*

$$\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_S)}(s) = \zeta^{\mathrm{irr}}_{\mathbf{G}(\mathbb{C})}(s)^{|K:\mathbb{Q}|} \prod_{\mathfrak{p} \in (\operatorname{Spec} \mathcal{O}) \setminus S} \zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_{\mathfrak{p}})}(s). \tag{4.9}$$

It is a problem of central importance to compute the abscissa of convergence $\alpha(\mathbf{G}(\mathcal{O}_S))$ of the representation zeta function $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_S)}(s)$. It is known that $\alpha(\mathbf{G}(\mathcal{O}_S))$ is always a rational number; see [2, Theorem 1.2] and compare Question 4.5.

The two types of factors of $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O})}(s)$ in (4.9) turn out to have quite distinct flavours. We discuss the archimedean local factors in Section 4.2.1, the non-archimedean local factors in 4.2.2, and return to global zeta functions of arithmetic groups in Section 4.2.3.

4.2.1 Witten zeta functions

In this section let $\Gamma = \mathbf{G}(\mathbb{C})$. For $n \in \mathbb{N}$ we denote by $r_n(\Gamma)$ the number of n-dimensional rational, irreducible complex representations of Γ. Let Φ be the root

system of $\mathbf{G}$ of rank $r = \mathrm{rk}(\Phi)$, let Φ^+ a choice of positive roots of Φ and set $\rho = \sum_{\alpha\in\Phi^+} \alpha$. We write $w_1, \ldots, w_r$ for the fundamental weights. The rational irreducible representations of Γ are all of the form W_λ, where $\lambda = \sum_{i=1}^r a_i w_i$ for $a_i \in \mathbb{N}_0$. The Weyl dimension formula asserts that

$$\dim W_\lambda = \prod_{\alpha\in\Phi^+} \frac{\langle \lambda+\rho, \alpha\rangle}{\langle \rho, \alpha\rangle}.$$

Note that the numerator is a product of $\kappa := |\Phi^+|$ affine linear functions $f_1, \ldots, f_\kappa$ in the integer coordinates of λ, whilst the denominator $C = \prod_{\alpha\in\Phi^+}\langle\rho,\alpha\rangle$ is a constant depending only on Φ. Thus

$$\zeta^{\mathrm{irr}}_\Gamma(s) = \sum_\lambda (\dim W_\lambda)^{-s} = C^s \sum_{a\in\mathbb{N}_0^r} \prod_{i=1}^{\kappa} f_i(a)^{-s}. \tag{4.10}$$

Example 4.9 Assume that $\mathbf{G}$ is of type G_2. Then $C = 120$, $r = 6$ and we may take

$$f_1 = f_2 = X_1 + 1, \qquad f_3 = X_1 + X_2 + 2, \qquad f_4 = X_1 + 2X_2 + 3,$$
$$f_5 = X_1 + 3X_2 + 4, \qquad f_6 = 2X_1 + 3X_2 + 5.$$

Theorem 4.10 ([33, Theorem 5.1]) *The abscissa of convergence of $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathbb{C})}(s)$ is r/κ.*

Multivariable generalisations of zeta functions like (4.10) have been considered by Matsumoto ([40]), among others. Functions of the form

$$\zeta(s_1, \ldots, s_r; \mathbf{G}) = \sum_{a\in\mathbb{N}_0^r} \prod_{i=1}^{r} f_i(a)^{-s_i},$$

where $s_1, \ldots, s_r$ are complex variables, are, in particular, known to have meromorphic continuation to the whole complex plane; cf. [40, Theorem 3].

Special values of Witten zeta functions are interpretable as volumes of moduli spaces of certain vector bundles; cf. [52, Section 7] and [51]. From (4.10), Zagier deduces:

Theorem 4.11 ([52]) *If $s \in 2\mathbb{N}$, then $\zeta^{\mathrm{irr}}_\Gamma(s) \in \mathbb{Q}\pi^{\kappa s}$.*

4.2.2 Representation zeta functions of compact p-adic analytic groups

Let Γ be a profinite group. For $n \in \mathbb{N}$ we denote by $r_n(\Gamma)$ the number ofcontinuous finite-dimensional irreducible complex representations of Γ. If Γ is finitely generated, then $r_n(\Gamma)$ is finite for all $n \in \mathbb{N}$ if and only if Γ is FAb, i.e., has the property that every open subgroup of Γ has finite abelianization.

Theorem 4.12 ([29, Theorem 1]) *Let p be an odd prime and Γ a FAb compact p-adic analytic group. Then there are natural numbers $n_1, \ldots, n_k$ and rational functions $W_1(Y), \ldots, W_k(Y) \in \mathbb{Q}(Y)$ such that*

$$\zeta^{\mathrm{irr}}_\Gamma(s) = \sum_{i=1}^{k} n_i^{-s} W_i(p^{-s}). \tag{4.11}$$

Example 4.13 Let R be a compact discrete valuation ring whose (finite) residue field $\mathbb{F}_q$ has odd characteristic. The representation zeta function of the group $\mathrm{SL}_2(R)$ was computed in [29, Section 7]:

$$\zeta^{\mathrm{irr}}_{\mathrm{SL}_2(R)}(s) = \zeta^{\mathrm{irr}}_{\mathrm{SL}_2(\mathbb{F}_q)}(s) + \frac{4q\left(\frac{q^2-1}{2}\right)^{-s} + \frac{q^2-1}{2}(q^2-q)^{-s} + \frac{(q-1)^2}{2}(q^2+q)^{-s}}{1-q^{1-s}}, \quad (4.12)$$

where

$$\zeta^{\mathrm{irr}}_{\mathrm{SL}_2(\mathbb{F}_q)}(s) = 1 + q^{-s} + \frac{q-3}{2}(q+1)^{-s} + \frac{q-1}{2}(q-1)^{-s} + 2\left(\frac{q+1}{2}\right)^{-s} + 2\left(\frac{q-1}{2}\right)^{-s}$$

is the representation zeta function of the finite group of Lie type $\mathrm{SL}_2(\mathbb{F}_q)$.

If R is a finite extension of $\mathbb{Z}_p$, the ring of p-adic integers, then (4.12) illustrates (4.11). It is remarkable that the same formula applies in the characteristic p case, that is if $R = \mathbb{F}_q[\![X]\!]$, the ring of formal power series over $\mathbb{F}_q$.

The proof of Theorem 4.12 utilizes the fact that a FAb compact p-adic analytic group Γ is virtually pro-p: it has an open normal subgroup N which one may assume to be uniformly powerful. The Kirillov orbit method for uniformly powerful groups and methods from model theory and the theory of definable p-adic integration may be used to describe the distribution of the representations of N. Clifford theory is then applied to extend the analysis for N to an analysis for Γ. The integers $n_1, \ldots, n_k$ are closely related to the dimensions of the irreducible representations of the finite group Γ/N.

Computing zeta functions of FAb compact p-adic analytic groups – such as the groups $\mathbf{G}(\mathcal{O}_\mathfrak{p})$ in (4.8) – explicitly is in general very difficult. The situation is more tractable for pro-p groups. Theorem 4.12 states that if Γ is a FAb compact p-adic analytic pro-p group, then $\zeta^{\mathrm{irr}}_\Gamma(s)$ is rational in p^{-s}. That this generating function is a power series in p^{-s} is obvious. Indeed, the irreducible *continuous* representations of a pro-p group Γ all have p-power dimensions, as they factorize over finite quotients of Γ, which are all finite p-groups.

Representation zeta functions of pro-p groups for which a version of the Kirillov orbit method is available may be computed in terms of p-adic integrals associated to polynomial mappings; see [4, Part 1] for details. These integrals are of a much simpler type than the general definable integrals used in the proof of Theorem 4.12. In the following we discuss some cases where this approach allows for an explicit computation of representation zeta functions.

We concentrate on groups of the form $\mathbf{G}(\mathfrak{o})$, where $\mathfrak{o}$ is a finite extension of $\mathcal{O}_\mathfrak{p}$ for some $\mathfrak{p} \in (\operatorname{Spec}\mathcal{O}) \setminus S$. Then $\mathfrak{o}$ is a compact discrete valuation ring of characteristic zero, with maximal ideal $\mathfrak{m}$, say, and finite residue field of characteristic p, where $\mathfrak{p} \mid p\mathcal{O}$. For $m \in \mathbb{N}$ we consider the *m-th principal congruence subgroup* $\mathbf{G}^m(\mathfrak{o})$, that is the kernel of the natural surjection

$$\mathbf{G}(\mathfrak{o}) \to \mathbf{G}(\mathfrak{o}/\mathfrak{m}^m).$$

The groups $\mathbf{G}^m(\mathfrak{o})$ are FAb p-adic analytic pro-p groups and, for sufficiently large $m \in \mathbb{N}$, the Kirillov orbit method is applicable. This follows from the fact that the

groups $\mathbf{G}^m(\mathfrak{o})$ are saturable and potent for $m \gg 0$; cf. [4, Proposition 2.3] and [25]. (In fact, if $\mathfrak{o}$ is an unramified extension of $\mathbb{Z}_p$, then $m = 1$ suffices.) One would like to understand the representation zeta functions $\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{o})}(s)$, and their variation with

- the prime ideal $\mathfrak{p} \in (\operatorname{Spec} \mathcal{O}) \setminus S$,
- the ring extension $\mathfrak{o}$, and
- the congruence level $m \in \mathbb{N}$.

The following result achieves much of this for the special linear groups $\mathrm{SL}_3(\mathfrak{o})$ and the special unitary groups $\mathrm{SU}_3(\mathfrak{o})$, assuming that $p \neq 3$. Here, the special unitary groups $\mathrm{SU}_3(\mathfrak{o})$ are defined in terms of the nontrivial Galois automorphism of the unramified quadratic extension of the field of fractions of $\mathfrak{o}$; see [4, Section 6] for details.

Theorem 4.14 ([4, Theorem E]) *Let $\mathfrak{o}$ be a compact discrete valuation ring of characteristic 0 whose residue field has cardinality q and characteristic different from 3. Let $\mathsf{G}(\mathfrak{o})$ be either $\mathrm{SL}_3(\mathfrak{o})$ or $\mathrm{SU}_3(\mathfrak{o})$. Then, for all sufficiently large $m \in \mathbb{N}$,*

$$\zeta^{\mathrm{irr}}_{\mathsf{G}^m(\mathfrak{o})}(s) = q^{8m} \frac{1 + u(q)q^{-3-2s} + u(q^{-1})q^{-2-3s} + q^{-5-5s}}{(1 - q^{1-2s})(1 - q^{2-3s})}, \tag{4.13}$$

where

$$u(X) = \begin{cases} X^3 + X^2 - X - 1 - X^{-1} & \text{if } \mathsf{G}(\mathfrak{o}) = \mathrm{SL}_3(\mathfrak{o}), \\ -X^3 + X^2 - X + 1 - X^{-1} & \text{if } \mathsf{G}(\mathfrak{o}) = \mathrm{SU}_3(\mathfrak{o}). \end{cases}$$

Furthermore, the following functional equation holds:

$$\zeta^{\mathrm{irr}}_{\mathsf{G}^m(\mathfrak{o})}(s)\Big|_{q \to q^{-1}} = q^{8(1-2m)} \zeta^{\mathrm{irr}}_{\mathsf{G}^m(\mathfrak{o})}(s).$$

Remark 4.15 We note that $\zeta^{\mathrm{irr}}_{\mathsf{G}^m(\mathfrak{o})}(s)$ is a rational function in q^{-s} whose coefficients are given by polynomials in q, that 8 is the dimension of the algebraic group SL_3, and that $\zeta^{\mathrm{irr}}_{\mathsf{G}^m(\mathfrak{o})}(s)/q^{8m}$ is independent of the congruence level m. Only a few signs in the numerators reflect the difference between special linear and unitary groups.

In general, one can give formulae for the representation zeta functions of groups of the form $\mathbf{G}^m(\mathfrak{o})$ — valid for all sufficiently large m and virtually independent of m — which are uniform both under variation of $\mathfrak{p}$ and $\mathfrak{o}$, and all but independent of m. More precisely, [4, Theorem A] implies the following result, which in turn generalizes Theorem 4.14.

Theorem 4.16 ([4, Theorem A]) *There exist a finite subset $T \subset (\operatorname{Spec} \mathcal{O}) \setminus S$, an integer $t \in \mathbb{N}$, and a rational function $R(X_1, \ldots, X_t, Y) \in \mathbb{Q}(X_1, \ldots, X_t, Y)$ such that, for every prime ideal $\mathfrak{p} \notin S \cup T$, the following holds.*

There exist algebraic integers $\lambda_1, \ldots, \lambda_t$, depending on $\mathfrak{p}$, such that, for all finite extensions $\mathfrak{O}$ of $\mathfrak{o} = \mathcal{O}_\mathfrak{p}$, and all sufficiently large $m \in \mathbb{N}$,

$$\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s) = q^{fdm} R(\lambda_1^f, \ldots, \lambda_t^f, q^{-fs}), \tag{4.14}$$

where $q = |\mathcal{O} : \mathfrak{p}|$, $|\mathfrak{O} : \mathfrak{P}| = q^f$, and $d = \dim \mathbf{G}$.

Furthermore, the following functional equation holds:

$$\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s)\Big|_{\substack{q \to q^{-1} \\ \lambda_i \to \lambda_i^{-1}}} = q^{fd(1-2m)} \zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s). \tag{4.15}$$

We note the close analogy between this result and Theorem 4.4, which it precedes. Generalizing points made in Remark 4.15, we further note that Theorem 4.16 implies that $\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s)$ is rational in q^{-fs} and $\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s)/q^{fdm}$ is independent of m. In general we do not expect that the coefficients of $\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathfrak{O})}(s)$ are given by polynomials in q^f. In fact, as in Theorem 4.4, the algebraic integers λ_i arise from formulae for the numbers of rational points of certain algebraic varieties over finite fields. One may ask, however, whether these numbers are given by polynomials for interesting classes of pro-p groups arising from classical groups, such as groups of the form $\mathrm{SL}_N^m(\mathfrak{o})$.

Question 4.17 Let $N, m \in \mathbb{N}$ and $\mathfrak{o}$ be a compact discrete valuation ring of characteristic 0 whose residue field has cardinality q and characteristic not dividing N. Does there exist a rational function $W_N(X,Y) \in \mathbb{Q}(X,Y)$ such that, for sufficiently large m,

$$\zeta^{\mathrm{irr}}_{\mathrm{SL}_N^m(\mathfrak{o})}(s) = q^{(N^2-1)m} W_N(q, q^{-s})?$$

The answer is "yes" in case $N = 2$ (cf. [3, Theorem 1.2]) and $N = 3$ (cf. Theorem 4.14).

The striking similarity between the formulae for the representation zeta functions of groups of the form $\mathrm{SL}_3^m(\mathfrak{o})$ and $\mathrm{SU}_3(\mathfrak{o})$ is reminiscent of Ennola duality for the characters of the finite groups $\mathrm{GL}_n(\mathbb{F}_q)$ and $\mathrm{GU}_n(\mathbb{F}_q)$; cf. [30]. I am not aware of such a duality in the realm of compact p-adic analytic groups, but read (4.13) as a strong indication for a connection like this.

Computing the representation zeta functions of the "full" p-adic analytic groups $\mathbf{G}(\mathcal{O}_\mathfrak{p})$ is significantly harder than those of their principal congruence subgroups. In principle, Clifford theory allows one to describe the representations of the former groups in terms of the representations of their open normal subgroups $\mathbf{G}^m(\mathcal{O}_\mathfrak{p})$. However, how to tie in explicit Clifford theory with the theory that leads to results like Theorem 4.16 in a way that is uniform in $\mathfrak{p}$ and $\mathfrak{o}$ is not clear in general.

The paper [6] contains formulae for the representation zeta functions of special linear groups of the form $\mathrm{SL}_3(\mathfrak{o})$ and special unitary groups of the form $\mathrm{SU}_3(\mathfrak{o})$, where $\mathfrak{o}$ is an unramified extension of $\mathbb{Z}_p$ and $p \neq 3$. The resulting formulae of the form (4.11) are significantly more complicated than the formulae (4.13) for the principal congruence subgroups, and are omitted here. We just record the fact that

$$(1 - q^{1-2s})(1 - q^{2-3s})$$

is a common denominator for the rational functions involved, just as in (4.13).

It is of great interest if these formulae also apply in characteristic p, i.e., for groups like $\mathrm{SL}_3(\mathbb{F}_q[\![X]\!])$. In contrast to the hands-on computations in [29], the computations in [6] do rely on the Kirillov orbit method for uniformly powerful subgroups of the relevant p-adic analytic groups, which is only available in characteristic 0.

In [6] we also compute the representation zeta functions of finite quotients of groups of the form

$$\mathrm{SL}_3(\mathfrak{o}), \mathrm{SU}_3(\mathfrak{o}), \mathrm{GL}_3(\mathfrak{o}), \mathrm{GU}_3(\mathfrak{o}),$$
$$\mathrm{SL}_3^m(\mathfrak{o}), \mathrm{SU}_3^m(\mathfrak{o}), \mathrm{GL}_3^m(\mathfrak{o}), \mathrm{GU}_3^m(\mathfrak{o})$$

by principal congruence subgroups, subject to some restrictions on the residue field characteristic p. Some further examples of representation zeta functions of p-adic analytic groups are contained in [3].

We close this section by mentioning a vanishing theorem for representation zeta functions.

Theorem 4.18 ([26]) *Let p be an odd prime and Γ an infinite FAb compact p-adic analytic group. Then $\zeta^{\mathrm{irr}}_{\Gamma}(-2) = 0$.*

The proof of this result uses the fact that, while the series $\zeta^{\mathrm{irr}}_{\Gamma}(s)$ does not converge in the usual topology for $s \in \mathbb{R}_{<0}$, the expressions $\zeta^{\mathrm{irr}}_{\Gamma}(e)$ do converge in the p-adic topology for all negative integers e.

4.2.3 Representation zeta functions of arithmetic lattices

We now return to the global representation zeta function of $\mathbf{G}(\mathcal{O}_S)$.

For the purpose of analyzing $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_S)}(s)$ via the Euler factorization (4.9), uniform formulae for zeta functions of the form $\zeta^{\mathrm{irr}}_{\mathbf{G}^m(\mathcal{O}_\mathfrak{p})}(s)$ — as provided, e.g., by Theorem 4.16 — are of limited value. Indeed, whilst the index of $\mathbf{G}^m(\mathcal{O}_\mathfrak{p})$ in $\mathbf{G}(\mathcal{O}_\mathfrak{p})$ is finite for each $\mathfrak{p}$ and all m, the representation zeta function of every finite index subgroup of $\mathbf{G}(\mathcal{O})$ will share all but finitely many of its non-archimedean factors with those of $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_S)}(s)$.

Essentially only for groups of type A_2 do we know how to use Clifford effectively to deduce explicit uniform formulae for the representation zeta functions $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_\mathfrak{p})}(s)$; cf. [6]. This allows for precise asymptotic results about the representation growth of arithmetic groups of type A_2.

Theorem 4.19 ([6]) *Let $\mathbf{G}$ be a connected, simply-connected absolutely simple algebraic group defined over K of type A_2, and assume that $\Gamma = \mathbf{G}(\mathcal{O}_S)$ has the strong Congruence Subgroup Property. Then $\alpha(\Gamma) = 1$. Moreover, $\zeta^{\mathrm{irr}}_{\Gamma}(s)$ admits meromorphic continuation to $\{s \in \mathbb{C} \mid \Re(s) > 5/6\}$. The continued function is analytic on this half-plane, except for a double pole at $s = 1$. Consequently, there exists a constant $c(\Gamma) \in \mathbb{R}_{>0}$, such that*

$$\sum_{i=1}^{n} r_i(\Gamma) \sim c(\Gamma) \cdot n \log n.$$

We comment briefly on the proof of Theorem 4.19. Let Γ be as in the theorem. It is a key fact that all but finitely many of the Euler factors of $\zeta^{\mathrm{irr}}_{\Gamma}(s)$ are of the form $\mathrm{SL}_3(\mathcal{O}_\mathfrak{p})$ or $\mathrm{SU}_3(\mathcal{O}_\mathfrak{p})$, where $\mathfrak{p}$ is a prime ideal of $\mathcal{O}$. To see that $\alpha(\Gamma) = 1$, it suffices to prove that the abscissa of convergence of the product over these factors is equal to 1. Indeed, [4, Theorem B] implies that the abscissa of convergence of the Euler factorization (4.9) remains unchanged by removing finitely many non-archimedean factors. The archimedean factors' abscissa of convergence is 2/3; cf. Theorem 4.10. To compute the abscissa of convergence of the Euler factorization of the factors of the form $\mathrm{SL}_3(\mathcal{O}_\mathfrak{p})$ or $\mathrm{SU}_3(\mathcal{O}_\mathfrak{p})$, one may either inspect the explicit formulae given in [6], or argue with "approximative Clifford theory" as in [4].

The existence of meromorphic continuation is evident from inspection of the explicit formulae for $\zeta^{\mathrm{irr}}_{\mathrm{SL}_3(\mathcal{O}_\mathfrak{p})}(s)$ and $\zeta^{\mathrm{irr}}_{\mathrm{SU}_3(\mathcal{O}_\mathfrak{p})}(s)$. The key here is that the relevant Euler factorization can be approximated by the following product of translates of (partial) Dedekind zeta functions:

$$\zeta_{K,S}(2s-1)\zeta_{K,S}(3s-2) = \prod_{\mathfrak{p}\in(\operatorname{Spec}\mathcal{O})\setminus S} \frac{1}{(1-|\mathcal{O}:\mathfrak{p}|^{1-2s})(1-|\mathcal{O}:\mathfrak{p}|^{2-3s})}. \tag{4.16}$$

Roughly speaking, dividing $\zeta^{\mathrm{irr}}_{\mathrm{SL}_3(\mathcal{O}_\mathfrak{p})}(s)$ or $\zeta^{\mathrm{irr}}_{\mathrm{SU}_3(\mathcal{O}_\mathfrak{p})}(s)$ by the appropriate local factor of (4.16) clears their common denominator $(1-q^{1-2s})(1-q^{2-3s})$, and the Euler factorization of the remaining numerators converges strictly better than the original Euler factorization.

Theorem 4.19 states, in particular, that the abscissa of convergence of the representation zeta function of an arithmetic group of type A_2 is always equal to 1: the degree of representation growth of very different groups — such as, for example, $\mathrm{SL}_3(\mathcal{O})$ and $\mathrm{SU}_3(\mathcal{O})$, for various number rings $\mathcal{O}$ — only depends on the root system of the underlying algebraic group. This remarkable fact is vastly generalized by the following result.

Theorem 4.20 ([5, Theorem 1.1]) *Let Φ be an irreducible root system. Then there exists a constant $\alpha_\Phi \in \mathbb{Q}$ such that, for every arithmetic group $\mathbf{G}(\mathcal{O}_S)$, where $\mathcal{O}_S$ is the ring of S-integers of a number field K with respect to a finite set of places S and $\mathbf{G}$ is a connected abolutely almost simple algebraic group over K with absolute root system Φ, the following holds: if $\mathbf{G}(\mathcal{O}_S)$ has the CSP, then $\alpha(\mathbf{G}(\mathcal{O}_S)) = \alpha_\Phi$.*

Theorem 4.20 reduces a conjecture of Larsen and Lubotzky on the invariance of representation growth of lattices in higher rank semisimple locally compact groups to a conjecture of Serre on the CSP; see [5, Theorem 1.3]. A key idea of its proof is to approximate the local factors of the representation zeta function $\zeta^{\mathrm{irr}}_{\mathbf{G}(\mathcal{O}_S)}(s)$ uniformly by certain definable integrals, in a way that leaves the abscissa of convergence unchanged. The proof uses deep, nonconstructive techniques from model theory, which hold little promise to yield an explicit description of the function $\Phi \mapsto \alpha_\Phi$. So far, the only explictly known values of this function are $\alpha_{A_1} = 2$ and $\alpha_{\mathsf{A}_2} = 1$. Recent results of Aizenbud and Avni imply that $\alpha_{A_\ell} \leq 22$ for all $\ell \in \mathbb{N}$; [1, Theorem A].

Question 4.21 What is the value of α_Φ in Theorem 4.20, for various root systems Φ?

4.3 Iterated wreath products and branch groups

Let Q be a finite group, acting on a finite set X of cardinality $|X| = d \geq 2$. We define iterated permutational wreath products as follows. Set $W(Q,0) := \{1\}$ and, for $k \in \mathbb{N}$, set $W(Q,k+1) = W(Q,k) \wr_X Q$. Passing to the inverse limit yields the profinite group $W(Q) := \varprojlim_k W(Q,k)$. Recall that, for a profinite group G, we denote by $r_n(G)$ the number of continuous n-dimensional irreducible complex representations of G up to isomorphism and that G is called rigid if $r_n(G) < \infty$ for all $n \in \mathbb{N}$.

Theorem 4.22 ([9]) *$W(Q)$ is rigid if and only if the group Q is perfect, i.e., $G = [G,G]$. In this case, the following hold.*

1. *The abscissa of convergence* $\alpha := \alpha(\zeta^{\mathrm{irr}}_{W(Q)}(s))$ *is positive and finite, i.e.,* $\alpha \in \mathbb{R}_{>0}$.
2. *Locally around* α*, the function* $\zeta^{\mathrm{irr}}_{W(Q)}(s)$ *allows for a Puiseux expansion of the form*
$$\sum_{n=0}^{\infty} c_n (s-\alpha)^{n/e}$$
for suitable $c_n \in \mathbb{C}$*,* $n \in \mathbb{N}$*, and* $e \in \{2, 3, \ldots, d\}$.
3. *Let* $p_1, \ldots, p_\ell$ *denote the primes dividing* $|Q|$*. There exists a nontrivial polynomial* $\Psi \in \mathbb{Q}[X_1, \ldots, X_d, Y_1, \ldots, Y_\ell]$ *such that*
$$\Psi(\zeta^{\mathrm{irr}}_{W(Q)}(s), \zeta^{\mathrm{irr}}_{W(Q)}(2s), \ldots, \zeta^{\mathrm{irr}}_{W(Q)}(ds), p_1^{-s}, \ldots, p_\ell^{-s}) = 0. \tag{4.17}$$

For examples illustrating in particular the functional equations (4.17), see [9]. For generalizations of these results to self-similar profinite branched groups, see [8].

Acknowledgements Over several years, my work has been generously supported by numerous funding bodies, including the EPSRC, the DFG, and the Nuffield Foundation. I am also grateful to the organizers of Groups St Andrews 2013.

References

[1] A. Aizenbud & N. Avni, Representation growth and rational singularities of the moduli space of local systems, arXiv:1307.0371, 2014.
[2] N. Avni, Arithmetic groups have rational representation growth, *Ann. of Math. (2)* **174** (2011), 1009–1056.
[3] N. Avni, B. Klopsch, U. Onn & C. Voll, Representation zeta functions of some compact p-adic analytic groups, *Zeta Functions in Algebra and Geometry*, 295–330, Contemp. Math. 566, Amer. Math. Soc., 2012.
[4] ———, Representation zeta functions of compact p-adic analytic groups and arithmetic groups, *Duke Math. J.* **162** (2013), 111–197.
[5] ———, Arithmetic groups, base change, and representation growth, arXiv:1110.6092, 2014.
[6] ———, Similarity classes of integral p-adic matrices and representation zeta functions of groups of type A_2, preprint, 2014.
[7] E. Avraham, Representation zeta functions of finitely generated, torsion free, nilpotent groups, Master's thesis, Ben-Gurion University of the Negev, Israel, 2012.
[8] L. Bartholdi, Representation zeta functions of self-similar branched groups, arXiv:1303.1805, 2013.
[9] L. Bartholdi & P. de la Harpe, Representation zeta functions of wreath products with finite groups, *Groups Geom. Dyn.* **4** (2010), 209–249.
[10] M.N. Berman, Uniformity and functional equations for local zeta functions of $\mathfrak{K}$-split algebraic groups, *Amer. J. Math.* **133** (2011), 1–27.
[11] M.N. Berman & B. Klopsch, A nilpotent group without local functional equations for pro-isomorphic subgroups, arXiv:1408.6669, 2014.
[12] B. Datskovsky & D.J. Wright, Density of discriminants of cubic extensions, *J. Reine Angew. Math.* **386** (1988), 116–138.
[13] J. Denef & F. Loeser, Caractéristiques d'Euler-Poincaré, fonctions zêta locales et modifications analytiques, *J. Amer. Math. Soc.* **5** (1992), 705–720.
[14] ———, Motivic Igusa zeta functions, *J. Algebraic Geom.* **7** (1998), 505–537.
[15] M.P.F. du Sautoy, Counting subgroups in nilpotent groups and points on elliptic curves, *J. Reine Angew. Math.* **549** (2002), 1–21.

[16] ———, Zeta functions of groups: The quest for order versus the flight from ennui, *Groups St Andrews 2001 in Oxford*, 150–189, CUP, 2003.
[17] M.P.F. du Sautoy & F.J. Grunewald, Analytic properties of zeta functions and subgroup growth, *Ann. of Math. (2)* **152** (2000), 793–833.
[18] ———, Zeta functions of groups and rings, *International Congress of Mathematicians, Vol. II*, 131–149, Eur. Math. Soc., 2006.
[19] M.P.F. du Sautoy & F. Loeser, Motivic zeta functions of infinite-dimensional Lie algebras, *Selecta Math. (N.S.)* **10** (2004), 253–303.
[20] M.P.F. du Sautoy & A. Lubotzky, Functional equations and uniformity for local zeta functions of nilpotent groups, *Amer. J. Math.* **118** (1996), 39–90.
[21] M.P.F. du Sautoy & L. Woodward, *Zeta Functions of Groups and Rings*, Lecture Notes Math. 1925, Springer-Verlag, 2008.
[22] A. Evseev, Reduced zeta functions of Lie algebras, *J. Reine Angew. Math.* **633** (2009), 197–211.
[23] S. Ezzat, Representation growth of finitely generated torsion-free nilpotent groups: Methods and examples, Ph.D. thesis, University of Canterbury, New Zealand, 2012.
[24] S. Ezzat, Counting irreducible representations of the Heisenberg group over the integers of a quadratic number field, *J. Algebra* **397** (2014), 609–624.
[25] J. González-Sánchez, Kirillov's orbit method for p-groups and pro-p groups, *Comm. Algebra* **37** (2009), 4476–4488.
[26] J. González-Sánchez, A. Jaikin-Zapirain & B. Klopsch, The representation zeta function of a FAb compact p-adic Lie group vanishes at -2, *Bull. London Math. Soc.* **46** (2014), 239–244.
[27] F.J. Grunewald, D. Segal & G.C. Smith, Subgroups of finite index in nilpotent groups, *Invent. Math.* **93** (1988), 185–223.
[28] E. Hrushovski, B. Martin, S. Rideau & R. Cluckers, Definable equivalence relations and zeta functions of groups, arXiv:math/0701011, 2014.
[29] A. Jaikin-Zapirain, Zeta function of representations of compact p-adic analytic groups, *J. Amer. Math. Soc.* **19** (2006), 91–118.
[30] N. Kawanaka, Generalized Gel′fand-Graev representations and Ennola duality, *Algebraic groups and related topics (Kyoto/Nagoya, 1983)*, 175–206, Adv. Stud. Pure Math. 6, North-Holland, 1985.
[31] B. Klopsch, Representation growth and representation zeta functions of groups, *Note Mat.* **33** (2013), 107–120.
[32] B. Klopsch, N. Nikolov & C. Voll, *Lectures on Profinite Topics in Group Theory*, London Math. Soc. Student Texts 77, CUP, 2011.
[33] M. Larsen & A. Lubotzky, Representation growth of linear groups, *J. Eur. Math. Soc.* **10** (2008), 351–390.
[34] M. Liebeck & A. Shalev, Character degrees and random walks in finite groups of Lie type, *Proc. London Math. Soc. (3)* **90** (2005), 61–86.
[35] R.I. Liu, Counting subrings of $\mathbb{Z}^n$ of index k, *J. Combin. Theory Ser. A* **114** (2007), 278–299.
[36] A. Lubotzky & A.R. Magid, Varieties of representations of finitely generated groups, *Mem. Amer. Math. Soc.* **58** (1985), no. 336.
[37] A. Lubotzky, A. Mann & D. Segal, Finitely generated groups of polynomial subgroup growth, *Israel J. Math.* **82** (1993), 363–371.
[38] A. Lubotzky & B. Martin, Polynomial representation growth and the congruence subgroup growth, *Israel J. Math.* **144** (2004), 293–316.
[39] A. Lubotzky & D. Segal, *Subgroup Growth*, Birkhäuser Verlag, 2003.
[40] K. Matsumoto, On Mordell-Tornheim and other multiple zeta-functions, *Proc. of the Session in Analytic Number Theory and Diophantine Equations*, Univ. Bonn, 2003.
[41] J. Nakagawa, Orders of a quartic field, *Mem. Amer. Math. Soc.* **122** (1996), no. 583.

[42] C. Nunley & A.R. Magid, Simple representations of the integral Heisenberg group, *Classical groups and related topics (Beijing, 1987)*, 89–96, Contemp. Math. 82, Amer. Math. Soc., 1989.
[43] T. Rossmann, Computing topological zeta functions of groups, algebras, and modules, I, arXiv:1405.5711, 2014.
[44] M.M. Schein & C. Voll, Normal zeta functions of the Heisenberg groups over number rings I – the unramified case, arXiv:1401.0173, 2013.
[45] ———, Normal zeta functions of the Heisenberg groups over number rings II – the non-split case, *Israel J. Math.*, to appear.
[46] R. Snocken, Zeta functions of groups and rings, Ph.D. thesis, Univ. of Southampton, 2013.
[47] R.P. Stanley, f-vectors and h-vectors of simplicial posets, *J. Pure Appl. Algebra* **71** (1991), 319–331.
[48] A. Stasinski & C. Voll, Representation zeta functions of nilpotent groups and generating functions for Weyl groups of type B, *Amer. J. Math.* **136** (2014), 501–550.
[49] C. Voll, Functional equations for local normal zeta functions of nilpotent groups, with an appendix by A. Beauville, *Geom. Func. Anal.* **15** (2005), 274–295,
[50] ———, Functional equations for zeta functions of groups and rings, *Ann. of Math. (2)* **172** (2010), 1181–1218.
[51] E. Witten, On quantum gauge theories in two dimensions, *Comm. Math. Phys.* **141** (1991), 153–209.
[52] D. Zagier, Values of zeta functions and their applications, *First European Congress of Mathematics, Vol. II (Paris, 1992)*, 497–512, Progr. Math. 120, Birkhäuser, 1994.