Spectral Properties of Finite Groups

Henry Bradford
St John's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2015

To my family and friends,
with love.

Abstract

This thesis concerns the diameter and spectral gap of finite groups. Our focus shall be on the asymptotic behaviour of these quantities for sequences of finite groups arising as quotients of a fixed infinite group.

In Chapter 3 we give new upper bounds for the diameters of finite groups which do not depend on a choice of generating set. Our method exploits the commutator structure of certain profinite groups, in a fashion analogous to the Solovay-Kitaev procedure from quantum computation. We obtain polylogarithmic upper bounds for the diameters of finite quotients of: groups with an analytic structure over a pro-p domain (with exponent depending on the dimension); Chevalley groups over a pro-p domain (with exponent independent of the dimension) and the Nottingham group of a finite field. We also discuss some consequences of our results for random walks on groups.

In Chapter 4 we construct new examples of expander Cayley graphs of finite groups, arising as congruence quotients of non-elementary subgroups of $\mathrm{SL}_2(\mathbb{F}_p[t])$ modulo certain square-free ideals. We describe some applications of our results to simple random walks on such subgroups, specifically giving bounds on the rate of escape from algebraic subvarieties, the set of squares and the set of elements with reducible characteristic polynomial in $\mathrm{SL}_2(\mathbb{F}_p[t])$.

Finally, in Chapter 5 we produce new expander congruence quotients of $\mathrm{SL}_2(\mathbb{Z}_p)$, generalising work of Bourgain and Gamburd [6]. The proof combines the Solovay-Kitaev procedure with a quantitative analysis of the algebraic geometry of these groups, which in turn relies on previously known examples of expanders.

Acknowledgements

First and foremost, it is my pleasure to acknowledge the fantastic support I have received throughout my DPhil from my supervisor, Marc Lackenby. Undertaking a research degree can be a stormy voyage at times, and Marc has always been a steadying hand at the tiller, and a consistent source of sound advice, encouragement and enthusiasm for my work.

Second, I would like to thank Mum, Dad, Jennifer and Polly, whose love, patience and support throughout my studies has meant the world to me.

Third, I would like to thank my dear friends Elizabeth Bennett; George Bray; Zoe Carroll; Jeremy Evans; Violet Kovacheva; Anna-Katharina Krüger; Michelle Leese; Jasmine Low; Lily Patchett; Corwin and Genni Pierce-Butler; Sandra Rankovic and Andrew Vourdas, for their unfailing care and affection.

Finally I am grateful to EPSRC, for providing the financial support which enabled me to conduct this research.

Statement of Originality

This thesis is entirely my own work. All results due to mathematicians other than myself, which are quoted in the text of this thesis, are acknowledged as such.

Henry Bradford
Oxford, 2015

Contents

Chapter 1

Introduction

In so far as this thesis has a single unifying theme, it is this: given a finite group G and a generating set $S \subset G$, what does it mean to say that S generates G "quickly" or "efficiently", and for which pairs (G, S) is such "efficient" generation achieved? Of course, how we answer the first question will determine the kind of answers we can expect to obtain to the second, and the route by which we shall arrive at them. Here we shall focus on two possible interpretations: the *diameter* and *spectral gap* of (G, S).

1.1 What Is Diameter?

The *diameter* of the pair (G, S) simply measures the length of words in the generators S needed to express elements of G. To be precise, we define:

$$\mathrm{diam}(G, S) = \min\{n \in \mathbb{N} : B_S(n) = G\}$$

where $B_S(n)$ is the (closed) ball of radius n in the word metric induced on G by S. That is:

$$B_S(n) = \{s_1 \cdots s_n : s_1, \ldots, s_n \in S \cup S^{-1} \cup \{1\}\}.$$

Alternatively, and for ease of visualisation, $\mathrm{diam}(G, S)$ is precisely the diameter of the *Cayley graph* $\mathrm{Cay}(G, S)$, equipped with the standard path metric.

It is intuitive to regard S as generating G "efficiently" if we can give a suitably small upper bound for $\mathrm{diam}(G, S)$. Such bounds are usually expressed as a function of the order $|G|$ of G.

For a given group G, $\mathrm{diam}(G, S)$ shall depend greatly on the choice of a generating set S: if $|S|$ is fixed then a simple counting argument shows that $\mathrm{diam}(G, S)$ is

bounded below by a logarithmic function of $|G|$. On the other hand, for any finite group G, $\mathrm{diam}(G, G) = 1$.

This example should serve as a warning: we cannot expect the quantity $\mathrm{diam}(G, S)$ to contain any interesting information about G if we permit ourselves to choose the generating set S arbitrarily. To avoid such arbitrary choices, among other reasons, it is desirable to define a notion of diameter for G which is an invariant of the group alone.

There are several reasonable ways to do this: one may for example fix the size of S and study the minimal possible value of $\mathrm{diam}(G, S)$, as S varies. One may also ask about the value of $\mathrm{diam}(G, S)$ for generic subsets S, if it indeed makes sense to talk about the generic behaviour of $\mathrm{diam}(G, S)$. The notion of diameter studied in this work, however, shall be the so-called *worst-case diameter* of G, defined by:

$$\mathrm{diam}(G) = \max\{\mathrm{diam}(G, S) : S \subseteq G, \langle S \rangle = G\}.$$

This definition has the obvious advantage that an upper bound for $\mathrm{diam}(G)$ contains some information about any particular generating set S with which we may happen to be working; it is in this sense always able to be turned to our needs. The downside is that the upper bound may not be optimal for our chosen generating set (or for any generating set at all, for that matter).

For which groups might one study $\mathrm{diam}(G)$? One class which has attracted substantial attention is the finite simple groups. The key conjecture in this area is the following, due to Babai.

Conjecture 1.1.1 ([1], 1992)**.** *There exists an absolute constant $C > 0$ such that if G is a non-abelian finite simple group, then:*

$$\mathrm{diam}(G) \leq C(\log|G|)^C.$$

This conjecture is still a long way from being proved, though substantial progress has been made in recent years. The first major case to fall was that of $\mathrm{PSL}_2(p)$ (p prime), which was dealt with by Helfgott. Indeed, his proof yielded a stronger result:

Theorem 1.1.2 ([36], 2005)**.** *For all $\delta > 0$, there exist $C(\delta), D(\delta), \epsilon(\delta) > 0$ such that, for any prime p and any generating set $S \subseteq \mathrm{PSL}_2(p)$, the following holds.*

(i) If $|S| < |\mathrm{PSL}_2(p)|^{1-\delta}$ then $|S \cdot S \cdot S| > C|S|^{1+\epsilon}$.

(ii) If $|S| > |\mathrm{PSL}_2(p)|^{\delta}$, then $\mathrm{diam}(\mathrm{PSL}_2(p), S) \leq D$. (In fact, Nikolov and Pyber [63], building on work of Gowers [34], observed that if $\delta > 8/9$, we may take $D = 3$.)

Why does this imply the desired polylogarithmic diameter bound? First, replacing S by $B_S(l_0)$ for l_0 an absolute constant, we may assume $|S|$ is bigger than any absolute constant. Second, we apply (i) iteratively to obtain $l \in \mathbb{N}$ such that $|B_S(l)| > |\mathrm{PSL}_2(p)|^{\delta_0}$ for $\delta_0 > 0$ an absolute constant (here l may be taken to be at most polylogarithmic in $|\mathrm{PSL}_2(p)|$). Finally, we apply (ii) to obtain $B_S(Dl) = \mathrm{PSL}_2(p)$.

Following the initial breakthrough of Helfgott's paper there was a flurry of results generalising his methods to other finite simple groups of Lie type, and versions of Theorem 1.1.2 were quickly provided for $\mathrm{PSL}_3(p)$ (by Helfgott in 2008 [37]) and for $\mathrm{PSL}_2(q)$ for q an arbitrary prime power (by Dinai in 2009 [27] and Varjú in 2010 [77]). The current state of the art in this direction is contained in independent works of Breuillard, Green and Tao [18] and Pyber and Szabó [66] from 2010. They showed that the analogue of Helfgott's result holds for any family of finite simple groups of Lie type of bounded Lie rank. Consequently, a weaker version of Babai's bound is now known to hold, in which the constant C depends on the Lie rank of G. This issue of dependence on Lie rank in diameter bounds is one to which we shall return in discussing our new results.

Meanwhile, the best currently known bounds for the diameters of the alternating groups A_n appeared in [39] in 2011, where the upper bound:

$$\mathrm{diam}(A_n) \leq \exp(C(\log n)^4 \log\log n)$$

was given (with C an absolute constant). This falls short of the polylogarithmic upper bound which is demanded by Conjecture 1.1.1 (and which has been expected for much longer: the particular case of Babai's conjecture for A_n is regarded as folkloric). It should however be noted that the much better bound $Cn^2(\log n)^C$ has since been proved for *random* generators [40].

It bears mention at this juncture that the aforementioned results on groups of Lie type did not develop in isolation. A key motivation driving the rapid development of this subject has been the remarkable applications such results on product growth have to expansion phenomena, first discovered by Bourgain and Gamburd [5]. We shall return to discuss these applications shortly.

1.2 The Solovay-Kitaev Procedure

Simple groups did not provide the *first* examples of families of groups with a poly-logarithmic upper bound for diam. These were due to Gamburd and Shahshahani, who observed that such a bound held for the sequence $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ (with p fixed and n varying).

Theorem 1.2.1 ([33], 2004)**.** *Let p be an odd prime. Then there exist $C_1(p) > 0$ and an absolute constant $C_2 > 0$ such that, for all $n \in \mathbb{N}$,*

$$\mathrm{diam}(\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})) \leq C_1(\log|SL_2(\mathbb{Z}/p^n\mathbb{Z})|)^{C_2}.$$

Even better, it is possible to give an explicit numerical estimate for C_2, of approximately 8.71.

When n is large, $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ is very much unlike a finite simple group such as $\mathrm{PSL}_2(\mathbb{F}_q)$: $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ has many large normal subgroups, arising as the kernels of congruence maps $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/p^m\mathbb{Z})$ for $m \leq n$. The presence of such *congruence kernels* is both a blessing and a curse. On the one hand, a clean statement about the growth of arbitrary subsets à la Theorem 1.1.2 becomes less accessible (there are in some sense too many subgroups in which a generating set may become partially trapped). On the other, the filtration by the congruence kernels opens the way to arguments by induction on the level of the filtration.

The proof of Gamburd and Shahshahani's result takes this form. The inspiration for their particular proof came from the world of quantum computation, and specifically from the *Solovay-Kitaev procedure* used in the construction of efficient quantum compilers. In group-theoretic terms, this is an algorithm which, given a finite symmetric subset S of $\mathrm{SU}(d)$ generating a dense subgroup, and an arbitrary element $g \in \mathrm{SU}(d)$, writes a *short* word in S approximating g up to small error. We expound further on the quantum compilation problem in Section 3.1.1.

The procedure relies only on a few key properties of $\mathrm{SU}(d)$, concerning commutator words. Gamburd and Shahshahani's insight was that these properties translate to the setting of the group $\mathrm{SL}_2(\mathbb{Z}_p)$, where we interpret "approximation up to small error" as agreement up to some congruence kernel $K_n = \mathrm{SL}_2(\mathbb{Z}_p) \cap (I_2 + p^n\mathbb{M}_2(\mathbb{Z}_p))$, for n sufficiently large. In this context, writing every element of $\mathrm{SL}_2(\mathbb{Z}_p)$ as a short word up to an error in K_n is equivalent to giving a small upper bound for $\mathrm{diam}(\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z}))$.

As hinted already, an additional virtue of diameter bounds proved via the Solovay-Kitaev method is that the implied constants are reasonably small and may be explicitly computed. The constants appearing in Theorem 1.2.1 were improved by Dinai

[28], using the theory of p-adic analytic groups. He then extended Theorem 1.2.1 to other Chevalley groups over $\mathbb{Z}_p$ [29].

Our contribution to this field is the development of an abstract group-theoretic formulation of the Solovay-Kitaev procedure for profinite groups, which is sufficiently flexible as to be applicable to a much broader range of examples than has been considered previously [12]. This shall be the subject of Chapter 3.

1.3 What Is Spectral Gap?

Meanwhile the *spectral gap* of the pair (G, S) is a quantitative measure of the *connectivity* of (G, S) and controls the equidistribution of the simple random walk on (G, S). In Chapter 2 we shall give a purely algebraic account of spectral gap, but for now it shall be a little easier to motivate with reference to the Cayley graph $\mathrm{Cay}(G, S)$. In common with any finite graph, $\mathrm{Cay}(G, S)$ has a (normalised) adjacency matrix $A_S \in \mathbb{M}_{|G|}(\mathbb{R})$. If $S \subseteq G$ is symmetric, then A_S is a non-negative symmetric matrix of operator norm 1; its spectrum $\sigma(A_S)$ is a finite subset of $[-1, 1]$. If S generates G, so that $\mathrm{Cay}(G, S)$ is connected, then 1 is a simple eigenvalue of A_S. In this case we define the *spectral gap* or *expansion* of (G, S) to be:

$$\epsilon(G, S) = \min\{1 - |\lambda| : \lambda \in \sigma(A_S) \setminus \{1\}\}.$$

We view S as generating G "efficiently" if $\epsilon(G, S)$ is large. How can we make this intuitive? It may be shown that, taking a simple random walk on $\mathrm{Cay}(G, S)$, the probability $\mu_S^{(l)}(g)$ of reaching a vertex g at time l differs from $1/|G|$ (that is, the weight assigned to g by the uniform distribution on the vertices) by at most $(1 - \epsilon)^l$. If the spectral gap is large, therefore, the sequence of probability distributions $(\mu_S^{(l)})_l$ arising from the random walk on $\mathrm{Cay}(G, S)$ converge quickly to the uniform distribution. In other words, for some reasonably small l, an approximately equal proportion of the (unreduced) words of length l in S are equal in G to each element g.

As with diameter, it is easy to "force" efficient generation by taking S to be very large (for instance taking $S = G$). A problem which has been of great interest to both pure mathematicians and computer scientists for many years is to produce sequences (G_n, S_n) where the G_n are of unbounded size, the S_n are of *bounded* size and the $\epsilon(G_n, S_n)$ remain large. If $\epsilon(G_n, S_n)$ may be bounded away from zero by a constant independent of n, then such a sequence is known as a *(two-sided) expander family* (see [41] and [54] for an overview of the diverse applications of expanders across mathematics).

The first examples of expander families were produced by Margulis [61], exploiting Kazhdan's property (T). We shall not define property (T) precisely, but it is a powerful representation-theoretic rigidity property of groups, which is satisfied by all lattices in higher-rank simple Lie groups over local fields. Margulis' observation was that if the G_n arise as finite quotients of a fixed property (T) group G, and the S_n are the images in these quotients of a fixed finite generating set S for G, then (G_n, S_n) is an expander family (for parity reasons, we assume for now that $1 \in S$, though in many cases this shall not be strictly necessary).

Example 1.3.1 (Margulis)**.** *Let $d \geq 3$ and let $S \subseteq \mathrm{SL}_d(\mathbb{Z})$ be a finite symmetric set, generating a finite index subgroup, and such that $I_d \in S$. Then there exists $q_0(S) \in \mathbb{N}$ such that $(\mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z}), \pi_q(S))_{(q,q_0)=1}$ is an expander family.*

Here $\pi_q : \mathrm{SL}_d(\mathbb{Z}) \twoheadrightarrow \mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z})$ is the congruence map. Meanwhile, expander congruence quotients of $\mathrm{SL}_2(\mathbb{Z})$ arose from Selberg's famous 3/16 Theorem on congruence covers of arithmetic manifolds [72]. Historically Selberg's result preceded Margulis' work on property (T) groups, but the connections with expansion were not noted until later [58].

Example 1.3.2 (Selberg)**.** *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a finite symmetric set, generating a finite index subgroup, and such that $I_2 \in S$. Then there exists $p_0(S) \in \mathbb{N}$ such that $(\mathrm{SL}_2(\mathbb{F}_p), \pi_p(S))_{p \geq p_0 \text{ prime}}$ is an expander family.*

1.4 The Age of the Machine

By the turn of the new millennium, property (T) and Selberg's Theorem had delivered numerous examples of expander families, but the mathematical community's understanding of expansion in groups was still unsatisfactory in a number of ways. The limitations of the available tools were neatly illustrated by Lubotzky in his *1-2-3 Problem* [53]. For $a \in \mathbb{Z}$, define:

$$x_a = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, y_a = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

and let $S_a = \{x_a^{\pm 1}, y_a^{\pm 1}, I_2\}$. For p prime let $\pi_p : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(p)$ be the congruence map. Then for $p \geq 5$, $\pi_p(S_a)$ generates $\mathrm{SL}_2(p)$.

(i) If $a = 1$, then $\langle S_a \rangle = \mathrm{SL}_2(\mathbb{Z})$, so by Selberg's Theorem $(\mathrm{SL}_2(p), \pi_p(S_1))_p$ is an expander family.

(ii) If $a = 2$, then $\langle S_a \rangle \leq \mathrm{SL}_2(\mathbb{Z})$ has finite index, so once again by Selberg's Theorem $(\mathrm{SL}_2(p), \pi_p(S_2))_p$ is an expander family.

(iii) If $a = 3$, then $\langle S_a \rangle \leq \mathrm{SL}_2(\mathbb{Z})$ has *infinite* index, so Selberg's Theorem cannot determine whether or not $(\mathrm{SL}_2(p), \pi_p(S_3))_p$ is an expander family.

This is in spite of the fact that, from the point of view of finite groups, $\pi_p(S_3)$ is almost exactly the same as $\pi_p(S_1)$ or $\pi_p(S_2)$: in each case x_a and y_a are generators for the same pair of unipotent subgroups.

Lubotzky's problem was eventually resolved in the breakthrough paper of Bourgain and Gamburd [5] (first announced in 2006), which has inspired much of the work on constructions of expanders since. The main result of that paper was:

Theorem 1.4.1. *For any $k \in \mathbb{N}_{\geq 2}$ and $\tau > 0$, there exists $\epsilon(k, \tau) > 0$ such that the following holds. Let p be a sufficiently large prime (depending on k and τ), and let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p)$ be of order $2k$. Suppose that:*

$$\mathrm{girth}(\mathrm{SL}_2(\mathbb{F}_p), S) \geq \tau \log p$$

Then the spectral gap of $(\mathrm{SL}_2(\mathbb{F}_p), S)$ is at least ϵ.

Here, for G a group and $S \subseteq G$, the *girth* of the pair (G, S) is the length of the shortest non-trivial reduced word in S which is equal to 1 in G. In other words, $\mathrm{girth}(G, S)$ is the length of the shortest embedded loop in the Cayley graph $\mathrm{Cay}(G, S)$ (or $\mathrm{girth}(G, S) = \infty$ if no such loop exists; in this case G is freely generated by S).

Now, if $S \subseteq \mathrm{SL}_2(\mathbb{Z})$ is a free generating set for a non-abelian free subgroup, it may be seen that the congruence images $\pi_p(S)$ satisfy the girth hypothesis in Theorem 1.4.1. For if w is a non-trivial reduced word in S, and $\pi_p(w) = 1$, then some non-zero coefficient of w is divisible by p. A simple estimate of euclidean norms shows that any such w must be of length $\gg_S \log p$. Using the Tits alternative, we can generalise the conclusion of Theorem 1.4.1 to congruence images of any subset generating a non-elementary subgroup, and thereby obtain:

Theorem 1.4.2. *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a finite symmetric subset. Suppose that $\langle S \rangle$ is Zariski-dense in $\mathrm{SL}_2(\mathbb{R})$. Then there exists $p_0(S) \in \mathbb{N}$ such that:*

$$(\mathrm{SL}_2(\mathbb{F}_p), \pi_p(S))_{p \geq p_0 \text{ prime}}$$

is an expander family.

Recall that in $\mathrm{SL}_2(\mathbb{R})$, a subgroup is Zariski-dense iff it is not virtually soluble. In particular, all the subsets S_a described in the 1-2-3 problem above have this property.

Significant though Theorem 1.4.1 may be in its own right, a greater part of the interest in [5] arose from the fact that the proof of Theorem 1.4.1 was based on a new procedure for constructing expanders, exploiting results from additive combinatorics, and which proved to be applicable to many other families of groups. This procedure, which has come to be known as the *Bourgain-Gamburd machine*, will be described in detail in Section 2.2. Roughly speaking though, the machine tells us that the expansion of a pair (G, S) may be guaranteed by three hypotheses. The first is that G should be highly *quasirandom*, meaning that G has no small-dimensional non-trivial complex representations. There are good classical estimates of quasirandomness for many familiar families of finite groups, including finite simple groups of Lie type. The combinatorics of quasirandom groups was studied by Gowers [34], who coined the term, but its connection with expansion was first noted by Sarnak and Xue [70]. Suppose A_S has an eigenvalue λ of modulus close to 1, so that the expansion is weak. The eigenspace of λ is a subrepresentation of the right-regular representation of G, so by quasirandomness has large dimension. This places a substantial lower bound on $\mathrm{tr}(A_S^{2l})$.

Proving expansion therefore reduces to showing that $\mathrm{tr}(A_S^{2l})$ decays quickly. Sufficient conditions for such decay come from the non-commutative Balog-Szemerédi-Gowers Theorem, due to Tao [75], which tells us that if decay fails, the measure $\mu_S^{(2l)}$ must concentrate somewhat on a small *approximate subgroup* A of G (that is, a symmetric subset containing 1 such that AA is covered by a small number of translates of A). Some papers utilising the Bourgain-Gamburd machine have tackled the problem of excluding this possibility head-on, but we can reduce the problem still further if G satisfies a *product theorem*. All finite groups have some obvious approximate subgroups: if A is already almost the whole of G, or is almost a proper subgroup of G, then A will not grow much under multiplication with itself. A product theorem says roughly that these are the only possibilities. Upon closer inspection, this is precisely what the results on growth in finite simple groups of Helfgott [36]; Breuillard-Green-Tao [18]; Pyber-Szabó [66] and the other authors discussed above (in the context of bounds on diameter) tell us: any subset which is not trapped in terms of a proper subgroup must grow quickly under multiplication with itself, and as such cannot be an approximate subgroup.

Expansion is thereby reduced to showing that $\mu_S^{(2l)}$ escapes quickly from proper subgroups of G. It should be noted that this is the only point at which the set S enters

the argument. We therefore usually expect some information about the geometry of S to be crucially involved in proving non-concentration in subgroups.

The years since Bourgain and Gamburd's initial paper have seen a flood of results which used the machine to construct new families of expanders. The construction of expander congruence quotients of linear algebraic groups has remained a very fruitful direction of enquiry in this context, following on from Theorem 1.4.2. In this vein, we now know that the following are also expander families:

(i) $(\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}), \pi_q(S))$, for $S \subseteq \mathrm{SL}_2(\mathbb{Z})$ generating a Zariski-dense subgroup, q ranging over all square-free positive integers coprime to some $q_0(S) \in \mathbb{N}$ (Bourgain-Gamburd-Sarnak, 2008 [10]);

(ii) $(\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z}), \pi_{p^n}(S))$, for $S \subseteq \mathrm{SL}_2(\mathbb{Z})$ generating a Zariski-dense subgroup, p a fixed sufficiently large prime and n ranging over $\mathbb{N}$ (Bourgain-Gamburd, 2008 [6]);

(iii) $(\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z}), \pi_{p^n}(S))$, for $d \geq 2$, $S \subseteq \mathrm{SL}_d(\mathbb{Z})$ generating a Zariski-dense subgroup, p a fixed sufficiently large prime and n ranging over $\mathbb{N}$; $(\mathrm{SL}_3(\mathbb{F}_p), \pi_p(S))$, for $S \subseteq \mathrm{SL}_3(\mathbb{Z})$ generating a Zariski-dense subgroup, p ranging over all sufficiently large primes (Bourgain-Gamburd, 2008 [7]);

(iv) $(\mathrm{SL}_2(\mathcal{O}_K/I), \pi_I(S))$, for $\mathcal{O}_K$ the ring of integers of some number field K, $S \subseteq \mathrm{SL}_2(\mathcal{O}_K)$ generating a Zariski-dense subgroup, $I \subseteq \mathcal{O}_K$ ranging over square-free ideals prime to some ideal $J(S) \subseteq \mathcal{O}_K$; $(\mathrm{SL}_3(\mathbb{Z}/q\mathbb{Z}), \pi_q(S))$, for $S \subseteq \mathrm{SL}_3(\mathbb{Z})$ generating a Zariski-dense subgroup, q ranging over all square-free positive integers coprime to some $q_0(S) \in \mathbb{N}$ (Varjú, 2010 [77]);

(v) $(\mathrm{SL}_d(\mathbb{F}_p), \pi_p(S))$, for $d \geq 2$, $S \subseteq \mathrm{SL}_d(\mathbb{Z})$ generating a Zariski-dense subgroup, p ranging over all sufficiently large primes (Breuillard-Green-Tao, 2010 [18]);

(vi) $(\mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z}), \pi_q(S))$, for $S \subseteq \mathrm{SL}_d(\mathbb{Z})$ generating a Zariski-dense subgroup, q ranging over all positive integers coprime to some $q_0(S) \in \mathbb{N}$ (Bourgain-Varjú, 2010 [11]);

(vii) $(\langle \pi_q(S) \rangle, \pi_q(S))$, for $d \geq 2$, $q_0 \in \mathbb{N}$, $S \subseteq \mathrm{GL}_d(\mathbb{Z}[1/q_0])$ generating a group whose Zariski-closure in $\mathrm{GL}_d(\mathbb{Q})$ has perfect connected component at identity, as q ranges over all square-free integers coprime to q_0 (Salehi Golsefidy-Varjú, 2011 [69]).

Note that it is implicit in (i)-(vi) above that the congruence images of S generate the specified finite groups. Even this claim is far from obvious; it follows from the work of Matthews-Vaserstein-Weisfeiler [62]; Nori [64] and Weisfeiler [79] on the phenomenon of *strong approximation* in linear algebraic groups. Suppose that $\mathbb{G}$ is a connected, simply connected, semisimple algebraic group defined over $\mathbb{Q}$, and suppose that $S \subseteq \mathbb{G}(\mathbb{Q})$ is a finite subset generating a Zariski dense subgroup. For all but finitely many primes p; there is a corresponding algebraic group $\mathbb{G}_p$ defined over $\mathbb{F}_p$. The *Strong Approximation Theorem* states that for all sufficiently large p, the congruence image $\pi_p(S)$ of S generates $\mathbb{G}_p(\mathbb{F}_p)$. Guided by the results above, we say that $\mathbb{G}$ satisfies *superstrong approximation* if for every such S, $(\mathbb{G}_p(\mathbb{F}_p), \pi_p(S))$ forms a family of expanders, as p ranges over all sufficiently large primes. From this perspective, Theorem 1.4.2 asserts that SL_2 satisfies superstrong approximation. Further, by combining strong approximation with (vii) above, we see that every $\mathbb{G}$ as in the Strong Approximation Theorem also satisfies superstrong approximation.

There is also a version of strong approximation for algebraic groups over fields of positive characteristic, due to Pink [65], but as yet the theory of superstrong approximation in positive characteristic is undeveloped. Our work in Chapter 4 represents a first step towards such a theory.

Although it does not relate directly to our concerns here, it is worth noting that the ideas underpinning the Bourgain-Gamburd machine are also relevant to spectral gap phenomena for group actions on compact real Lie groups. The first such result was also due to Bourgain and Gamburd ([8], 2006), showing that the action on $L^2(\mathrm{SU}(2))$ of a finite rank free subgroup F of $\mathrm{SU}(2)$ satisfying an appropriate *Diophantine* condition has the spectral gap property. The exact definition of the Diophantine condition shall not concern us for now, so suffice it to say that it requires that a non-trivial short reduced word in F does not lie too close to the identity in matrix norm, and that it is satisfied by subgroups consisting of matrices with algebraic entries.

This result on $\mathrm{SU}(2)$ was extended to $\mathrm{SU}(d)$ (Bourgain and Gamburd [9], 2011) and to arbitrary simple real Lie groups (Benoist and de Saxcé [2], 2014), again under appropriate hypotheses on the subgroup acting. We shall return to the theme of Diophantine conditions on groups in our discussion of spectral gap results for compact Lie groups over *non-archimedean* fields in Chapter 5.

1.5 Sieving in Groups

Beyond their interpretation in terms of random walks on finite groups, expanders are also valuable in the analysis of random walks on *infinite* groups, via the *group sieve method*.

Let Γ be an infinite group, generated by a finite subset S, and let $X \subseteq \Gamma$. We would like to show that X is *sparse*, in the sense that the probability that the simple random walk on $\mathrm{Cay}(\Gamma, S)$ lies in X at time l decays rapidly as a function of l. Suppose we have a sequence of finite quotients $\pi_n : \Gamma \twoheadrightarrow G_n$, with $|G_n|$ tending to infinity. A first, easy observation is that for every n, the probability $\mu_S^{(l)}(X)$ of lying in X at time l is bounded above by the probability $\mu_{\pi_n(S)}^{(l)}(\pi_n(X))$ of the random walk on $\mathrm{Cay}(G_n, \pi_n(S))$ lying in $\pi_n(X)$. If l is sufficiently large that the latter random walk has already equidistributed (so that $\mu_{\pi_n(S)}^{(l)}$ is almost uniform), then $\mu_{\pi_n(S)}^{(l)}(\pi_n(X))$ is approximately $|\pi_n(X)|/|G_n|$. The point is that in many examples, it is much easier in practice to estimate the *size* of $\pi_n(X)$ than to compute $\mu_{\pi_n(S)}^{(l)}(\pi_n(X))$ directly. If X really is sparse, then it is reasonable to hope that $|\pi_n(X)|/|G_n|$ will tend rapidly to zero.

This method is most powerful when the spectral gap of the $(G_n, \pi_n(S))$ is large, so that for given l, the bound $\mu_S^{(l)}(X) \ll |\pi_n(X)|/|G_n|$ is applicable for a large corresponding value of n, so decays rapidly as a function of l and not just of n.

Here are some examples of groups Γ and subsets X which have been analysed using the group sieve method in results from the literature. In all these cases, the probability of return to the subset X decays exponentially fast in time.

(i) $\Gamma \leq \mathrm{GL}_d(K)$ is finitely generated, not virtually soluble, for K a field of characteristic zero; $X = \Gamma \cap \mathcal{V}$, for $\mathcal{V}$ an algebraic subvariety in GL_d such that $\dim(\mathcal{R}(\mathcal{V} \cap \mathbb{G})) < \dim(\mathbb{G})$, where $\mathbb{G}$ is the Zariski closure of Γ and $\mathcal{R}$ is its soluble radical (based on Salehi Golsefidy-Varjú [69], see [17]);

(ii) $\Gamma \leq \mathrm{GL}_d(\mathbb{C})$ is finitely generated, not virtually soluble; $X \subseteq \Gamma$ is the set of proper powers in Γ (Lubotzky-Meiri [55]);

(iii) $\Gamma \leq \mathbb{G}(K)$ is finitely generated Zariski dense, for $\mathbb{G}$ a connected semisimple algebraic group defined and split over K a numberfield; $X \subseteq \Gamma$ is the set of matrices without full Galois group (Jouve-Kowalski-Zywina [45], Lubotzky-Rosenzweig [59]);

(iv) $\Gamma = \mathrm{MCG}(\Sigma)$, the mapping class group of the closed orientable surface Σ of genus $g \geq 1$; $X \subseteq \Gamma$ is the set of non-pseudo-Anosov mapping classes (Rivin [67]);

(v) The same result, with $\mathrm{MCG}(\Sigma)$ replaced by $\mathcal{T}(\Sigma)$, the Torelli group of Σ (Lubotzky-Meiri [56]);

(vi) $\Gamma = \mathrm{Aut}(F_n)$, the mapping class group of the free group of rank $n \geq 2$; $X \subseteq \Gamma$ is the set of non-iwip automorphisms (Rivin [67]);

(vii) The same result, with $\mathrm{Aut}(F_n)$ replaced by IA_n, the kernel of the action of $\mathrm{Aut}(F_n)$ on first homology (Lubotzky-Meiri [57]).

We shall use the group sieve to prove new results with a similar flavour to these for subgroups of $\mathrm{SL}_2(\mathbb{F}_p[t])$ in Chapter 4. Sieving results in their turn may be applied in the construction of other families of expanders, having been used for instance in [7] to show escape of the random walk from proper subgroups of $\mathrm{SL}_3(\mathbb{F}_p)$. In a similar vein (i) above shall be an important tool used in our work on expansion in $\mathrm{SL}_2(\mathbb{Z}_p)$ in Chapter 5.

1.6 New Results and Structure of the Thesis

Chapter 2 shall be an exposition on the technical background which underpins the proofs of our new results. The basic definitions of an expander family; the diameter of a finite group and the simple random walk on a finitely generated group are presented in Section 2.1, along with some basic results following from these definitions. This material is basic to all of our new work. Section 2.2 lays out the Bourgain-Gamburd machine, which is essential to the results of Chapters 4 and 5. Section 2.3 develops some of the basic theory of profinite groups and of groups with an analytic structure over a pro-p domain with discrete valuation. Section 2.4 is concerned with the construction of Chevalley groups over arbitrary rings, with a particular emphasis on Chevalley groups over pro-p domains (as these shall be an important source of examples of analytic groups over pro-p domains). Section 2.5 defines and explores the Nottingham group of a finite field. These last three sections are used in Chapter 3. Finally, Section 2.6 investigates some consequences of Bourgain's sum-product theorem for the ring $\mathbb{Z}/p^n\mathbb{Z}$, which shall be important to our work in Chapter 5.

We emphasize that Chapter 2 is purely expository: the contents of each individual section fall well within the knowledge base of the experts in the relevant field; however,

since the symmetric difference of these various knowledge bases is substantial, we hope that assembling all of this material in one place shall aid the comprehensibility of the subsequent chapters, regardless of the reader's mathematical background.

Our new results are presented in Chapters 3, 4 and 5, as follows.

1.6.1 New uniform diameter bounds in pro-p groups

Chapter 3 shall describe our results on diameters of finite groups obtained from the Solovay-Kitaev procedure. These results were first presented in [12].

Our first result concerns compact groups with a compatible structure as an R-analytic manifold, where $(R, \mathcal{M})$ is a pro-p domain with discrete valuation. Every such group has an open subgroup with an especially simple R-analytic structure, called an *R-standard group*. Precise definitions are given in Section 2.3. In a d-dimensional R-analytic group G, an R-standard subgroup may be identified (as a space) with the product space $\mathcal{M}^{(d)}$; the balls $(\mathcal{M}^n)^{(d)}$ around 0 form a filtration by open normal subgroups. The commutator structure in $\mathcal{M}^{(d)}$ is controlled by the Lie algebra $\mathcal{L}_G$. Recall that a Lie algebra $\mathcal{L}$ is called *perfect* if $\mathcal{L}$ is equal to its derived subalgebra $(\mathcal{L}, \mathcal{L})$.

Theorem 3.1.3. *Let $(R, \mathcal{M})$ be a commutative unital discrete valuation pro-p domain. Let G be a d-dimensional R-standard group, $K_n = (\mathcal{M}^n)^{(d)} \lhd_o G$. Suppose $\mathcal{L}_G$ is perfect. Then there exist $C_1(G)$, $C_2(d) > 0$ such that for all $n \in \mathbb{N}$:*

$$\operatorname{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}.$$

In the case $R = \mathbb{Z}_p$, we can say more, exploiting the concept of a *uniform* subgroup and associated additional features of the Lie theory (explained in detail in Section 3.4.1). Every $\mathbb{Z}_p$-standard group is uniform and every compact p-adic analytic group has an open characteristic uniform subgroup. In a $\mathbb{Z}_p$-standard group G, the balls K_n described in Theorem 3.1.3 coincide with the terms of the lower central p-series for G. Recall that a profinite group G is *FAb* if every open subgroup has finite abelianisation.

Theorem 3.1.5. *Let $p \geq 3$. Let G be a d-dimensional compact p-adic analytic group. Let $K_1 \leq G$ be an open characteristic uniform subgroup; $(K_n)_n$ its lower central p-series. If G is FAb then there exist $C_1(G), C_2(d) > 0$ such that for all $n \in \mathbb{N}$:*

$$\operatorname{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}. \tag{1.1}$$

For $G = K_1$ then conversely: if there exist $C_1, C_2 > 0$ such that (1.1) holds for all $n \in \mathbb{N}$ then G is FAb.

One familiar family of R-analytic groups is the class of Chevalley linear algebraic groups over R. Here we have a stronger conclusion than that available in the general setting of Theorem 3.1.3: the degree C_2 in the diameter bound may be taken to be *independent of the dimension.*

Theorem 3.1.6. *Let $(R, \mathcal{M})$ be a commutative unital discrete valuation pro-p domain, with $\mathcal{M}$ generated by $\mathcal{P}$. Let $G \leq \mathrm{GL}_d(R)$ be the adjoint Chevalley group of type $X_l \in \{A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4, G_2\}$ over R (here d is the dimension of the associated Lie algebra). Suppose:*

$$(X_l, p) \notin \{(A_1, 2), (B_l, 2), (C_l, 2), (D_l, 2)\}.$$

Let $K_n = G \cap (I_d + \mathcal{P}^n \mathbb{M}_d(R))$. Then there exist $C_1(G) > 0$ and an absolute constant $C_2 > 0$ such that for all $n \in \mathbb{N}$:

$$\mathrm{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}.$$

Moreover, the same bound holds for $G = \mathrm{SL}_d(R), \mathrm{SO}_d(R)$ or $\mathrm{Sp}_d(R)$ provided $p \geq 3$, and for $G = \mathrm{SL}_d(R)$ with $p = 2$ provided $d \geq 3$.

Recall the correspondence between the classical root system of type X_l and the associated adjoint Chevalley group G: if $X_l = A_l$ then $G = \mathrm{PSL}_d(R)$, and in particular if $X_l = A_1$ then $G = \mathrm{PSL}_2(R)$; if $X_l = B_l$ or D_l then $G = \mathrm{PSO}_d(R)$ (with the dichotomy between B_l and D_l corresponding to the parity of d); if $X_l = C_l$ then $G = \mathrm{PSp}_d(R)$. In the case $R = \mathbb{Z}_p$, this result was proved by Dinai [29], under the additional hypothesis $p > \max\{\frac{l+2}{2}, 19\}$.

Note the contrast between the diameter bounds obtained in Theorem 3.1.6 and the currently available results for the Chevalley groups over finite fields coming from [18] and [66]: there the degree of the polylogarithm in the upper bound depended on the Lie rank of the group, whereas in our work the degree C_2 is an absolute constant. On the one hand this resilience in the face of increasing rank indicates the power of the Solovay-Kitaev procedure; on the other hand it may be seen as an additional piece of evidence in favour of Babai's Conjecture 1.1.1.

Finally we consider a class of non-linear examples. Recall that, for R a commutative unital ring, the *Nottingham group* $\mathcal{N}(R)$ of R is the set of formal power series over R of the form:

$$f(t) = t + \sum_{k=2}^{\infty} \lambda_k t^k$$

for some $\lambda_k \in R$, with group operation the formal composition of power series. We take $R = \mathbb{F}_q$ a finite field (for q a power of the prime p) and write $\mathcal{N}_q$ for $\mathcal{N}(\mathbb{F}_q)$. We shall be concerned with the filtration by *congruence* subgroups:

$$K_n = \left\{ t + \sum_{k=n+1}^{\infty} \lambda_k t^k \in \mathcal{N}_q \right\} \lhd_o \mathcal{N}_q.$$

Theorem 3.1.7. *Suppose $p \geq 3$. Then there exist $C_1(q) > 0$ and an absolute constant $C_2 > 0$ such that for all $n \in \mathbb{N}$:*

$$\mathrm{diam}(\mathcal{N}_q / K_n) \leq C_1 (\log |\mathcal{N}_q / K_n|)^{C_2}.$$

As an application of these results, we make some observations about mixing times of random walks in the finite groups we study. The direct relationship between diameter and spectral gap was recently exploited by Varjú, who in [78] used a representation-theoretic argument to produce uniform weak spectral gap estimates for $SL_d(\mathbb{Z}/p^n\mathbb{Z})$, and deduced polylogarithmic diameter bounds. Here we reverse the direction of the argument, and deduce weak spectral gap estimates from uniform diameter bounds.

For Γ a countable group and $S \subseteq \Gamma$ a finite symmetric set, let $X_1, X_2, \ldots$ be a sequence of independent random variables, each with law:

$$\frac{1}{|S|} \chi_S \in \ell^2(\Gamma).$$

For $l \in \mathbb{N}$, the *simple random walk on* (Γ, S) *at time* l is the random variable $Y_l = X_1 \cdots X_l$.

For $(R, \mathcal{M})$ a discrete valuation pro-p domain; G a d-dimensional R-standard group; $x_1, \ldots, x_d$ an R-basis for $\mathcal{M}^{(d)}$ (a set of so-called "co-ordinates of the first kind") and $S \subseteq G$ a finite symmetric subset, we may express the simple random walk on $(\langle S \rangle, S)$ by:

$$Y_l = L_1^{(l)} x_1 + \cdots + L_d^{(l)} x_d$$

for some random variables $L_1^{(l)}, \ldots, L_d^{(l)}$ supported on R.

Corollary 3.1.8. *Suppose $\mathcal{L}_G$ is perfect and $S \subseteq G$ generates a dense subgroup. Then there exists $C(d) > 0$, such that for any $C' > 0$ there exists $C''(G, |S|, C') > 0$ and $C'''(d, |R/\mathcal{M}|, C') > 0$ such that, for any $(\lambda_1, \ldots, \lambda_d) \in R^{(d)}$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}\left[\|L_1^{(l)} - \lambda_1\|, \ldots, \|L_d^{(l)} - \lambda_d\| \leq c^{N+1} \right] - \frac{1}{|R/\mathcal{M}|^{dN}} \right| \leq e^{-C''' N^{C'}}$$

whenever $l \geq C'' N^{C+C'}$.

In other words, for such l the probability that Y_l is close to any element of $\mathcal{M}^{(d)}$ is nearly constant, with error at most $e^{-C'''N^{C'}}$. Here $c \in (0,1)$ is the norm of a generator $\mathcal{P}$ for the maximal ideal $\mathcal{M}$ of R.

For a d-dimensional uniform pro-p group G, an alternative representation for elements is available: for any $g \in G$ and any minimal (ordered) generating set $a_1, \dots, a_d$ for G, there exist $\mu_1, \dots, \mu_d \in \mathbb{Z}_p$ such that $g = a_1^{\mu_1} \cdots a_d^{\mu_d}$ (so-called "co-ordinates of the second kind"). We therefore have:

$$Y_l = a_1^{M_1^{(l)}} \cdots a_d^{M_d^{(l)}}$$

for some random variables $M_1^{(l)}, \dots, M_d^{(l)}$ supported on $\mathbb{Z}_p$.

Corollary 3.1.9. *Let $p \geq 3$. Suppose G is uniform and FAb and $S \subseteq G$ generates a dense subgroup. Then there exists $C(d) > 0$, such that for any $C' > 0$ there exists $C''(G, |S|, C') > 0$ and $C'''(d, p, C') > 0$ such that, for any $\mu_1, \dots, \mu_d \in \mathbb{Z}_p$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}\left[\|M_1^{(l)} - \mu_1\|, \dots, \|M_d^{(l)} - \mu_d\| \leq p^{-N-1} \right] - \frac{1}{p^{dN}} \right| \leq e^{-C'''N^{C'}}$$

whenever $l \geq C''N^{C+C'}$.

In the Nottingham group $\mathcal{N}_q$, the question of mixing times for the groups $\mathcal{N}_q/K_n$ was raised by Diaconis [25]. We may express:

$$Y_l = t + \sum_{i=2}^{\infty} A_i^{(l)} t^i$$

for some random variables $A_i^{(l)}$ supported on $\mathbb{F}_q$.

Corollary 3.1.10. *Let $p \geq 3$. Suppose $S \subseteq \mathcal{N}_q$ generates a dense subgroup. Then there exists an absolute constant $C > 0$, such that for any $C' > 0$ there exists $C''(q, |S|, C') > 0$ and $C'''(q, C') > 0$ such that, for any sequence $(\alpha_i)_i$ in $\mathbb{F}_q$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}[A_2^{(l)} = \alpha_2, \dots, A_N^{(l)} = \alpha_N] - \frac{1}{q^{N-1}} \right| \leq e^{-C'''N^{C'}}$$

whenever $l \geq C''N^{C+C'}$.

1.6.2 Expansion, random walks and sieving in $\mathrm{SL}_2(\mathbb{F}_p[t])$

Chapter 4 shall be devoted to our results on constructing expander congruence quotients of $\mathrm{SL}_2(\mathbb{F}_p[t])$, and associated bounds on return probabilities for simple random walks on $\mathrm{SL}_2(\mathbb{F}_p[t])$. These results were first presented in [13].

In spite of the major strides forward that have been made on the phenomenon of superstrong approximation in linear groups following Bourgain and Gamburd's work (and to which we have alluded already) it is notable that the work of recent years has focused entirely on the characteristic zero case, with a theory of expansion in linear groups over fields of positive characteristic remaining largely undeveloped. The results of Chapter 4 commence the development of such a theory.

Fix a prime $p \geq 3$. Our main results concern the escape of random walks on $\mathrm{SL}_2(\mathbb{F}_p[t])$ from subsets X of various types. All of the escape results are proved by the group sieve method described above: an upper bound for the probability of the random walk lying in X is given by the probability of the random walk on a congruence quotient lying in the image of X. This in turn may be bounded above in terms of the *size* of the image of X, by our results on expander congruence quotients. Our first result on random walks demonstrates exponentially fast escape from proper algebraic subvarieties.

Theorem 4.1.1. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric subset, generating a non-elementary subgroup. Let $F : \mathbb{M}_2(\mathbb{F}_p[t])^r \to \mathbb{F}_p[t]$ be a polynomial over $\mathbb{F}_p[t]$ which does not vanish on $\mathrm{SL}_2(\mathbb{F}_p[t])^r$. Then there exist $C_1(F)$, $C_2(S) > 0$ such that, letting $V(F) \subseteq \mathbb{M}_2(\mathbb{F}_p[t])^r$ be the affine algebraic subvariety of $\mathrm{SL}_2(\mathbb{F}_p[t])^r$ defined by F,*

$$(\times_{i=1}^r \mu_S^{(l)})(V(F)) \leq C_1 e^{-C_2 l}.$$

Here $\times_{i=1}^r \mu_S^{(l)}$ is the product measure.

Second, we turn to proper powers in $\mathrm{SL}_2(\mathbb{F}_p[t])$. In the characteristic zero case, the work of Lubotzky and Meiri on the group sieve provides a very general non-concentration result [55], as we have seen. In positive characteristic we have:

Theorem 4.1.2. *Let S be as in Theorem 4.1.1. There exist $C_1, C_2(S) > 0$ such that:*

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : g = h^2 \text{ for some } h \in \mathrm{SL}_2(\mathbb{F}_p[t])\}) \leq C_1 e^{-C_2\sqrt{l/\log l}}.$$

Finally we prove a non-concentration estimate for elements with reducible characteristic polynomial.

Theorem 4.1.4. *Let S be as in Theorem 4.1.1. There exist $C_1, C_2(S) > 0$ such that:*

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : \chi_g \text{ is reducible } \}) \leq C_1 e^{-C_2\sqrt{l/\log l}}.$$

It is very likely that bounds on return probabilities of random walks to other subsets of $\mathrm{SL}_2(\mathbb{F}_p[t])$ may be proved by the same method, and we emphasize that our results are best viewed as sample, rather than an exhaustive list, of the applications of this theory.

We now turn to our results on expanders.

Definition 1.6.1. *For $M > 0$, an integer $n > 1$ will be called M-*rough *if n has no prime factor less than M. A polynomial $f \in \mathbb{F}_p[t]$ will be called M-rough if the degree of every irreducible factor of f is a M-rough integer.*

Our main result on constructing expanders is:

Theorem 4.1.7. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric subset, generating a non-elementary subgroup. Suppose every entry of every element of S has degree at most D. Let $(f_i)_i \subseteq \mathbb{F}_p[t]$ be a sequence of distinct polynomials. Then there exists $M > 0$ (depending on D and p) such that, if $(f_i)_i$ are M-rough then for $i_0 \in \mathbb{N}$ sufficiently large (depending on D, p), $(\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i)), \pi_{f_i}(S))_{i \geq i_0}$ is a two-sided expander family, provided one of the following holds:*

(i) The f_i are irreducible;

(ii) The f_i are square-free, every irreducible factor of every f_i has prime degree, and no two irreducible factors of any f_i have the same degree.

Here, and throughout, for $\mathbb{G}$ a linear algebraic group defined over $\mathbb{F}_p$ and $f \in \mathbb{F}_p[t]$, $\pi_f : \mathbb{G}(\mathbb{F}_p[t]) \to \mathbb{G}(\mathbb{F}_p[t]/(f))$ shall denote the congruence map.

One of the keys to the proof of Theorem 4.1.7 shall be an analysis of Cayley graphs of large girth. For G a finite group generated by a subset S, recall that the girth of (G, S) is the length of the shortest non-trivial reduced word in S which equals 1 in G, or equivalently the length of the shortest non-trivial embedded loop in the Cayley graph $\mathrm{Cay}(G, S)$. In the course of our analysis, we also obtain the following result, which applies to generating sets which may not be congruence images of a fixed subset in $\mathrm{SL}_2(\mathbb{F}_p[t])$.

Theorem 4.1.8. *For any $C > 0$ and any $k \in \mathbb{N}_{\geq 2}$, there exists $M > 0$ (depending on k, p and C) such that, if $(n_i)_i$ a sequence of M-rough positive integers, $S_{n_i} \subseteq \mathrm{SL}_2(p^{n_i})$ is symmetric with $|S_{n_i}| = 2k$, and $\mathrm{girth}(\mathrm{SL}_2(p^{n_i}), S_{n_i}) \geq Cn_i$, for all $i \in \mathbb{N}$, then for i_0 sufficiently large (depending on C, k), $(\mathrm{SL}_2(p^{n_i}), S_{n_i})_{i \geq i_0}$ is a two-sided expander family.*

The lower bound on girth is a natural hypothesis: for instance it is satisfied by *generic* subsets of $\mathrm{SL}_2(p^n)$. Indeed large girth is a key component of the proof [19] that random pairs of generators in $\mathrm{SL}_2(p^n)$ yield expanders.

Beyond the concrete results which we obtain, our work in Chapter 4 should be instructive in indicating the path that the development of a theory of superstrong approximation in positive characteristic could take, and specifically in highlighting the potential obstructions to applying the methods of Bourgain and Gamburd in such a project. We shall fully describe these obstructions, which also account for the particular hypotheses of Theorem 4.1.7.

1.6.3 Spectral gap in $\mathrm{SL}_2(\mathbb{Z}_p)$

The goal of Chapter 5 shall be to prove:

Theorem 5.1.1. *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z}_p)$ be a finite symmetric set, generating a subgroup Γ whose closure $\overline{\Gamma}$ in $\mathrm{SL}_2(\mathbb{Z}_p)$ is open. Let $A \subseteq \mathbb{Z}_p$ be the set of entries occurring in elements of S. Suppose that for every $a \in A$, there exists $f_a(X) \in \mathbb{Q}[X]$ such that $f_a(a) = 0$. Then $(\pi_{p^n}(\Gamma), \pi_{p^n}(S))_n$ is a family of two-sided expanders, where $\pi_{p^n} : \mathrm{SL}_2(\mathbb{Z}_p) \to \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ is the congruence map.*

This is a generalisation of the main result of [6]. We may equivalently view Theorem 5.1.1 either as a result on expanders (as we have stated it here) or as a result about the spectral gap of the action of Γ on the p-adic Lie group $\mathrm{SL}_2(\mathbb{Z}_p)$. In this second guise, Theorem 5.1.1 is a non-archimedean cousin of the aforementioned spectral gap results for compact real Lie groups due to Bourgain and Gamburd ([8] and [9]) and Benoist and de Saxcé [2].

Chapter 2

Background Material

2.1 Diameter, Expansion and Random Walks

2.1.1 Random walks and spectral gap

In this section G is a finite group. We take $\ell^2(G)$ to be the Hilbert space of complex-valued functions on G, equipped with the inner product given by:

$$\langle \phi, \psi \rangle = \sum_{g \in G} \phi(g)\overline{\psi(g)}$$

for $\phi, \psi \in \ell^2(G)$. For $A \subseteq G$, we define $\chi_A \in \ell^2(G)$ to be the indicator function of A. For $g \in G$, we may also denote $\chi_{\{g\}}$ by χ_g. Observe that $\{\chi_g\}_{g \in G}$ is an orthonormal basis of $\ell^2(G)$. The *convolution* $\phi * \psi \in \ell^2(G)$ of ϕ and ψ is given by:

$$(\phi * \psi)(g) = \sum_{h \in G} \phi(h)\psi(h^{-1}g).$$

Equipped with $*$, $\ell^2(G)$ is a complex $|G|$-dimensional algebra: it is isomorphic to the complex group algebra $\mathbb{C}G$ of G. For $l \in \mathbb{N}$, we define the *convolution power* $\phi^{(l)}$ recursively via:

$$\phi^{(0)} = \chi_1;\ \phi^{(l+1)} = \phi^{(l)} * \phi.$$

We may alternatively regard non-negative real-valued functions $\phi \in \ell^2(G)$ as measures on G: for $A \subseteq G$ we write $\phi(A) = \sum_{a \in A} \phi(a)$. Given two such measures ϕ, ψ, $\phi * \psi$ is equal to the pushforward of the product measure $\phi \times \psi$ under the product map $G \times G \to G$. In particular, if ϕ, ψ are probability measures on G and X, Y are independent random variables with law ϕ, ψ, respectively, then $\phi * \psi$ is the law of XY.

For $S \subseteq G$ let $\mu_S = \frac{1}{|S|}\chi_S \in \ell^2(G)$.

Remark 2.1.1. *Let $S \subseteq G$ be symmetric.*

(i) Let W_l be the number of loops in the Cayley graph $Cay(G, S)$ of length l based at the identity. Then $\mu_S^{(l)}(1) = W_l/|S|^l$.

(ii) Since S is symmetric,

$$\mu_S^{(2l)}(1) = \sum_{g\in G} \mu_S^{(l)}(g)\mu_S^{(l)}(g^{-1}) = \sum_{g\in G} \mu_S^{(l)}(g)^2 = \|\mu_S^{(l)}\|_2^2.$$

(iii) Moreover, for arbitrary $g \in G$,

$$\mu_S^{(2l)}(g) = \sum_{h\in G} \mu_S^{(l)}(hg)\mu_S^{(l)}(h^{-1}) \leq \|\mu_S^{(l)}\|_2^2$$

by the Cauchy-Schwarz inequality, so $\mu_S^{(2l)}(1) = \max_{g\in G} \mu_S^{(2l)}(g)$ (by (ii)).

(iv) Finally,

$$\begin{aligned}
\mu_S^{(2l+2)}(1) &= \sum_{g\in G} \mu_S^{(2l)}(g)\mu_S^{(2)}(g^{-1}) \\
&\leq (\max_{g\in G} \mu_S^{(2l)}(g)) \cdot \sum_{h\in G} \mu_S^{(2)}(h) \\
&= \max_{g\in G} \mu_S^{(2l)}(g) \\
&= \mu_S^{(2l)}(1) \text{ (by (iii)).}
\end{aligned}$$

Define a linear operator $A_S : \ell^2(G) \to \ell^2(G)$ (called the *adjacency operator*) by:

$$A_S(f) = \mu_S * f.$$

A_S is self-adjoint of operator norm 1; let its spectrum be:

$$1 = \lambda_1 \geq \lambda_2 \geq \ldots \lambda_{|G|} \geq -1$$

with the eigenvalue $\lambda_1 = 1$ corresponding to the constant functionals on G. More generally, the 1-eigenspace of A_S is generated by the indicator functions of the right cosets of $\langle S\rangle \leq G$. In particular $\lambda_1 > \lambda_2$ iff S generates G.

Lemma 2.1.2 (Trace Formula)**.** *Let W_{2l} be as in Remark 2.1.1 (i). Then:*

$$W_{2l} = \frac{|S|^{2l}}{|G|} \sum_{i=1}^{|G|} \lambda_i^{2l}.$$

Let $\ell^2_0(G) \leq \ell^2(G)$ be the space of functions of mean zero on G (that is, the orthogonal complement of the constant functions), and note that $\ell^2_0(G)$ is preserved by A_S. Let $\rho = \max(|\lambda_2|, |\lambda_{|G|}|)$, the norm of the restriction $A_S \mid_{\ell^2_0(G)}$ in the Banach space $B(\ell^2_0(G))$ of bounded linear operators on $\ell^2_0(G)$.

Lemma 2.1.3. *For any $l \in \mathbb{N}$; $g, h \in G$,*

$$|\langle A_S^l \chi_g, \chi_h \rangle - \tfrac{1}{|G|}| \leq \rho^l.$$

Proof. Noting that $\chi_g - \frac{1}{|G|}\chi_G \in \ell^2_0(G)$,

$$\begin{aligned} |\langle A_S^l \chi_g, \chi_h \rangle - \frac{1}{|G|}| &= |\langle A_S^l (\chi_g - \frac{1}{|G|}\chi_G), \chi_h \rangle| \\ &\leq \|A_S^l (\chi_g - \frac{1}{|G|}\chi_G)\|_2 \end{aligned}$$

by the Cauchy-Schwarz inequality. The result follows, since:

$$\|\chi_g - \frac{1}{|G|}\chi_G\|_2 \leq 1.$$

□

Lemma 2.1.3 has an obvious interpretation in terms of random walks. Let $X_1, X_2, \ldots$ be a sequence of independent random variables, each with law μ_S. For $l \in \mathbb{N}$, the *simple random walk* $Y_l = X_1 \cdots X_l$ on (G, S) at time l has law $\mu_S^{(l)}$. For $g, h \in G$, the probability that a simple random walk starting at g will be at h at time l is:

$$\mathbb{P}\big[Y_l g = h\big] = \langle A_S^l \chi_g, \chi_h \rangle.$$

Lemma 2.1.3 then gives an explicit bound on the rate at which the distribution of Y_l converges to the uniform distribution on G (assuming S generates G).

2.1.2 Diameters of finite groups

Given a finite group G and a generating set $S \subseteq G$, the *diameter* of the pair (G, S) is given by:

$$\operatorname{diam}(G, S) = \min\{n \in \mathbb{N} : B_S(n) = G\}$$

where $B_S(n)$ is the (closed) *word-ball* of radius n, given by:

$$B_S(n) = \{s_1 \cdots s_n : s_1, \ldots, s_n \in S \cup S^{-1} \cup \{1\}\}.$$

Example 2.1.4. *(i) It is easy to see that, if $A \subseteq G$ and $n \in \mathbb{N}$ satisfy $B_A(n) = B_A(n+1)$, then $B_A(n) = B_A(n+m)$ for all $m \in \mathbb{N}$. In particular, for $S \subseteq G$ a generating set, and $n < \mathrm{diam}(G,S)$, $|B_S(n)| < |B_S(n+1)|$. It follows that:*

$$\mathrm{diam}(G,S) \leq |G| - 1$$

so there is an absolute linear upper bound for the diameter. Moreover,

$$\mathrm{diam}(\mathbb{Z}/n\mathbb{Z}, \{1\}) = \lfloor \tfrac{n}{2} \rfloor$$

so this linear upper bound is attained (at least up to constants) for some pairs (G,S).

(ii) For all $A \subseteq G$ and $n \in \mathbb{N}$,

$$|B_A(n)| \leq 2(2|A| - 1)^n$$

(by a simple count of words), provided $|A| \geq 2$. It follows that for $S \subseteq G$ a generating set satisfying $|S| \geq 2$,

$$\mathrm{diam}(G,S) \geq (\log|G| - \log(2))/\log(2|S| - 1).$$

Examples attaining this logarithmic lower bound may also be found (again, up to constants), but this is a far more delicate matter, to which we shall return later in our discussion of expanders.

(iii) The dependence on $|S|$ in (ii) is necessary, as for any G,

$$\mathrm{diam}(G,G) = 1.$$

Our interest shall be in upper bounds for the diameter which do not depend on the generators. With this in mind, we define for a finite group G:

$$\mathrm{diam}(G) = \max\{\mathrm{diam}(G,S) : S \subseteq G, \langle S \rangle = G\}.$$

This quantity is referred to by many authors as the *worst-case diameter* for G.

One advantage of working with diam is that it behaves well with respect to extensions.

Lemma 2.1.5. *Let G be a finite group, $K \lhd G$. Then:*

(i) $\operatorname{diam}(G/K) \leq \operatorname{diam}(G)$*;*

(ii) $\operatorname{diam}(G) \leq 2\operatorname{diam}(G/K)\operatorname{diam}(K) + \operatorname{diam}(G/K) + \operatorname{diam}(K)$.

Proof. (i) is straightforward.

(ii) Let $S \subseteq G$ be a generating set. Then $B_S(\operatorname{diam}(G/K))$ contains a transversal T to K in G, with $1 \in T$. By the Reidemeister-Schreier process (see for instance [44], [60]), $B_S(2 \cdot \operatorname{diam}(G/K) + 1)$ contains a generating set for K. Hence:

$$\operatorname{diam}(G, S) \leq \operatorname{diam}(G/K) + \operatorname{diam}(K) \cdot (2 \cdot \operatorname{diam}(G/K) + 1)$$

as required.

□

There is also a close connection between diameter and random walks. It seems clear that a pair (G, S) on which the simple random walk approaches uniformity quickly should have small diameter: if the probability of the simple random walk reaching an element g at time l is non-zero, then there exists at least one word in S of length l equal to g in G. Less obviously, there is a corresponding converse inequality, a proof of which may be found in [26]. In summary we have:

Proposition 2.1.6. *Let ρ be as in Lemma 2.1.3. Then:*

$$\frac{\operatorname{diam}(G, S) - 1}{\log|G|} \leq \frac{1}{1 - \rho} \leq |S| \operatorname{diam}(G, S)^2.$$

2.1.3 Expanders

Definition 2.1.7. *For $\epsilon > 0$, the pair (G, S) is a (two-sided)* ϵ-expander *if $\rho \leq 1 - \epsilon$. A sequence $(G_n, S_n)_{n \in \mathbb{N}}$ is called an* ϵ-expander family *if (G_n, S_n) is an ϵ-expander for every $n \in \mathbb{N}$, or just an* expander family *if there exists $\epsilon > 0$ such that $(G_n, S_n)_{n \in \mathbb{N}}$ is an ϵ-expander family.*

The two-sided version of expansion (also known as *absolute expansion*) that we use here is stronger than the one-sided version which will be more familiar to many readers, and which is equivalent to the combinatorial notion of expansion defined in terms of the discrete Cheeger constant. For the most part, however, the distinction need not concern us, thanks to a recent result of Breuillard, Green, Guralnick and Tao [19]:

Theorem 2.1.8. *For any $\epsilon > 0, k \in \mathbb{N}$, there exists $\delta(\epsilon, k) > 0$ such that, if (G, S) is a one-sided ϵ-expander with $|S| = k$, then one of the following holds:*

(i) (G, S) is a two-sided δ-expander;

(ii) There exists $H \leq G$ with $|G : H| = 2$ and $S \cap H = \emptyset$.

We recall some facts about expanders which will be used in what follows. Those readers more familiar with the one-sided version of expansion may note that these results about two-sided expanders follow from their one-sided analogues together with Theorem 2.1.8. Note that condition (ii) of Theorem 2.1.8 is equivalent to $\mathrm{Cay}(G, S)$ being bipartite.

Lemma 2.1.9. *Let Γ be a finitely generated group; $(K_n)_n$ be a sequence of finite index normal subgroups of Γ; $\pi_n : \Gamma \twoheadrightarrow \Gamma/K_n$ be the natural epimorphism. Let $S, T \subseteq \Gamma$ be finite symmetric subsets, with $\langle S \rangle = \langle T \rangle = \Gamma$. Suppose $\mathrm{Cay}(\Gamma/K_n, \pi_n(T))$ is not bipartite, for all $n \in \mathbb{N}$. If $(\Gamma/K_n, \pi_n(S))_n$ is an expander family then so is $(\Gamma/K_n, \pi_n(T))_n$.*

Lemma 2.1.10. *Let Γ; $(K_n)_n$; π_n be as in Lemma 2.1.9 and let $H \leq \Gamma$ be a finitely generated subgroup. Suppose $\pi_n(H) = \Gamma/K_n$ for all $n \in \mathbb{N}$. Let $S \subseteq \Gamma$, $T \subseteq H$ be finite symmetric subsets, with $\langle S \rangle = \Gamma$, $\langle T \rangle = H$. If $(\Gamma/K_n, \pi_n(T))_n$ is an expander family, and $\mathrm{Cay}(\Gamma/K_n, \pi_n(S))$ is not bipartite, then $(\Gamma/K_n, \pi_n(S))_n$ is an expander family.*

In all cases in which we shall use these results, the finite groups concerned shall have no subgroups of index two, so the associated Cayley graphs shall never be bipartite.

Finally, expanders exhibit rapid mixing of the random walk and small diameter:

Proposition 2.1.11. *Let $(G_n, S_n)_n$ be an expander family.*

(i) For any $\alpha > 0$ there exists $C_1 > 0$ such that for all $n \in \mathbb{N}$ and $l \geq C_1 \log|G_n|$,

$$|\langle A_S^l \chi_g, \chi_h \rangle - \frac{1}{|G|}| \leq |G_n|^{-\alpha}.$$

(ii) There exists $C_2 > 0$ such that for all $n \in \mathbb{N}$, $\mathrm{diam}(G_n, S_n) \leq C_2 \log|G_n|$.

Proof. Immediate from Lemma 2.1.3 and Proposition 2.1.6. □

2.2 The Bourgain-Gamburd Machine

This section describes the set of tools for constructing expanders which have come collectively to be know as the *Bourgain-Gamburd machine*, following their introduction in [5]. They shall form the core of our constructions of expanders in Chapters 4 and 5. The emphasis of our discussion here is tailored to the requirements of our own work; we do not claim to include all the possible ways in which Bourgain and Gamburd's ideas may be applied.

2.2.1 Quasirandomness

The notion of a *quasirandom finite group* was introduced by Gowers in [34]. The name is somewhat misleading: typically in combinatorics an object in some class is described as *quasirandom* if it possesses some combination of properties that are shared by generic objects according to some model of random members of the class. A finite graph Γ of density p, for instance, is termed quasirandom if, whenever $A, B \subseteq V(\Gamma)$, the number of edges between A and B is approximately $p|A||B|$ (as one would expect for a random graph according to, say, the Erdős-Renyi model). Quasirandom finite groups do not fit readily into this picture: there is to the author's knowledge no well-accepted model for random finite groups. Instead, the name comes from the fact that the "bipartite Cayley graph" of a quasirandom group is a "quasirandom bipartite graph" (it shall not be of concern to us precisely what this means).

The condition Gowers exploited for a group to give rise to such quasirandom graphs was a lower bound on the dimension of non-trivial complex representations. Here we give two formulations, one of which is also relevant to infinite groups.

Definition 2.2.1. *Let $C, \alpha > 0$.*

(i) Let G be a finite group. We say G is (C, α)-quasirandom *if every non-trivial complex representation of G has dimension at least $C|G|^{\alpha}$.*

(ii) More generally let G be an arbitrary group; $\mathcal{N}$ a family of finite-index normal subgroups of G. We say G is weakly (C, α)-quasirandom *with respect to $\mathcal{N}$ if, whenever ρ is a finite complex representation of G such that* $\ker(\rho) \in \mathcal{N}$,

$$\dim(\rho) \geq C|G : \ker(\rho)|^{\alpha}.$$

Our terminology here is non-standard: for instance Tao [76] defines a finite group G to be *d-quasirandom* if every non-trivial complex representation of G has dimension

at least d (hence, that which we call (C, α)-*quasirandom* would in Tao's terminology be called $C|G|^{\alpha}$-*quasirandom*). Gowers' definition [34] is given in terms of a condition on the space of complex functionals on G: it is polynomially equivalent to Tao's.

Our own definition is informed by the fact that bounds on dimensions given in terms of a fractional power of the group order are most relevant to applications concerning expansion. This fact has already been somewhat reflected by the terminology in the literature: Varjú [78] refers to groups satisfying condition (ii) of Definition 2.2.1 as (C, α)-*quasirandom*. We refer to such groups as *weakly* quasirandom because we shall be dealing with both parts of Definition 2.2.1 in this section and it shall be helpful to be able to distinguish easily between them. Of course, according to our nomenclature, a finite group G which is (C, α)-quasirandom is also weakly (C, α)-quasirandom with respect to the family of all normal subgroups.

Let us begin by noting some elementary properties of group representations.

Lemma 2.2.2. *If G is a finite group and $H \lneq G$ then the permutational representation of G on the coset space $(G : H)$ induces a non-trivial complex representation of G of dimension $|G : H|$. In particular, if G is (C, α)-quasirandom then every $H \lneq G$ satisfies $|G : H| \geq C|G|^{\alpha}$.*

Lemma 2.2.3. *Let $1 \to N \to G \to Q \to 1$ be a short exact sequence of finite groups.*

(i) If every non-trivial complex representation of G has dimension $\geq d$, then the same holds for Q. In particular, if G is (C, α)-quasirandom, then Q is $(C|N|^{\alpha}, \alpha)$-quasirandom.

(ii) If every non-trivial complex representation of either N or Q has dimension $\geq d$, then the same holds for G. In particular, if N is (C_1, α_1)-quasirandom and Q is (C_2, α_2)-quasirandom, then G is $(\min(C_1/|Q|^{\alpha_1}, C_2/|N|^{\alpha_2}), \min(\alpha_1, \alpha_2))$-quasirandom.

Proposition 2.2.4. *Let G be a group, $H \leq_f G$ and $C, \alpha > 0$.*

(i) If G is finite (C, α)-quasirandom, then H is $(C|G : H|^{\alpha-1}, \alpha)$-quasirandom.

(ii) If G is weakly (C, α)-quasirandom with respect to the family of all finite-index normal subgroups then H is weakly $(C|G : H|^{\alpha-1}, \alpha)$-quasirandom with respect to the family of all finite-index normal subgroups.

Proof. Let ρ be a non-trivial finite d-dimensional complex representation of H. Recall that there is an *induced representation* $\tilde{\rho} = \mathrm{Ind}_H^G(\rho)$ of G. This is a finite $|G : H|d$-dimensional complex representation of G, satisfying $\ker(\tilde{\rho}) \leq \ker(\rho)$.

(i) If G is finite (C, α)-quasirandom, then:

$$|G : H|d \geq C|G|^{\alpha}$$

so $d \geq C|G : H|^{\alpha-1}|H|^{\alpha}$.

(ii) If G is weakly (C, α)-quasirandom with respect to the family of all finite-index normal subgroups, then:

$$|G : H|d \geq C|G : \ker(\tilde{\rho})|^{\alpha} \geq C|G : \ker(\rho)|^{\alpha}$$

so $d \geq C|G : H|^{\alpha-1}|H : \ker(\rho)|^{\alpha}$.

□

Which groups exhibit strong quasirandomness properties, and which do not?

Example 2.2.5. *Every non-trivial finite abelian group has a non-trivial one-dimensional complex representation. Conversely, every one-dimensional complex representation has abelian image, so every finite group without a non-trivial one-dimensional complex representation is perfect.*

In particular, for all $C, \alpha > 0$, if G is a (C, α)-quasirandom finite group satisfying $|G| > 1/C^{\frac{1}{\alpha}}$, then G is perfect.

Theorem 2.2.6 (Landazuri-Seitz, [50])**.** *For all $l \in \mathbb{N}$, there exist $C(l), \alpha(l) > 0$ such that if G is a finite simple group of Lie type of rank l, then G is (C, α)-quasirandom.*

Example 2.2.7. *The standard permutational representation of A_n induces a non-trivial complex representation of dimension n. In particular there do not exist $C, \alpha > 0$ such that A_n is (C, α)-quasirandom for all $n \in \mathbb{N}$.*

The relevance of quasirandomness-type conditions to our purposes is that they allow us to reduce expansion to an *a priori* weaker condition:

Definition 2.2.8. *Let G be a finite group with symmetric generating set S and let $C : (0, \infty) \to (0, \infty)$. We say that (G, S)* satisfies the ℓ^2-flattening condition with function C *if, for all $\epsilon > 0$, there exists $l \leq C(\epsilon) \log|G|$ such that:*

$$\mu_S^{(2l)}(1) < |G|^{\epsilon-1}.$$

This upper bound on $\mu_S^{(2l)}(1)$ is exactly what we would expect to hold if (G, S) were a good expander: if after logarithmic time $\mu_S^{(2l)}$ assigns a mass of approximately $1/|G|$ to every element of G, then the identity (which by Remark 2.1.1 (iii) is the element at which the greatest mass accumulates) cannot receive a mass much greater than this. Of course, in principle (G, S) could still fail to be an expander when the ℓ^2-flattening condition holds, if most elements were assigned a mass of approximately $|G|^{\epsilon-1}$ but a few "hard-to-reach" elements received a much smaller mass.

The observation that ℓ^2-flattening combines with quasirandomness to yield expansion goes back to Sarnak and Xue [70]; the basic idea is that G_n acts on the eigenspace of the second largest eigenvalue of the adjacency operator A_{S_n}, so that by quasirandomness this eigenvalue has high multiplicity. If the expansion of (G_n, S_n) is poor, then it follows that the trace of $A_{S_n}^{2l}$ is large (being the sum of the $(2l)$th powers of the eigenvalues). However this contradicts ℓ^2-flattening.

The most basic result of this form is the following:

Proposition 2.2.9. *Let G be a finite group with symmetric generating set S, let $C_1, \alpha > 0$ and let $C_2 : (0, \infty) \to (0, \infty)$. Suppose:*

(i) G is (C_1, α)-quasirandom;

(ii) (G, S) satisfies the ℓ^2-flattening condition with function C_2.

Then there exists $\delta(\alpha, C_1, C_2, |S|) > 0$ such that (G, S) is a δ-expander, provided $|G|$ is sufficiently large, depending on $\alpha, C_1, C_2, |S|$.

Proof. First note that $\mu_S^{(2l)}(1) = \langle A_S^{2l}\chi_g, \chi_g\rangle$ for every $g \in G$. Hence:

$$\mu_S^{(2l)}(1)|G| = \mathrm{tr}(A_S^{2l}) = \sum_{i=1}^{|G|} \lambda_i^{2l} \tag{2.1}$$

where $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_{|G|}$ are the eigenvalues of A_S. Next, it is clear that A_S commutes with the right regular representation $(R, \ell^2(G))$ of G. In particular, R preserves the eigenspaces of A_S. Now let $\lambda = \lambda_i$ for some $i \geq 2$ and let $V_\lambda \leq \ell^2(G)$ be the λ-eigenspace of A_S. (R, V_λ) is a representation of G, and is non-trivial, because the λ-eigenfunctions of A_S are non-constant. By quasirandomness of G, $\dim(V_\lambda) \geq C_1|G|^\alpha$, so that, by (2.1):

$$\mu_S^{(2l)}(1)|G| \geq \dim(V_\lambda)\lambda^{2l} \geq C_1|G|^\alpha\lambda^{2l}.$$

Let $\epsilon \in (0, \alpha)$. By the ℓ^2-flattening condition, we may choose $l \leq C_2(\epsilon) \log|G|$ as in Definition 2.2.8, so that:

$$|G|^{\epsilon-\alpha} \geq C_1 \lambda^{2l} \geq C_1 \lambda^{2C_2(\epsilon)\log|G|}.$$

For $|\lambda| > e^{\frac{\epsilon-\alpha}{2C_2(\epsilon)}}$, and for $|G|$ sufficiently large, this is a contradiction. □

In fact for the results of Chapters 4 and 5, we shall require two stronger versions of this result. In the first, the hypothesis is weakened by replacing the quasirandom groups G_n with direct products of quasirandom groups.

Proposition 2.2.10. *Let G be a finite group with symmetric generating set S, let $C_1, \alpha > 0$ and let $C_2 : (0, \infty) \to (0, \infty)$. Suppose:*

(i) There exist finite groups $G_1, \ldots, G_n$ with $G = G_1 \times \cdots \times G_n$;

(ii) G_i is (C_1, α)-quasirandom for $i = 1, \ldots, n$;

(iii) For every $J \subseteq \{1, \ldots, n\}$, $(\prod_{j \in J} G_j, S)$ satisfies the ℓ^2-flattening condition with function C_2.

Then there exists $\delta(\alpha, C_1, C_2, |S|) > 0$ such that (G, S) is a δ-expander, provided each $|G_i|$ is sufficiently large, depending on $\alpha, C_1, C_2, |S|$.

The key point here is that the expansion of (G, S) is independent of n. To prove this result we recall a basic fact from the representation theory of finite groups.

Lemma 2.2.11. *Let G_1, G_2 be finite groups, and let ρ be an irreducible complex representation of $G_1 \times G_2$. Then there exist irreducible complex representations ρ_1, ρ_2 of G_1, G_2, respectively, such that $\rho \cong \rho_1 \otimes \rho_2$.*

Proof of Proposition 2.2.10. First note that, by shrinking α and taking $|G_i|$ sufficiently large, we may take $C_1 = 1$. We proceed by induction on n. The base case $n = 1$ is Proposition 2.2.9.

Suppose the claim holds for smaller n. Let $\lambda < 1$ be an eigenvalue of A_S with corresponding eigenfunction f. For each i let V_i be the subspace of $\ell^2(G)$ spanned by $\{\chi_{gG_i} : g \in G\}$, and let W_i be its orthogonal complement (consisting of functions with integral zero on each left coset of G_i). Note that V_i, W_i are preserved under left-translation by G (so are preserved by A_S). Suppose that $f \notin W_i$ for some i. Let $\pi_i : \ell^2(G) \twoheadrightarrow \ell^2(G/G_i)$ extend the quotient map $G \twoheadrightarrow G/G_i$. Then $\pi_i \circ A_S = A_{\pi_i(S)} \circ \pi_i$, so that $\pi_i(f)$ is a non-zero λ-eigenfunction of $A_{\pi_i(S)}$, and the claim follows by induction.

We may therefore assume that $f \in W_i$ for all i. Let $U \leq \ell^2(G)$ be the subspace spanned by all right-translates of f, so that U is a right G-representation. Note that:

(a) U is contained in the λ-eigenspace of A_S, since A_S commutes with right translation;

(b) $U \leq W_i$ for all i.

Let U' be an irreducible G-subrepresentation of U. By Lemma 2.2.11, there exist irreducible G_i-representations U_i such that $U' \cong U_1 \otimes \cdots \otimes U_n$.

We note that for all i, U_i is a *non-trivial* representation of G_i. For suppose U_i were trivial. Then the action of G_i on U' would be trivial, so every element of U' would be constant on the left cosets of G_i, and $U' \leq V_i$. But by (b), $U' \leq W_i$ as well, and $V_i \cap W_i = \{0\}$, a contradiction.

By quasirandomness of G_i, $\dim(U_i) \geq |G_i|^\alpha$. Thus:

$$\dim(U) \geq \dim(U') = \prod_{i=1}^{n} \dim(U_i) \geq \prod_{i=1}^{n} |G_i|^\alpha = |G|^\alpha.$$

By (a) the λ-eigenspace V_λ of A_S contains U, so we may argue as in Proposition 2.2.9: by (2.1) V_λ satisfies:

$$\mu_S^{(2l)}(1)|G| \geq \dim(V_\lambda)\lambda^{2l} \geq |G|^\alpha \lambda^{2l}.$$

Let $\epsilon \in (0, \alpha)$. By the ℓ^2-flattening condition, we may choose $l \leq C_2(\epsilon)\log|G|$ as in Definition 2.2.8, so that:

$$|G|^{\epsilon-\alpha} \geq \lambda^{2l} \geq \lambda^{2C_2(\epsilon)\log|G|}.$$

For $|\lambda| > e^{\frac{\epsilon-\alpha}{2C_2(\epsilon)}}$, and for $|G|$ sufficiently large, this is a contradiction. □

In our second strengthening of Proposition 2.2.9, we replace the sequence of quasirandom groups with a family of finite quotients of a weakly quasirandom group.

Proposition 2.2.12. *Let G be a group generated by a finite symmetric set S. Let $\mathcal{N}$ be family of finite-index normal subgroups of G which is closed under both finite intersections and passing to larger normal subgroups. Let $C_1, \alpha > 0$ and let $C_2 : (0, \infty) \to (0, \infty)$. Suppose:*

(i) G is weakly (C_1, α)-weakly quasirandom with respect to $\mathcal{N}$;

(ii) For every $N \in \mathcal{N}$, $(G/N, S)$ satisfies the ℓ^2-flattening condition with function C_2.

Then there exists $\delta(\alpha, C_1, C_2, |S|) > 0$ such that $(G/N, S)$ is a δ-expander whenever $N \in \mathcal{N}$ is such that $|G/N|$ is sufficiently large, depending on $\alpha, C_1, C_2, |S|$.

Proof. We proceed by induction on the lattice $\mathcal{N}$. For $N \in \mathcal{N}$ recall that the regular representation of G/N decomposes as a direct sum of irreducible representations. Each eigenvalue λ of A_S (acting on $\ell^2(G/N)$) is an eigenvalue of one of these representations. By induction, it suffices to prove spectral gap for λ lying in the spectrum of a *faithful* representation of G/N (for any non-faithful representation is a faithful representation of G/M for some $M \geq N$ and the result follows by induction).

We argue as in Proposition 2.2.9. By (2.1) the λ-eigenspace $V_\lambda \leq \ell^2(G/N)$ of A_S satisfies:

$$\mu_S^{(2l)}(1)|G/N| \geq \dim(V_\lambda)\lambda^{2l} \geq |G|^\alpha \lambda^{2l}.$$

Let $\epsilon \in (0, \alpha)$. By the ℓ^2-flattening condition, we may choose $l \leq C_2(\epsilon)\log|G|$ as in Definition 2.2.8, so that:

$$|G/N|^{\epsilon-\alpha} \geq \lambda^{2l} \geq \lambda^{2C_2(\epsilon)\log|G/N|}.$$

For $|\lambda| > e^{\frac{\epsilon-\alpha}{2C_2(\epsilon)}}$, and for $|G|$ sufficiently large, this is a contradiction. The result follows by induction. □

2.2.2 The ℓ^2-flattening lemma

In the presence of quasirandomness then, we have reduced the problem of constructing expanders to verification of the ℓ^2-flattening condition. How might this be achieved? Certainly, there are some obvious algebraic obstructions to ℓ^2-flattening. For instance, if most of the mass of $\mu_S^{(l)}$ becomes concentrated in a small proper subgroup H of G (or a coset thereof), then much of this mass will be recycled back into H again and again, so will not spread out efficiently over G. More generally, if H is not a genuine subgroup but is a subset with *slow growth* under multiplication with itself (in other words, a subset which is not necessarily closed under multiplication in G, but which is *almost* closed, in some quantitative sense) then the mass of $\mu_S^{(l)}$ will still tend to be recycled back into a small subset of G, and ℓ^2-flattening will fail. The centrepiece of the machine is Bourgain and Gamburd's remarkable observation that such sets of slow growth constitute essentially the only obstruction to ℓ^2-flattening. That is to say, if μ is a symmetric probability measure on G such that $(\mu * \mu)(1)$ is not appreciably smaller than $\mu(1)$, then there exists a set of small growth contained in the support of $\mu * \mu$, upon which $\mu * \mu$ is concentrated.

This idea is encapsulated in the *ℓ^2-flattening Lemma*, which in turn draws upon the *non-commutative Balog-Szemerédi-Gowers Theorem*, due to Tao [75].

What do we mean by a subset with *slow growth*? The term has several plausible interpretations. Here we present versions of the ℓ^2-flattening Lemma based on two of them, namely on the concept of a *set of small tripling* and on that of an *approximate subgroup*. The two notions are closely related, as we shall see.

Definition 2.2.13. *Let G be an arbitrary group; $A \subseteq G$ be finite and $K \geq 1$. A is said to have* tripling at most K *if:*

$$|A \cdot A \cdot A| \leq K|A|.$$

*A is said to be a K-*approximate subgroup of G *if:*

(i) $1 \in A$;

(ii) *A is symmetric;*

(iii) *There exists $X \subseteq G$ such that $|X| \leq K$ and $A \cdot A \subseteq A \cdot X$.*

Remark 2.2.14. *(i) If $\phi : G_1 \to G_2$ is a homomorphism of groups and $A \subseteq G_1$ is a K-approximate subgroup then so is $\phi(A) \subseteq G_2$.*

(ii) If $A, X \subseteq G$ are such that $A \cdot A \subseteq A \cdot X$, then for any $r \geq 2$,

$$A^{(r)} \subseteq A \cdot X^{(r-1)}. \tag{2.2}$$

In particular, if A is a K-approximate subgroup of G, then $A^{(r)}$ is a K^{2r-1}-approximate subgroup of G. Likewise, if A has tripling at most K, then $A^{(r)}$ has tripling at most K^r.

(iii) Taking $r = 3$ in (2.2), we see that every K-approximate subgroup has tripling at most K^2.

(iv) Conversely (though this is much less obvious), there exists an absolute constant $C > 0$ such that if $A \subseteq G$ has tripling at most K, then $A' = (A \cup A^{-1} \cup \{1\})^2$ is a CK^C-approximate subgroup of G satisfying $|A'| \leq CK^C|A|$.

Many basic results about finite groups which rely on counting arguments can be generalised to the setting of approximate subgroups [38]. To give just one example, we have an approximate analogue of (one direction of) the class equation. Recall that for $g \in G$, $\mathrm{ccl}_G(g)$ is the conjugacy class of g in G.

Lemma 2.2.15. *Let $A \subseteq G$ be non-empty finite symmetric, $l \geq 1$. Then for every $g \in A^{(l)}$,*

$$|A^{(2)} \cap C_G(g)| \geq |A|/|A^{(l+2)} \cap \mathrm{ccl}_G(g)|. \tag{2.3}$$

In particular, if A is a K-approximate subgroup, if $U \subseteq A^{(l)}$ consists of pairwise non-G-conjugate elements, and $|U| \geq M$, then for some $g \in U$:

$$|A^{(2)} \cap C_G(g)| \geq M/K^{l+1}. \tag{2.4}$$

This shall follow from the following, which may be viewed as a version of (one direction of) the Orbit-Stabiliser Theorem valid for arbitrary finite subsets.

Lemma 2.2.16. *Let G be a group, acting on a set X. Let $x \in X$, let $\mathrm{Stab}_G(x)$ be the stabiliser of x in G and let $A \subseteq G$ be finite non-empty. Then:*

$$|(A^{-1}A) \cap \mathrm{Stab}_G(x)| \geq |A|/|A \cdot x|.$$

Proof. For $y \in A \cdot x$, define $B_y = \{a \in A : a \cdot x = y\}$. By the pigeonhole principle, for some y, $|B_y| \geq |A|/|A \cdot x|$.

For any $b \in B_y$, $b^{-1}B_y \subseteq (A^{-1}A) \cap \mathrm{Stab}_G(x)$, so $|(A^{-1}A) \cap \mathrm{Stab}_G(x)| \geq |B_y|$, as required. □

Proof of Lemma 2.2.15. (2.3) is immediate from Lemma 2.2.16, applied to the action of G on itself by conjugation, with $x = g$, and noting that the set of A-conjugates of g is contained in $A^{(l+2)} \cap \mathrm{ccl}_G(g)$. Suppose (2.4) fails for all $g \in U$. Then:

$$\begin{aligned}
K^{l+1}|A| &\geq |A^{(l+2)}| \\
&\geq \sum_{g \in U} |A^{(l+2)} \cap \mathrm{ccl}_G(g)| \text{ (as the } g \text{ are pairwise non-conjugate)} \\
&\geq |A| \sum_{g \in U} (1/|A^{(2)} \cap C_G(g)|) \text{ (by (2.3))} \\
&> K^{l+1}|A||U|/M \\
&\geq K^{l+1}|A|, \text{ a contradiction.}
\end{aligned}$$

□

Theorem 2.2.17 (ℓ^2-flattening Lemma with sets of small tripling; Lemma 15 from [77]). *There exists an absolute constant $C > 0$ such that the following holds. Let G be an arbitrary group; μ, ν be finitely supported probability measures on G and $K > 2$. Suppose:*

$$\|\mu * \nu\|_2 > \|\mu\|_2^{\frac{1}{2}} \|\nu\|_2^{\frac{1}{2}} / K.$$

Then there exists a symmetric subset $A \subseteq G$ such that:

(i) $1/CK^C \|\mu\|_2^2 \leq |A| \leq CK^C/\|\mu\|_2^2$;

(ii) A *has tripling at most* CK^C;

(iii) For all $g \in A$, $(\tilde{\mu} * \mu)(g) \geq 1/CK^C|A|$.

Here $\tilde{\mu}$ is given by $\tilde{\mu}(g) = \mu(g^{-1})$, so that $\tilde{\mu} = \mu$ iff μ is symmetric. Recall (Remark 2.1.1 (ii)) that in case μ is symmetric, $(\mu * \mu)(1) = \|\mu\|_2^2$. Notice also that the hypothesis is even weaker than was intimated above: we do not require that the measure μ collides substantially with *itself* under convolution, only that it collides with *some* measure ν.

Theorem 2.2.18 (ℓ^2-flattening Lemma with approximate subgroups; Lemma 4.0.1 from [76]). *There exist absolute constants $C > 0$ such that the following holds: let ν be a finitely supported symmetric probability measure on G and let $K \geq 1$. Then one of the following holds:*

(i) $\|\nu * \nu\|_2 \leq \frac{1}{K}\|\nu\|_2$;

(ii) G has an CK^C-approximate subgroup H such that:

$$|H| \leq CK^C/\|\nu\|_2^2$$

and $\nu(gH) \geq CK^{-C}$ for some $g \in G$.

Remark 2.2.19. *Letting H be as in Theorem 2.2.18 (ii),*

$$|H \cdot H| \leq CK^C|H| \leq C^2K^{2C}/\|\nu\|_2^2,$$

and by Remark 2.2.14 (ii), $H \cdot H$ is a C^3K^{3C}-approximate subgroup of G. Moreover,

$$(\nu * \nu)(H \cdot H) \geq \nu(Hg^{-1})\nu(gH) = \nu(gH)^2 \geq C^2K^{-2C}$$

since ν is symmetric. In other words, up to polynomial losses in the constants involved and replacing ν by $\nu * \nu$, we can take $g = 1$ in Theorem 2.2.18 (ii).

2.2.3 Product theorems and non-concentration

With the (appropriate version of the) ℓ^2-flattening Lemma in hand, it remains only to show that the $\mu_S^{(l)}$ are not too concentrated on approximate subgroups or sets of small tripling in G. In the setting of Chapter 5, we shall tackle this problem head on, showing that any approximate subgroup in which $\mu_S^{(l)}$ becomes too trapped must already be bigger than the conclusion of Theorem 2.2.18 permits.

However it is often possible to reduce the problem further, by classifying the slowly growing subsets of G. We have noted already that proper subgroups do not grow. We shall say that G *satisfies a product theorem* if a partial converse to this observation holds: namely that every subset of slow growth is controlled by a coset of a subgroup. To capture this property of groups, we make an auxiliary definition.

Definition 2.2.20. *Let G be a finite group. Let $C > 0$ and let $\beta : (0, \infty) \to (0, \infty)$. We say that G* satisfies the product theorem for sets of small tripling with respect to (C, β) *if, for all $\epsilon > 0$ and $A \subseteq G$ symmetric satisfying:*

(i) $|A| < |G|^{1-\epsilon}$;

(ii) For all $g \in G$ and $H \lneq G$,

$$|gH \cap A|/|A| < |G : H|^{-\epsilon}|G|^{\beta(\epsilon)},$$

then $|A \cdot A \cdot A| \geq C|A|^{1+\beta(\epsilon)}$.

Lemma 2.2.21. *Let G be a finite group which satisfies the product theorem for sets of small tripling with respect to (C_1, β). For all $\epsilon > 0$ there exists $\delta(C_1, \beta, \epsilon) > 0$ such that for μ, ν probability measures on G, if:*

(i) $\|\mu\|_2 > |G|^{-\frac{1}{2}+\epsilon}$;

(ii) For all $g \in G$ and $H \lneq G$, $\mu(gH) < |G : H|^{-\epsilon}$

then $\|\mu * \nu\|_2 \leq \|\mu\|_2^{\frac{1}{2}+\delta}\|\nu\|_2^{\frac{1}{2}}$, provided $|G|$ is sufficiently large, depending on C_1, β.

Proof. Suppose (for a contradiction) that there exists $\epsilon > 0$ such that for all $\delta > 0$ there exist probability measures μ, ν on G such that:

(i) $\|\mu\|_2 > |G|^{-\frac{1}{2}+\epsilon}$;

(ii) For all $g \in G$ and $H \lneq G$, $\mu(gH) < |G : H|^{-\epsilon}$;

(iii) $\|\mu * \nu\|_2 > \|\mu\|_2^{\frac{1}{2}+\delta}\|\nu\|_2^{\frac{1}{2}}$.

By (iii), Theorem 2.2.17 is applicable with $K = \|\mu\|_2^{-\delta}$. Let A be as in Theorem 2.2.17. First, by Theorem 2.2.17 (i),

$$|A| \leq CK^C/\|\mu\|_2^2 = C\|\mu\|_2^{-2-\delta C} < C|G|^{1-2\epsilon+\delta(\frac{C}{2}-\epsilon C)}$$

(by (i) above). Second, for all $g \in G$ and $H \lneq G$,

$$\begin{aligned}
|gH \cap A|/|A| &\leq CK^C|gH \cap A| \min_{a\in A}(\tilde{\mu} * \mu)(a) \text{ (by Theorem 2.2.17 (iii))} \\
&\leq CK^C(\tilde{\mu} * \mu)(gH) \\
&\leq CK^C \max_{g'\in G} \mu(g'H) \\
&\leq CK^C|G : H|^{-\epsilon} \text{ (by (ii) above)} \\
&= C\|\mu\|_2^{-\delta C}|G : H|^{-\epsilon} \text{ (by choice of } K) \\
&< C|G|^{\frac{\delta C}{2}}|G : H|^{-\epsilon} \text{ (by (i) above).}
\end{aligned}$$

Taking δ sufficiently small that $\delta C/2 < \beta(\epsilon)$, we satisfy the hypotheses of Definition 2.2.20. Therefore:

$$|A \cdot A \cdot A| \geq C_1|A|^{1+\beta(\epsilon)}.$$

But by Theorem 2.2.17 (ii), A has tripling at most $C\|\mu\|_2^{-C\delta}$, so taking δ sufficiently small, we have the required contradiction. □

We summarise one of the paths to expansion explored in this section in the following theorem. This is the version of the Bourgain-Gamburd machine which is applied in [77], and is also the version of the machine from which our constructions of expanders in Chapter 4 shall follow.

Theorem 2.2.22. *Let G be a finite group; $S \subseteq G$ a symmetric subset. Suppose:*

(i) (Quasirandomness) There exist $C_1, \alpha > 0$ such that, for some finite groups $G_1, \ldots, G_n$, $G = \prod_{i=1}^n G_i$ and G_i is (C_1, α)-quasirandom for every i.

(ii) (Product theorem) There exist $C_2 > 0$ and $\beta : (0, \infty) \to (0, \infty)$ such that G satisfies the product theorem for sets of small tripling (Definition 2.2.20) with respect to (C_2, β).

(iii) (Non-concentration) There exists $\gamma > 0$ and $C_3 > 0$ such that for some $l \leq C_3 \log|G|$ and every $H \lneq G$,

$$\mu_S^{(2l)}(H) \leq |G : H|^{-\gamma}.$$

Then there exists $\epsilon > 0$ (depending on $\alpha, \beta(\cdot), \gamma, C_1, C_2, C_3, |S|$) such that (G, S) is a two-sided ϵ-expander, provided the $|G_i|$ are sufficiently large depending on these constants.

Proof. By Proposition 2.2.10 and condition (i), it suffices to check that (G, S) satisfies the ℓ^2-flattening condition with respect to some C.

Fix $\epsilon > 0$. Applying (iii) to $H = 1$, we have $\mu_S^{(2l_0)}(1) \leq |G|^{-\gamma}$ for some $l_0 \leq C_0 \log|G|$. By (ii) and (iii), Lemma 2.2.21 is applicable, assuming $\|\mu_S^{(l_0)}\|_2 > |G|^{-\frac{1}{2}+\frac{\epsilon}{2}}$ (if at any point in what follows this assumption becomes invalid, then ℓ^2-flattening is already satisfied, as $\mu_S^{(2l)}(1) = \|\mu_S^{(l)}\|_2^2$ by Remark 2.1.1 (ii)).

Let $\delta > 0$ be as in Lemma 2.2.21. Applying the conclusion of Lemma 2.2.21 with $\mu = \nu = \mu_S^{(l_0)}$, we have:

$$\|\mu_S^{(2l_0)}\|_2 \leq \|\mu_S^{(l_0)}\|_2^{(1+\delta)}.$$

Replacing l_0 by $2^k l_0$ and iteratively applying this last inequality (again, assuming that ℓ^2-flattening is not already satisfied, so that Lemma 2.2.21 continues to apply),

$$\|\mu_S^{(2^k l_0)}\|_2 \leq \|\mu_S^{(l_0)}\|_2^{(1+\delta)^k} \leq |G|^{-\frac{\gamma}{2}(1+\delta)^k}.$$

Taking $k \geq (\log(\frac{1}{2} - \frac{\epsilon}{2}) - \log\frac{\gamma}{2})/\log(1+\delta)$, it follows that $\mu_S^{(2l)}(1) \leq |G|^{\epsilon-1}$ for some $l \leq 2^k l_0 \leq 2^k C_0 \log|G|$. Hence we satisfy the ℓ^2-flattening condition for some function $C(\epsilon) \leq 2^k C_0$, and expansion follows. □

Remark 2.2.23. *For $H \leq G$, and $\phi \in \ell^2(G)$, define $\overline{\phi} \in \ell^2(G/H)$ by:*

$$\overline{\phi}(gH) = \sum_{h \in H} \phi(gh).$$

Then since S is symmetric, for $l \in \mathbb{N}$,

$$\mu_S^{(2l)}(H) = \|\overline{\mu_S^{(l)}}\|_2^2.$$

Now define $\overline{A_S} : \ell^2(G/H) \to \ell^2(G/H)$ *by:*

$$\overline{A_S}(F)(gH) = \frac{1}{|S|} \sum_{s \in S} F(sgH).$$

Then $\overline{A_S}$ *is a linear operator; it is a contraction (being the adjacency operator on the Schreier graph of* $(G/H, S)$*) and satisfies, for* $\phi \in \ell^2(G)$*,*

$$\overline{A_S}(\overline{\phi}) = \overline{\mu_S * \phi}.$$

It follows that $\mu_S^{(2l)}(H)$ *is a decreasing function of* l*. Hypothesis (iii) of Theorem 2.2.22 therefore follows from an apparently weaker variant, in which our* $l \leq C_3 \log|G|$ *is permitted to depend on the subgroup* H*.*

2.3 Analytic Pro-p Groups

2.3.1 Pro-p groups

In this section we recall a (very) few well-known properties of pro-p groups. All proofs may be found in [30] Chapter 1.

Definition 2.3.1. *A topological group* G *is* profinite *if it is compact Hausdorff and has a basis of neighbourhoods of the identity consisting of open subgroups. For* p *a prime,* G *is* pro-p *if it is profinite and every open subgroup has index a power of* p*.*

It should be noted that every open subgroup in a profinite group G has finite index, so that every open subgroup contains an open normal subgroup of G. This means that we may always take a neighbourhood basis at the identity consisting of open normal subgroups.

Proposition 2.3.2. *Let* G *be a profinite group and let* $\mathcal{N}$ *be a family of open normal subgroups forming a neighbourhood basis at the identity. Then* G *is isomorphic to the inverse limit*

$$\varprojlim_{N \in \mathcal{N}} G/N$$

equipped with the Tychonoff topology, and where $\mathcal{N}$ *is ordered by reverse-inclusion. Conversely, the inverse limit of any inverse system of finite discrete groups (respectively finite discrete p-groups) is profinite (respectively pro-p).*

Definition 2.3.3. *Let G be a topological group. A* (topological) generating set *for G is a subset $X \subseteq G$ such that the subgroup $\langle X \rangle$ of G generated (abstractly) by X is dense in G. G is* (topologically) finitely generated *if G has a finite (topological) generating set.*

Theorem 2.3.4 (Serre)**.** *Let G be a finitely generated pro-p group. The every subgroup of G of finite index is open in G and the derived subgroup $[G, G]$ is closed in G.*

Corollary 2.3.5. *Every abstract homomorphism from a finitely generated pro-p group to a profinite group is continuous.*

These results were later extended to finitely generated *profinite* groups in the celebrated work of Nikolov and Segal.

Definition 2.3.6. *Let G be a pro-p group. The* lower central p-series *of G is the sequence $(G_n)_n$ defined recursively by:*

$$G_1 = G;\ G_{n+1} = \overline{G_n^p [G_n, G]}.$$

Proposition 2.3.7. *If G is a finitely generated pro-p group then G_n is open in G for every n, and $(G_n)_n$ forms a neighbourhood basis at the identity in G. In particular:*

$$G \cong \varprojlim_{n \geq 1} G/G_n.$$

Finally we explain the concept of a p-adic power in a pro-p group.

Lemma 2.3.8. *Let G be a pro-p group, let $g \in G$ and let $(a_n)_n, (b_n)_n \subseteq \mathbb{Z}$ be sequences which converge in $\mathbb{Z}_p$ to the same limit. Then $(g^{a_n})_n, (g^{b_n})_n$ converge in G to the same limit.*

We may therefore well-define:

Definition 2.3.9. *Let G be a pro-p group, let $g \in G$ and let $\lambda \in \mathbb{Z}_p$. Then:*

$$g^\lambda = \lim_{n \to \infty} g^{a_n}$$

where $(a_n)_n \subseteq \mathbb{Z}$ is a sequence converging in $\mathbb{Z}_p$ to λ.

Proposition 2.3.10. *Let G be a pro-p group, let $g, h \in G$ and let $\lambda, \mu \in \mathbb{Z}_p$. Then:*

(i) $g^{\lambda+\mu} = g^\lambda g^\mu$;

(ii) $g^{\lambda\mu} = (g^\lambda)^\mu$;

(iii) If $gh = hg$ *then* $(gh)^\lambda = g^\lambda h^\lambda$;

(iv) The map $\nu \mapsto g^\nu$ *defines a (continuous) homomorphism of* $\mathbb{Z}_p$ *onto* $\overline{\langle g \rangle}$.

2.3.2 Pro-p domains

Fix p prime. Let R be a commutative unital Noetherian ring. Recall that R is called a *local ring* if R has a unique non-zero maximal ideal $\mathcal{M}$ (we shall refer to *the local ring* $(R, \mathcal{M})$). The quotient $R/\mathcal{M}$ is called the *residue field* of R.

Notation 2.3.11. *Following [30], here and in Sections 2.3.3; 2.3.4 and 2.4 we adopt the convention that for R a local ring; $\mathcal{I} \lhd R$ and $n \in \mathbb{N}$, $\mathcal{I}^n$ shall denote the nth power of $\mathcal{I}$ (an ideal in R), while $\mathcal{I}^{(n)}$ shall denote the n-fold Cartesian product of $\mathcal{I}$.*

There is a topology on R, called the *$\mathcal{M}$-adic topology*, induced by declaring the filtration $(\mathcal{M}^n)_n$ to be a basis for the neighbourhoods of 0.

Definition 2.3.12. *The local ring $(R, \mathcal{M})$ is called a* pro-p ring *if:*

(i) The residue field of R is finite of characteristic p;

(ii) R is complete with respect to the $\mathcal{M}$-adic topology.

A pro-p ring $(R, \mathcal{M})$ where $\mathcal{M}$ is principal is called a discrete valuation pro-p ring *and a pro-p ring which is an integral domain will be called a* pro-p domain.

Theorem 2.3.13 (Krull Intersection Theorem)**.** *If $(R, \mathcal{M})$ is a pro-p domain, then $\bigcap_{i=1}^{\infty} \mathcal{M}^i = \{0\}$, so that the $\mathcal{M}$-adic topology on R is Hausdorff.*

Let $\mathbb{K}$ be the field of fractions of R. Fix $c \in (0, 1)$ and define a norm $\|\cdot\|$ on R (compatible with the $\mathcal{M}$-adic topology) by:

$$\|a\| = c^n \text{ for } a \in \mathcal{M}^n \setminus \mathcal{M}^{n+1};\ \|0\| = 0.$$

In particular if $(R, \mathcal{M})$ is a discrete valuation ring, with $\mathcal{P} \in \mathcal{M}$ such that $\mathcal{M} = (\mathcal{P})$, then $\|\mathcal{P}\| = c$. In this case, we extend $\|\cdot\|$ to $\mathbb{K}$ via:

$$\|a\| = \|a\mathcal{P}^n\|c^{-n} \text{ for } n \text{ sufficiently large that } a\mathcal{P}^n \in R.$$

Example 2.3.14. *$\mathbb{Z}_p$ and $\mathbb{F}_q[[t]]$ (for q a power of p) are discrete valuation pro-p domains, with maximal ideals (p) and (t), respectively.*

Theorem 2.3.15 (Cohen [23])**.** *Every discrete valuation pro-p domain is a finitely generated free module over a subring of the form $\mathbb{Z}_p$ and $\mathbb{F}_p[[t]]$.*

2.3.3 R-analytic groups

In this section $(R, \mathcal{M})$ is a discrete valuation pro-p domain, with $\mathcal{M}$ generated by $\mathcal{P} \in \mathcal{M}$. For proofs of results quoted, refer to Chapter 13 of [30].

Definition 2.3.16. *Write* $\underline{X} = (X_1, \ldots, X_d)$, $\underline{Y} = (Y_1, \ldots, Y_d)$. *Denote by* $R[[\underline{X}, \underline{Y}]]$ *the ring of formal non-commuting power series in the* $2d$ *variables* $X_1, \ldots, X_d$, $Y_1, \ldots, Y_d$. *For* $i = 1, \ldots, d$, *let* $F_i(\underline{X}, \underline{Y}) \in R[[\underline{X}, \underline{Y}]]$. *Then* $\underline{F} = (F_1, \ldots, F_d)$ *is a* formal group law, *of dimension* d *over* R, *if:*

(i) $\underline{F}(\underline{X}, \underline{0}) = \underline{X}$ *and* $\underline{F}(\underline{0}, \underline{Y}) = \underline{Y}$;

(ii) $\underline{F}(\underline{X}, \underline{F}(\underline{Y}, \underline{Z})) = \underline{F}(\underline{F}(\underline{X}, \underline{Y}), \underline{Z})$.

Proposition 2.3.17 (13.16 in [30])**.** *Let* $\underline{F}$ *be a formal group law. There exist power series* $\underline{B}(\underline{X}, \underline{Y})$, $\underline{I}(\underline{X})$, $\underline{O}(\underline{X}, \underline{Y})$, $\underline{P}(\underline{X})$, $\underline{Q}(\underline{X}, \underline{Y})$, *with* $\underline{B}$ *bilinear in* $\underline{X}$ *and* $\underline{Y}$; *every term of* $\underline{O}, \underline{P}, \underline{Q}$ *having total degree at least* 3 *and every term of* $\underline{O}, \underline{Q}$ *having degree at least* 1 *in each of* $\underline{X}, \underline{Y}$, *such that:*

(i) $\underline{F}(\underline{X}, \underline{Y}) = \underline{X} + \underline{Y} + \underline{B}(\underline{X}, \underline{Y}) + \underline{O}(\underline{X}, \underline{Y})$;

(ii) $\underline{I}(\underline{X}) = -\underline{X} + \underline{B}(\underline{X}, \underline{X}) + \underline{P}(\underline{X})$ *and* $\underline{F}(\underline{X}, \underline{I}(\underline{X})) = \underline{0} = \underline{F}(\underline{I}(\underline{X}), \underline{X})$;

(iii) $\underline{F}((\underline{I} \circ \underline{F})(\underline{Y}, \underline{X}), \underline{F}(\underline{X}, \underline{Y})) = \underline{B}(\underline{X}, \underline{Y}) - \underline{B}(\underline{Y}, \underline{X}) + \underline{Q}(\underline{X}, \underline{Y})$.

Definition 2.3.18. *An* R-standard group *of dimension* d *is a topological group* $(G, \cdot)$ *with underlying space* $G = \mathcal{M}^{(d)}$ *such that there exists a formal group law* $\underline{F}$ *of dimension* d *such that, for all* $g, h \in G$,

$$g \cdot h = \underline{F}(g, h).$$

Note that, for $\underline{B}, \underline{I}, \underline{Q}$ *as in Proposition 2.3.17, we have:*

$$g^{-1} = \underline{I}(g),\ [g, h] = \underline{B}(g, h) - \underline{B}(h, g) + \underline{Q}(g, h).$$

Recall that for R a ring, $\mathbb{M}_d(R)$ is the ring of d-by-d matrices over R.

Example 2.3.19. *(i)* $(\mathcal{M}^{(d)}, +)$ *is an* R*-standard group of dimension* d.

(ii) Let $\mathrm{GL}_d^1(R) = I_d + \mathcal{P}\mathbb{M}_d(R)$. *Then* $\mathrm{GL}_d^1(R) \leq \mathrm{GL}_d(R)$ *and, identifying* $\mathrm{GL}_d^1(R)$ *with* $\mathcal{M}^{(d^2)}$ *in the obvious way, multiplication in* $\mathrm{GL}_d^1(R)$ *is given by a formal group law of dimension* d^2.

(iii) *Let* $\mathrm{SL}_d^1(R) = \mathrm{SL}_d(R) \cap \mathrm{GL}_d^1(R)$ *be the kernel of the congruence map* $\mathrm{SL}_d(R) \twoheadrightarrow \mathrm{SL}_d(R/\mathcal{M})$. *Then we may identify* $\mathrm{SL}_d^1(R)$ *with* $\mathcal{M}^{(d^2-1)}$ *via* $A \mapsto ((A - I_d)_{i,j})_{(i,j)\neq(d,d)}$ *(since these* $d^2 - 1$ *co-ordinates together with the determinant condition uniquely determine* $A_{d,d}$*). Under this identification, multiplication in* $\mathrm{SL}_d^1(R)$ *is given by a formal group law of dimension* $d^2 - 1$.

Proposition 2.3.20 (13.22 in [30])**.** *For* $n \in \mathbb{N}$ *define* $K_n = (\mathcal{M}^n)^{(d)} \subseteq G$. *Then for all* $n, m \in \mathbb{N}$*:*

(i) $K_n \lhd_o K_1 = G$*;*

(ii) $[K_n, K_m] \subseteq K_{n+m}$*;*

(iii) If $m \leq n$, K_n/K_{n+m} *is isomorphic to the additive group* $(\mathcal{M}^n/\mathcal{M}^{n+m})^{(d)}$*;*

(iv) $G \cong \varprojlim G/K_n$ *is a pro-p group.*

R-standard groups arise as subgroups of *R-analytic groups*, in which they may be seen as providing an especially nice chart at identity.

Definition 2.3.21. *A topological group* G *is a* R-analytic group *if* G *has the structure of an* R*-analytic manifold such that the functions:*

(i) $f : G \times G \to G$ *given by* $(x, y) \mapsto xy$*,*

(ii) $i : G \to G$ *given by* $x \mapsto x^{-1}$

are analytic.

Of course, this definition begs the further questions of how we define the concepts of a *R-analytic manifold structure* and an *analytic function*. We do not answer these questions here (and instead refer the interested reader to [30]).

Theorem 2.3.22 (13.20 in [30])**.** *Let* G *be an* R*-analytic group. Then* G *has an open* R*-standard subgroup.*

Proposition 2.3.23 (13.24 in [30])**.** *For* $v, w \in \mathcal{M}^{(d)}$*, define:*

$$(v, w) = \underline{B}(v, w) - \underline{B}(w, v).$$

Then $L(G) = (\mathcal{M}^{(d)}, +, (\cdot, \cdot))$ *is a* R*-Lie ring. That is,* $(\cdot, \cdot)$ *satisfies the Jacobi identity (and is obviously* R*-bilinear antisymmetric).*

Remark 2.3.24. *For each n, $\mathcal{P}^n L(G)$ is a Lie subring of $L(G)$. As a set it is equal to K_{n+1}. Moreover by Proposition 2.3.17, the additive cosets of $\mathcal{P}^n L(G)$ in $L(G)$ are the same as the multiplicative cosets of K_{n+1} in G.*

Definition 2.3.25. *The* Lie algebra *of G is $\mathcal{L}_G = L(G) \bigotimes_R \mathbb{K}$, where $\mathbb{K}$ is the field of fractions of R.*

Example 2.3.26. *(i) For $G = (\mathcal{M}^{(d)}, +)$, $\mathcal{L}_G$ is the d-dimensional abelian $\mathbb{K}$-Lie algebra.*

(ii) $\mathcal{L}_{\mathrm{GL}_d^1(R)} = \mathfrak{gl}_d(\mathbb{K})$.

(iii) $\mathcal{L}_{\mathrm{SL}_d^1(R)} = \mathfrak{sl}_d(\mathbb{K})$.

2.3.4 p-adic analytic groups

The most familiar definition of a p-adic analytic group is the specific case of Definition 2.3.21 for which $R = \mathbb{Z}_p$. Once again, we do not define the notions of a *p-adic analytic manifold structure* and an *analytic function* here, as Definition 2.3.21 is intended only to fix our terminology; we shall not make much direct use of it.

A second characterisation of p-adic analytic groups is available, which is useful for visualising the class and producing examples:

Theorem 2.3.27 (5.2, 7.19 and 8.34 in [30])**.** *A topological group G is p-adic analytic if and only if G has an open pro-p subgroup H such that H embeds as a closed subgroup of $\mathrm{GL}_d(\mathbb{Z}_p)$, for some d.*

A third characterization, utilising the concept of a pro-p group of *finite rank*, elucidates some of the subgroup structure of p-adic analytic groups:

Definition 2.3.28. *For G a pro-p group, the* rank $r(G)$ *is given by:*

$$r(H) = \sup\{d(K) : K \leq_c H\}.$$

Theorem 2.3.29 (8.33 in [30])**.** *A topological group G is p-adic analytic if and only if G has an open pro-p subgroup H such that $r(H)$ is finite.*

However, the characterization which will be of most use to us, and which will provide an alternative approach to constructing the Lie algebra of G, is based on the concept of a *uniform* subgroup. Uniform pro-p groups are close relatives of $\mathbb{Z}_p$-*standard* groups (which are just the special case of the definitions and results in Section 2.3.3 for $R = \mathbb{Z}_p$), but will be more convenient for our purposes.

Let $p \geq 3$ be prime; G be a pro-p group and $(G_n)_n$ be its lower central p-series.

Definition 2.3.30. *G is* powerful *if $G/\overline{G^p}$ is abelian. G is* uniform *if it is finitely generated, powerful and torsion-free. The* dimension *of a uniform group G is the minimal size of a topological generating set.*

The relationship between powerful groups and p-adic analytic groups was first described by Lazard [52].

Theorem 2.3.31 (3.13 in [30])**.** *Let G be a pro-p group. Then $r(G)$ is finite if and only if G is finitely generated and G has an open powerful characteristic subgroup.*

Theorem 2.3.32 (4.2 in [30])**.** *Let G be a finitely generated powerful pro-p group. Then G_n is uniform for all sufficiently large n.*

Theorem 2.3.33 (4.6 in [30])**.** *If G is a pro-p group and H, K are open uniform subgroups of G, then H and K have the same dimension.*

By Theorems 2.3.29, 2.3.31 and 2.3.32 a topological group G is p-adic analytic if and only if it has an open uniform pro-p subgroup. Indeed, for G compact we can say more: letting $H \leq G$ be as in Theorem 2.3.29, $|G : H|$ is finite, so H contains an open subgroup K which is characteristic in G. K is still pro-p of finite rank, so by Theorem 2.3.31 contains an open characteristic powerful subgroup L. Finally, for n sufficiently large, L_n is uniform by Theorem 2.3.32. Recalling that the terms of the lower central p-series are characteristic subgroups, we deduce:

Corollary 2.3.34. *A compact topological group G is p-adic analytic if and only if G is a profinite group with an open characteristic uniform pro-p subgroup.*

Example 2.3.35. *Every $\mathbb{Z}_p$-standard group of dimension d is a uniform pro-p group of dimension d (8.31 of [30]). Conversely, if G is a d-dimensional uniform pro-p group, then G_2 is a d-dimensional $\mathbb{Z}_p$-standard group (8.23 (iii) of [30]). In particular, every compact p-adic analytic group has an open characteristic $\mathbb{Z}_p$-standard subgroup. We describe the formal group law on G_2 below (Remark 2.3.40).*

We recall some properties of uniform groups. Unless otherwise specified, let G be a d-dimensional uniform pro-p group.

Theorem 2.3.36 (3.6, 4.9 in [30])**.** *Let $\{a_1, \ldots, a_d\}$ be a topological generating set for G; $n, m \in \mathbb{N}$.*

(i) $(\lambda_1, \ldots, \lambda_d) \mapsto a_1^{\lambda_1} \cdots a_d^{\lambda_d}$ *defines a homeomorphism $\mathbb{Z}_p^d \to G$.*

(ii) G_{n+1} *is uniform of dimension* d.

(iii) $(G_{n+1})_{m+1} = G_{m+n+1}$.

(iv) $G_{n+1} = \{x^{p^n} : x \in G\}$.

(v) $\{a_1^{p^n}, \ldots, a_d^{p^n}\}$ *is a topological generating set for* G_{n+1}.

There is a complete normed $\mathbb{Q}_p$-algebra $\hat{A}$, an embedding $G \hookrightarrow \hat{A}^*$ satisfying:

$$\forall g \in G, g - 1 \in \hat{A}_0, \text{ where } \hat{A}_0 = \{x \in \hat{A} : \|x\| \leq p^{-1}\}$$

and mutually inverse analytic functions:

$$\log : 1 + \hat{A}_0 \to \hat{A}_0,$$
$$\exp : \hat{A}_0 \to 1 + \hat{A}_0.$$

$\hat{A}$ is naturally a $\mathbb{Q}_p$-Lie algebra with Lie bracket:

$$(x, y) = xy - yx.$$

$\log(G)$ is a free d-dimensional $\mathbb{Z}_p$-module and a $\mathbb{Z}_p$-Lie subalgebra of $\hat{A}$.

Lemma 2.3.37 (6.25 and 7.12 from [30])**.** *Let* $x \in \hat{A}_0$, $n \in \mathbb{Z}$.

(i) $\exp(nx) = \exp(x)^n$.

(ii) $\log((1 + x)^n) = n \log(1 + x)$.

(iii) $(\log(G), \log(G)) \subseteq p \log(G)$.

Moreover, for $g \in G$, $\lambda \in \mathbb{Z}_p$, $\lambda \log(g) = \log(g^\lambda)$.

Combining this Lemma with Theorem 2.3.36 (iv), we have:

Corollary 2.3.38. *For all* $n \in \mathbb{N}$, $p^n \log(G) = \log(G_{n+1})$.

Proposition 2.3.39 (6.27 and 6.28 in [30])**.** *There are formal non-commutative power series* $\Phi(X, Y)$, $\Psi(X, Y)$ *satisfying:*

$$\Phi(X, Y) = X + Y + \tfrac{1}{2}(XY - YX) + h.o.(X, Y)$$
$$\Psi(X, Y) = (XY - YX) + h.o.(X, Y)$$

(with $h.o.(X, Y)$ *denoting terms composed of brackets of length at least three) such that, for* $x, y \in \hat{A}_0$,

(i) $\Phi(x,y)$ *converges to* $\log(\exp(x)\exp(y))$*;*

(ii) $\Psi(x,y)$ *converges to* $\log(\exp(-x)\exp(-y)\exp(x)\exp(y))$*.*

Remark 2.3.40. *Let* $x_1,\dots,x_d \in \log(G)$ *be a* $\mathbb{Z}_p$*-basis for* $\log(G)$*. Identify* $\mathbb{Z}_p^{(d)}$ *with* $\log(G)$ *via:*

$$\theta : (\alpha_i)_{i=1}^d \mapsto \sum_{i=1}^d \alpha_i x_i.$$

Then, identifying $\mathbb{Z}_p^{(d)}$ *with* G *via* $\exp\circ\theta$*, multiplication in* G *corresponds to the formal group law:*

$$(\underline{a},\underline{b}) \mapsto \theta^{-1}(\Phi(\theta(\underline{a}),\theta(\underline{b})))$$

on $\mathbb{Z}_p^{(d)}$*. Moreover, under this identification the subgroup* G_{n+1} *corresponds to* $p^n\log(G) = \theta((p^n\mathbb{Z}_p)^{(d)})$*, by Corollary 2.3.38. In particular,* $G_2 \cong (p\mathbb{Z}_p)^{(d)}$ *is a* $\mathbb{Z}_p$*-standard subgroup.*

Proposition 2.3.41 (4.8 and 4.31 in [30])**.** *Let* H *be a uniform closed subgroup of* G*;* $N \lhd G$ *be closed such that* G/N *is uniform.*

(i) $\log(H)$ *is a* $\mathbb{Z}_p$*-subalgebra of* $\log(G)$*.*

(ii) N *is uniform, with* $\dim(N) = \dim(G) - \dim(G/N)$*.*

(iii) $\log(N)$ *is an ideal in* $\log(G)$*, and* $\log(G/N) \cong \log(G)/\log(N)$*.*

Proposition 2.3.42 (7.15 in [30])**.** *Let* S *be a* $\mathbb{Z}_p$*-Lie subalgebra of* $\log(G)$ *such that the* $\mathbb{Z}_p$*-module* $\log(G)/S$ *is torsion-free.*

(i) $\exp(S)$ *is a closed uniform subgroup of* G*.*

(ii) If S *is an ideal of* $\log(G)$*, then* $\exp(S) \lhd G$ *and* $G/\exp(S)$ *is uniform.*

We may define $\mathcal{L}_G = \mathrm{span}_{\mathbb{Q}_p}(\log(G))$, a d-dimensional $\mathbb{Q}_p$-Lie algebra. By Remark 2.3.40, this is isomorphic to the Lie algebra described in Definition 2.3.25.

2.4 Chevalley Groups

Unless otherwise stated, proofs of assertions left unproven in this section may be found in [22].

Let Φ be a root system of type $X_l \in \{A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4, G_2\}$, $\Pi \subseteq \Phi$ be a fundamental system of roots. In a moment, we shall describe the *Lie ring of type* X_l over a ring. This shall be done by specifying a basis and expressing the brackets between basis vectors as explicit linear combinations of the basis. First however we should indicate how to construct the coefficients appearing in these linear combinations. We give the construction in full, not because its precise features will be required in what follows (they won't) but to reassure the reader that a suitable choice of coefficients exists and may be explicitly constructed, since to do so is a little delicate and involves a certain amount of "definition-chasing".

Definition 2.4.1. *A* set of structure constants of type X_l *is a set of integers* $\{N_{\alpha,\beta}\}_{\alpha,\beta\in\Phi}$ *such that, for all* $\alpha, \beta, \gamma, \delta \in \Phi$*:*

(0) *If* $\alpha + \beta \notin \Phi$ *then* $N_{\alpha,\beta} = 0$*, and if* $\alpha + \beta \in \Phi$ *then* $N_{\alpha,\beta} \in \{\pm(p+1)\}$*, where* $p \in \mathbb{N}$ *is maximal such that* $\beta - p\alpha \in \Phi$*;*

(i) $N_{\alpha,\beta} = -N_{\beta,\alpha}$*;*

(ii) $\frac{N_{\alpha,\beta}}{(\gamma,\gamma)} = \frac{N_{\beta,\gamma}}{(\alpha,\alpha)} = \frac{N_{\gamma,\alpha}}{(\beta,\beta)}$*, provided* $\alpha + \beta + \gamma = 0$*;*

(iii) $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$*;*

(iv) $\frac{N_{\alpha,\beta}N_{\gamma,\delta}}{(\alpha+\beta,\alpha+\beta)} = \frac{N_{\beta,\gamma}N_{\alpha,\delta}}{(\beta+\gamma,\beta+\gamma)} = \frac{N_{\gamma,\alpha}N_{\beta,\delta}}{(\gamma+\alpha,\gamma+\alpha)}$*, provided no two of* $\alpha, \beta, \gamma, \delta$ *are opposite, and* $\alpha + \beta + \gamma + \delta = 0$*.*

To describe the possible set of structure constants, we make an auxiliary definition.

Definition 2.4.2. *Fix a total order* $\prec$ *on the space spanned by* Φ*. A pair* $(\alpha, \beta) \in \Phi^2$ *is* extraspecial *if:*

(i) $\alpha + \beta \in \Phi$*;*

(ii) $0 \prec \alpha \prec \beta$*;*

(iii) *If* $(\alpha', \beta') \in \Phi^2$ *is another pair satisfying (i), (ii) and such that* $\alpha + \beta = \alpha' + \beta'$ *then* $\alpha \preccurlyeq \alpha'$*.*

Proposition 2.4.3. *Let $\Psi \subseteq \Phi^2$ be the set of extraspecial pairs. For every function $F : \Psi \to \{\pm 1\}$ there exists a unique set of structure constants $\{N_{\alpha,\beta}\}_{\alpha,\beta\in\Phi}$ of type X_l satisfying:*

$$N_{\alpha,\beta} = F(\alpha,\beta)(q+1),$$

for every $(\alpha,\beta) \in \Psi$, where $q \in \mathbb{N}$ is maximal such that $\beta - q\alpha \in \Phi$.

In other words, sets of structure constants are exactly parametrised by their values at the extraspecial pairs. Now let S be a commutative unital ring.

Definition 2.4.4. *Let $\{N_{\alpha,\beta}\}_{\alpha,\beta\in\Phi}$ be a set of structure constants of type X_l. The* Lie ring of type X_l over S, *denoted $\mathcal{L}_S(X_l)$, is the free S-module on the basis:*

$$\{E_\alpha\}_{\alpha\in\Phi} \sqcup \{H_\beta\}_{\beta\in\Pi}$$

with defining Lie brackets:

(i) $(H_\alpha, H_\beta) = 0$;

(ii) $(H_\alpha, E_\beta) = A_{\alpha,\beta}E_\beta$;

(iii) $(E_\alpha, E_{-\alpha}) = H_\alpha$;

(iv) $(E_\alpha, E_\beta) = 0$ *for* $\alpha+\beta \notin \Phi$;

(v) $(E_\alpha, E_\beta) = N_{\alpha,\beta}E_{\alpha+\beta}$ *for* $\alpha+\beta \in \Phi$, *where* $A_{\alpha,\beta} = \frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ *is the* Cartan integer.

Proposition 2.4.5. *For fixed X_l, $\mathcal{L}_S(X_l)$ is independent (up to isomorphism) of the choice of structure constants $\{N_{\alpha,\beta}\}_{\alpha,\beta\in\Phi}$ in Definition 2.4.4.*

Proposition 2.4.6. *If K is a field then exactly one of the following holds:*

(i) $\mathcal{L}_K(X_l)$ *is a perfect Lie algebra,*

(ii) $\mathrm{char}(K) = 2$ *and* $X_l = A_1$ *or* C_l.

Definition 2.4.7. *We define the* adjoint Chevalley group of type X_l over S *to be the subgroup $\mathcal{G}_S^{\mathrm{ad}}(X_l) \leq \mathrm{Aut}_S(\mathcal{L}_S(X_l))$ generated by the set $\{\rho(x_\alpha(t))\}_{\alpha\in\Phi;t\in S}$, defined by:*

(i) $\rho(x_\alpha(t))(E_\alpha) = E_\alpha$;

(ii) $\rho(x_\alpha(t))(E_{-\alpha}) = E_{-\alpha} + tH_\alpha - t^2E_\alpha$;

(iii) $\rho(x_\alpha(t))(H_\alpha) = H_\alpha - 2tE_\alpha$;

(iv) $\rho(x_\alpha(t))(H_\beta) = H_\beta - A_{\beta,\alpha}tE_\alpha$, *where* $A_{\beta,\alpha} = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$ *is the Cartan integer;*

(v) $\rho(x_\alpha(t))(E_\beta) = E_\beta + \sum_{i=1}^{q} M_{\alpha,\beta,i}t^iE_{i\alpha+\beta}$, *where* $M_{\alpha,\beta,i} = \frac{1}{i!}\prod_{j=0}^{i-1} N_{\alpha,j\alpha+\beta}$ *and* q *is maximal such that* $q\alpha + \beta \in \Phi$,

for any $\alpha, \beta \in \Phi$ *linearly independent and* $t \in S$.

Remark 2.4.8. *Letting* $p \in \mathbb{N}$ *be maximal such that* $\beta - p\alpha \in \Phi$, *we have* $N_{\alpha,j\alpha+\beta} \in \{\pm(p+j+1)\}$, *so that:*

$$M_{\alpha,\beta,i} \in \{\pm \begin{pmatrix} p+i \\ i \end{pmatrix}\} \subseteq \mathbb{Z}.$$

Specifically, the identity (v) in the above definition makes sense for arbitrary S.

Proposition 2.4.9. $\mathcal{G}_S^{\mathrm{ad}}(X_l)$ *preserves the Lie bracket on* $\mathcal{L}_S(X_l)$.

Proposition 2.4.10. $\mathcal{G}_S^{\mathrm{ad}}(X_l) \leq \mathrm{GL}(\mathcal{L}_S(X_l))$ *is independent of a choice of structure constants.*

It was no coincidence that we chose the (ostensibly unnecessarily complicated) notation $\rho(x_\alpha(t))$ for the generators of $\mathcal{G}_S^{\mathrm{ad}}(X_l)$. They shall lie in the image of a representation ρ on $\mathcal{L}_S(X_l)$ of another group, which we define now.

Definition 2.4.11. *We define the* universal Chevalley group of type X_l over S *to be the group* $\mathcal{G}_S(X_l)$ *abstractly generated by the symbols* $\{x_\alpha(t)\}_{\alpha\in\Phi;t\in S}$, *subject to the following* Steinberg relations*:*

(i) For any $\alpha \in \Phi$ *and any* $t_1, t_2 \in S$,

$$x_\alpha(t_1)x_\alpha(t_2) = x_\alpha(t_1 + t_2).$$

(ii) For any $\alpha, \beta \in \Phi$ *linearly independent and any* $s, t \in S$,

$$[x_\beta(s), x_\alpha(t)] = \prod_{i,j>0} x_{i\alpha+j\beta}(N_{i,j;\alpha,\beta}(-t)^i s^j),$$

where the products are ordered such that $i + j$ *is increasing; and* $N_{i,j;\alpha,\beta} \in \{\pm1, \pm2, \pm3\}$ *are defined by:*

$$N_{i,1;\alpha,\beta} = M_{\alpha,\beta,i},\ N_{1,j;\alpha,\beta} = (-1)^j M_{\alpha,\beta,j},\ N_{3,2;\alpha,\beta} = \tfrac{1}{3}M_{\alpha+\beta,\alpha,2},$$
$$N_{2,3;\alpha,\beta} = -\tfrac{2}{3}M_{\alpha+\beta,\beta,2}.$$

(iii) For any $\alpha \in \Phi$ and any $s, t \in S^*$,

$$c_\alpha(s)c_\alpha(t) = c_\alpha(st),$$

where $c_\alpha(s) = x_\alpha(s)x_{-\alpha}(-s^{-1})x_\alpha(s)(x_\alpha(1)x_{-\alpha}(-1)x_\alpha(1))^{-1}$.

Immediately from these definitions we have:

Lemma 2.4.12. *(i) If S' is a subring of S, then the inclusion of $\{x_\alpha(t)\}_{\alpha\in\Phi;t\in S'}$ into $\{x_\alpha(t)\}_{\alpha\in\Phi;t\in S}$ induces a homomorphism $\psi : \mathcal{G}_{S'}(X_l) \to \mathcal{G}_S(X_l)$ (in other words, for each Φ, every Steinberg relation over S' is also a Steinberg relation over S).*

(ii) $c_\alpha(1) = 1$.

Theorem 2.4.13 (Steinberg)**.** *Let K be a field. There is a group homomorphism $\rho : \mathcal{G}_K(X_l) \to \mathrm{GL}_d(K)$ (where $d = |\Phi| + |\Pi|$ is the dimension of $\mathcal{L}_K(X_l)$), such that $\{\rho(x_\alpha(t))\}_{\alpha\in\Phi;t\in S}$ are as in Definition 2.4.7. Moreover $\ker(\rho)$ is finite and central in $\mathcal{G}_K(X_l)$.*

Note that for S a subring of K and $\psi : \mathcal{G}_S(X_l) \to \mathcal{G}_K(X_l)$ as in Lemma 2.4.12, $\rho(\psi(\mathcal{G}_S(X_l))) = \mathcal{G}_S^{\mathrm{ad}}(X_l) \leq \mathrm{GL}_d(S)$.

Of particular interest to us shall be Chevalley groups over pro-p domains, because these acquire an analytic structure.

Theorem 2.4.14 (Exercise 13.11 in [30])**.** *Let $(R, \mathcal{M})$ be a pro-p domain. For each $n \geq 1$, let $G_n \leq \mathcal{G}_R(X_l)$ be the subgroup generated by the set:*

$$\{x_\alpha(t)\}_{\alpha\in\Phi;t\in\mathcal{M}^n} \cup \{c_\beta(1+s)\}_{\beta\in\Pi;s\in\mathcal{M}^n}.$$

(i) $G_n \lhd_f \mathcal{G}_R(X_l)$, for all $n \geq 1$.

(ii) The map $\theta_n : (\mathcal{M}^n)^{(|\Phi|+|\Pi|)} \to G_n$, given by:

$$\theta(\underline{t}) = (\prod_{\alpha\in\Phi^+} x_\alpha(t_\alpha))(\prod_{\beta\in\Pi} c_\beta(1+t_\beta))(\prod_{\alpha\in\Phi^-} x_\alpha(t_\alpha))$$

(with the products ordered by the height function induced on Φ by Π) is a bijection, for every $n \geq 1$. Identifying G_1 with $\mathcal{M}^{(|\Phi|+|\Pi|)}$ via θ_1, G_1 is an R-standard group of dimension $|\Phi| + |\Pi|$.

(iii) $\mathcal{L}_{G_1}$ is perfect, unless $p = 2$ and $X_l = A_1$ or C_l. Indeed $\mathcal{L}_{G_1} \cong \mathcal{L}_{\mathbb{K}}(X_l)$.

Now let $\psi : \mathcal{G}_R(X_l) \to \mathcal{G}_{\mathbb{K}}(X_l)$ be as described in Lemma 2.4.12 (i). It is clear from Definition 2.4.7 that:

(i) For any $\alpha \in \Phi; s, t \in R$ and $n \geq 1$, if $s \equiv t \mod \mathcal{M}^n$ then $\rho(x_\alpha(s)) \equiv \rho(x_\alpha(t)) \mod \mathcal{M}^n$.

(ii) In particular, for $t \in \mathcal{M}^n$, $\rho(x_\alpha(t)) \equiv I_d \mod \mathcal{M}^n$.

From Lemma 2.4.12 (ii), it follows that for any $\beta \in \Phi, s \in \mathcal{M}^n$, $\rho(c_\beta(1+s)) \equiv I_d \mod \mathcal{M}^n$. Thus we have:

Corollary 2.4.15. $\rho(\psi(G_n)) \leq K_n := \mathcal{G}^{\mathrm{ad}}_{\mathbb{K}}(X_l) \cap (I_d + \mathbb{M}_d(\mathcal{M}^n))$.

2.5 The Nottingham Group

Another class of finitely generated pro-p groups of note is the Nottingham groups of the finite fields. These groups have attracted considerable interest for three main reasons. First, it is relatively easy to do explicit computations in them, so they serve as a useful first testing ground for more general techniques and conjectures in pro-p group theory. Second, they arise naturally in number theory as the groups of wild automorphisms of $\mathbb{F}_q((t))$. Third, they are exotic groups in their own right, and exhibit a number of extreme properties.

The purpose of this section is to lay the groundwork for our analysis of the diameters of finite quotients of the Nottingham group in Chapter 3. Our exposition is largely based upon that appearing in [21], and we refer the reader there for further details, together with [49] and [80].

2.5.1 Definition and first properties

Fix p a prime and let $q = p^a$ be a power of p. Let $\mathbb{F}_q$ denote the finite field of order q.

Definition 2.5.1. *The* Nottingham group $\mathcal{N}_q$ *over* $\mathbb{F}_q$ *is the group with underlying set:*

$$\mathcal{N}_q = \left\{ t + \sum_{i=2}^{\infty} \lambda_i t^i : (\lambda_i)_i \in \mathbb{F}_q^{\mathbb{N}} \right\} \subseteq \mathbb{F}_q[[t]]$$

and group operation given by formal substitution of variables; that is, for $f = t + \sum_{i=2}^{\infty} \lambda_i t^i, g = t + \sum_{i=2}^{\infty} \mu_i t^i \in \mathcal{N}_q$,

$$f \cdot g = f + \sum_{i=2}^{\infty} \mu_i f^i$$
$$(= t + (\lambda_2 + \mu_2)t^2 + (\lambda_3 + 2\lambda_2\mu_2 + \mu_3)t^3 + \ldots)$$

$\mathcal{N}_q$ is a compact topological group when endowed with the degree topology inherited from $\mathbb{F}_q[[t]]$.

We define the sequence of *congruence subgroups* $(K_n)_n$ of $\mathcal{N}_q$ to be the subsets:

$$K_n = \left\{ t + \sum_{k=n+1}^{\infty} \lambda_k t^k \in \mathcal{N}_q \right\}$$

(so that in particular $K_1 = \mathcal{N}_q$).

Lemma 2.5.2. *$(K_n)_n$ satisfies:*

(i) $K_n \lhd_o \mathcal{N}_q$*;*

(ii) $|K_n : K_{n+1}| = q$ *for all* $n \in \mathbb{N}$*;*

(iii) $\mathcal{N}_q \cong \varprojlim_n \mathcal{N}_q/K_n$ *as topological groups.*

In particular, $\mathcal{N}_q$ is a pro-p group and the K_n form a neighbourhood basis at identity.

For $n \geq 1$ and $\lambda \in \mathbb{F}_q$, define:

$$e_{n,\lambda}(t) = t + \lambda t^{n+1} \in K_n.$$

Of course, for $\lambda \neq 0$, $e_{n,\lambda} \in K_n \setminus K_{n+1}$. The elements $e_{n,\lambda}$ form an infinite topological generating set for $\mathcal{N}_q$, as follows:

Lemma 2.5.3. *(i) For any $n \geq 1$ $\lambda, \mu \in \mathbb{F}_q$,*

$$e_{n,\lambda} \cdot e_{n,\mu} \equiv e_{n,\lambda+\mu} \mod K_{2n}$$

(so in particular $e_{n,\lambda}^k \equiv e_{n,k\lambda} \mod K_{2n}$ for all $k \in \mathbb{N}$).

(ii) Let $\{\gamma_1, \ldots, \gamma_a\}$ be an $\mathbb{F}_p$-basis for $\mathbb{F}_q$. Then:

$$K_n/K_{n+1} = \langle e_{n,\gamma_1} K_{n+1}, \ldots, e_{n,\gamma_a} K_{n+1} \rangle \cong C_p^a.$$

Hence:

$$\mathcal{N}_q = \{ e_{1,\lambda_1} \cdot e_{2,\lambda_2} \cdots : (\lambda_k)_k \in \mathbb{F}_q^{\mathbb{N}} \}.$$

Of especial interest to us in Chapter 3 shall be the commutator structure of $\mathcal{N}_q$, which is fortunately very well-behaved.

Lemma 2.5.4. *Let $m, n \in \mathbb{N}$.*

(i) Let $g = t + \sum_{k=n+1}^{\infty} \lambda_k t^k \in K_n \setminus K_{n+1}$, $h = t + \sum_{k=m+1}^{\infty} \mu_k t^k \in K_m \setminus K_{m+1}$, so that $\lambda_{n+1}, \mu_{m+1} \neq 0$. Then:

$$[g, h] \equiv t + \lambda_{n+1}\mu_{m+1}(n-m)t^{m+n+1} \mod K_{m+n+1}.$$

(ii) For any $\lambda, \mu \in \mathbb{F}_q$,

$$[e_{n,\lambda}, e_{m,\mu}] \equiv e_{m+n,\lambda\mu}^{n-m} \mod K_{\min(m+2n,2m+n)}.$$

(iii) For $p \geq 3$, if $p \nmid (n-m)$ (respectively $p \mid (n-m)$), then $[K_n, K_m] = K_{m+n}$ (respectively $[K_n, K_m] = K_{m+n+1}$).

2.5.2 Further group-theoretic properties

Henceforth we suppose $p \geq 3$. The results quoted here are not really used anywhere else in this work, but to the reader unfamiliar with the Nottingham group they should amply justify its inclusion in the cast of major characters studied in pro-p group theory.

Proposition 2.5.5. *$\mathcal{N}_q$ is hereditarily just-infinite. In other words, for any $H \leq_o \mathcal{N}_q$ and any $K \lhd_c H$, $K \leq_o H$.*

Proposition 2.5.6. *Let $(\gamma_n(\mathcal{N}_q))_n$ be the lower central series of $\mathcal{N}_q$. Then:*

$$|\gamma_n(\mathcal{N}_q)/\gamma_{n+1}(\mathcal{N}_q)| \leq q^2$$

for all $n \in \mathbb{N}$. In particular, $\mathcal{N}_q$ has finite width.

Hereditarily just-infinite pro-p groups (or HJI groups for short) have a status in pro-p group theory similar to that enjoyed by finite simple groups in finite group theory. Unfortunately, they are too large and wild a class for there to be any realistic prospect of fully classifying them. Within the class of HJI groups, those of finite width have drawn particular attention, in part because they form a class which may be sufficiently restricted as to be amenable to classification. It is still far from clear what such a classification should look like, but the last two results suggest that the Nottingham group (and its relatives) could represent an important case of it.

Other examples of HJI groups arise from analytic groups over pro-p domains (with simple Lie algebra), but the Nottingham group is very much unlike an analytic pro-p group.

Theorem 2.5.7 (Leedham-Green, Weiss (Theorem 10 in [21])). *Every finite p-group embeds into $\mathcal{N}_q$.*

Corollary 2.5.8. *$\mathcal{N}_q$ is not linear over any field.*

Theorem 2.5.9 (Camina [21], Fesenko [31]). *Every countably based pro-p group embeds as a closed subgroup of $\mathcal{N}_q$.*

Theorem 2.5.10 (Segal, Shalev (Theorem 5.3 of [73])). *Let R be a pro-p domain (as defined in Section 2.3.2). Then $\mathcal{N}_q$ is not analytic over R.*

It remains to the author's knowledge an open question whether $\mathcal{N}_q$ is a finitely presented pro-p group.

2.6 The Sum-Product Phenomenon in $\mathbb{Z}/p^n\mathbb{Z}$

In this section, we marshal some results related to the sum-product phenomenon in the ring $\mathbb{Z}/p^n\mathbb{Z}$, originally presented in [3] and [6].

For a general (commutative, unital) ring R, a *sum-product theorem* for R states that for an arbitrary non-empty finite subset $A \subseteq R$, at least one of the sumset $A+A$ or the product set $A \cdot A$ is significantly larger than A, provided certain obvious obstructions do not apply.

Sum-product theorems are in a similar genre to the product theorems for finite groups we encountered in Section 2.2. Indeed the proof of product theorems for linear groups over finite fields such as Helfgott's [36] involves associating to a generating set for the group a subset of the underlying field, in such a way that growth of sum-product in the field guarantees growth of the generating set under multiplication.

The precise nature of the obstructions to growth we wish to avoid depends on the ring R in question. For $R = \mathbb{Z}/p^n\mathbb{Z}$ it turns out that if growth of sum-product fails for A, then A is either: (i) already almost all of $\mathbb{Z}/p^n\mathbb{Z}$ or (ii) contained in a few cosets of a subring $p^m\mathbb{Z}/p^n\mathbb{Z}$. To be precise, we have the following result of Bourgain:

Theorem 2.6.1 (Theorem 1 in [3]). *For all $\delta_1, \delta_2 > 0$ there exist $\gamma, \epsilon > 0$ such that, if $A \subseteq \mathbb{Z}/p^n\mathbb{Z}$ satisfies:*

(i) $|A| < p^{(1-\delta_1)n}$;

(ii) $|\pi_{p^m}(A)| > p^{\delta_2 m}$ *whenever* $n \geq m > \epsilon n$

then:

$$|A + A| + |A \cdot A| > p^{\gamma n}|A|. \tag{2.5}$$

Our interest in this result is that it allows us to produce sets A which efficiently generate a large part of $\mathbb{Z}/p^n\mathbb{Z}$, by taking iterated sum- and product-sets and difference sets (see Proposition 2.6.3 and Corollary 2.6.4 below for precise statements). The ability to produce such sets shall be crucial to the results of Chapter 5.

First we have an auxiliary Fourier-analytic result, which shows that families of subsets $A_i, B_i \subseteq \mathbb{Z}/p^n\mathbb{Z}$ which are sufficiently *equidistributed* (in the sense that none of them is too concentrated on any coset of a subring $p^m\mathbb{Z}/p^n\mathbb{Z}$) cover $\mathbb{Z}/p^n\mathbb{Z}$ upon taking the product sets $A_i \cdot B_i$ and summing.

Lemma 2.6.2. *Let $\beta_1, \beta_2 \in (0,1)$; $K \in \mathbb{Z}_{>0}$ be such that $\beta_1 + \beta_2 > 3/2$ and $K > \frac{2}{\beta_1+\beta_2-3/2}$. Let $A_i, B_i \subseteq \mathbb{Z}/p^n\mathbb{Z}$ for $1 \leq i \leq K$, satisfying:*

(i) $\max_{x \in \mathbb{Z}/p^m\mathbb{Z}} |A_i \cap \pi_{p^m}^{-1}(x)| < p^{-\beta_1 m}|A_i|$

(ii) $\max_{y \in \mathbb{Z}/p^m\mathbb{Z}} |B_i \cap \pi_{p^m}^{-1}(y)| < p^{-\beta_2 m}|B_i|$

for all $0 < m \leq n$. Then for any $z \in \mathbb{Z}/p^n\mathbb{Z}$, there exist $x_i \in A_i, y_i \in B_i$ such that $z = \sum_{i=1}^K x_i y_i$.

Proof. Denote $\underline{A} = \prod_{i=1}^K A_i$, $\underline{B} = \prod_{i=1}^K B_i$. Define $\phi : \mathbb{Z}/p^n\mathbb{Z} \to [0, \infty)$ by:

$$\phi(w) = \left| \left\{ (\underline{x}, \underline{y}) \in \underline{A} \times \underline{B} : \sum_{i=1}^K x_i y_i = w \right\} \right| / (|\underline{A}||\underline{B}|).$$

We prove that for any $w \in \mathbb{Z}/p^n\mathbb{Z}$, $|\phi(w) - \frac{1}{p^n}| < \frac{1}{p^{n+1}}$. In particular this implies that $\phi(w)$ is always non-zero, and so yields the required result.

We first observe:

$$\phi(w) = \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{\underline{x} \in \underline{A}} \sum_{\underline{y} \in \underline{B}} \sum_{0 \leq z < p^n} e_{p^n}(z(w - \sum_{i=1}^K x_i y_i))$$

(where for $x \in \mathbb{Z}/p^n\mathbb{Z}$, $e_{p^n}(x) = e^{\frac{2\pi i x}{p^n}}$). This is because, for fixed $\underline{x} \in \underline{A}$ and $\underline{y} \in \underline{B}$, the contribution to the exponential sum as z varies is p^n when $(\underline{x}, \underline{y})$ is a solution to $\sum_{i=1}^K x_i y_i = w$, and is 0 if it is not. Thus:

$$\begin{aligned}
|\phi(w) - \frac{1}{p^n}| &\leq \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{0 < z < p^n} \Big| \sum_{\underline{x} \in \underline{A}} \sum_{\underline{y} \in \underline{B}} e_{p^n}(z(w - \sum_{i=1}^K x_i y_i)) \Big| \\
&= \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{0 < z < p^n} \Big| \sum_{\underline{x} \in \underline{A}} \sum_{\underline{y} \in \underline{B}} e_{p^n}(z \sum_{i=1}^K x_i y_i) \Big| \\
&= \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{0 < z < p^n} \prod_{i=1}^K \Big| \sum_{x \in A_i} \sum_{y \in B_i} e_{p^n}(zxy) \Big|.
\end{aligned}$$

For any $z \in \mathbb{Z}/p^n\mathbb{Z}$, there exist unique $0 \leq n' < n$ and $v \in \mathbb{Z}/p^{n-n'}\mathbb{Z}$ with $(p, v) = 1$, such that $z = vp^{n'}$, so that:

$$\Big|\sum_{x\in A_i}\sum_{y\in B_i} e_{p^n}(zxy)\Big| = \Big|\sum_{x\in A_i}\sum_{y\in B_i} e_{p^{n-n'}}(vxy)\Big|.$$

For $X \subseteq \mathbb{Z}/p^n\mathbb{Z}$, define $\psi_X : \mathbb{Z}/p^{n-n'}\mathbb{Z} \to \mathbb{C}$ by:

$$\psi_X(y) = |\{x \in X : \pi_{p^{n-n'}}(x) = y\}|.$$

By hypothesis, for all $y \in \mathbb{Z}/p^{n-n'}\mathbb{Z}$ and for $1 \leq i \leq K$,

$$\psi_{A_i}(y) < p^{-\beta_1(n-n')}|A_i|; \psi_{B_i}(y) < p^{-\beta_2(n-n')}|B_i|. \tag{2.6}$$

Now,

$$\begin{aligned}\Big|\sum_{x\in A_i}\sum_{y\in B_i} e_{p^{n-n'}}(vxy)\Big| &= \Big|\sum_{x,y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{A_i}(x)\psi_{B_i}(y)e_{p^{n-n'}}(vxy)\Big| \\ &\leq \Big(\sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{A_i}(x)^2\Big)^{\frac{1}{2}}\Big(\sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}}\Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(vxy)\Big|^2\Big)^{\frac{1}{2}}\end{aligned}$$

by the Cauchy-Schwarz inequality. Note that:

$$\sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}}\Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(vxy)\Big|^2 = \sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}}\Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(xy)\Big|^2$$

and for every $x \in \mathbb{Z}/p^{n-n'}\mathbb{Z}$:

$$\Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(xy)\Big| = \Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(-xy)\Big| = |\hat{\psi_{B_i}}(x)|,$$

where $\hat{\psi_{B_i}}$ is the Fourier transform of ψ_{B_i}. By the Plancherel Theorem,

$$\sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}}\Big|\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)e_{p^{n-n'}}(vxy)\Big|^2 = \|\hat{\psi_{B_i}}\|_2^2 = p^{n-n'}\|\psi_{B_i}\|_2^2.$$

It follows that:

$$\begin{aligned}|\sum_{x\in A_i}\sum_{y\in B_i} e_{p^{n-n'}}(vxy)| &\leq \Big(\sum_{x\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{A_i}(x)^2\Big)^{\frac{1}{2}}\Big(\sum_{y\in\mathbb{Z}/p^{n-n'}\mathbb{Z}} \psi_{B_i}(y)^2\Big)^{\frac{1}{2}}p^{\frac{n-n'}{2}} \\ &\leq p^{(\frac{3}{2}-\beta_1-\beta_2)(n-n')}|A_i||B_i|\end{aligned}$$

by (2.6). Therefore:

$$\begin{aligned}
|\phi(w) - \frac{1}{p^n}| &\leq \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{0\leq n'<n} \sum_{v\in(\mathbb{Z}/p^{n-n'}\mathbb{Z})^*} \prod_{i=1}^{K} |\sum_{x\in A_i}\sum_{y\in B_i} e_{p^{n-n'}}(vxy)| \\
&\leq \frac{1}{p^n|\underline{A}||\underline{B}|} \sum_{0\leq n'<n} \sum_{v\in(\mathbb{Z}/p^{n-n'}\mathbb{Z})^*} p^{K(\frac{3}{2}-\beta_1-\beta_2)(n-n')}|\underline{A}||\underline{B}| \\
&\leq \frac{1}{p^n} \sum_{0\leq n'<n} (1-\frac{1}{p}) p^{(1+K(\frac{3}{2}-\beta_1-\beta_2))(n-n')} \\
&< \frac{1}{p^{n+1}}
\end{aligned}$$

since $1 + K(\frac{3}{2} - \beta_1 - \beta_2) < -1$. □

We now combine Lemma 2.6.2 with Theorem 2.6.1 to prove efficient generation of a large subring.

Proposition 2.6.3. *For all $\alpha_1, \alpha_2 > 0$, there exists $\epsilon > 0$; $r, s \in \mathbb{Z}_{>0}$ such that, for $A \subseteq \mathbb{Z}/p^n\mathbb{Z}$ satisfying:*

$$|\pi_{p^m}(A)| > p^{\alpha_1 m} \text{ whenever } n \geq m > \epsilon n,$$

there exists $0 \leq k < \alpha_2 n$ such that:

$$p^k\mathbb{Z}/p^n\mathbb{Z} \subseteq \Sigma_r A^{(s)} - \Sigma_r A^{(s)}.$$

Proof. Iteratively applying (2.5), for any $\delta_1 > 0$ there exist $R, S \in \mathbb{Z}_{>0}$ such that $B := \Sigma_R A^{(S)}$ satisfies $|B| \geq p^{(1-\delta_1)n}$.

For $\alpha \in (0, 1)$, define:

$$S_\alpha = \{m \leq n : \max_{x\in\mathbb{Z}/p^m\mathbb{Z}} |B \cap \pi_{p^m}^{-1}(x)| \geq p^{-(1-\alpha)m}|B|\}.$$

Note that $0 \in S_\alpha$ and that (provided $\delta_1 < \alpha$), $n \notin S_\alpha$. If $m > \frac{\delta_1 n}{\alpha}$, then $n - m < (1 - \delta_1)n - (1 - \alpha)m$ so that $p^{n-m} < p^{-(1-\alpha)m}|B|$. But for any $x \in \mathbb{Z}/p^m\mathbb{Z}$, $|\pi_{p^m}^{-1}(x)| \leq p^{n-m}$, so that $m \notin S_\alpha$.

Let $M = \max(S_\alpha)$, so that $M \leq \frac{\delta_1 n}{\alpha}$. There exists $x \in \mathbb{Z}/p^M\mathbb{Z}$ such that $C := B \cap \pi_{p^M}^{-1}(x)$ satisfies:

$$|C| \geq p^{-(1-\alpha)M}|B|.$$

For any $\overline{x} \in C$, there exists $D \subseteq \mathbb{Z}/p^{n-M}\mathbb{Z}$ such that $|D| = |C|$ with:

$$C = \overline{x} + p^M D.$$

Let $0 < k \leq n - M$. By maximality of M,

$$\begin{aligned} p^{-(1-\alpha)k}|D| &= p^{-(1-\alpha)k}|C| \\ &\geq p^{-(1-\alpha)(k+M)}|B| \\ &> \max_{y \in \mathbb{Z}/p^{k+M}\mathbb{Z}} |B \cap \pi_{p^{k+M}}^{-1}(y)| \\ &\geq \max_{y \in \mathbb{Z}/p^{k}\mathbb{Z}} |D \cap \pi_{p^{k}}^{-1}(y)|. \end{aligned}$$

We may now apply Lemma 2.6.2 with $\beta_1 = \beta_2 \in (\frac{7}{8}, 1)$, $K = 8$ and $A_i = B_i = D$. Setting $\alpha \in (0, 1 - \beta_1)$,

$$\max_{y \in \mathbb{Z}/p^{k}\mathbb{Z}} |D \cap \pi_{p^{k}}^{-1}(y)| < p^{-\beta_1 k}|D|$$

for all $0 < k \leq n - M$, so that by Lemma 2.6.2, $\mathbb{Z}/p^{n-M}\mathbb{Z} \subseteq \Sigma_8 D^{(2)}$. Hence:

$$\begin{aligned} p^{2M}\mathbb{Z}/p^{n}\mathbb{Z} &\subseteq \Sigma_8 (p^M D)^{(2)} \\ &\subseteq \Sigma_8 (C - \overline{x})^{(2)} \\ &\subseteq \Sigma_{16} C^{(2)} - \Sigma_{16} C^{(2)} \\ &\subseteq \Sigma_{16} B^{(2)} - \Sigma_{16} B^{(2)} \\ &= \Sigma_{16R^2} A^{(2S)} - \Sigma_{16R^2} A^{(2S)}. \end{aligned}$$

We require $2M < \alpha_2 n$. Recall that $M \leq \frac{\delta_1 n}{\alpha}$. It therefore suffices to set $\delta_1 \in (0, \frac{\alpha_2 \alpha}{2})$. □

In our applications we shall use the following variant of Proposition 2.6.3, in which the hypothesis has been weakened: we no longer require that A fails to be contained in a few cosets of $p^m\mathbb{Z}/p^n\mathbb{Z}$ at *all* sufficiently large scales m; only that it should not be so contained at *some* sufficiently small scale.

Corollary 2.6.4. *For all $\alpha > 0$, there exist $\epsilon > 0$; $r, s \in \mathbb{Z}_{>0}$ such that, if $A \subseteq \mathbb{Z}/p^n\mathbb{Z}$ is such that:*

$$|\pi_{p^m}(A)| > p^{\alpha m} \text{ for some } m < \epsilon n,$$

then for some $j \leq k \leq n$ satisfying $k - j \geq \frac{\alpha m}{4}$ and $k \ll_\alpha m$,

$$p^j\mathbb{Z}/p^k\mathbb{Z} \subseteq \pi_{p^k}(\Sigma_r A^{(s)} - \Sigma_r A^{(s)}).$$

Proof. Let $k_1 < m$ be maximal such that:

$$\max_{x \in \mathbb{Z}/p^{k_1}\mathbb{Z}} |\{y \in \pi_{p^m}(A) : \pi_{p^{k_1}}(y) = x\}| \geq p^{\frac{-\alpha k_1}{2}} |\pi_{p^m}(A)|. \tag{2.7}$$

Such k_1 clearly exists: we have equality in the above for $k_1 = 0$, for instance.

For any $x \in \mathbb{Z}/p^{k_1}\mathbb{Z}$,

$$|\{y \in \mathbb{Z}/p^m\mathbb{Z} : \pi_{p^{k_1}}(y) = x\}| = p^{m-k_1}.$$

Hence:

$$|\{y \in \pi_{p^m}(A) : \pi_{p^{k_1}}(y) = x\}| \leq p^{m-k_1},$$

so that:

$$p^{m-k_1} \geq p^{\frac{-\alpha k_1}{2}} |\pi_{p^m}(A)| > p^{\alpha m - \frac{\alpha k_1}{2}} \geq p^{\frac{\alpha m}{2}} \tag{2.8}$$

(by (2.7)). If $0 < k_2 \leq m - k_1$, then:

$$\begin{aligned} p^{\frac{-\alpha(k_1+k_2)}{2}} |\pi_{p^m}(A)| &> \max_{x \in \mathbb{Z}/p^{k_1+k_2}\mathbb{Z}} |\{y \in \pi_{p^m}(A) : \pi_{p^{k_1+k_2}}(y) = x\}| \\ &\geq \max_{x \in \mathbb{Z}/p^{k_1+k_2}\mathbb{Z} : \pi_{p^{k_1}}(x) = z} |\{y \in \pi_{p^m}(A) : \pi_{p^{k_1+k_2}}(y) = x\}| \\ &\geq |\{y \in \pi_{p^m}(A) : \pi_{p^{k_1}}(y) = z\}| / |\{\pi_{p^{k_1+k_2}}(y) : y \in \pi_{p^m}(A), \pi_{p^{k_1}}(y) = z\}| \end{aligned}$$

for any $z \in \mathbb{Z}/p^{k_1}\mathbb{Z}$. If z attains the maximum in (2.7), then:

$$|\{y \in \pi_{p^m}(A) : \pi_{p^{k_1}}(y) = z\}| \geq p^{\frac{-\alpha k_1}{2}} |\pi_{p^m}(A)|$$

so:

$$|\{\pi_{p^{k_1+k_2}}(y) : y \in \pi_{p^m}(A), \pi_{p^{k_1}}(y) = z\}| > p^{\frac{-\alpha k_2}{2}}. \tag{2.9}$$

Define $B = \{x \in \mathbb{Z}/p^{n-k_1}\mathbb{Z} : p^{k_1}x \in A - A\}$. We claim that, for $k_2 \leq m - k_1$, $|\pi_{p^{k_2}}(B)| > p^{\frac{\alpha k_2}{2}}$. Let $C_z = \{x \in A : \pi_{p^{k_1}}(x) = z\}$. For $x, y \in C_z$,

$$\pi_{p^{k_1}}(x) - \pi_{p^{k_1}}(y) = z - z = 0 \text{ (in } \mathbb{Z}/p^{k_1}\mathbb{Z})$$

so there exists $w \in \mathbb{Z}/p^n\mathbb{Z}$ such that $x - y = p^{k_1}w$. Fix some $x \in C_z$, and let:

$$D = \pi_{p^{n-k_1}}(\{p^{-k_1}(x - y) : y \in C_z\}) \subseteq B.$$

Then $|\pi_{p^{k_2}}(B)| \geq |\pi_{p^{k_2}}(D)| \geq |\pi_{p^{k_2}}(C)|$ (as $y \mapsto p^{-k_1}(x - y)$ is injective), and $|\pi_{p^{k_2}}(C)| > p^{\frac{\alpha k_2}{2}}$, by (2.9), and we have the desired claim.

In other words, $\pi_{p^{m-k_1}}(B)$ satisfies the hypothesis of Proposition 2.6.3, with n replaced by $m - k_1$, $\alpha_1 = \alpha/2$, and ϵ arbitrarily small. Setting $\alpha_2 = 1/2$, we conclude that there exist $0 \leq i \leq \frac{m-k_1}{2}$; $r, s \in \mathbb{Z}_{>0}$ such that:

$$\pi_{p^{m-k_1}}(\Sigma_r B^{(s)} - \Sigma_r B^{(s)}) = \Sigma_r(\pi_{p^{m-k_1}}(B))^{(s)} - \Sigma_r(\pi_{p^{m-k_1}}(B))^{(s)} \supseteq p^i\mathbb{Z}/p^{m-k_1}\mathbb{Z}. \tag{2.10}$$

By definition of B,

$$\begin{aligned}\pi_{p^{m+(s-1)k_1}}(\Sigma_{2^s r}A^{(s)} - \Sigma_{2^s r}A^{(s)}) &\supseteq \pi_{p^{m+(s-1)k_1}}(\Sigma_r(A-A)^{(s)} - \Sigma_r(A-A)^{(s)}) \\ &\supseteq \pi_{p^{m+(s-1)k_1}}(\Sigma_r(p^{k_1}B)^{(s)} - \Sigma_r(p^{k_1}B)^{(s)}) \\ &= \pi_{p^{m+(s-1)k_1}}(p^{sk_1}(\Sigma_r(B)^{(s)} - \Sigma_r(B)^{(s)})) \\ &\supseteq p^{i+sk_1}\mathbb{Z}/p^{m+(s-1)k_1}\mathbb{Z}\end{aligned}$$

(by (2.10)). We now take $k = m + (s-1)k_1$, $j = i + sk_1$. First, $k < sm < s\epsilon n$, so stipulating $\epsilon < 1/s$ we have $k \leq n$. Second,

$$k - j = m - k_1 - i \geq (m - k_1)/2 \geq \alpha m/4 \text{ (by (2.8))}.$$

Finally, since $k < sm$ and s depends only on α, $k \ll_\alpha m$. □

Chapter 3

New Uniform Diameter Bounds in Pro-p Groups

3.1 Introduction

This chapter shall describe our results on diameters of finite groups obtained from the Solovay-Kitaev procedure. These results were first presented in [12].

3.1.1 The Solovay-Kitaev procedure in quantum computation and beyond

Let us begin by sketching some of the development of the Solovay-Kitaev procedure, from its original implementation in the context of quantum computer science to the more recent applications to diameters of finite groups. Our exposition here is based in large part upon [24].

In classical computation the basic unit of information is the *bit*. A bit can exist in one of precisely two states: "on" or "off". By contrast in quantum computation the basic unit of memory is the quantum bit or *qubit*, which can exist in a *superposition* of these states. If we represent the two classical states of a bit as a pair of orthonormal vectors e_1 and e_2, then a superposition of the two memory states is represented by a direction in the complex Hilbert space spanned by e_1 and e_2, that is by an element of $\mathbb{C}P^1$.

For simplicity, consider a quantum computer with a memory consisting only of a single qubit. A programme run on such a computer takes the form of a finite sequence of operations (known as *quantum logic gates*) which act on the memory state as elements of SU(2). At the end of the programme, the memory state is *observed*, and one of the two classical states is outputted: if the memory state after

the final gate is represented by the unit vector $\alpha e_1 + \beta e_2$, then the output is e_1 with probability $|\alpha|^2$ (respectively e_2 with probability $|\beta|^2$).

A major part of the power of quantum computation is the greater flexibility that the ability to act on superpositions of the classical states affords: instead of the finite number of operations that a classical computer can perform on its memory, we have a continuum of options among the elements of SU(2). This flexibility has led to the design of quantum algorithms which give much better performance than their classical competitors: we have for instance the well-known quantum database-search algorithm of Grover [35], which is quadratically faster than the best known classical search algorithms, or Shor's algorithm [74], which achieves polynomial-time integer factorization.

There is a cost, though. A quantum computer, if built, could only ever have a finite number of gates at its disposal, whereas a programme run on it could call for any of the uncountable number of gates in SU(2). The best that can be hoped for is that the gates $s_1, \ldots, s_r$ available to the computer (known as *instructions*) generate a subgroup of SU(2) which is *dense*, so that the arbitrary gate called for by the programme can be approximated, up to arbitrarily small error, by a finite sequence of instructions (here, and throughout, we make the simplifying assumption that the instruction set may be taken to be symmetric). The computer could then employ a *quantum compiler*: an additional (classical) algorithm which would take as input the arbitrary quantum gates from the programme, and output approximating sequences of instructions which the computer was actually able to perform. The fact that the result of the sequence of instructions may not exactly equal what was demanded by the programme is not a major problem: the result outputted upon observation is only a probable correct result anyway, and observation of a quantum state differing by a small error will give the same result with high probability (the errors in our approximations to elements of SU(2) being measured by distance in operator norm).

What *is* a big problem, however, is the possibility that the quantum compiler may be inefficient, in the sense that the number of instructions it requires to approximate an arbitrary gate up to error ϵ grows too quickly as ϵ shrinks. If this happens, and the computer must work too hard to implement the programme, then any gains in speed the programme enjoys over classical rivals may be lost.

We would like to know, then, that given a finite instruction set S, an arbitrary gate $g \in \mathrm{SU}(2)$, and $\epsilon > 0$, there exists a word w in S, of length bounded by a slowly growing function of $1/\epsilon$, which approximates g up to an error of size at most ϵ. Moreover we would like a fast classical algorithm to write w, given S, g and

ϵ. This is precisely the problem which Solovay-Kitaev procedure solves: indeed it delivers a bound on both the length of w and the time taken to write it, which is a polylogarithmic function of $1/\epsilon$. The procedure also works for $\mathrm{SU}(d)$ with larger d, and specifically for $\mathrm{SU}(2^n)$, which is the group of quantum logic gates for a computer with a memory consisting of n qubits (2^n being the dimension of the space of possible quantum states of n particles).

How does the Solovay-Kitaev procedure work? First, we naïvely generate $\mathrm{SU}(d)$ up to a first, fairly large error: since the instruction set S generates a dense subgroup of $\mathrm{SU}(d)$, for any $\epsilon_0 > 0$ there exists $l_0(S) \in \mathbb{N}$ such that $B_S(l_0)$ contains an ϵ_0-net. ϵ_0 shall be a sufficiently small absolute constant, to be determined.

We now inductively construct sequences $(l_n) \subseteq \mathbb{N}$ and $(\epsilon_n) \subseteq (0,1)$ such that $B_S(l_n)$ contains an ϵ_n-net. The construction relies on two facts about the group $\mathrm{SU}(d)$, which hold for all sufficiently small $\delta, \epsilon > 0$. The first is an elementary computation and tells us that there is some cancellation among error terms upon taking commutators.

Lemma 3.1.1. *Let $u, \tilde{u}, v, \tilde{v} \in \mathrm{SU}(d)$. Suppose:*

$$\|u - I_d\|, \|v - I_d\| < \delta;\ \|u - \tilde{u}\|, \|v - \tilde{v}\| < \epsilon.$$

Then $\|[u,v] - [\tilde{u}, \tilde{v}]\| = O_d(\delta\epsilon + \epsilon^2)$.

Once again, $\|\cdot\|$ is the operator norm. The second observation is that every element of $\mathrm{SU}(d)$ of sufficiently small operator norm may be approximated by a commutator.

Lemma 3.1.2. *Let $k \in \mathrm{SU}(d)$, satisfying $\|k - I_d\| < \epsilon$. Then there exist $u, v \in \mathrm{SU}(d)$ satisfying:*

$$\|u - I_d\|, \|v - I_d\| = O_d(\epsilon^{\frac{1}{2}});\ \|[u,v] - k\| = O_d(\epsilon^{\frac{3}{2}}).$$

Lemma 3.1.2 may be viewed as the Lie group analogue of a corresponding statement about Lie algebras: that for every traceless Hermitian matrix H, iH may be written as the Lie bracket of two Hermitian matrices of the appropriate norm. Lemma 3.1.2 follows from this latter claim by exponentiating.

We now construct the l_n, ϵ_n. Let $g \in \mathrm{SU}(d)$ be arbitrary. Given ϵ_n, there exists $w_n \in B_S(l_n)$ such that $\|g - w_n\| < \epsilon_n$, by induction. We apply Lemma 3.1.2 with $k = g w_n^{-1}$, to produce u and v. By induction again we have $\tilde{u}, \tilde{v} \in B_S(l_n)$ with $\|u - \tilde{u}\|, \|v - \tilde{v}\| < \epsilon_n$. It follows from Lemma 3.1.2 that $w_{n+1} = [\tilde{u}, \tilde{v}] w_n \in B_S(5l_n)$ satisfies $\|g - w_{n+1}\| < \epsilon_{n+1}$ with $\epsilon_{n+1} = O_d(\epsilon_n^{\frac{3}{2}})$, so we may take $l_{n+1} = 5l_n$.

For ϵ_0 sufficiently small $(\epsilon_n)_n$ decays doubly exponentially fast, at the cost of $(l_n)_n$ growing exponentially fast, so that l_n is polylogarithmically large in ϵ_n^{-1}.

What does all this have to do with diameters of finite groups? We can pose the problem of "approximating" arbitrary elements by short words up to "small error" in other topological groups, as well as compact real Lie groups like $\mathrm{SU}(d)$. In the case of a profinite group G, the natural neighbourhood basis for the topology is provided be the cosets of open normal subgroups, instead of the balls which we used in $\mathrm{SU}(d)$. Given a descending sequence $(K_n)_n$ of open normal subgroups and supposing we have a finite subset $S \subseteq G$ generating a dense subgroup, our problem becomes to find a sequence $(l_n)_n$ such that every element of G is approximable up to an error in K_n, by a word of length at most l_n. In other words, $G = K_n B_S(l_n)$, or equivalently $\mathrm{diam}(G/K_n, S) \leq l_n$.

Further, one can describe hypotheses on G, K_n which imitate the Lemmata 3.1.1 and 3.1.2, in such a way that the inductive procedure for constructing approximating words, described above for $\mathrm{SU}(d)$, translates to the profinite setting. It was this translation, achieved in the specific case $G = \mathrm{SL}_2(\mathbb{Z}_p)$, with $(K_n)_n$ the congruence subgroups, which formed the basis of Gamburd and Shahshahani's bound on $\mathrm{diam}(\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z}))$ [33]. The refinements made by Dinai [28] related to the translation of Lemma 3.1.2, and introduced the exponential map as a tool for computing commutators in $\mathrm{SL}_2(\mathbb{Z}_p)$ in terms of brackets in the associated Lie algebra. The presence of a supporting Lie algebra shall be of use in many of our results too, though an exponential map shall not always be available.

It is worth noting that in the profinite setting, the first step of approximating up to an initial error ϵ_0 corresponds to generating G/K_{n_0}, for n_0 an absolute constant. For any $S \subseteq G$ generating a dense subgroup, $\mathrm{diam}(G/K_{n_0}, S) \leq |G/K_{n_0}|$; in particular the lengths of words in S required for this first approximation may be bounded independent of S. Noting also that any generating set for any of the G/K_n lifts to a finite subset of G generating a dense subgroup, it follows that the Solovay-Kitaev algorithm produces diameter bounds independent of the choice of generating set, that is bounds for $\mathrm{diam}(G/K_n)$.

3.1.2 Statement of results

Let us recall the statements of our results on diameter from the Introduction.

Theorem 3.1.3. *Let $(R, \mathcal{M})$ be a commutative unital discrete valuation pro-p domain. Let G be a d-dimensional R-standard group, $K_n = (\mathcal{M}^n)^{(d)} \lhd_o G$. Suppose $\mathcal{L}_G$ is perfect. Then there exist $C_1(G)$, $C_2(d) > 0$ such that for all $n \in \mathbb{N}$:*

$$\mathrm{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}.$$

Definition 3.1.4. *Let G be a profinite group. G is* FAb *if every open subgroup has finite abelianisation.*

Theorem 3.1.5. *Let $p \geq 3$. Let G be a d-dimensional compact p-adic analytic group. Let $K_1 \leq G$ be an open characteristic uniform subgroup; $(K_n)_n$ its lower central p-series. If G is FAb then there exist $C_1(G), C_2(d) > 0$ such that for all $n \in \mathbb{N}$:*

$$\mathrm{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}. \tag{3.1}$$

For $G = K_1$ then conversely: if there exist $C_1, C_2 > 0$ such that (3.1) holds for all $n \in \mathbb{N}$ then G is FAb.

Recall that the definitions of a discrete valuation pro-p domain $(R, \mathcal{M})$; an R-analytic group and its associated Lie algebra and a uniform pro-p group were given in Section 2.3. We shall also see (Proposition 3.4.2) that a uniform pro-p group is FAb if and only if the associated Lie algebra is perfect.

Theorem 3.1.6. *Let $(R, \mathcal{M})$ be a commutative unital discrete valuation pro-p domain, with $\mathcal{M}$ generated by $\mathcal{P}$. Let $G \leq \mathrm{GL}_d(R)$ be the adjoint Chevalley group of type $X_l \in \{A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4, G_2\}$ over R (here d is the dimension of the associated Lie algebra). Suppose:*

$$(X_l, p) \notin \{(A_1, 2), (B_l, 2), (C_l, 2), (D_l, 2)\}.$$

Let $K_n = G \cap (I_d + \mathcal{P}^n \mathbb{M}_d(R))$. Then there exist $C_1(G) > 0$ and an absolute constant $C_2 > 0$ such that for all $n \in \mathbb{N}$:

$$\mathrm{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}.$$

Moreover, the same bound holds for $G = \mathrm{SL}_d(R), \mathrm{SO}_d(R)$ or $\mathrm{Sp}_d(R)$ provided $p \geq 3$, and for $G = \mathrm{SL}_d(R)$ with $p = 2$ provided $d \geq 3$.

Recall that the definitions of the adjoint Chevalley groups were given in Section 2.4. It is also important to note that, even though we have already seen that the groups occurring in Theorem 3.1.6 are R-analytic with perfect Lie algebra (Theorem 2.4.14), the conclusion of Theorem 3.1.6 does not follow directly from Theorem 3.1.3. For in Theorem 3.1.3, the degree C_2 of the polylogarithm in the diameter bound depends on the dimension of the group, whereas in Theorem 3.1.6 it is independent of the dimension.

Next, recall from Section 2.5 the definitions of the Nottingham group $\mathcal{N}_q$ over the finite field of order q and characteristic p, and of the congruence subgroups K_n.

Theorem 3.1.7. *Suppose $p \geq 3$. Then there exist $C_1(q) > 0$ and an absolute constant $C_2 > 0$ such that for all $n \in \mathbb{N}$:*

$$\operatorname{diam}(\mathcal{N}_q/K_n) \leq C_1(\log|\mathcal{N}_q/K_n|)^{C_2}.$$

Recall that in our results on random walks $S \subseteq G$ is a finite symmetric set, $X_1, X_2, \ldots$ is a sequence of independent random variables, each uniformly distributed on S. For $l \in \mathbb{N}$, $Y_l = X_1 \cdots X_l$.

For $(R, \mathcal{M})$ a discrete valuation pro-p domain; G a d-dimensional R-standard group and $x_1, \ldots, x_d$ an R-basis for $\mathcal{M}^{(d)}$ we may write:

$$Y_l = L_1^{(l)} x_1 + \cdots + L_d^{(l)} x_d$$

for some random variables $L_1^{(l)}, \ldots, L_d^{(l)}$ supported on R.

Corollary 3.1.8. *Suppose $\mathcal{L}_G$ is perfect and $S \subseteq G$ generates a dense subgroup. Then there exists $C(d) > 0$, such that for any $C' > 0$ there exists $C''(G, |S|, C') > 0$ and $C'''(d, |R/\mathcal{M}|, C') > 0$ such that, for any $(\lambda_1, \ldots, \lambda_d) \in R^{(d)}$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}\left[\|L_1^{(l)} - \lambda_1\|, \ldots, \|L_d^{(l)} - \lambda_d\| \leq c^{N+1} \right] - \frac{1}{|R/\mathcal{M}|^{dN}} \right| \leq e^{-C''' N^{C'}}$$

whenever $l \geq C'' N^{C+C'}$.

Here $c \in (0,1)$ is the norm of a generator $\mathcal{P}$ for the maximal ideal $\mathcal{M}$ of R.

For G a d-dimensional uniform pro-p group and any minimal (ordered) generating set $a_1, \ldots, a_d$ for G, recall that by Theorem 2.3.36 (i) we have:

$$Y_l = a_1^{M_1^{(l)}} \cdots a_d^{M_d^{(l)}}$$

for some random variables $M_1^{(l)}, \ldots, M_d^{(l)}$ supported on $\mathbb{Z}_p$.

Corollary 3.1.9. *Let $p \geq 3$. Suppose G is uniform and FAb and $S \subseteq G$ generates a dense subgroup. Then there exists $C(d) > 0$, such that for any $C' > 0$ there exists $C''(G, |S|, C') > 0$ and $C'''(d, p, C') > 0$ such that, for any $\mu_1, \ldots, \mu_d \in \mathbb{Z}_p$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}\left[\|M_1^{(l)} - \mu_1\|, \ldots, \|M_d^{(l)} - \mu_d\| \leq p^{-N-1} \right] - \frac{1}{p^{dN}} \right| \leq e^{-C''' N^{C'}}$$

whenever $l \geq C'' N^{C+C'}$.

For $G = \mathcal{N}_q$ the Nottingham group of the finite field of order q we may express:

$$Y_l = t + \sum_{i=2}^{\infty} A_i^{(l)} t^i$$

for some random variables $A_i^{(l)}$ supported on $\mathbb{F}_q$.

Corollary 3.1.10. *Let $p \geq 3$. Suppose $S \subseteq \mathcal{N}_q$ generates a dense subgroup. Then there exists an absolute constant $C > 0$, such that for any $C' > 0$ there exists $C''(q, |S|, C') > 0$ and $C'''(q, C') > 0$ such that, for any sequence $(\alpha_i)_i$ in $\mathbb{F}_q$, and for any $N \in \mathbb{N}$, we have:*

$$\left| \mathbb{P}[A_2^{(l)} = \alpha_2, \ldots, A_N^{(l)} = \alpha_N] - \frac{1}{q^{N-1}} \right| \leq e^{-C''' N^{C'}}$$

whenever $l \geq C'' N^{C+C'}$.

One of the virtues of the method by which we arrive at our diameter bounds is that the constants appearing are in principle computable. Indeed, Propositions 3.2.1 and 3.2.5 shall allow us to read off the degree C_2 of the polylogarithm giving our upper bound on $\mathrm{diam}(G/K_n)$. For ease of reference, let us record some cases here:

G	K_n	C_2
$\mathrm{SL}_d(R)$ (with $p \geq 3$ for $d = 2$)	$\{g \in G : g \equiv I \mod \mathcal{P}^n\}$	≈ 5.46
$\mathrm{Sp}_{2d}(R)$ (with $p \geq 3$) or $\mathrm{SO}_d(R)$	$\{g \in G : g \equiv I \mod \mathcal{P}^n\}$	≈ 6.49
Adjoint Chevalley group of type E_8 over R	$\{g \in G : g \equiv I \mod \mathcal{P}^n\}$	≈ 37.69
d-dimensional R-standard; $\mathcal{L}_G$ perfect	As in Theorem 3.1.3	$6\log(4d+1)/\log(3)$
d-dimensional FAb p-adic analytic (with $p \geq 3$)	As in Theorem 3.1.5	$6\log(4d+1)/\log(3)$
$\mathcal{N}_q$ (with $p \geq 3$)	As in Theorem 3.1.7	12

Here, as elsewhere, R is a discrete valuation pro-p domain with maximal ideal generated by $\mathcal{P}$ and q is a power of p. The case of the E_8-group gives the largest value of C_2 among the Chevalley groups. It is likely that this value is far from optimal, as we shall discuss later.

The chapter is structured as follows. In Section 3.2 we discuss analogues of the Solovay-Kitaev procedure for profinite groups upon which all our results will be based. In Section 3.3 we prove Theorem 3.1.6 in the case of classical groups. This is achieved via a very concrete analysis of the Lie algebras of these groups, in their standard matrix representation, and does not require any understanding of the associated root systems. In Section 3.4 we study the Lie algebras of R-analytic groups, prove Theorem 3.1.3 and deduce both Theorem 3.1.5 and the exceptional case of Theorem 3.1.6. In Section 3.5 we prove Theorem 3.1.7. Consequences of these results for mixing times of random walks are explained in Section 3.6.

3.2 The Profinite Solovay-Kitaev Procedure

In this section we give bounds on the diameters of finite quotients of a general profinite group G under some hypotheses on the behaviour of commutators in G. Our hypotheses are based on the conditions underlying the Solovay-Kitaev procedure outlined in Section 3.1.1. The proofs of Theorems 3.1.3, 3.1.5, 3.1.6 and 3.1.7 will thereby be reduced to a verification that commutators in the groups concerned satisfy these hypotheses. Our first result in this direction, which will also serve as a warm-up for the more general technical result required for some applications, is:

Proposition 3.2.1. *Let G be a profinite group, $(K_n)_{n\geq 1}$ a descending sequence of open normal subgroups of G. Suppose:*

(i) For all $m, n \geq 1$, $[K_m, K_n] \subseteq K_{m+n}$;

(ii) There exists $n_0 \geq 1$ such that for all $m, n \geq n_0$ satisfying $n \leq m \leq 2n$, and all $g \in K_{n+m}$, there exist:

$$g_1, \ldots, g_A \in K_n \ , \ h_1, \ldots, h_A \in K_m$$

such that $[g_1, h_1] \cdots [g_A, h_A] g^{-1} \in K_{2n+m}$.

Then $G/\bigcap_{n=1}^{\infty} K_n$ is finitely generated and there exists $C > 0$ (depending only on A, $|G/K_{2n_0}|$) such that for all $n \geq 1$,

$$\operatorname{diam}(G/K_n) \leq C n^{\frac{\log(8A^2+6A)}{\log 2}}.$$

Remark 3.2.2. *In all the examples we consider below, we will have in addition that the sequence* $(|K_i/K_{i+1}|)_i$ *is constant, so that a bound for* $\mathrm{diam}(G/K_n)$*, which is polynomial in* n*, is polylogarithmic in* $|G/K_n|$*.*

In fact, rather than hypothesis (i) itself the proof uses a reformulation (i'), as explained in the following Lemma, which is more recognizable as a profinite translation of Lemma 3.1.1.

Lemma 3.2.3. *Let* G *be a profinite group,* $(K_n)_{n\geq 1}$ *a descending sequence of open normal subgroups. The following conditions are equivalent:*

(i) For all $m, n \geq 1$, $[K_m, K_n] \subseteq K_{m+n}$*;*

(i') For all $m, m', n, n' \geq 1$*, with* $m \leq m'$, $n \leq n'$*, and for all* $g, g' \in K_n$*;* $h, h' \in K_m$ *with* $g^{-1}g' \in K_{n'}$*;* $h^{-1}h' \in K_{m'}$*,*

$$[g,h]^{-1}[g',h'] \in K_{\min(m+n',m'+n)}.$$

Proof. Assuming (i), write $\tilde{g} = g^{-1}g', \tilde{h} = h^{-1}h'$. Then we may express $[g', h']$ as:

$$[g',h'] = [g,\tilde{h}][g,h][[g,h],\tilde{h}][[g,h\tilde{h}],\tilde{g}][\tilde{g},h\tilde{h}]$$

by standard commutator identities. Now:

$$[g,\tilde{h}] \in K_{n+m'};\ [[g,h],\tilde{h}] \in K_{n+m+m'};\ [[g,h\tilde{h}],\tilde{g}] \in K_{n+n'+m};\ [\tilde{g},h\tilde{h}] \in K_{n'+m}$$

by (i), so that $[g,h] \equiv [g',h'] \mod K_{\min(m+n',m'+n)}$.

Conversely, assuming (i'), let $g \in K_n$, $h \in K_m$. We may assume $n \leq m$. Then $g^{-1}h \in K_n$. Taking $n' = n$, $m' > m$ in (i'), we have $\min(m+n', m'+n) = n+m$, so we may set $g' = h' = h$ to obtain:

$$1 = [h,h] \equiv [g,h] \mod K_{n+m}.$$

In other words, $[g,h] \in K_{n+m}$, as required. □

The diameter bound will come from the following Lemma, the conditions of which we shall verify in the setting of Proposition 3.2.1.

Lemma 3.2.4. *Let* G *be a profinite group,* $(K_n)_{n\geq 1}$ *a descending sequence of open normal subgroups. Suppose there exist* $n_0, B, D \in \mathbb{N}$*, with* $D \geq 2$*, such that, for every* $n \geq n_0$ *and every* $X \subseteq G$*,*

$$K_n/K_{Dn} \subseteq K_{Dn}X/K_{Dn} \Rightarrow K_{Dn}/K_{D^2n} \subseteq K_{D^2n}X^B/K_{D^2n}. \tag{3.2}$$

Then $G/\bigcap_{n=1}^{\infty} K_n$ is finitely generated and there exists $C > 0$ (depending only on $B, |G/K_{Dn_0}|$) such that for any $n \in \mathbb{N}$,

$$\operatorname{diam}(G/K_n) \leq Cn^{\frac{\log B}{\log D}}.$$

Proof. Let $S \subseteq G$, and suppose the restriction of the quotient $\pi_{Dn_0} : G \twoheadrightarrow G/K_{Dn_0}$ to $\langle S \rangle$ is surjective. Then for some $l_0 \in \mathbb{N}$ (independent of S),

$$K_{Dn_0} B_S(l_0)/K_{Dn_0} = G/K_{Dn_0}$$

(we may always take $l_0 \leq |G/K_{Dn_0}|$). In particular we have:

$$K_{n_0}/K_{Dn_0} \subseteq K_{Dn_0} B_S(l_0)/K_{Dn_0}.$$

By an easy induction involving (3.2), we have for any $i \in \mathbb{N}$,

$$K_{D^i n_0}/K_{D^{i+1} n_0} \subseteq K_{D^{i+1} n_0} B_S(B^i l_0)/K_{D^{i+1} n_0}.$$

It follows that, for any $n \leq D^i n_0$,

$$\operatorname{diam}(G/K_n, S) \ll_{B,l_0} B^i.$$

Hence for arbitrary n, choosing i such that $D^{i-1} n_0 \leq n \leq D^i n_0$,

$$\operatorname{diam}(G/K_n, S) \ll_{B,l_0} B^{\frac{\log n}{\log D}} = n^{\frac{\log B}{\log D}}.$$

Now let $\overline{S} \subseteq G/K_n$ and suppose $\langle \overline{S} \rangle = G/K_n$. If $n \leq Dn_0$, then $\operatorname{diam}(G/K_n, \overline{S}) \leq l_0$. Otherwise, the image of $\overline{S}$ in G/K_{Dn_0} is a generating set, and the preceding argument applies.

In particular, let $\tilde{S} \subseteq G$ be finite with image in G/K_{Dn_0} a generating set. Then for every n, $\tilde{S}$ generates G modulo K_n, so $\tilde{S}$ maps to a topological generating set in $G/\bigcap_{n=1}^{\infty} K_n$. □

Proof of Proposition 3.2.1. Let $n \geq n_0$. Suppose $X \subseteq G$ is such that:

$$K_n/K_{2n} \subseteq K_{2n} X/K_{2n}. \tag{3.3}$$

Let $g \in K_{2n}$. By hypothesis (ii) there exist $g_1, \ldots, g_A, h_1, \ldots, h_A \in K_n$ such that:

$$g \equiv [g_1, h_1] \cdots [g_A, h_A] \mod K_{3n}.$$

By (3.3) there exist $g_1', \ldots, g_A', h_1', \ldots, h_A' \in X$ with $g_i \equiv g_i', h_i \equiv h_i' \mod K_{2n}$ for $i = 1, \ldots, A$. By hypothesis (i') from Lemma 3.2.3, $[g_i, h_i] \equiv [g_i', h_i'] \mod K_{3n}$. Hence $g \equiv [g_1', h_1'] \cdots [g_A', h_A'] \mod K_{3n}$, so that:

$$K_{2n}/K_{3n} \subseteq K_{3n} X^{4A}/K_{3n}. \tag{3.4}$$

Likewise, let $g \in K_{3n}$. There exist $g_1, \ldots, g_A \in K_n$, $h_1, \ldots, h_A \in K_{2n}$ such that:

$$g \equiv [g_1, h_1] \cdots [g_A, h_A] \mod K_{4n}.$$

By (3.3) and (3.4) there exist $g'_1, \ldots, g'_A \in X$ and $h'_1, \ldots, h'_A \in X^{4A}$ such that $g_i \equiv g'_i \mod K_{2n}$ and $h_i \equiv h'_i \mod K_{3n}$, so that $[g_i, h_i] \equiv [g'_i, h'_i] \mod K_{4n}$ for $i = 1, \ldots, A$ and:

$$g \equiv [g'_1, h'_1] \cdots [g'_A, h'_A] \mod K_{4n}.$$

Hence:

$$K_{3n}/K_{4n} \subseteq K_{4n}X^{8A^2+2A}/K_{4n}. \tag{3.5}$$

Combining (3.3), (3.4) and (3.5), we obtain $K_{2n}/K_{4n} \subseteq K_{4n}X^{8A^2+6A}/K_{4n}$. The required result now follows from Lemma 3.2.4, applied with $B = 8A^2+6A$, $D = 2$. □

Proposition 3.2.1 suffices to prove Theorem 3.1.6 in the case of classical groups over pro-p rings. For general analytic pro-p groups and for the Nottingham group, however, generating elements as products of commutators is more difficult. For example, $[K_n, K_m]$ may not be the whole of K_{n+m} (as will always be the case in Proposition 3.2.1) but some deeper subgroup K_{n+m+k} (with $k \geq 1$ bounded independent of m, n, say). To circumvent these and other complexities of the general case, we prove a stronger version of Proposition 3.2.1, in which hypothesis (ii) has been weakened:

Proposition 3.2.5. *Let G be a profinite group, $(K_n)_{n\geq 1}$ a descending sequence of open normal subgroups of G. Suppose:*

(i) For all $m, n \geq 1$, $[K_m, K_n] \subseteq K_{m+n}$;

(ii) There exists $\epsilon \in (0, 1)$; $A, M_1, M_2 \in \mathbb{N}$ such that for all $n \geq M_1$, there exist $n_i, m_i \in \mathbb{N}$ (for $i = 1, 2, 3$) with:

$$\tfrac{n}{3}(2 + i + \epsilon) \leq n_i \leq m_i \leq \tfrac{2n}{3}(2 + i);\ n_i + m_i = (2 + i)n - M_2$$

and for all $g \in K_{(2+i)n}$, there exist:

$$g_1, \ldots, g_A \in K_{n_i}\ ,\ h_1, \ldots, h_A \in K_{m_i}$$

such that $[g_1, h_1] \cdots [g_A, h_A]g^{-1} \in K_{(2+i)n+n_i-M_2} = K_{2n_i+m_i}$.

Then $G/\bigcap_{n=1}^{\infty} K_n$ is finitely generated and there exists $C > 0$ (depending on A, $|G/K_{3n_0}|$, where $n_0 = \max\{2M_1, \lceil \frac{3M_2}{\epsilon} \rceil\}$) such that:

$$\operatorname{diam}(G/K_n) \leq Cn^{\frac{6\log(4A+1)}{\log 3}}.$$

Proof. First claim that for any $n \geq \max\{2M_1, \frac{3M_2}{\epsilon}\}$ and any $X \subseteq G$,

$$K_n/K_{3n} \subseteq K_{3n}X/K_{3n} \Rightarrow K_n/K_{6n} \subseteq K_{6n}X^{(4A+1)^3}/K_{6n}. \tag{3.6}$$

Let $g \in K_{3n}$. By hypothesis (ii), we have $g_1, \ldots, g_A \in K_{n_1}$, $h_1, \ldots, h_A \in K_{m_1}$ such that:

$$g \equiv [g_1, h_1] \cdots [g_A, h_A] \mod K_{3n+n_1-M_2}.$$

By assumption, there exist $g_1', \ldots, g_A', h_1', \ldots, h_A' \in X$ such that $g_i \equiv g_i'$, $h_i \equiv h_i' \mod K_{3n}$, so that $g_i' \in K_{n_1}$, $h_i' \in K_{m_1}$. By Lemma 3.2.3,

$$[g_i, h_i] \equiv [g_i', h_i'] \mod K_{3n+n_1}.$$

Hence $g \equiv [g_1', h_1'] \cdots [g_A', h_A'] \mod K_{3n+n_1-M_2}$. Therefore:

$$K_{3n}/K_{3n+n_1-M_2} \subseteq K_{3n+n_1-M_2}X^{4A}/K_{3n+n_1-M_2}$$

and, combining with the hypothesis $K_n/K_{3n} \subseteq K_{3n}X/K_{3n}$,

$$K_n/K_{3n+n_1-M_2} \subseteq K_{3n+n_1-M_2}X^{4A+1}/K_{3n+n_1-M_2}.$$

In particular, since $n_1 \geq n + \frac{\epsilon n}{3} \geq n + M_2$, $K_n/K_{4n} \subseteq K_{4n}X^{4A+1}/K_{4n}$.
We now simply repeat the same procedure: let $n_2, m_2 \in \mathbb{N}$ be as above. We deduce:

$$K_{4n}/K_{4n+n_2-M_2} \subseteq K_{4n+n_2-M_2}X^{4A(4A+1)}/K_{4n+n_2-M_2}.$$

Combining this estimate with $K_n/K_{4n} \subseteq K_{4n}X^{4A+1}/K_{4n}$, and since $4n+n_2-M_2 \geq 5n$, we have:

$$K_n/K_{5n} \subseteq K_{5n}X^{(4A+1)^2}/K_{5n}.$$

Finally let $n_3, m_3 \in \mathbb{N}$ be as above. We have:

$$K_{5n}/K_{5n+n_3-M_2} \subseteq K_{5n+n_3-M_2}X^{4A(4A+1)^2}/K_{5n+n_3-M_2}.$$

Combining with $K_n/K_{5n} \subseteq K_{5n}X^{(4A+1)^2}/K_{5n}$, since $5n + n_3 \geq 6n$, the claim follows.
Using (3.6) we have for $n \geq \max\{2M_1, \frac{3M_2}{\epsilon}\}$,

$$K_n/K_{3n} \subseteq K_{3n}X/K_{3n} \implies K_{2n}/K_{6n} \subseteq K_{6n}X^{(4A+1)^3}/K_{6n}.$$

Applying (3.6) again, with n replaced by $2n$ and X replaced by $X^{(4A+1)^3}$,

$$K_{2n}/K_{12n} \subseteq K_{12n}X^{(4A+1)^6}/K_{12n}$$

so that in particular, $K_{3n}/K_{9n} \subseteq K_{9n}X^{(4A+1)^6}/K_{9n}$. The result now follows from Lemma 3.2.4, applied with $B = (4A+1)^6$, $D = 3$. □

The proof of Proposition 3.2.5 is sufficiently robust that qualitatively similar (though quantitatively worse) diameter bounds should be available under even weaker hypotheses. We shall not pursue such results here, as the level of generality already achieved is sufficient for all the examples we shall consider. We conclude this section by noting some cases in which hypothesis (i) of Propositions 3.2.1 and 3.2.5 is always satisfied.

Example 3.2.6. *(i) Let G be* any *pro-p group; K_n be the nth term of the lower central p-series for G.*

(ii) Let R be a unital profinite ring; $G \leq R^*$; $I \lhd R$ a proper two-sided open ideal. Define $K_n = G \cap (1 + I^n) \lhd G$. Let $n, m \in \mathbb{N}$ with $n \leq m$ and let $g \in K_n$, $h \in K_m$. Let $a, \tilde{a} \in I^m$, $b, \tilde{b} \in I^n$ be such that:

$$g = 1 + a,\ g^{-1} = 1 + \tilde{a},\ h = 1 + b,\ h^{-1} = 1 + \tilde{b}.$$

Then $a + \tilde{a} + \tilde{a}a = b + \tilde{b} + \tilde{b}b = 0$, so:

$$[g, h] \equiv 1 + ab + \tilde{a}b + \tilde{a}\tilde{b} + \tilde{b}a \equiv 1 + ab - ba \mod I^{2n+m}.$$

In particular, $[g, h] \in K_{n+m}$.

(iii) As a particular case of (ii), letting $R = \mathbb{F}_p G$ and $I \lhd R$ be the augmentation ideal, K_n is the nth mod-p dimension subgroup of G.

3.3 Diameter in Classical Groups

In this section we prove Theorem 3.1.6 in the case for which X_l is classical, so that the associated adjoint Chevalley group over R is one of $\mathrm{PSL}_d(R)$, $\mathrm{PSO}_d(R)$, or $\mathrm{PSp}_d(R)$ (with d even in the latter case). To be more precise, we prove the diameter bound for $G = \mathrm{SL}_d(R), \mathrm{SO}_d(R)$ or $\mathrm{Sp}_d(R)$; $K_n = G \cap (I_d + \mathcal{P}^n \mathbb{M}_d(R))$. The required result for the adjoint form then follows from Lemma 2.1.5: letting $\rho : G \to \mathrm{GL}_D(R)$ be the adjoint representation of G on the associated Lie algebra (of dimension D), for any $g \in G$, if $g \equiv I_d \mod \mathcal{P}^n$ then $\rho(g) \equiv I_D \mod \mathcal{P}^n$. Thus letting $K_n = G \cap (I_d + \mathcal{P}^n \mathbb{M}_d(R))$, $L_n = \rho(G) \cap (I_D + \mathcal{P}^n \mathbb{M}_D(R))$, ρ descends to an epimorphism $G/K_n \twoheadrightarrow \rho(G)/L_n$. By Lemma 2.1.5 (i),

$$\mathrm{diam}(\rho(G)/L_n) \leq \mathrm{diam}(G/K_n) \leq C_1(\log|G/K_n|)^{C_2}.$$

The polylogarithmic diameter bound in $|G/K_n|$ then translates to a polylogarithmic bound in $|\rho(G)/L_n|$ (with possibly larger constant C_1). For $|G/K_n| \ll |R/\mathcal{M}|^{d^2n}$ and $|\rho(G)/L_n| \gg |R/\mathcal{M}|^n$.

We verify the hypotheses of Proposition 3.2.1 for $G = \mathrm{SL}_d(R)$, $\mathrm{SO}_d(R)$, or $\mathrm{Sp}_d(R)$. Recall that we permit ourselves the assumption that $p \geq 3$ unless $G = \mathrm{SL}_d(R)$ and $d \geq 3$. Hypothesis (i) follows immediately from Example 3.2.6 (ii). Moreover, for $g \in K_n$, $h \in K_m$, with $n \leq m \leq 2n$, writing:

$$g = I_d + \mathcal{P}^n X;\ h = I_d + \mathcal{P}^m Y$$

for some $X, Y \in \mathbb{M}_d(R)$, we have:

$$[g, h] \equiv I_d + \mathcal{P}^{m+n}(X, Y) \mod \mathcal{P}^{m+2n}$$

where $(X, Y) = XY - YX$ is the Lie bracket. Hence for $g_1, \ldots, g_A \in K_n$, $h_1, \ldots, h_A \in K_m$, writing $g_i = I_d + \mathcal{P}^n X_i$, $h_i = I_d + \mathcal{P}^m Y_i$, we have:

$$[g_1, h_1] \cdots [g_A, h_A] \equiv I_d + \mathcal{P}^{m+n}((X_1, Y_1) + \ldots + (X_A, Y_A)) \mod \mathcal{P}^{m+2n}.$$

To verify hypothesis (ii) of Proposition 3.2.1, it therefore suffices to find $A \in \mathbb{N}$ (independent of G) such that, for any $g \in K_{m+n}$, we can find $X_1, \ldots X_A$, $Y_1, \ldots, Y_A \in \mathbb{M}_d(R)$ such that:

(a) $g - I_d \equiv \mathcal{P}^{m+n}((X_1, Y_1) + \ldots + (X_A, Y_A)) \mod \mathcal{P}^{m+2n}$;

(b) There exist $g_1, \ldots, g_A \in K_n$, $h_1, \ldots, h_A \in K_m$ such that

$$g_i - I_d \equiv \mathcal{P}^n X_i \mod \mathcal{P}^{2n},\ h_i - I_d \equiv \mathcal{P}^m Y_i \mod \mathcal{P}^{2m}$$

for $1 \leq i \leq A$.

As in the statement of Proposition 3.2.1, finding A independent of G yields an exponent C_2 in Theorem 3.1.6 independent of X_l. For $G = \mathrm{SL}_d(R)$, $\mathrm{SO}_d(R)$ or $\mathrm{Sp}_d(R)$, let $\mathfrak{g} = \mathfrak{sl}_d(R)$, $\mathfrak{so}_d(R)$ or $\mathfrak{sp}_d(R)$ be the associated Lie ring over R. Conditions (a), (b) above will follow straightforwardly from the following, which we verify for each group scheme in turn:

(a') For every $n \in \mathbb{N}$ and every $g \in K_n$, there exists $X \in \mathfrak{g}$ such that such that $g - I_d \equiv \mathcal{P}^n X \mod \mathcal{P}^{2n}$.

(b') There exists $A \in \mathbb{N}$ (independent of $\mathfrak{g}$) such that every element of $\mathfrak{g}$ is the sum of at most A brackets in $\mathfrak{g}$ (as we shall see, it suffices to take $A = 2$ for $\mathfrak{g} = \mathfrak{sl}_d(R)$ and $A = 3$ for $\mathfrak{g} = \mathfrak{so}_d(R)$ or $\mathfrak{sp}_d(R)$).

(c') There exists $\mathcal{B} \subseteq \mathfrak{g}$, generating $\mathfrak{g}$ as a $\mathbb{Z}$-module, such that for every $n \in \mathbb{N}$ and every $X \in \mathcal{B}$, there exists $g \in K_n$ such that such that $g - I_d \equiv \mathcal{P}^n X \mod \mathcal{P}^{2n}$.

For, given $g \in K_{n+m}$, we immediately produce X_i, Y_i as in (a) by applying (a'), (b') to g. Now writing an arbitrary element $Z \in \mathfrak{g}$ as $\sum_{i=1}^{r} Z_i$, for $Z_i \in \mathcal{B}$, and letting $k_1, \ldots, k_r \in K_l$ be such that $k_i - I_d \equiv \mathcal{P}^l Z_i \mod \mathcal{P}^{2l}$ as in (c'), we have:

$$I_d + \mathcal{P}^l Z \equiv k_1 \cdots k_r \in K_l \mod \mathcal{P}^{2l}.$$

Applying this observation to X_i, Y_i with $l = n, m$ respectively, we obtain g_i, h_i as in (b).

3.3.1 SL_d

Let $\mathfrak{sl}_d(R)$ denote the space of traceless $d \times d$ matrices over R; it is spanned over R by the matrices $E_{i,j}$, $D_{a,b}$, for $i \neq j$, $a < b$, where:

$$(E_{i,j})_{r,s} = \delta_{i,r}\delta_{j,s},\ (D_{a,b})_{r,s} = \delta_{a,r}\delta_{a,s} - \delta_{b,r}\delta_{b,s}.$$

(a') Let $g \in K_n$. Write $g = I_d + \mathcal{P}^n X$, for some $X \in \mathbb{M}_d(R)$. Then:

$$1 = \det(g) \equiv 1 + \mathcal{P}^n tr(X) \mod \mathcal{P}^{2n}$$

so $\mathrm{tr}(X) \equiv 0 \mod \mathcal{P}^n$. Hence there exists $X' \in \mathfrak{sl}_d(R)$ such that $X \equiv X' \mod \mathcal{P}^n$.

(b') First suppose $d \geq 3$. Define R-module endomorphisms $T_1, T_2 : \mathfrak{sl}_d(R) \to \mathfrak{sl}_d(R)$ by:

$$T_1(X) = \big(X, \sum_{i=1}^{d-1} E_{i+1,i}\big)$$
$$T_2(X) = \big(X, \sum_{i=1}^{d-1} E_{i,i+1}\big).$$

Then:

$$D_{j,j+1} = T_1(E_{j,j+1}) \text{ for } j = 1, \ldots, d-1,$$
$$E_{i,j-1} - E_{i+1,j} = T_1(E_{i,j}) \text{ for } 1 \leq i \leq d-1,\ i+2 \leq j \leq d,$$
$$E_{1,i+1} = T_1(-E_{i,1}) \text{ for } 2 \leq i \leq d-1,$$
$$E_{3,2} - 2E_{2,1} = T_1(D_{1,2}).$$

Transposing, we also have:

$$\{E_{i-1,j} - E_{i,j+1} : 1 \leq j \leq d-1,\, j+2 \leq i \leq d\}$$
$$\cup \{E_{j+1,1} : 2 \leq j \leq d-1\} \cup \{E_{2,3} - 2E_{1,2}\} \subseteq \mathrm{im}(T_2).$$

It may therefore be seen that $\mathrm{im}(T_1) \cup \mathrm{im}(T_2)$ contains an R-basis for $\mathfrak{sl}_d(R)$, so $\mathfrak{sl}_d(R) = \mathrm{im}(T_1) + \mathrm{im}(T_2)$. Now suppose $d = 2$ and $p > 2$. Then for any $a, b, c, \in R$,

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = (\begin{pmatrix} 0 & -b \\ c & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}) + (\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}).$$

(c') Let $\mathcal{B} = \{xE_{i,j} : x \in R, i \neq j\} \cup \{x(D_{a,b} + E_{a,b} - E_{b,a}) : x \in R, a \leq b\}$. Then $\mathcal{B}$ clearly spans $\mathfrak{sl}_d(R)$ and, for any $n \in \mathbb{N}$, $X \in \mathcal{B}$, $\det(I_d + \mathcal{P}^n X) = 1$.

Remark 3.3.1. *The preceding argument breaks down for $d = 2$, $p = 2$. Let $X, Y \in \mathbb{M}_2(R)$ with $\mathrm{tr}(X) = \mathrm{tr}(Y) = 0$. Then:*

$$(X, Y) \equiv (X_{12}Y_{21} - X_{21}Y_{12}) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mod \mathcal{P}.$$

Hence we cannot express an arbitrary traceless matrix as a sum of brackets, as we do above in higher rank or characteristic other than two.

3.3.2 SO_d

Denote by $\mathfrak{so}_d(R)$ the space of skew-symmetric $d \times d$ matrices over R; it is spanned over R by the matrices $X_{i,j} = E_{i,j} - E_{j,i}$, for $1 \leq i < j \leq d$.

(a') Let $g \in K_n$. Write $g = I_d + \mathcal{P}^n X$, for some $X \in \mathbb{M}_d(R)$. Then:

$$I_d = (I_d + \mathcal{P}^n X)(I_d + \mathcal{P}^n X^T) \equiv I_d + \mathcal{P}^n (X + X^T) \mod \mathcal{P}^{2n}$$

so $X^T \equiv -X \mod \mathcal{P}^n$. Hence there exists $X' \in \mathfrak{so}_d$ such that $X \equiv X' \mod \mathcal{P}^n$.

(b') Define the R-module endomorphisms $T_1, T_2, T_3 : \mathfrak{so}_d(R) \to \mathfrak{so}_d(R)$ by:

$$T_1(X) = (X, \sum_{i=1}^{d-1} X_{i,i+1})$$
$$T_2(X) = (X, X_{1,d-1} + X_{1,d} + X_{2,d})$$
$$T_3(X) = (X, X_{1,2}).$$

Then for $1 < i < d-1$,

$$X_{i,i+2} - X_{i-1,i+1} = T_1(X_{i,i+1});\ X_{i+1,d} - X_{i-1,d} - X_{i,d-1} = T_1(X_{i,d}).$$

For $1 < j < d-1$,

$$X_{2,j} + X_{1,j+1} - X_{1,j-1} = T_1(X_{1,j}).$$

For $1 < i, j < d$, with $i+1 < j$,

$$X_{i+1,j} + X_{i,j+1} - X_{i-1,j} - X_{i,j-1}.$$

For $3 \leq j \leq d-2$,

$$X_{1,j} = T_2(X_{j,d-1});\ X_{j,d} = T_2(-X_{2,j})$$

and:

$$X_{1,2} = T_2(-X_{2,d});\ X_{d-1,d} = T_2(X_{1,d-1});\ X_{1,d-1} = T_3(-X_{2,d-1});$$
$$X_{1,d} = T_3(-X_{2,d});\ X_{2,d} = T_3(-X_{1,d}).$$

Therefore $\mathrm{im}(T_1) \cup \mathrm{im}(T_2) \cup \mathrm{im}(T_3)$ contains an R-basis for $\mathfrak{so}_d(R)$, so $\mathfrak{so}_d(R) = \mathrm{im}(T_1) + \mathrm{im}(T_2) + \mathrm{im}(T_3)$.

(c') For $\alpha \in R$, $l \in \mathbb{N}$, consider the polynomial $f(X) = X^2 - (1 - \alpha^2\mathcal{P}^{2l})$. Then $f(1) = \alpha^2\mathcal{P}^{2l} \equiv 0 \mod \mathcal{P}^{2l}$ but $f'(1) = 2 \not\equiv 0 \mod \mathcal{P}$. By Hensel's Lemma, there exists $\beta \in R$ such that $f(\beta) = 0$ and $\beta \equiv 1 \mod \mathcal{P}^{2l}$. Hence for any $i \neq j$,

$$g_{i,j}^{(l)}(\alpha) := I_d + \alpha\mathcal{P}^l(E_{i,j} - E_{j,i}) + (\beta - 1)(E_{i,i} + E_{j,j}) \in K_l$$

and $g_{i,j}^{(l)}(\alpha) \equiv I_d + \alpha\mathcal{P}^l(E_{i,j} - E_{j,i}) \mod \mathcal{P}^{2l}$.

Remark 3.3.2. *In contrast to the cases of* $\mathrm{SL}_d(R)$ *and* $\mathrm{Sp}_d(R)$, $\mathrm{SO}_d(R)$ *is not in general the simply connected form of the Chevalley group of its type; this is rather a proper central extension of* $\mathrm{SO}_d(R)$ *by a finite group. Increasing the constant* C_1 *in Theorem 3.1.6, the diameter bounds obtained above for* SO_d *extend to the simply connected form by Lemma 2.1.5 (ii).*

3.3.3 Sp_d

Let $d = 2g$ and let $\Omega = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}$, so that $\mathfrak{sp}_d(R)$ is the set of $d \times d$ matrices X over R satisfying the relation $X^T\Omega + \Omega X = 0$. Suppose $p > 2$.

For $1 \leq i, j \leq g$, define the matrices:

$$A_{i,j} = \begin{pmatrix} E_{i,j} & 0 \\ 0 & -E_{j,i} \end{pmatrix}, \; B_{i,j} = \begin{pmatrix} 0 & E_{i,j} + E_{j,i} \\ 0 & 0 \end{pmatrix},$$
$$C_{i,j} = \begin{pmatrix} 0 & 0 \\ E_{i,j} + E_{j,i} & 0 \end{pmatrix} \in \mathfrak{sp}_d(R).$$

We have:

$$(A_{i,j}, A_{k,l}) = \delta_{j,k}A_{i,l} - \delta_{i,l}A_{k,j},$$
$$(A_{i,j}, B_{k,l}) = \delta_{j,k}B_{i,l} + \delta_{j,l}B_{i,k},$$
$$(A_{i,j}, C_{k,l}) = -\delta_{i,l}C_{j,k} - \delta_{i,k}C_{j,l},$$
$$(B_{i,j}, C_{k,l}) = \delta_{j,k}A_{i,l} + \delta_{j,l}A_{i,k} + \delta_{i,k}A_{j,l} + \delta_{i,l}A_{j,k}.$$

Hence:

$$A_{i,j} = (A_{i,j}, A_{j,j}), \text{ for } i \neq j,$$
$$A_{i,i} = (\tfrac{1}{2}B_{i,i}, \tfrac{1}{2}C_{i,i}),$$
$$B_{i,j} = (A_{i,k}, B_{k,j}), \text{ for } i \neq k \neq j,$$
$$C_{i,j} = (-A_{k,i}, C_{j,k}), \text{ for } i \neq k \neq j.$$

(a') Let $g \in K_n$. Write $g = I_d + \mathcal{P}^n X$, for some $X \in \mathbb{M}_d(R)$. Then:

$$\begin{aligned} \Omega &= g^T\Omega g \\ &= \Omega + \mathcal{P}^n(\Omega X + X^T\Omega) + \mathcal{P}^{2n}X^T\Omega X \\ &\equiv \Omega + \mathcal{P}^n(\Omega X + X^T\Omega) \mod \mathcal{P}^{2n} \end{aligned}$$

so $\Omega X + X^T\Omega \equiv 0 \mod \mathcal{P}^n$. Hence there exists $X' \in \mathfrak{sp}_d(R)$ such that $X \equiv X' \mod \mathcal{P}^n$.

(b') Define the R-module endomorphisms $U_1, U_2 : \mathfrak{sp}_d(R) \to \mathfrak{sp}_d(R)$ by:

$$U_1(X) = \big(X, \sum_{i=1}^{g} A_{i,i}\big)$$
$$U_2(X) = \big(X, \sum_{i=1}^{g} (B_{i,i} + C_{i,i})\big).$$

Then for any $1 \leq i, j \leq g$, $B_{i,j}, C_{i,j} \in \mathrm{im}(U_1)$, $A_{i,j} + A_{j,i} \in \mathrm{im}(U_2)$. Define the R-Lie subring $V \leq \mathfrak{sp}_d(R)$:

$$V = \{\begin{pmatrix} X & 0 \\ 0 & -X^T \end{pmatrix} : X \in \mathfrak{gl}_g(R)\}.$$

We show that, for any $X \in \mathfrak{so}_g(R)$, there exist $v_1, v_2 \in V$ and symmetric $Z \in \mathfrak{gl}_g(R)$ such that:

$$\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} = (v_1, v_2) + \begin{pmatrix} Z & 0 \\ 0 & -Z \end{pmatrix}.$$

Now for an arbitrary element $v \in \mathfrak{sp}_d(R)$ there exist $X \in \mathfrak{so}_g(R)$ and symmetric $B, C, Y \in \mathfrak{gl}_g(R)$ such that:

$$\begin{aligned} v &= \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} + \begin{pmatrix} Y & 0 \\ 0 & -Y \end{pmatrix} + \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \\ &= (v_1, v_2) + \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} + \begin{pmatrix} Y+Z & 0 \\ 0 & -(Y+Z) \end{pmatrix} \end{aligned}$$

and $\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \in \mathrm{im}(U_1)$, $\begin{pmatrix} Y+Z & 0 \\ 0 & -(Y+Z) \end{pmatrix} \in \mathrm{im}(U_2)$, so that every element of $\mathfrak{sp}_d(R)$ is expressible as a sum of three brackets.

It will suffice to check that any element of $\mathfrak{so}_g(R)$ is expressible as the sum of a bracket in $\mathfrak{gl}_g(R)$ and a symmetric matrix. Define the R-module endomorphisms $S_1, S_2 : \mathfrak{gl}_g(R) \to \mathfrak{gl}_g(R)$ by:

$$\begin{aligned} S_1(X) &= (X, E_{1,1}) \\ S_2(X) &= \big(X, \sum_{i=1}^{d-1}(E_{i,i+1} - E_{i+1,i})\big) \end{aligned}$$

and for $X \in \mathfrak{gl}_g(R)$, write $X = X_1 + X_2$, with X_1 symmetric, X_2 skew-symmetric. Then:

$$\big(X, E_{1,1} + \sum_{i=1}^{d-1}(E_{i,i+1} - E_{i+1,i})\big) - S_1(X_1) - S_2(X_2)$$

is symmetric. We already described the image of $S_2 \mid_{\mathfrak{so}_g(R)}$ (in the guise of T_1 in our analysis of SO_d). For $2 \leq i \leq g$,

$$S_1(E_{1,i} + E_{i,1}) = E_{i,1} - E_{1,i}.$$

These elements, together with $im(S_2 \mid_{\mathfrak{so}_g(R)})$, span $\mathfrak{so}_g(R)$ over R, and the result follows.

(c') For any $\alpha \in R$ and any $l \in \mathbb{N}$ we have:

$$I_d + \alpha\mathcal{P}^l B_{i,j}, I_d + \alpha\mathcal{P}^l C_{i,j} \in K_l \text{ for any } 1 \leq i, j, \leq d.$$
$$I_d + \alpha\mathcal{P}^l A_{i,j} \in K_l \text{ provided } i \neq j.$$

Finally, $(1 + \alpha\mathcal{P}^l)^{-1} \equiv 1 - \alpha\mathcal{P}^l \mod \mathcal{P}^{2l}$, so:

$$K_l \ni I + \begin{pmatrix} \alpha\mathcal{P}^l E_{i,i} & 0 \\ 0 & ((1 + \alpha\mathcal{P}^l)^{-1} - 1)E_{i,i} \end{pmatrix} \equiv I + \alpha t^{\mathcal{P}} A_{i,i} \mod \mathcal{P}^{2l}.$$

Remark 3.3.3. *For $R = \mathbb{Z}_p$, the value $A = 3$ was achieved in [29], under the additional assumption that $p \geq \frac{l+2}{2}$, where l is the rank of the associated Chevalley group scheme. This assumption was necessary in the specific manipulations of the root systems which were applied in Dinai's argument. Hence even in the p-adic case, the results of this section are new in large rank for small p.*

3.4 Diameter in Analytic Pro-p Groups

In this section we prove Theorem 3.1.3, using the results of Section 2.3.3. Then we deduce Theorems 3.1.5 and 3.1.6. First we require:

Proposition 3.4.1. *Let G be a d-dimensional R-standard group. Suppose $\mathcal{L}_G$ is perfect. There exists $k \in \mathbb{N}$ such that every element of $(\mathcal{M}^k)^{(d)}$ is expressible as a sum of at most d brackets in L_G.*

Proof. Let $\{x_1, \ldots, x_d\}$ be a R-basis for $L(G)$. Then there exist $r_i, s_i \in \{1, \ldots, d\}$ such that $\{(x_{r_1}, x_{s_1}), \ldots, (x_{r_d}, x_{s_d})\}$ is a $\mathbb{K}$-basis for $\mathcal{L}_G$. Let $\lambda_{i,j} \in \mathbb{K}$ be such that:

$$x_i = \sum_{j=1}^{d} \lambda_{i,j}(x_{r_j}, x_{s_j}).$$

Let $k \in \mathbb{N}$ be defined by:

$$\|\mathcal{P}\|^{-k} = \max(\{1\} \cup \{\|\lambda_{i,j}\| : 1 \leq i, j \leq d\}).$$

Then for any $1 \leq i, j \leq d$, $\mathcal{P}^k \lambda_{i,j} \in R$. Hence for any $x \in L(G)$, there exist $\mu_1, \ldots, \mu_d \in R$ such that:

$$\begin{aligned}
\mathcal{P}^k x &= \mathcal{P}^k \sum_{i=1}^{d} \mu_i x_i \\
&= \mathcal{P}^k \sum_{i=1}^{d} \mu_i \sum_{j=1}^{d} \lambda_{i,j}(x_{r_j}, x_{s_j}) \\
&= \sum_{j=1}^{d} (\sum_{i=1}^{d} \mu_i \mathcal{P}^k \lambda_{i,j} x_{r_j}, x_{s_j})
\end{aligned}$$

as required. □

Proof of Theorem 3.1.3. We verify the hypotheses of Proposition 3.2.5. Hypothesis (i) is Proposition 2.3.20 (ii). For hypothesis (ii), we take ϵ arbitrary; $A = d$; $M_1 \geq \max\{\frac{k}{3} + 1, 2\}$; $M_2 = k$, where k is as in Proposition 3.4.1. For $i = 1, 2, 3$ choose $\frac{n}{3}(2 + i + \epsilon) \leq n_i \leq m_i \leq \frac{2n}{3}(2 + i)$ such that $n_i + m_i = (2 + i)n - M_2$ (this is possible by our choice of M_1, M_2).

Let $g \in K_{(2+i)n}$. Let $h \in K_{M_2}$ be such that $g = \mathcal{P}^{n_i + m_i} h$. Then by Proposition 3.4.1 there exist $g_1, \ldots, g_d, h_1, \ldots, h_d \in G$ such that:

$$h = \sum_{i=1}^{d} (g_i, h_i)$$

so that:

$$\begin{aligned}
g &= \sum_{i=1}^{d} (\mathcal{P}^{n_i} g_i, \mathcal{P}^{m_i} h_i) \\
&\equiv \sum_{i=1}^{d} [\mathcal{P}^{n_i} g_i, \mathcal{P}^{m_i} h_i] \mod \mathcal{P}^{2n_i + m_i} \text{ (by Proposition 2.3.17 (iii))} \\
&\equiv [\mathcal{P}^{n_i} g_1, \mathcal{P}^{m_i} h_1] \cdots [\mathcal{P}^{n_i} g_d, \mathcal{P}^{m_i} h_d] \mod \mathcal{P}^{2n_i + 2m_i} \text{ (by Proposition 2.3.17 (i))}.
\end{aligned}$$

Since $2n_i + m_i = (2 + i)n + n_i - M_2$, we are done. □

3.4.1 FAb p-adic analytic groups

Proposition 3.4.2. *Let G be a d-dimensional uniform pro-p group. The following are equivalent:*

(i) G has finite abelianisation;

(ii) G is FAb;

(iii) $\mathcal{L}_G$ is perfect.

Recall that a profinite group G is defined to be *FAb* if it has an open subgroup with infinite abelianisation. This proposition appears to be well-known to the experts, but the author is not aware of an existing reference for the proof.

Proof. $(ii) \Rightarrow (i)$ is clear.

For $(iii) \Rightarrow (ii)$, suppose $H \leq_o G$ is such that there exists an epimorphism $\phi : H \twoheadrightarrow \mathbb{Z}_p$. We may suppose that $H = G^{p^n}$ for some $n \in \mathbb{N}$. For if $h \in H$ is such that $\mathbb{Z}_p = \overline{\langle \phi(h) \rangle}$, and $n \in \mathbb{N}$ is such that $G^{p^n} \leq H$, then $h^{p^n} \in G^{p^n}$, and $p^n\mathbb{Z}_p = \overline{\langle \phi(h^{p^n}) \rangle} \leq \phi(G^{p^n}) \leq \mathbb{Z}_p$, so $\phi(G^{p^n}) \leq_o \mathbb{Z}_p$, and $\phi(G^{p^n}) \cong \mathbb{Z}_p$.

Now let $N = \ker(\phi)$, so that by Proposition 2.3.41 (ii), $N \lhd_c H$ is uniform of dimension $d-1$; $\log(H) = p^n \log(G)$ and $\log(H)/\log(N) \cong \mathbb{Z}_p$.

Hence $\mathcal{L}_H = \mathcal{L}_G$ so $\mathcal{L}_G/\mathcal{L}_N \cong \mathbb{Q}_p$, and $\mathcal{L}_G$ is not perfect.

For $(i) \Rightarrow (iii)$, suppose $\mathcal{I} \lhd \mathcal{L}_G$, with $\dim(\mathcal{I}) = d-1$. Let $I = \log(G) \cap \mathcal{I} \lhd \log(G)$ (so that $\mathcal{I} = \mathrm{span}_{\mathbb{Q}_p}(I)$). Let $v \in \log(G)$, and suppose there exists $\lambda \in \mathbb{Z}_p \setminus \{0\}$ such that $\lambda v \in I$. Then $v = \lambda^{-1}(\lambda v) \in \mathcal{I}$, so $v \in \mathcal{I} \cap \log(G) = I$. Thus $\log(G)/I$ is torsion-free, so by Proposition 2.3.42, $\exp(I) \lhd G$ is uniform and $G/\exp(I)$ is uniform, with:

$$\begin{aligned} \dim(G/\exp(I)) &= \dim(G) - \dim(\exp(I)) \\ &= \mathrm{rk}(\log(G)) - \mathrm{rk}(I) \\ &= \dim(\mathcal{L}_G) - \dim(\mathcal{I}) \\ &= 1. \end{aligned}$$

and a 1-dimensional uniform group is by definition infinite procyclic, so $G/\exp(I) \cong \mathbb{Z}_p$. □

Proof of Theorem 3.1.5. First suppose that G is a FAb compact p-adic analytic group. As noted in Corollary 2.3.34, G has an open characteristic uniform subgroup H. By Remark 2.3.40, H_2 is $\mathbb{Z}_p$-standard. Let $K_n \lhd_o H_2$ be as in Theorem 3.1.3. Then by Remark 2.3.40 and Theorem 2.3.36 (iii),

$$K_n = (H_2)_n = H_{n+1}$$

and H_{n+1} is a characteristic subgroup of H. In particular, $K_n \lhd_o G$. Since G is FAb, so is H_2, hence by Proposition 3.4.2 and Theorem 2.3.36 (ii), $\mathcal{L}_{H_2}$ is perfect. As in the proof of Theorem 3.1.3, $(K_n)_n$ satisfies the hypotheses of Proposition 3.2.5 and the result follows.

Now suppose that G is uniform and not FAb. By Proposition 3.4.2, there exists an epimorphism $\phi : G \twoheadrightarrow \mathbb{Z}_p$. By Proposition 2.3.41, $N = \ker(\phi)$ is uniform of dimension $d-1$. We may therefore choose a generating set $S = \{a_1, \ldots, a_d\}$ for G such that $\{a_1, \ldots, a_{d-1}\}$ is a generating set for N and $\overline{\langle \phi(a_d) \rangle} = \mathbb{Z}_p$.

Let $\pi_n : \mathbb{Z}_p \twoheadrightarrow \mathbb{Z}/p^n\mathbb{Z}$ be the natural projection. Then $G_{n+1} \subseteq \ker(\pi_n \circ \phi)$, so:

$$\operatorname{diam}(G/G_{n+1}, S) \geq \operatorname{diam}(\mathbb{Z}/p^n\mathbb{Z}, \{(\pi_n \circ \phi)(a_d)\}) \geq Cp^n = C|G/G_{n+1}|^{\frac{1}{d}}.$$

In particular $\operatorname{diam}(G/G_{n+1}, S)$ is not polylogarithmic in $|G/G_{n+1}|$. □

3.4.2 Exceptional groups

With Theorem 3.1.3 in hand we may complete the proof of Theorem 3.1.6. To avoid cluttered notation, we write $G_{\mathrm{sc}} = \mathcal{G}_R(X_l)$ and $G_{\mathrm{ad}} = \mathcal{G}_R^{\mathrm{ad}}(X_l)$.

Proof of Theorem 3.1.6. If $X_l \in \{A_l, B_l, C_l, D_l\}$, G_{ad} is one of $\mathrm{PSL}_d(R)$, $\mathrm{PSO}_d(R)$ or $\mathrm{PSp}_d(R)$. The result then follows as in Section 3.3. If not, then letting G_1 be as in Theorem 2.4.14, G_1 satisfies the hypothesis of Theorem 3.1.3, so that for some $\tilde{C}_1, C_2 > 0$,

$$\operatorname{diam}(G_{\mathrm{sc}}/G_n) \leq \tilde{C}_1(\log|G_{\mathrm{sc}}/G_n|)^{C_2}.$$

The map $\rho \circ \psi : G_{\mathrm{sc}} \twoheadrightarrow G_{\mathrm{ad}}$ descends, by Corollary 2.4.15, to an epimorphism $G_{\mathrm{sc}}/G_n \twoheadrightarrow G_{\mathrm{ad}}/K_n$. By Lemma 2.1.5 (i),

$$\operatorname{diam}(G_{\mathrm{ad}}/K_n) \leq \operatorname{diam}(G_{\mathrm{sc}}/G_n).$$

Finally, $|G_{\mathrm{sc}}/G_n| \ll_{R,X_l} |R/\mathcal{M}|^{dn}$ and $|G_{\mathrm{ad}}/K_n| \geq |R/\mathcal{M}|^n$, so

$$(\log|G_{\mathrm{sc}}/G_n|)^{C_2} \ll_{R,X_l} (\log|G_{\mathrm{ad}}/K_n|)^{C_2}$$

and the result follows (replacing $\tilde{C}_1$ by some larger constant C_1).

The bound we thus obtain for C_2 is independent of X_l, since we need only apply Theorem 3.1.3 for finitely many types X_l. □

Remark 3.4.3. *(i) The method of this section is also applicable to the classical Chevalley groups, though does not yield uniformity in the exponent C_2. In particular we obtain a diameter bound in the case $(X_l, p) = (B_l, 2)$ or $(D_l, 2)$, which does not fall under the purview of Theorem 3.1.6. The case $(X_l, p) = (A_1, 2)$ or $(C_l, 2)$ is beyond the scope of our methods, however, because the associated Lie algebras may not be perfect.*

(ii) The best degree C_2 in Theorem 3.1.6 which we can obtain by the above method is based on taking $A = 248$ in Proposition 3.2.5, because 248 is the dimension of $\mathcal{G}_R(X_l)$ as an R-analytic group in the case $X_l = E_8$. It is likely that this is far from optimal, and that a much lower degree could be obtained via a more direct analysis of the Lie algebras of the exceptional groups, akin to that employed for the classical groups in Section 3.3. In the case $R = \mathbb{Z}_p$, this has already largely been achieved by Dinai in [29]: he showed that for $p > 19$, every element of the $\mathbb{Z}_p$-Lie ring associated to an exceptional group can be expressed as the sum of three brackets.

3.5 Diameter in the Nottingham Group

We deduce Theorem 3.1.7 from Proposition 3.2.5. Hypothesis (i) of Proposition 3.2.5 is immediate from Lemma 2.5.4 (iii).

We shall show that, provided $p \geq 3$, for $n \leq m \leq 2n$ satisfying $p \nmid (m - n)$ every element of K_{m+n} may be expressed, modulo K_{m+2n}, as $[g_1, h_1][g_2, h_2]$ for some $g_i \in K_m$, $h_i \in K_n$. Now, for any $\epsilon \in (0, 1)$, $n \geq 5$ and $i = 1, 2, 3$, there exist $n_i, m_i \in \mathbb{N}$ such that $n_i + m_i = (2 + i)n$; $\frac{n}{3}(2 + i + \epsilon) \leq n_i \leq m_i \leq \frac{2n}{3}(2 + i)$ and $m_i - n_i \in \{1, 2\}$. We therefore satisfy hypothesis (ii) of Proposition 3.2.5 with ϵ arbitrary; $A = 2$; $M_1 = 5$; $M_2 = 0$.

For any $\lambda_i, \nu \in \mathbb{F}_q$; $K, M, N \in \mathbb{N}$ with $N \leq M$; applying Lemma 2.5.4 (iii) and an easy induction, we have:

$$[g, e_{M,\nu}] \equiv [e_{N,\lambda_1}, e_{M,\nu}] \cdots [e_{N+K-1,\lambda_K}, e_{M,\nu}] \mod K_{2N+M+1}$$

where $g = e_{N,\lambda_1} \cdots e_{N+K-1,\lambda_K}$. Moreover, by Lemma 2.5.3 (i) and Lemma 2.5.4 (ii),

$$\begin{aligned}[e_{N+i,\lambda_{i+1}}, e_{M,\nu}] &\equiv (e_{M+N+i,\lambda_{i+1}\nu})^{(N-M)+i} \mod K_{2N+M+2i}K_{N+2M+i} \\ &\equiv e_{M+N+i,\lambda_{i+1}\nu((N-M)+i)} \mod K_{2M+2N+2i}.\end{aligned}$$

Hence for any $\lambda_i, \mu_i \in \mathbb{F}_q$, setting:

$$g_1 = e_{n,\lambda_1} \cdots e_{2n-1,\lambda_n},\ g_2 = e_{n,\mu_1} \cdots e_{2n-2,\mu_{n-1}}$$

we have:

$$\begin{aligned}[g_1, e_{m,1}][g_2, e_{m+1,1}] &\equiv (\prod_{i=0}^{n-1} e_{n+m+i,\lambda_{i+1}(n+i-m)})(\prod_{i=1}^{n-1} e_{n+m+i,\mu_i(n-m-2+i)}) \\ &\equiv e_{n+m,\lambda_1(n-m)}(\prod_{i=1}^{n-1} e_{n+m+1,\lambda_{i+1}(n-m+i)+\mu_i(n-m-2+i)}) \mod K_{2n+m}\end{aligned}$$

since K_{n+m}/K_{2n+m} is abelian. $p \nmid (n-m)$, and since $p \geq 3$, for each $1 \leq i \leq n-1$, p divides at most one of $n-m+i, n-m-2+i$. Hence by varying the λ_i and μ_i, using the form described in Lemma 2.5.3 (ii), we can express any element of K_{n+m} modulo K_{2n+m}.

3.6 Limit Theorems for Random Walks

The purpose of this section is to prove Corollaries 3.1.8, 3.1.9 and 3.1.10. Recall the relationships between diameter, mixing of random walks and the spectral gap $1-\rho$ of the adjacency operator A_S for a pair (G, S) given in Lemma 2.1.3 and Proposition 2.1.6:

$$|\langle A_S^l \chi_g, \chi_h \rangle - \frac{1}{|G|}| \leq \rho^l \tag{3.7}$$

$$\frac{\operatorname{diam}(G,S)-1}{\log|G|} \leq \frac{1}{1-\rho} \leq |S| \operatorname{diam}(G,S)^2 \tag{3.8}$$

In particular, for $\operatorname{diam}(G,S) \leq C_1 \log^{C_2}|G|$,

$$1-\rho \geq \frac{1}{|S|C_1^2 \log^{2C_2}|G|}$$

by (3.8). Setting $C_3 = |S|C_1^2$, and applying (3.7), we have:

$$|\langle A_S^l \chi_g, \chi_h \rangle - \frac{1}{|G|}| \leq (1 - \frac{1}{C_3 \log^{2C_2}|G|})^l.$$

Recall that $(1-\frac{1}{x})^x$ is an increasing function for $x > 1$, converging to e^{-1} as $x \to \infty$. Hence, for any $C_4 > 0$ and taking $l \geq C_3 \log^{2C_2+C_4}|G|$, we deduce:

$$|\langle A_S^l \chi_g, \chi_h \rangle - \frac{1}{|G|}| \leq e^{-\log^{C_4}|G|}.$$

Proof of Corollary 3.1.8. We may identify:

$$G/K_{N+1} \cong \{\lambda_1 x_1 + \cdots + \lambda_d x_d : \lambda_1, \ldots, \lambda_d \in R/\mathcal{M}^N\} \cong (R/\mathcal{M}^N)^d,$$

as a set, so $|G/K_{N+1}| = |R/\mathcal{M}|^{dN}$ and:

$$\left|\mathbb{P}\big[\|L_1^{(l)} - \lambda_1\|, \ldots, \|L_d^{(l)} - \lambda_d\| \leq c^{N+1}\big] - \frac{1}{|R/\mathcal{M}|^{dN}}\right| = \left|\langle A_S^l \chi_e, \chi_g\rangle - \frac{1}{|G/K_{N+1}|}\right|$$

where $g = \lambda_1 x_1 + \cdots + \lambda_d x_d \in G/K_{N+1}$. The result is now a consequence of Theorem 3.1.3 and the discussion following (3.7) and (3.8), taking:

$$C = 2C_2, C' = C_4, C'' = C_3(d \cdot \log|R/\mathcal{M}|)^{2C_2+C_4}, C''' = (d \cdot \log|R/\mathcal{M}|)^{C_4}.$$

□

Proof of Corollary 3.1.9. By Theorem 2.3.36,

$$G/K_{N+1} \cong \langle K_{N+1} a_1\rangle \times \ldots \times \langle K_{N+1} a_d\rangle \cong (\mathbb{Z}/p^N\mathbb{Z})^d,$$

as a set, so $|G/K_{N+1}| = p^{dN}$ and:

$$\left|\mathbb{P}\big[\|M_1^{(l)} - \mu_1\|, \ldots, \|M_d^{(l)} - \mu_d\| \leq p^{-N-1}\big] - \frac{1}{p^{dN}}\right| = \left|\langle A_S^l \chi_e, \chi_g\rangle - \frac{1}{|G/K_{N+1}|}\right|$$

where $g = K_{N+1} a_1^{\mu_1} \cdots a_d^{\mu_d} \in G/K_{N+1}$. The result now follows from Theorem 3.1.5 and the discussion following (3.7) and (3.8), taking:

$$C = 2C_2, C' = C_4, C'' = C_3(d \cdot \log p)^{2C_2+C_4}, C''' = (d \cdot \log p)^{C_4}.$$

□

Proof of Corollary 3.1.10. Letting $G_N = \mathcal{N}_q/K_N$, $|G_N| = q^{N-1}$, so:

$$\left|\mathbb{P}\big[A_2^{(l)} = \alpha_2, \ldots, A_N^{(l)} = \alpha_N\big] - \frac{1}{q^{N-1}}\right| = \left|\langle A_S^l \chi_e, \chi_g\rangle - \frac{1}{|G_N|}\right|,$$

where $g = t + \sum_{i=2}^N \alpha_i t^i$. The result follows from Theorem 3.1.7 and the discussion following (3.7) and (3.8), taking:

$$C = 2C_2, C' = C_4, C'' = C_3(\log q)^{2C_2+C_4}, C''' = (\log q)^{C_4}.$$

□

Chapter 4

Expansion, Random Walks and Sieving in $\mathrm{SL}_2(\mathbb{F}_p[t])$

4.1 Introduction

The goal of this chapter shall be to prove our results on the escape of random walks on $\mathrm{SL}_2(\mathbb{F}_p[t])$ from subsets of various types, and the constructions of expander congruence quotients which underpin them.

4.1.1 Statement of results

Recall that we have bounds on the return probability to algebraic subvarieties; the set of squares and the set of elements with reducible characteristic polynomial, as follows:

Theorem 4.1.1. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric subset, generating a non-elementary subgroup. Let $F : \mathbb{M}_2(\mathbb{F}_p[t])^r \to \mathbb{F}_p[t]$ be a polynomial over $\mathbb{F}_p[t]$ which does not vanish on $\mathrm{SL}_2(\mathbb{F}_p[t])^r$. Then there exist $C_1(F)$, $C_2(S) > 0$ such that, letting $V(F) \subseteq \mathbb{M}_2(\mathbb{F}_p[t])^r$ be the affine algebraic subvariety of $\mathrm{SL}_2(\mathbb{F}_p[t])^r$ defined by F,*

$$(\times_{i=1}^{r} \mu_S^{(l)})(V(F)) \leq C_1 e^{-C_2 l}.$$

Theorem 4.1.2. *Let S be as in Theorem 4.1.1. There exist $C_1, C_2(S) > 0$ such that:*

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : g = h^2 \text{ for some } h \in \mathrm{SL}_2(\mathbb{F}_p[t])\}) \leq C_1 e^{-C_2\sqrt{l/\log l}}.$$

Remark 4.1.3. *More could be said about proper powers. We could, for instance, easily strengthen the proof of Theorem 4.1.2 to show that $\mu_S^{(l)}$ escapes from the sets of mth powers in $SL_2(\mathbb{F}_p[t])$, for all $m \in \mathbb{N}$ satisfying $p \equiv 1 \mod m$, simultaneously. However, absent an application, we shall not rehearse the details of such an argument.*

Theorem 4.1.4. *Let S be as in Theorem 4.1.1. There exist $C_1, C_2(S) > 0$ such that:*

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : \chi_g \text{ is reducible } \}) \leq C_1 e^{-C_2\sqrt{l/\log l}}.$$

Now recall our results on expanders:

Definition 4.1.5. *For $M > 0$, an integer $n > 1$ will be called M-*rough *if n has no prime factor less than M. A polynomial $f \in \mathbb{F}_p[t]$ will be called M-rough if the degree of every irreducible factor of f is a M-rough integer.*

Example 4.1.6. *Let $M > 0$.*

(i) Every prime $> M$ is M-rough.

(ii) There is a sequence $(n_i)_i$ of M-rough integers growing linearly in i. For, given M, let π be the set of all primes up to M. Let $n_i = Ni + 1$, where $N = \prod_{P\in\pi} P$. It will be significant in the applications in Section 4.3 that the set of rough integers is sufficiently dense.

Theorem 4.1.7. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric subset, generating a non-elementary subgroup. Suppose every entry of every element of S has degree at most D. Let $(f_i)_i \subseteq \mathbb{F}_p[t]$ be a sequence of distinct polynomials. Then there exists $M > 0$ (depending on D and p) such that, if $(f_i)_i$ are M-rough then for $i_0 \in \mathbb{N}$ sufficiently large (depending on D, p), $(\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i)), \pi_{f_i}(S))_{i\geq i_0}$ is a two-sided expander family, provided one of the following holds:*

(i) The f_i are irreducible;

(ii) The f_i are square-free, every irreducible factor of every f_i has prime degree, and no two irreducible factors of any f_i have the same degree.

Theorem 4.1.8. *For any $C > 0$ and any $k \in \mathbb{N}_{\geq 2}$, there exists $M > 0$ (depending on k, p and C) such that, if $(n_i)_i$ a sequence of M-rough positive integers, $S_{n_i} \subseteq \mathrm{SL}_2(p^{n_i})$ is symmetric with $|S_{n_i}| = 2k$, and $\mathrm{girth}(\mathrm{SL}_2(p^{n_i}), S_{n_i}) \geq Cn_i$, for all $i \in \mathbb{N}$, then for i_0 sufficiently large (depending on C, k), $(\mathrm{SL}_2(p^{n_i}), S_{n_i})_{i\geq i_0}$ is a two-sided expander family.*

Let us at this point say a quick word about the roughness hypothesis in Theorems 4.1.7 and 4.1.8. These results shall both be proved using the Bourgain-Gamburd machine; specifically they shall follow from Theorem 2.2.22. Recall that one key component of the machine is that the random walk should escape rapidly from proper subgroups. One particular family of subgroups of $\mathrm{SL}_2(p^n)$ with which we must contend

is the family of *subfield subgroups*: those obtained by restricting coefficients to a proper subfield of $\mathbb{F}_{p^n}$. The purpose of imposing roughness on the integer n is that it ensures all proper subfields of $\mathbb{F}_{p^n}$ are small, so that the associated subfield subgroups are themselves too small to contain the random walk. This hypothesis places a substantial restriction on the finite groups which are susceptible to our methods, but as noted in Example 4.1.6, there are still enough rough polynomials to "sieve out" the subsets from which the random walk is to escape in Theorems 4.1.1, 4.1.2 and 4.1.4.

The chapter is structured as follows: in Section 4.2 we prove Theorems 4.1.7 and 4.1.8. Specifically, Section 4.2.1 shall deal with hypotheses (i) and (ii) of Theorem 2.2.22 and further reduce Theorem 4.1.7 to the case of non-abelian free subgroups. We then turn to hypothesis (iii) of Theorem 2.2.22. In Section 4.2.2 it is verified for Cayley graphs of $\mathrm{SL}_2(p^n)$ with large girth under the roughness hypothesis. This yields Theorem 4.1.7 (i) and Theorem 4.1.8. The generalisation of this argument required for Theorem 4.1.7 (ii) is explained in Section 4.2.3.

We discuss the applications to random walks in $\mathrm{SL}_2(\mathbb{F}_p[t])$ in Section 4.3. In Section 4.3.1 we explain in general terms how non-concentration results in infinite groups can be obtained via the group sieve method, using expansion results on finite quotients. Theorems 4.1.1, 4.1.2 and 4.1.4 are proved in the subsequent three sections.

4.1.2 Further questions

It is natural to try to weaken the roughness hypothesis in Theorems 4.1.7 and 4.1.8. However there are some significant obstacles to doing so. For instance it is clear that Theorem 4.1.8 does not remain true for arbitrary sequences $(n_i)_i$:

Example 4.1.9. *Let $n_i = 2^i$. Then we may identify $\mathbb{F}_{p^{n_i}}$ with a proper subfield of $\mathbb{F}_{p^{n_{i+1}}}$, and hence embed $\mathrm{SL}_2(p^{n_i}) \hookrightarrow \mathrm{SL}_2(p^{n_{i+1}})$. For i even, let S_{n_i} be a generating set for $\mathrm{SL}_2(p^{n_i})$ satisfying $\mathrm{girth}(\mathrm{SL}_2(p^{n_i}), S_{n_i}) \gg n_i$. For i odd, let $S_{n_i} = S_{n_{i-1}}$, so that $\langle S_{n_i} \rangle \lneq \mathrm{SL}_2(p^{n_i})$. Then for every i, $\mathrm{girth}(\mathrm{SL}_2(p^{n_i}), S_{n_i}) \gg n_i$, but $\{(\mathrm{SL}_2(p^{n_i}), S_{n_i})\}_{i \geq j}$ is not an expander family for any j.*

So the presence of large subfield subgroups presents a genuine obstruction to expansion of subsets. It should be noted however that Example 4.1.9 exhibits an obstruction to expansion which is *qualitative*, rather than *quantitative* in nature. That is to say, expansion in $(\mathrm{SL}_2(p^{n_i}), S_{n_i})$ fails simply by virtue of the fact that $\langle S_{n_i} \rangle \neq \mathrm{SL}_2(p^{n_i})$ for infinitely many i. This leads to the question of whether this is the *only* obstruction to expansion in these groups. Specifically:

Question 4.1.10. *Let $S_n \subseteq \mathrm{SL}_2(p^n)$ with $\mathrm{girth}(\mathrm{SL}_2(p^n), S_n) \gg n$. Does there exist $\epsilon > 0$ such that $(\langle S_n \rangle, S_n)$ is an ϵ-expander for all n sufficiently large?*

Question 4.1.11. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric set generating a non-elementary subgroup. Let $(f_i)_i \subseteq \mathbb{F}_p[t]$ be a sequence of distinct irreducible polynomials. Does there exist $\epsilon > 0$ such that $(\langle \pi_{f_i}(S) \rangle, \pi_{f_i}(S))$ is an ϵ-expander for all i sufficiently large?*

A second way in which Theorem 4.1.7 (ii) might be extended would be relax the assumption that no two irreducible factors of f_i have the same degree. As a model case, let $f, g \in \mathbb{F}_p[t]$ be distinct irreducibles of degree n, and consider the group $\mathrm{SL}_2(\mathbb{F}_p[t]/(f \cdot g))$. By the Chinese Remainder Theorem, this may be identified with $\mathrm{SL}_2(p^n) \times \mathrm{SL}_2(p^n)$. A potential obstruction to expansion in this group comes from proper subdirect products of $\mathrm{SL}_2(p^n) \times \mathrm{SL}_2(p^n)$, which arise as the graphs of automorphisms of $\mathrm{SL}_2(p^n)$. It remains an open question how to demonstrate non-concentration in such subgroups, as would be required for hypothesis (iii) of Theorem 2.2.22.

The primality assumption in Theorem 4.1.7 (ii) comes from hypothesis (ii) of Theorem 2.2.22, which in our setting is satisfied by results of Varjú [77]. The applicability of these results shall be discussed in more detail in Section 4.2. Roughly speaking though, for the product theorem to apply to reductions modulo polynomials with unboundedly many irreducible factors (so that the corresponding congruence quotients decompose as products with unboundedly many quasisimple factors), the subgroup structure of the quasisimple factors must be highly restricted. It seems plausible that a generalisation of Varjú's product theorem which relaxes these restrictions may be discovered, and the primality assumption thereby removed.

An expansion result for reductions modulo arbitrary square-free polynomials seems even further out of reach. For then the decompositions of the congruence quotients into products of quasisimple groups contain unboundedly many isomorphic factors, so Varjú's product theorem fails even more dramatically. It may be that the fastest route to a result on expansion in this general setting is to tackle the question of concentration in approximate subgroups directly.

Even an expansion result in the case of two irreducible factors of the same degree would have useful consequences for sieving in $\mathrm{SL}_2(\mathbb{F}_p[t])$. For in the presence of such a result (and the relevant strengthening of the product theorem indicated above) we could substitute the group sieve of Lubotzky-Meiri for Proposition 4.3.3 in the proofs of Theorems 4.1.2 and 4.1.4, thereby improving the upper bounds in those two results from $e^{-C\sqrt{l/\log l}}$ to e^{-Cl}.

4.2 Constructing the Expanders

As a notational convenience, for $n \in \mathbb{N}$ we set $Q_n = \mathrm{SL}_2(p^n)$.

4.2.1 Reduction to non-concentration for free generators

In this section we reduce the proof of Theorem 4.1.7 to the following Proposition:

Proposition 4.2.1. *Let $T \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be the symmetric closure of a finite subset, freely generating a non-abelian free subgroup. Suppose every entry of every element of T has degree at most $\tilde{D}$. Then there exists $C, M, \gamma > 0$ (depending on $\tilde{D}$, $|T|$ and p) such that the following holds. Let $f \in \mathbb{F}_p[t]$ be an M-rough square-free polynomial with no two irreducible factors having the same degree. Then for every $H \lneq G = \mathrm{SL}_2(\mathbb{F}_p[t]/(f))$, there exists $l \leq C \log|G|$ such that:*

$$\mu_T^{(2l)}(H) \leq |G : H|^{-\gamma}.$$

The reduction shall be via Theorem 2.2.22. We reference known results which cover hypotheses (i) and (ii) of Theorem 2.2.22. We then use the general results about expanders from Section 2.1.3 to reduce the question of expansion for arbitrary sets S as in the Statement of Theorem 4.1.7 to expansion for finite sets $T \subseteq \langle S \rangle$ freely generating $\langle T \rangle$. This shall be via a Tits alternative.

The quasirandomness condition in our setting is classical (see for instance [50]):

Theorem 4.2.2. *There is an absolute constant $C > 0$ such that every non-trivial complex representation of Q_n has dimension at least Cp^n.*

Let f be as in Proposition 4.2.1 and let $p_1, \ldots, p_N$ be the irreducible factors of f, of degrees $n_1, \ldots, n_N$ respectively. It follows from the Chinese Remainder Theorem that:

Lemma 4.2.3. *The natural map:*

$$(\prod_{j=1}^N \pi_{p_j}) : \mathrm{SL}_2(\mathbb{F}_p[t]/(f)) \to \prod_{j=1}^N Q_{n_j}$$

is an isomorphism.

We turn next to the product theorem. In the setting of Theorem 4.1.7 (i), this is due to Dinai [27]. For Theorem 4.1.7 (ii) we use Proposition 14 of [77], which we quote in full:

Theorem 4.2.4. *For all $C > 0$, $L \in \mathbb{N}$ and $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ there exist $C'(C, L, \beta) > 0$ and $\beta'_{C,L,\beta} : \mathbb{R}^+ \to \mathbb{R}^+$ such that the following holds. Let G be a finite group, $G_1, \ldots, G_N$ be finite groups such that $G \cong G_1 \times \cdots \times G_N$. Suppose:*

(i) For any finite group F, $|\{i \in \{1, \ldots, N\} : G_i \cong F\}| \leq L$;

(ii) For $1 \leq i \leq N$, G_i is quasisimple and $|Z(G_i)| \leq L$;

(iii) For $1 \leq i \leq N$, any non-trivial complex representation of G_i has dimension at least $|G_i|^{\frac{1}{L}}$;

(iv) For $1 \leq i \leq N$ and for some $m < L$, there are classes $\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_m$ of subgroups of G_i satisfying:

(a) $\mathcal{H}_0 = \{Z(G_i)\}$,

(b) Each $\mathcal{H}_j$ is closed under conjugation in G_i,

(c) For each $H < G_i$ there is $1 \leq j \leq m$ and $H^\sharp \in \mathcal{H}_j$ such that $|H : H \cap H^\sharp| \leq L$,

(d) For $1 \leq j \leq m$ and for each $H_1, H_2 \in \mathcal{H}_j$ with $H_1 \neq H_2$, there exists $j' < j$ and $H^\sharp \in \mathcal{H}_{j'}$ such that $|H_1 \cap H_2 : H_1 \cap H_2 \cap H^\sharp| \leq L$.

If $G_1, \ldots, G_N$ satisfy the conditions of Definition 2.2.20 (the product theorem for sets of small tripling) with respect to (C, β), then G satisfies the conditions of Definition 2.2.20 with respect to $(C', \beta'_{C,L,\beta})$.

We check that this result applies to $G = \mathrm{SL}_2(\mathbb{F}_p[t]/(f_i))$, for f_i as in Theorem 4.1.7 (ii). The decomposition as a product is given by Lemma 4.2.3. (i) follows from the assumption that no two irreducible factors of f_i have the same degree. (ii) is well-known for $G_i = Q_{n_i}$. (iii) is Theorem 4.2.2. For (iv), we recall the classification of subgroups of Q_n (see for instance [47]).

Proposition 4.2.5. *For $\mathbb{F}_q$ the finite field of order q and characteristic $p \geq 3$, any proper subgroup H of $\mathrm{SL}_2(\mathbb{F}_q)$ satisfies one of the following:*

(i) H fixes a point in the projective line $\mathbb{F}_{q^2}\mathbb{P}^1$ over the quadratic extension $\mathbb{F}_{q^2}$ of $\mathbb{F}_q$. In particular H is metabelian.

(ii) $H \leq \mathrm{SL}_2(\mathbb{F}_5)$.

(iii) H *is conjugate in* $\mathrm{SL}_2(\mathbb{F}_q)$ *to a subgroup of* $\mathrm{SL}_2(\mathbb{F}')$ *for some proper subfield* $\mathbb{F}'$ *of* $\mathbb{F}_q$.

Define $\mathcal{H}_1$ to be the set of stabilisers in Q_n of pairs of distinct points in $\mathbb{F}_{p^{2n}}\mathbb{P}^1$, and $\mathcal{H}_2$ to be the set of stabilisers in Q_n of points in $\mathbb{F}_{p^{2n}}\mathbb{P}^1$. We check that the conditions of Theorem 4.2.4 (iv) are satisfied by $\mathcal{H}_0 = \{Z(Q_n)\}, \mathcal{H}_1, \mathcal{H}_2$, in the case for which n is prime. (a), (b) are obvious, and (c) is immediate from Proposition 4.2.5, since by primality of n, the only proper subfield subgroups of Q_n are the conjugates of Q_1, which are of bounded size. (d) is a consequence of the following elementary fact from linear algebra:

Lemma 4.2.6. *Suppose* $g \in Q_n$ *has at least three distinct fixed points in* $\mathbb{F}_{q^2}\mathbb{P}^1$. *Then* $g \in Z(Q_n)$.

Now let S be as in the statement of Theorem 4.1.7. We produce a pair of words in S freely generating a non-abelian free subgroup. In the classical Tits alternative, the lengths of our free generators as words in S, and hence the degrees of their entries, depend on S and not just on D. However, we can obtain a bound depending only on D by utilising the following result of Breuillard:

Theorem 4.2.7 (Uniform Tits Alternative [16])**.** *For every* $d \geq 2$, *there exists* $N(d) > 0$ *such that, for any field* K, *and* $S \subseteq \mathrm{SL}_d(K)$ *finite symmetric, either* $\langle S \rangle$ *is virtually soluble or the ball* $B_S(N(d))$ *of radius* $N(d)$ *in the word metric contains two elements which freely generate a non-abelian free subgroup of* $\mathrm{SL}_d(K)$.

Proof of Theorem 4.1.7. Let $N = N(2)$ be as in Theorem 4.2.7 and let $x, y \in B_S(N)$ freely generate a non-abelian free group. Every entry of every member of $T = \{x^{\pm 1}, y^{\pm 1}\}$ is expressible as a sum of monomials of degree at most N in the entries of the elements of S, hence has degree at most DN. We now apply Theorem 2.2.22 to $(\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i))), \pi_{f_i}(T))$.

We verify the conditions of Theorem 2.2.22. Note that, taking M sufficiently large in Theorem 4.1.7, we may assume in Lemma 4.2.3 that the $|Q_{n_j}|$ are larger than any given constant. Hypothesis (i) is immediate from Theorem 4.2.2 and Lemma 4.2.3. Hypothesis (ii) follows from [27] and Theorem 4.2.4. Hypothesis (iii) follows from Proposition 4.2.1, applied with $f = f_i$ and $\tilde{D} = DN$, and Remark 2.2.23.

We conclude that $(\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i)), \pi_{f_i}(T))_i$ is an expander family. By Lemma 2.1.10, $(\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i)), \pi_{f_i}(B_S(N)))_i$ is an expander family. The required result follows from Lemma 2.1.9, since $\langle S \rangle = \langle B_S(N) \rangle$. □

Remark 4.2.8. *The constants M and i_0 in the statement of Theorem 4.1.7 could in principle be computed, by keeping track of the bounds arising in the proof of Proposition 4.2.1 below. They shall involve both the constant N from the statement of the Uniform Tits Alternative and the known spectral radius $\sqrt{3}/2$ for the simple random walk on $\{x^{\pm 1}, y^{\pm 1}\}$ in $F(x, y)$ (see [46] Theorem 3).*

Moreover, the proof of the Uniform Tits Alternative is effective, so N could in principle be computed (though to our knowledge this has not been done). To be more precise, although the published proof of Theorem 4.2.7 does make use of compactness arguments, they apply only to one stage of the argument (Lemma 2.1 of [15]), are even there only relevant to the case of matrices over archimedean fields, and may be avoided altogether by employing an alternative argument (Remark 1.12 of [14]).

In computing M and i_0, one would obtain an explicit description of the degrees of reductions which would give rise to families of expanders, in terms only of the degrees of the entries of elements of S, $|S|$ and p.

4.2.2 Non-concentration: the irreducible case

In this section we warm up to the proof of Proposition 4.2.1 by examining the case for which the polynomials f_i are irreducible, so that $\mathrm{SL}_2(\mathbb{F}_p[t]/(f_i)) = Q_{\deg(f_i)}$. The proof of this case shall contain all the key ideas of the general case (to be discussed in the following section) but is technically simpler. Indeed, more generally we shall prove:

Proposition 4.2.9. *For any $C_1 > 0$ and any $k \in \mathbb{N}$ with $k \geq 2$, there exists $C_2, C_3, \gamma > 0$ (depending on $C_1, p, |S|$) such that, if n is a C_2-rough positive integer, $S_n \subseteq Q_n$ is symmetric with $|S_n| = 2k$, and $\mathrm{girth}(Q_n, S_n) \geq C_1 n$, then for n sufficiently large and for all $H_n \lneq Q_n$, there exists $l \leq C_3 \log|Q_n|$ such that:*

$$\mu_{S_n}^{(2l)}(H_n) \leq |Q_n|^{-\gamma}.$$

The relevant case of Proposition 4.2.1 follows immediately from Proposition 4.2.9 and the following Lemma:

Lemma 4.2.10. *let T be as in Proposition 4.2.1 and $f \in \mathbb{F}_p[t]$ be of degree n. Then:*

$$\mathrm{girth}(\mathrm{SL}_2(\mathbb{F}_p[t]/(f)), \pi_f(T)) \geq n/\tilde{D}.$$

Proof. Let w be a non-trivial reduced word in T of length l. Every entry of every element of T has degree at most $\tilde{D}$, so every entry of w has degree at most $\tilde{D}l$. Now suppose $\pi_f(w) = 1$, so that $w \in (I_2 + f \cdot \mathbb{M}_2(\mathbb{F}_p[t])) \setminus \{I_2\}$. Then at least one entry of w has degree at least n, so $l \geq n/\tilde{D}$, as required. □

Given the discussion in Section 4.2.1, Proposition 4.2.9 also immediately implies Theorem 4.1.8.

Proof of Theorem 4.1.8. As in the proof of Theorem 4.1.7, we apply Theorem 2.2.22. Taking M sufficiently large in Theorem 4.1.8, we may assume the n_i to be larger than any given constant. Hypothesis (i) of Theorem 2.2.22 is Theorem 4.2.2; hypothesis (ii) is [27] and hypothesis (iii) follows from Proposition 4.2.9 and Remark 2.2.23. □

We now turn to the proof of Proposition 4.2.9. Once again we exploit the classification of subgroups of Q_n.

Informally, in all cases, the girth hypothesis and Kesten's Theorem will reduce the problem of bounding $\mu_{S_n}^{(2l)}(H_n)$ to providing an upper bound for $|H_n \cap B_{S_n}(2l)|$. For $H_n \leq \mathrm{SL}_2(\mathbb{F}_5)$ this is immediate. The roughness hypothesis on n will guarantee that any proper subfield subgroup is too small to fill $|B_{S_n}(2l)|$. Non-concentration in metabelian subgroups will be achieved by the same combinatorial argument as was used for the corresponding case in [6]: a metabelian group satisfies a short group law, so that if $|H_n \cap B_{S_n}(2l)|$ is large, there will be many short relations between the elements of S_n. However the girth hypothesis guarantees that this will not happen.

First we recall:

Theorem 4.2.11 (Kesten)**.** *Let X be a finite set. Then there exists $C_4(|X|) > 0$ such that $\mu_X \in \ell^2(F(X))$ satisfies:*

$$\mu_X^{(2l)}(g) \ll_{|X|} e^{-C_4 l}$$

for all $g \in F(X)$.

The technicalities of non-concentration in subgroups are contained in the following general Lemma.

Lemma 4.2.12. *Let G be a finite group, $S \subseteq G$ symmetric with $|S| = 2k$ and let $C_1 > 0$ be such that $\mathrm{girth}(G, S) > C_1 \log|G|$. Let $C_4(k) > 0$ be the constant from Kesten's Theorem. Let $H \lneq G$. Let $\gamma > 0$ and let $C_5 \in (0, C_1)$. Suppose $|G|$ is sufficiently large (depending on C_1, k).*

(i) Suppose H is metabelian. Suppose $C_5 \log|G| \leq 2l \leq \frac{C_1}{32} \log|G|$. Suppose $\gamma < C_4 C_5/2$. Then $\mu_S^{(2l)}(H) \leq |G|^{-\gamma}$.

(ii) Let $C_6 > 0$ and suppose $|H| \leq C_6 |G|^{\frac{1}{C_2}}$ Suppose $C_5 \log|G| \leq 2l \leq C_1 \log|G|$. Suppose $\gamma + 1/C_2 < C_4 C_5/2$. Then $\mu_S^{(2l)}(H) \leq |G|^{-\gamma}$.

Proof. (i) Define a homomorphism $\theta : F \to G$ from a non-abelian free group F on free basis X, such that θ maps $X \cup X^{-1}$ bijectively onto S. Then θ maps $B_X(32l)$ bijectively onto $B_S(32l)$.

Consider $Y = B_X(2l) \cap \theta^{-1}(H)$. Then for any $a, b, c, d \in Y$, $\theta([[a, b], [c, d]]) = 1$ (since H is metabelian) so $[[a, b], [c, d]] = 1$ (since $[[a, b], [c, d]] \in B_X(32l)$).

Recall that the centraliser of every non-trivial element of a free group is cyclic. Hence there exists $x \in F$ such that for all $a, b \in Y$,

$$[a, b] \in Z := \langle x \rangle \cap B_X(8l) \tag{4.1}$$

so that $|Z| \leq 16l + 1$. Now for $a \in Y$ and $z \in Z$, define:

$$W_{a,z} = \{b \in Y : [a, b] = z\}.$$

Then $W_{a,z}$ is contained in a single coset of the centraliser of a, and in $B_X(2l)$, so that $|W_{a,z}| \leq 4l + 1$. Fix $a \in Y$. By (4.1),

$$Y \subseteq \bigcup_{z \in Z} W_{a,z}.$$

We conclude that:

$$|H \cap B_S(2l)| \leq |Y| \leq (16l + 1)(4l + 1).$$

By Kesten's Theorem and the girth hypothesis,

$$\mu_S^{(2l)}(H) \ll_k e^{-C_4 l} |H \cap B_S(2l)| \ll l^2 e^{-C_4 l},$$

so decays exponentially fast.

(ii) Suppose (for a contradiction) that for some $C_5 \log|G| \leq 2l \leq C_1 \log|G|$,

$$|G|^{-\gamma} < \mu_S^{(2l)}(H) \ll_k e^{-C_4 l} |H| \leq C_6 e^{-C_4 l} |G|^{\frac{1}{C_2}}.$$

(the second inequality being by Kesten's Theorem and the girth hypothesis). Hence:

$$|G|^{\frac{1}{C_2} + \gamma} \gg_{k, C_6} e^{C_4 l}.$$

But $e^{C_4 l} \geq |G|^{\frac{C_4 C_5}{2}}$ so we have the required contradiction by choice of C_2 and γ.

□

Proof of Proposition 4.2.9. Suppose $C_5 n \leq 2l \leq \frac{C_1}{32} n$, for some $C_5 \in (0, \frac{C_1}{32})$. We consider each case of Proposition 4.2.5 separately:

(i) Suppose H_n is metabelian. Choosing $\gamma \in (0, C_4 C_5/2)$, the required result follows from Lemma 4.2.12 (i).

(ii) If $H_n \leq \mathrm{SL}_2(\mathbb{F}_5)$, then $|H_n| \leq 120$, so $\mu_{S_n}^{(2l)}(H_n) \ll_k e^{-C_4 l}$ for any $2l \leq C_1 n$, by Kesten's Theorem and the girth hypothesis.

(iii) Suppose that there exists a proper subfield $\mathbb{F}' < \mathbb{F}_{p^n}$ such that H_n is contained in (some conjugate of) $\mathrm{SL}_2(\mathbb{F}')$. Recall that there exists $m \mid n$ such that $\mathbb{F}' = \mathbb{F}_{p^m}$. By the roughness hypothesis, $m \leq n/C_2$ so:

$$|H_n| \leq |Q_m| \leq p^{3m} \leq (p^{3n})^{\frac{1}{C_2}} \ll_p |Q_n|^{\frac{1}{C_2}}.$$

Choosing γ sufficiently small and C_2 sufficiently large, we may suppose $\gamma + 1/C_2 < C_4 C_5/2$, and the result follows from Lemma 4.2.12 (ii).

□

4.2.3 Non-concentration: the general case

In this section we complete the proof of Proposition 4.2.1. The proof shall be very similar in spirit to that of the special case discussed in Section 4.2.2: recall that there, Proposition 4.2.5 guaranteed that every proper subgroup of $\mathrm{SL}_2(\mathbb{F}_p[t]/(f))$ was either metabelian (Case (i)) or small (Cases (ii) and (iii)), so fell within reach of Lemma 4.2.12. Something similar is true in general, but to apply Lemma 4.2.12 we first need to use the product decomposition of $\mathrm{SL}_2(\mathbb{F}_p[t]/(f))$ from Lemma 4.2.3, and project down to either the factors on which the image of our proper subgroup is metabelian, or those on which it is small, depending on which make up the larger part of the product.

Recall the notation of Section 4.2.1: $f \in \mathbb{F}_p[t]$ is an M-rough square-free polynomial with no two irreducible factors having the same degree. $G = \mathrm{SL}_2(\mathbb{F}_p[t]/(f))$ and $H \lneq G$. Let $p_1, \ldots, p_N$ be the irreducible factors of f, of degrees $n_1, \ldots, n_N$ respectively.

Recall (Lemma 4.2.3) that:

$$(\prod_{j=1}^{N} \pi_{p_j}) : \mathrm{SL}_2(\mathbb{F}_p[t]/(f)) \to \prod_{j=1}^{N} Q_{n_j}$$

is an isomorphism.

Corollary 4.2.13. *$\pi_{p_j}(H) \lneq Q_{n_j}$ for some $1 \leq j \leq N$.*

Proof. We proceed by induction on N (the case $N = 1$ being trivial). Suppose (for a contradiction) that the projections π_{p_j} of H to Q_{n_j} are all surjective. Denote $F = \prod_{j=1}^{N-1} Q_{n_j}$, so that by Lemma 4.2.3, $G \cong F \times Q_{n_N}$. Define:

$$K_1 = \{g \in F : (g, 1) \in H\},\ K_2 = \{g \in Q_{n_N} : (1, g) \in H\}.$$

By induction the projections of H to F and Q_{n_N} are surjective. By Goursat's Lemma, $K_1 \lhd F$, $K_2 \lhd Q_{n_N}$ and $F/K_1 \cong Q_{n_N}/K_2$.

If $K_2 \neq Q_{n_N}$ then F has $\mathrm{PSL}_2(p^{n_N})$ as a composition factor. But this is not the case, as the n_j are all distinct. Hence $K_2 = Q_{n_N}$ and $K_1 = F$, so $H = G$. □

Up to a reordering of the p_i, there exist $k, m, n \in \mathbb{N}$ with $k + m + n = N$ such that:

(i) $\pi_{p_i}(H) = Q_{n_i}$ for $1 \leq i \leq k$;

(ii) $\pi_{p_i}(H)$ is metabelian for $k + 1 \leq i \leq k + m$;

(iii) $\pi_{p_i}(H) \lneq Q_{n_i}$ is not metabelian for $k + m + 1 \leq i \leq N$.

Let $C_2, \gamma > 0$ be constants satisfying the conditions of Lemma 4.2.12. For M sufficiently large, by the roughness hypothesis and Proposition 4.2.5, $|\pi_{p_i}(H)| \leq |Q_{n_i}|^{\frac{1}{C_2}}$ for $k+m+1 \leq i \leq N$. Moreover by Corollary 4.2.13, at least one of m, n is non-zero.

Write $F_1 = \prod_{i=1}^{k} p_i$, $F_2 = \prod_{i=k+1}^{k+m} p_i$, $F_3 = \prod_{i=k+m+1}^{N} p_i$, so that $f = F_1 \cdot F_2 \cdot F_3$. Applying Lemma 4.2.3 with f replaced by F_1, F_2, F_3 respectively, we have:

Lemma 4.2.14. *(i) $\pi_{F_1}(H) = \prod_{i=1}^{k} \mathrm{SL}_2(p^{n_i})$.*

(ii) $\pi_{F_2}(H)$ is metabelian.

(iii) $|\pi_{F_3}(H)| \leq |\pi_{F_3}(G)|^{\frac{1}{C_2}}$.

Finally, we are ready to complete:

Proof of Proposition 4.2.1. $|H| \geq |\pi_{F_1}(H)| = |\pi_{F_1}(G)|$, so:

$$|G : H| \leq |G|/|\pi_{F_1}(G)| = |\pi_{F_2F_3}(G)|.$$

Case 1: $\deg(F_2) \geq \deg(F_3)$:

We have $|\pi_{F_2F_3}(G)|^{\frac{1}{2}} \ll_p |\pi_{F_2}(G)| \leq |\pi_{F_2F_3}(G)|$. By Lemma 4.2.10,

$$\text{girth}(\pi_{F_2}(G), S) \geq \frac{\deg(F_2)}{\tilde{D}} \geq \frac{\log|\pi_{F_2}(G)|}{3\tilde{D}\log p},$$

so that by Lemma 4.2.12 (i), if:

$$C_5 \log|\pi_{F_2}(G)| \leq 2l \leq \frac{\log|\pi_{F_2}(G)|}{96\tilde{D}\log p},$$

then:

$$\mu_S^{(2l)}(H) \leq \mu_S^{(2l)}(\pi_{F_2}(H)) \leq |\pi_{F_2}(G)|^{-\gamma} \ll_p |\pi_{F_2F_3}(G)|^{\frac{-\gamma}{2}} \leq |G : H|^{\frac{-\gamma}{2}}$$

and

$$2l \leq \frac{\log|\pi_{F_2}(G)|}{96\tilde{D}\log p} \leq \frac{\log|\pi_{F_2F_3}(G)|}{96\tilde{D}\log p}.$$

Case 2: $\deg(F_3) \geq \deg(F_2)$:

We have $|\pi_{F_2F_3}(G)|^{\frac{1}{2}} \ll_p |\pi_{F_3}(G)| \leq |\pi_{F_2F_3}(G)|$. By Lemma 4.2.10,

$$\text{girth}(\pi_{F_3}(G), S) \geq \frac{\deg(F_3)}{\tilde{D}} \geq \frac{\log|\pi_{F_3}(G)|}{3\tilde{D}\log p},$$

so that by Lemma 4.2.12 (ii), if:

$$C_5 \log|\pi_{F_3}(G)| \leq 2l \leq \frac{\log|\pi_{F_3}(G)|}{3\tilde{D}\log p},$$

then:

$$\mu_S^{(2l)}(H) \leq \mu_S^{(2l)}(\pi_{F_3}(H)) \leq |\pi_{F_3}(G)|^{-\gamma} \ll_p |\pi_{F_2F_3}(G)|^{\frac{-\gamma}{2}} \leq |G : H|^{\frac{-\gamma}{2}}$$

and

$$2l \leq \frac{\log|\pi_{F_3}(G)|}{3\tilde{D}\log p} \leq \log|\pi_{F_2F_3}(G)|.$$

The required result follows. □

4.3 Non-Concentration Results

4.3.1 Two different sieves

Let us recall the basic concept of the group sieve method. We start with a simple observation:

Lemma 4.3.1. *Let G be a discrete countable group; H a finite group and $\phi : G \to H$ an epimorphism. Let ν be a probability measure on G and $X \subseteq G$. Then $\nu(X) \leq (\phi\nu)(\phi(X)) \leq |\phi(X)| \cdot \max_{x \in X}(\phi\nu)(\phi(x))$.*

Here $\phi\nu$ is the pushforward measure defined, for $Y \subseteq H$, by:

$$(\phi\nu)(Y) = \nu(\phi^{-1}(Y)).$$

Though straightforward, this bound can be very useful: when $\nu = \mu_S^{(l)}$, for $S \subseteq G$ symmetric, with $\phi(S)$ generating H, then for l sufficiently large and even, $\phi\nu$ is almost uniform on H, so that:

$$(\phi\nu)(\phi(X)) \ll |\phi(X)|/|H|. \tag{4.2}$$

Moreover if $(H, \phi(S))$ is a good expander, equidistribution occurs for l sufficiently *small* that (4.2) gives a non-trivial bound on the rate at which $\mu_S^{(l)}$ escapes from X.

The present section contains two different instantiations of this philosophy for the group $\mathrm{SL}_2(\mathbb{F}_p[t])$, taking (H, ϕ) to be one of the congruence quotients from Theorem 4.1.7. In the first of these it shall be sufficient to consider congruences modulo irreducible polynomials. We define, for G a countable discrete group and $\nu_1, \ldots, \nu_r$ finitely supported probability measures on G, the product measure $\times_{i=1}^r \nu_i$ on G by:

$$(\times_{i=1}^r \nu_i)(X) = \sum_{x \in X} \prod_{i=1}^r \nu_i(x_i), \text{ for } X \subseteq G^r.$$

Proposition 4.3.2. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$, $M > 0$ be as in Theorem 4.1.7; let $(n_i)_i$ be as in Example 4.1.6 (ii) and let $f_i \in \mathbb{F}_p[t]$ be irreducible of degree n_i. Let $X \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])^r$ and suppose there exists $\alpha, C > 0$ such that for all i sufficiently large,*

$$|\pi_{f_i}(X)|/|Q_{n_i}|^r \leq Cp^{-\alpha n_i}.$$

Then there exist $C_1(C, r), C_2(\alpha, p, S) > 0$ such that for all $l \in \mathbb{N}$,

$$(\times_{i=1}^r \mu_S^{(l)})(X) \leq C_1 e^{-C_2 l}.$$

Proof. By Theorem 4.1.7 and Lemma 2.1.3, there exists $c > 0$ such that, for $i \geq i_0$, $l \geq cn_i$ and any $x \in Q_{n_i}$, $(\pi_{f_i}\mu_S^{(l)})(x) \leq 2/|Q_{n_i}|$. Fix $\delta \in (0,1)$, so that for l sufficiently large, $\exists i \geq i_0$ such that $l \geq cn_i \geq \delta l$. Then for i sufficiently large,

$$\begin{aligned}(\times_{j=1}^r \mu_S^{(l)})(X) &\leq (\times_{j=1}^r \pi_{f_i}\mu_S^{(l)})(\pi_{f_i}(X)) \\ &\leq 2^r |\pi_{f_i}(X)|/|Q_{n_i}|^r \text{ (by Lemma 4.3.1)} \\ &\leq 2^r C p^{-\alpha n_i} \text{ (by hypothesis)} \\ &\leq 2^r C e^{-\frac{\alpha\delta l \log p}{c}}\end{aligned}$$

as required. □

Proposition 4.3.2 is very useful for proving escape of the random walk from such subsets as proper algebraic subvarieties, which have small image in congruence quotients, as we shall see. However, Proposition 4.3.2 is powerless in the face of subsets X whose images modulo f_i are of order $\sim \gamma|Q_{n_i}|$, for $\gamma \in (0,1)$, say. This difficulty may be partially resolved by considering, instead of individual congruence quotients Q_{n_i}, large products $Q_{n_i} \times \ldots \times Q_{n_{i+k}}$. The image of X in such a quotient will be of order $\sim \gamma^k |Q_{n_i}| \cdots |Q_{n_{i+k}}|$, so by allowing k to grow and applying Theorem 4.1.7, we may recover a good non-concentration estimate. As discussed in the Introduction, Theorem 4.1.7 is not powerful enough to retain exponentially fast escape from such X. However we still have:

Proposition 4.3.3. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$, $M > 0$ be as in Theorem 4.1.7; let $(n_i)_i$ be the sequence of all primes greater than M (arranged in ascending order) and let $f_i \in \mathbb{F}_p[t]$ be irreducible of degree n_i. Let $X \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])^r$ and suppose there exists $\gamma \in (0,1)$ and $i_1 \in \mathbb{N}$ such that for all $i \geq i_1$,*

$$|\pi_{f_i}(X)|/|Q_{n_i}|^r \leq \gamma.$$

Then there exist $C_1(r), C_2(\gamma, p, S) > 0$ such that for all $l \in \mathbb{N}$,

$$(\times_{i=1}^r \mu_S^{(l)})(X) \leq C_1 e^{-C_2\sqrt{l/\log l}}.$$

Proof. Define $g_i = \prod_{k=i_2}^{i_2+i-1} f_k \in \mathbb{F}_p[t]$, with i_2 sufficiently large (to be determined). Then:

$$|\pi_{g_i}(X)| \leq \gamma^i \prod_{k=i_2}^{i_2+i-1} |Q_{n_k}|^r$$

(provided $i_2 \geq i_1$). By Theorem 4.1.7, there exists $c > 0$ such that, provided i_2 is sufficiently large, for $l \geq c\sum_{k=i_2}^{i_2+i-1} n_k$ and for any $g \in \mathrm{SL}_2(\mathbb{F}_p[t]/(g_i))$,

$$(\pi_{g_i}\mu_S^{(l)})(g) \leq 2/\prod_{k=i_2}^{i_2+i-1} |Q_{n_k}|.$$

For such l,

$$\begin{aligned}(\times_{j=1}^r \mu_S^{(l)})(X) &\leq (\times_{j=1}^r \pi_{g_i}\mu_S^{(l)})(\pi_{g_i}(X)) \\ &\leq |\pi_{g_i}(X)|(2/\prod_{k=i_2}^{i_2+i-1} |Q_{n_k}|)^r \\ &\leq 2^r\gamma^i.\end{aligned}$$

Recalling that n_k grows like $k\log k$, we have $\sum_{k=i_2}^{i_2+i-1} n_k \asymp i^2 \log i$. Choosing $i \asymp \sqrt{l/\log l}$, $i^2 \log i \gg l$ and the result follows. □

4.3.2 Escape from subvarieties

We are now ready to prove Theorem 4.1.1. In view of Proposition 4.3.2, it will suffice to bound the size of projections of subvarieties to congruence quotients. We use:

Theorem 4.3.4 (Schwarz-Zippel [51])**.** *Let $\mathbb{F}$ be a finite field; $\overline{\mathbb{F}}$ be its algebraic closure. Let V be an affine algebraic subvariety of $\mathbb{F}^d$, defined by A polynomials in $\overline{\mathbb{F}}[x_1, \ldots, x_d]$, each of total degree at most B. Then:*

$$|V| \ll_{A,B,d} |\mathbb{F}|^{\dim(V)}.$$

Proof of Theorem 4.1.1. SL_2^r is irreducible of dimension $3r$, so by Theorem 4.3.4,

$$|\pi_{f_i}(V(F))| \ll_F p^{(3r-1)n_i} \ll p^{-n_i}|Q_{n_i}|^r.$$

The result now follows from Proposition 4.3.2. □

Example 4.3.5. *Under the hypotheses of Theorem 4.1.1:*

(i) Zero entries are rare: let $F_1 : \mathbb{M}_2(\mathbb{F}_p[t]) \to \mathbb{F}_p[t]$ be given by $F_1\begin{pmatrix} a & b \\ c & d \end{pmatrix} = abcd$. Then there exist $C_1, C_2 > 0$ such that:

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : g \text{ has a zero entry}\}) = \mu_S^{(l)}(V(F_1)) \leq C_1 e^{-C_2 l}.$$

(ii) Matrices with a particular trace are rare: fix $\alpha \in \mathbb{F}_p[t]$ and let $F_\alpha : \mathbb{M}_2(\mathbb{F}_p[t]) \to \mathbb{F}_p[t]$ be given by $F_\alpha(A) = \mathrm{tr}(A) - \alpha$.

Then there exist $C_1, C_2 > 0$ such that:

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : \mathrm{tr}(g) = \alpha\}) = \mu_S^{(l)}(V(F_\alpha)) \leq C_1 e^{-C_2 l}.$$

(iii) Torsion elements are rare: Let $g \in \mathrm{SL}_2(\mathbb{F}_p[t])$. Conjugate g, possibly over a quadratic extension, to an upper triangular matrix $\tilde{g} = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}$. Suppose there exists $n \in \mathbb{N}$ such that $g^n = I_2$. Then $a^n = 1$. This is only possible if a lies in a quadratic extension of $\mathbb{F}_p$. In particular $\mathrm{tr}(g) \in \mathbb{F}_p$, so g satisfies one of the bounded set of polynomials F_α as in (ii) above, for $\alpha \in \mathbb{F}_p$. Hence there exist $C_1, C_2 > 0$ such that:

$$\begin{aligned}\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : g \text{ has finite order}\}) &\leq \sum_{\alpha \in \mathbb{F}_p} \mu_S^{(l)}(V(F_\alpha)) \\ &\leq C_1 e^{-C_2 l}.\end{aligned}$$

(iv) Elements fixing a point in the adjoint representation are rare: Recall that $\mathrm{SL}_2(\mathbb{F}_p[t])$ acts linearly on $\mathfrak{sl}_2(\mathbb{F}_p[t])$ by conjugation. Given $g \in \mathrm{SL}_2(\mathbb{F}_p[t])$, let $\mathrm{Ad}(g) \in \mathrm{GL}_3(\mathbb{F}_p[t])$ be the matrix of the associated linear transformation with respect to some (fixed) $\mathbb{F}_p[t]$-basis for $\mathfrak{sl}_2(\mathbb{F}_p[t])$.

Now recall that, given polynomials $F_1(X), F_2(X)$ over some field K, there is a polynomial function $\mathrm{Res}(F_1(X), F_2(X))$ of their coefficients (defined over $\mathbb{Z}$ and depending only on the degrees of F_1, F_2) which vanishes precisely when F_1, F_2 have a common root in some extension of K. In particular, $F(g) = \mathrm{Res}(\chi_{\mathrm{Ad}(g)}(X), X - 1)$ is a polynomial in the entries of g which vanishes precisely when g has a non-zero fixed point in $\mathfrak{sl}_2(\mathbb{F}_p[t])$. Moreover $F(g)$ does not vanish identically on $\mathrm{SL}_2(\mathbb{F}_p[t])$: $F\begin{pmatrix} 1+t & 2+t \\ t & 1+t \end{pmatrix} \neq 0$, for instance. We conclude that there exist $C_1, C_2 > 0$ such that:

$$\mu_S^{(l)}(\{g \in \mathrm{SL}_2(\mathbb{F}_p[t]) : \exists X \in \mathfrak{sl}_2(\mathbb{F}_p[t]) \setminus \{0\} \text{ s.t. } X^g = X\}) \leq C_1 e^{-C_2 l}.$$

4.3.3 Squares in $\mathrm{SL}_2(\mathbb{F}_p[t])$ are rare

In this section we prove Theorem 4.1.2. Let $X \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be the set of squares. In light of Proposition 4.3.3, it suffices to bound the sizes of images $\pi_{f_i}(X)$. We note some elementary facts about $\mathrm{SL}_2(Q)$, for Q an arbitrary odd prime power. Let $D(Q) \leq \mathrm{SL}_2(Q)$ be the subgroup of diagonal matrices. Recall that $D(Q)$ is cyclic of order $Q - 1$.

Lemma 4.3.6. *Let $g \in D(Q)$ be non-central in $\mathrm{SL}_2(Q)$. Then:*

(i) $C_{\mathrm{SL}_2(Q)}(g) = D(Q)$*;*

(ii) $|\mathrm{ccl}_{\mathrm{SL}_2(Q)}(g) \cap D(Q)| = 2$*;*

(iii) If g is a square in $\mathrm{SL}_2(Q)$ then it is a square in $D(Q)$.

Now $2 \mid (Q-1)$, so the set of squares in $D(Q)$ is of order $\frac{Q-1}{2}$. $Z(\mathrm{SL}_2(Q)) = \{\pm I_2\}$ consists of squares in $\mathrm{SL}_2(Q)$, so that by Lemma 4.3.6 (iii), there is a subset $\{g_i\}_{i=1}^{\frac{Q-1}{2}} \subseteq D(Q)$ consisting entirely of non-squares in $\mathrm{SL}_2(Q)$. If $g \in \mathrm{SL}_2(Q)$ is not a square, then $\mathrm{ccl}_{\mathrm{SL}_2(Q)}(g)$ consists entirely of non-squares, and by Lemma 4.3.6 (i), $|\mathrm{ccl}_{\mathrm{SL}_2(Q)}(g)| = Q(Q+1)$. Hence:

$$\begin{aligned}
|\{\text{non-squares in } \mathrm{SL}_2(Q)\}| &\geq |\bigcup_{i=1}^{\frac{Q-1}{2}} \mathrm{ccl}_{\mathrm{SL}_2(Q)}(g_i)| \\
&\geq \frac{1}{2}\sum_{i=1}^{\frac{Q-1}{2}} |\mathrm{ccl}_{\mathrm{SL}_2(Q)}(g_i)| \text{ (by Lemma 4.3.6 (ii))} \\
&\geq \frac{1}{4}(Q-1)Q(Q+1) \\
&= \frac{1}{4}|\mathrm{SL}_2(Q)|.
\end{aligned}$$

Theorem 4.1.2 is now immediate from Proposition 4.3.3, taking $\gamma = \frac{3}{4}$.

4.3.4 Reducible characteristic polynomials in $\mathrm{SL}_2(\mathbb{F}_p[t])$ are rare

In this section we prove Theorem 4.1.4. Let $Y \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be the set of elements with reducible characteristic polynomial. Once again, we bound $|\pi_{f_i}(Y)|$. Let $g \in Y$ and let $f \in \mathbb{F}_p[t]$ be irreducible of degree n. Since $\chi_g \in \mathbb{F}_p[t][X]$ splits over $\mathbb{F}_p[t]$, $\chi_{\pi_f(g)} \in \mathbb{F}_{p^n}[X]$ splits over $\mathbb{F}_{p^n}$. Let Q be an arbitrary odd prime power. It will suffice to bound the set of elements $g \in \mathrm{SL}_2(Q)$ with reducible characteristic polynomial. We distinguish two cases and prove exponential decay in each:

Case 1: $\mathrm{tr}(g) \neq \pm 2$.

χ_g does not have a repeated root, so is diagonalisable in $\mathrm{SL}_2(Q)$. Hence there exists non-central $h \in D(Q)$ such that $\mathrm{ccl}_{\mathrm{SL}_2(Q)}(g) = \mathrm{ccl}_{\mathrm{SL}_2(Q)}(h)$. There are $Q-3$ non-central elements $h \in D(Q)$, and each has conjugacy class in $\mathrm{SL}_2(Q)$

of order $Q(Q+1)$, by Lemma 4.3.6 (i). Therefore the number of non-central diagonalisable elements g is at most:

$$\begin{aligned}\Big|\bigcup_{h\in D(Q)\setminus Z(\mathrm{SL}_2(Q))} \mathrm{ccl}_{\mathrm{SL}_2(Q)}(h)\Big| &\leq \frac{1}{2}\sum_{h\in D(Q)\setminus Z(\mathrm{SL}_2(Q))} |\mathrm{ccl}_{\mathrm{SL}_2(Q)}(h)| \text{ (by Lemma 4.3.6 (ii))}\\ &\leq \frac{1}{2}(Q-3)Q(Q+1)\\ &\leq \frac{1}{2}|\mathrm{SL}_2(Q)|.\end{aligned}$$

Case 2: $\mathrm{tr}(g) = \pm 2$ is immediate from Example 4.3.5 (ii).

Theorem 4.1.4 follows from Proposition 4.3.3, with any $\gamma > \frac{1}{2}$.

Chapter 5

Spectral Gap in $\mathrm{SL}_2(\mathbb{Z}_p)$

5.1 Introduction

Fix $p \geq 3$ prime. The goal of this chapter shall be to prove:

Theorem 5.1.1. *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z}_p)$ be a finite symmetric set, generating a subgroup Γ whose closure $\overline{\Gamma}$ in $\mathrm{SL}_2(\mathbb{Z}_p)$ is open. Let $A \subseteq \mathbb{Z}_p$ be the set of entries occurring in elements of S. Suppose that for every $a \in A$, there exists $f_a(X) \in \mathbb{Q}[X]$ such that $f_a(a) = 0$. Then $(\pi_{p^n}(\Gamma), \pi_{p^n}(S))_n$ is a family of two-sided expanders, where $\pi_{p^n} : \mathrm{SL}_2(\mathbb{Z}_p) \to \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ is the congruence map.*

Recall that:

Proposition 5.1.2. *Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z}_p)$. Then the following are equivalent:*

(i) $\overline{\Gamma} \leq_o \mathrm{SL}_2(\mathbb{Z}_p)$;

(ii) Γ *is Zariski-dense in* $\mathrm{SL}_2(\mathbb{Q}_p)$;

(iii) Γ *is not virtually soluble;*

(iv) $F_2 \hookrightarrow \Gamma$.

The main result of Bourgain and Gamburd's paper [6] is the special case of Theorem 5.1.1 for which the matrices in S are supported over $\mathbb{Z}$. In fact their result is stated in a weaker form, where p is chosen such that the image of S in $\mathrm{SL}_2(p)$ is a generating set, so that S generates a dense subgroup of $\mathrm{SL}_2(\mathbb{Z}_p)$ (if S generates a non-elementary subgroup of $\mathrm{SL}_2(\mathbb{Z})$ this will hold for all sufficiently large p). Nevertheless, their argument yields the same conclusion in the more general setting of S generating a subgroup whose closure in $\mathrm{SL}_2(\mathbb{Z}_p)$ is open (but not necessarily dense).

The proof shall in a sense be a marriage of the techniques employed in the last two chapters: expansion shall follow from a version of the Bourgain-Gamburd machine. After applying the tools of quasirandomness and the ℓ^2-flattening Lemma, it suffices to check that the random walk escapes quickly from approximate subgroups. The proof then proceeds by running the Solovay-Kitaev procedure *inside* an approximate subgroup. The upshot shall be that any approximate subgroup on which the random walk concentrates must efficiently generate a large subgroup of $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$, and must therefore be too big to satisfy the conclusion of the ℓ^2-flattening Lemma.

Applying the Solovay-Kitaev procedure entirely within an approximate subgroup $H \subseteq \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ presents some additional technical difficulties which were not present in the setting of Chapter 3. First, in generating the Lie ring $\mathfrak{sl}_2(\mathbb{Z}/p^n\mathbb{Z})$ we do not *a priori* have access to all elements of $\mathbb{Z}/p^n\mathbb{Z}$ to use as coefficients, but only a certain subset arising from the coefficients of the elements in H (exactly how this set is produced shall be explained in Section 5.4). To produce arbitrary coefficients we need the sum-product results of Bourgain from Section 2.6.

Second, to run the Solovay-Kitaev procedure in H at all, H must resemble the ambient group $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ from the perspective of algebraic geometry. In other words, the preimage of H in $\mathrm{SL}_2(\mathbb{Z}_p)$ must not be controlled by small sublevel sets of non-trivial polynomials on $\mathrm{SL}_2(\mathbb{Z}_p)$. Because the random walk may be hypothesised to concentrate on H, it suffices to check that the random walk does not concentrate on such sublevel sets. Using the fact that the coefficients are algebraic, this latter claim may be reduced still further to the problem of proving non-concentration on *subvarieties*.

In [6] this non-concentration was achieved by using the result of [5] to "sieve out" proper subvarieties in congruence quotients already known to exhibit expansion. We have already made use of this same idea, in the proof of Theorem 4.1.1. What allows us to handle a broader range of coefficients than the integers is the proliferation of new results on expander congruence quotients in the years since the publication of [6], and in particular the results of [69].

We should emphasize that the novelty of our work in this chapter is entirely contained in the stage of the argument dealing with the escape of the random walk on $\mathrm{SL}_2(\mathbb{Z}_p)$ from algebraic subvarieties, and the upgrading of these non-concentration estimates to sublevel sets of the relevant polynomials. Once the problem has been successfully "pushed down" into the finite quotients, the argument is essentially identical to that presented in [6]. We do deviate from their presentation in Section 5.3.5 (corresponding to Proposition 4.3 in [6]), where we use a different proof which is (to

us) more intuitive and more amenable to generalisation to other settings. There are also minor differences between the argument of Section 5 and the corresponding Sections 5 and 6 of [6], where we have been unable to exactly replicate some of the steps of Bourgain and Gamburd's argument; nevertheless the architecture of the proof is theirs.

Given all that, why have we opted to present the proof in full detail? The motivation is threefold. First, there is no natural result that corresponds to the parts of the proof unchanged from [6] which we could use as a black box, so we have felt it necessary to reproduce the complete argument to verify to the reader's satisfaction that these parts are indeed unchanged. Second, we aim to give a treatment of Bourgain and Gamburd's ideas emerging from [6] which will be as accessible and well-motivated as possible, and thereby aid in the comprehension of some important and interesting, but technically difficult mathematics. Finally, we wish to use the proof of Theorem 5.1.1 as a roadmap to other results for which a similar proof strategy would appear to be relevant. Such results (actual and potential) are described in Section 5.5.

We also acknowledge the results of [68], which allow one to prove Theorem 5.1.1 using a different method. In that paper, expansion was proved for the $\mod p^n$ congruence quotients of finitely generated subgroups of $\mathrm{GL}_d(\mathbb{Q})$ whose Zariski closures have connected component at identity which is perfect. By restriction of scalars this conclusion may be extended to subgroups supported on p-adic integers which are algebraic over $\mathbb{Q}$.

The chapter is structured as follows: in Section 5.2 we prove weak quasirandomness for $\mathrm{SL}_2(\mathbb{Z}_p)$ and thereby reduce Theorem 5.1.1 to non-concentration in approximate subgroups. Specifically, we state Theorem 5.2.11, which asserts that any approximate subgroup in which the random walk becomes highly concentrated has iterated product set (of controlled length) containing a large congruence kernel, contradiction the ℓ^2-flattening Lemma. In Section 5.3 we investigate the algebraic geometry of such approximate subgroups, and show that it resembles that of the ambient group. Section 5.4 combines the results of Section 5.3 with the sum-product results from Section 2.6 to set up the Solovay-Kitaev procedure in our approximate subgroup and deduce Theorem 5.2.11. Finally, in Section 5.5 we sketch the generalisation of Theorem 5.1.1 to $\mathrm{SL}_d(\mathbb{Z}_p)$, discuss a potential weakening of our hypothesis on the generating set and explore the possibility of proving a version for $\mathrm{SL}_2(\mathbb{F}_p[[t]])$.

5.2 Reduction to Non-Concentration in Approximate Subgroups

The proof of Theorem 5.1.1 shall follow from the Bourgain-Gamburd machine. The goal of this section shall be to implement the aspects of the machine which reduce the problem of proving expansion to that of proving non-concentration in approximate subgroups.

In Section 5.2.1 we describe the normal subgroups of $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$, then prove weak quasirandomness of $\mathrm{SL}_2(\mathbb{Z}_p)$ in Section 5.2.2. The results of these sections are well-known: weak quasirandomness of $\mathrm{SL}_2(\mathbb{Z}_p)$ is provided by Lemma 5.1 of [71], for instance. We include a full proof here only in the interests of completeness. Finally, we shall apply the ℓ^2-flattening lemma in Section 5.2.3.

For ease of notation, in the sequel we shall write $G_n = \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ and $K_n = \ker(\pi_{p^n}) = \mathrm{SL}_2(\mathbb{Z}_p) \cap (I_2 + p^n\mathbb{M}_2(\mathbb{Z}_p))$. For $m \leq n$ we shall allow ourselves to identify K_m and its image in K_n under π_{p^n}, when to do so will not cause excessive confusion.

5.2.1 Normal subgroups of $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$

In this section we prove the following:

Proposition 5.2.1. *Let $N \lhd G_n$ be a proper normal subgroup. Then there exists $m \leq n$, $W \leq Z(G_n)$ such that $N = K_m W$.*

Lemma 5.2.2. *$(K_i/K_n)_{1\leq i\leq n}$ is the upper central series for K_1/K_n; that is:*

$$K_i/K_{i+1} = Z(K_1/K_{i+1}) \text{ for every } 1 \leq i \leq n.$$

Proof. Let $g, h \in K_1$. Direct computation yields that if $h \equiv I_2 \mod p^i$, then $gh \equiv hg \mod p^{i+1}$. Indeed, we already carried out the required computation in Section 3.3.

Conversely, if $h \in K_1$ satisfies $gh \equiv hg \mod p^{i+1}$ for every $g \in K_1$, then setting in turn:

$$g = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ p & 1 \end{pmatrix},$$

we have $h_{1,1} - h_{2,2}, h_{1,2}, h_{2,1} \in p^i\mathbb{Z}_p$. We may therefore write:

$$h = (1+\mu)I_2 + p^i Z$$

for some $\mu \in p\mathbb{Z}_p$, $Z \in \mathbb{M}_2(\mathbb{Z}_p)$.

But $\det(h) = 1$, so:

$$\mu(\mu + 2) + p^i(Z_{1,1} + Z_{2,2}) \in p^{i+1}\mathbb{Z}_p$$

and since $p > 2$, $\mu + 2 \notin p\mathbb{Z}_p$, $\mu \in p^i\mathbb{Z}_p$ and $h \in K_i$. □

Lemma 5.2.3. *For $m \leq n \leq 2m$, the map:*

$$(a, b, c) \mapsto I_2 + p^m \begin{pmatrix} a & b \\ c & -a \end{pmatrix} + K_n$$

is an isomorphism $(\mathbb{Z}/p^{n-m}\mathbb{Z})^3 \cong K_m/K_n$. In particular, for any $m \geq 1$, $\mathbb{F}_p^3 \cong K_m/K_{m+1}$.

Proof. Direct verification. □

Corollary 5.2.4. $G_n = p^{3n-2}(p^2 - 1)$.

By Lemma 5.2.2, the action of $\mathrm{SL}_2(\mathbb{Z}_p)$ on K_i by conjugation descends to an action of $\mathrm{SL}_2(\mathbb{F}_p)$ on K_i/K_{i+1}, which, by the above isomorphism, yields a three-dimensional representation of $\mathrm{SL}_2(\mathbb{F}_p)$.

Lemma 5.2.5. *The action of $\mathrm{SL}_2(\mathbb{F}_p)$ on K_i/K_{i+1} thus defined is irreducible.*

Proof. Note that it suffices to check that the image of $\mathrm{SL}_2(\mathbb{F}_p)$ under the representation spans $\mathbb{M}_3(\mathbb{F}_p)$. For if the image were to preserve a non-trivial proper subspace of K_i/K_{i+1}, then so too would any linear combination of elements in the image. In particular the image would not span $\mathbb{M}_3(\mathbb{F}_p)$.

$g = \begin{pmatrix} w & x \\ y & z \end{pmatrix} \in \mathrm{SL}_2(\mathbb{F}_p)$ acts under the representation via the matrix:

$$\begin{pmatrix} wz + xy & yz & -wx \\ 2xz & z^2 & -x^2 \\ -2wy & -y^2 & w^2 \end{pmatrix}$$

We may verify directly that a spanning set for $\mathbb{M}_3(\mathbb{F}_p)$ may be produced by varying g. For instance:

$$I_2, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

maps to a spanning set, since $p \geq 3$. □

Lemma 5.2.6. *Let P be a finite p-group. Every normal subgroup of P intersects $Z(P)$ non-trivially.*

Proof of Proposition 5.2.1. We classify the normal subgroups of G_n, proceeding by induction on n. If $n = 1$, $G_n = \mathrm{SL}_2(\mathbb{F}_p)$ is quasisimple. Let $n > 1$ and let $N \lhd G_n$. We distinguish two cases:

If $N \cap K_1 = 1$, then $N \cong NK_1/K_1 \lhd G_1$, so that if $N \cong NK_1/K_1$ is not central, $\pi_p|_N : N \to G_1$ is an isomorphism and the extension:

$$1 \to K_1 \hookrightarrow G_n \twoheadrightarrow G_1 \to 1$$

splits as a direct product. In particular N centralises K_1. However there exists $k \in K_1$ such that $g = k \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in N$, and for $\mu \in p^{n-1}\mathbb{Z}_p$,

$$\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}^g = \begin{pmatrix} 1+\mu & \mu \\ -\mu & 1-\mu \end{pmatrix} \neq \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix},$$

a contradiction. Hence NK_1/K_1 is central, so $N \leq Z(G_n)$ (as $\pi_p|_N$ is injective).

On the other hand, if $N \cap K_1 \neq 1$, then, by Lemma 5.2.2 and Lemma 5.2.6, $N \cap K_{n-1} \neq 1$, so that, by Lemma 5.2.5, $K_{n-1} \leq N$. Letting $1 \leq m \leq n-1$ be minimal such that $K_m \leq N$, apply the inductive hypothesis to $N/K_m \lhd G_m$, to deduce that $N/K_m \leq Z(G_m)$. Then $N \leq K_m Z(G_n)$, as required. □

5.2.2 Quasirandomness for $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$

In this section, we prove the following result:

Proposition 5.2.7. *Let ρ be an irreducible complex representation of G_n, such that $\ker(\rho) \leq Z(G_n)$. Then $\dim(\rho) \gg_p p^n$.*

It follows immediately from Proposition 5.2.1 and Corollary 5.2.4 that there exists $C(p) > 0$ such that any dense subgroup Γ of $\mathrm{SL}_2(\mathbb{Z}_p) = \varprojlim_n G_n$ is $(C, \frac{1}{3})$-weakly quasirandom with respect to the family of all finite index normal subgroups containing some $K_n \cap \Gamma$. Indeed by Proposition 2.2.4, the same conclusion holds if we merely assume that the closure of Γ in $\mathrm{SL}_2(\mathbb{Z}_p)$ is open.

Proposition 5.2.7 shall follow from bounds on the orbits of G_n, acting on the Pontryagin dual of an abelian normal subgroup. Our exposition of this argument is based on that appearing in [6] (though the proof there contains one oversight, which we indicate at the corresponding stage of our proof). See also [71] for a more general result. The author is not aware of any earlier proof of Proposition 5.2.7, but it is possible that one exists in the literature: the idea of understanding representations of a p-adic analytic group in terms of its action on the dual Lie algebra goes back to Howe [42], and draws on the work of Kirillov on nilpotent (real) Lie groups [48].

Recall that, for G a finite group, $N \lhd G$, G acts on irreducible complex representations of N as follows: for $g \in G$ and θ an irrep of N,

$$\theta^g(n) = \theta(n^g) \text{ for all } n \in N.$$

Theorem 5.2.8 (Clifford)**.** *Let ρ be an irrep of G. Let $\rho\mid_N = \sum_{i=1}^{d} \theta_i$ be the decomposition of $\rho\mid_N$ as a direct sum of irreps of N. Then $\theta_1, \ldots, \theta_d$ lie in a single orbit of the G-action, and all members of the orbit occur.*

Now let $n/2 \leq m \leq n$. Our goal shall be to apply Clifford's Theorem to $K_m \lhd G_n$. Define:

$$e_1 = I_2 + p^m \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, e_2 = I_2 + p^m \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, e_3 = I_2 + p^m \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in K_m.$$

As was observed in Lemma 5.2.3, $(a, b, c) \mapsto e_1^a e_2^b e_3^c$ defines an isomorphism $(\mathbb{Z}/p^{n-m}\mathbb{Z}) \to K_m$. The Pontryagin dual $\hat{K_m}$ is a free $(\mathbb{Z}/p^{n-m}\mathbb{Z})$-module with basis $\hat{e_1}, \hat{e_2}, \hat{e_3}$. It is clear that the action of G_n on $\hat{K_m}$ described above is via $(\mathbb{Z}/p^{n-m}\mathbb{Z})$-module automorphisms; under this action the matrix of $g = \begin{pmatrix} w & x \\ y & z \end{pmatrix} \in G_n$ with respect to the given basis is:

$$\begin{pmatrix} wz + xy & 2xz & -2wy \\ yz & z^2 & -y^2 \\ -wx & -x^2 & w^2 \end{pmatrix}. \tag{5.1}$$

In other words, it is precisely the transpose of the image of g, under the action of G_n by conjugation on the $(\mathbb{Z}/p^{n-m}\mathbb{Z})$-module K_m, with respect to basis e_1, e_2, e_3.

Proposition 5.2.9. *There exists $C(p) > 0$ such that for any $\chi \in \hat{K_m} \setminus p\hat{K_m}$,*

$$|\mathrm{Stab}_{G_n}(\chi)| \leq Cp^{2m+n}.$$

Proof of Proposition 5.2.7. Let $\rho\mid_N = \sum_{i=1}^{d} \theta_i$ be as in Clifford's Theorem. Then $\dim(\rho) \geq d$. Let $x \in K_{n-1}$ be non-central in G_n. Then, since $\ker(\rho) \leq Z(G_n)$, $\theta_i(x) \neq 0$ for some i. Note that for all $\chi \in p\hat{K_m}$, $\chi(K_{n-1}) = 0$. Hence $\theta_i \in \hat{K_m} \setminus p\hat{K_m}$. By Proposition 5.2.9 and Clifford's Theorem,

$$d = |G_n|/|\mathrm{Stab}_{G_n}(\theta_i)| \gg_p p^{2n-2m} \gg_p p^n,$$

as required. □

Remark 5.2.10. *It is at this stage that an error appears in the proof of Proposition 5.2.7 presented in [6]. There it is asserted that the required lower bound on the lengths of orbits of irreps appearing in Clifford's Theorem corresponds to an upper bound on the size of the stabilizers in the action of G_n on K_m, rather than the action on $\hat{K_m}$.*

Proof of Proposition 5.2.9. First note that K_{n-m} acts trivially on $\hat{K_m}$, so the action of G_n on $\hat{K_m}$ descends to an action of G_{n-m} and:

$$|\mathrm{Stab}_{G_n}(\chi)| = |\mathrm{Stab}_{G_{n-m}}(\chi)||K_{n-m}| = p^{3m}|\mathrm{Stab}_{G_{n-m}}(\chi)|.$$

It therefore suffices to check $|\mathrm{Stab}_{G_{n-m}}(\chi)| \ll_p p^{n-m}$. Moreover since

$$|\mathrm{Stab}_{G_{n-m}}(\chi) : K_1 \cap \mathrm{Stab}_{G_{n-m}}(\chi)| \leq |G_1 : K_1| = p^3 - p,$$

it suffices to bound $|K_1 \cap \mathrm{Stab}_{G_{n-m}}(\chi)|$. Let $g \in K_1 \cap \mathrm{Stab}_{G_{n-m}}(\chi)$ and let $w, x, y, z \in \mathbb{Z}/p^{n-m}\mathbb{Z}$ be such that:

$$g = I_2 + p \begin{pmatrix} w & x \\ y & z \end{pmatrix}.$$

There exist $a, b, c \in \mathbb{Z}/p^{n-m}\mathbb{Z}$ such that $\chi = a\hat{e_1} + b\hat{e_2} + c\hat{e_3}$ and p does not divide at least one of a, b, c. Using the matrix representation (5.1), the condition $\chi^g = \chi$ becomes the three constraints:

(i) $f_1(w, x, y, z) := (a(w + z) + 2bx - 2cy) + (a(wz + xy) + 2bxz - 2cwy)p \equiv 0,$

(ii) $f_2(w, x, y, z) := (ay + 2bz) + (ayz + bz^2 - cy^2)p \equiv 0,$

(iii) $f_3(w, x, y, z) := (-ax + 2cw) + (-awx - bx^2 + w^2)p \equiv 0,$

while the requirement that $\det(g) = 1$ becomes the constraint:

(iv) $f_4(w, x, y, z) := (w + z) + (wz - xy)p \equiv 0$

(with all these equivalences being modulo p^{n-m-1}). We seek to bound the number of solutions $(w, x, y, z) \in \mathbb{Z}/p^{n-m-1}\mathbb{Z}$. Write:

$$w = \sum_{i=0}^{n-m-2} w_i p^i; x = \sum_{i=0}^{n-m-2} x_i p^i; y = \sum_{i=0}^{n-m-2} y_i p^i; z = \sum_{i=0}^{n-m-2} z_i p^i,$$

with $w_i, x_i, y_i, z_i \in \{0, 1, \ldots, p-1\}$. Reducing (i)-(iv) modulo p, we obtain:

(i') $a(w_0 + z_0) + 2bx_0 + 2cy_0 \equiv 0 \mod p,$

(ii') $ay_0 + 2bz_0 \equiv 0 \mod p,$

(iii') $-ax_0 + 2cy_0 \equiv 0 \mod p,$

(iv') $w_0 + z_0 \equiv 0 \mod p.$

Applying (ii'), (iii') and (iv') if $a \not\equiv 0 \mod p$ (respectively (i'), (ii') and (iv') if $b \not\equiv 0 \mod p$; (i'), (iii') and (iv') if $c \not\equiv 0 \mod p$), there are at most p choices for (w_0, x_0, y_0, z_0).

We now iteratively apply this observation: at the $(k+1)$th stage suppose (w_i, x_i, y_i, z_i) is determined for $0 \leq i \leq k$, so that:

$$f_j\Big(\sum_{i=0}^{k} w_i p^i, \sum_{i=0}^{k} x_i p^i, \sum_{i=0}^{k} y_i p^i, \sum_{i=0}^{k} z_i p^i\Big) \equiv 0 \mod p^{k+1} \text{ for } j = 1, 2, 3, 4.$$

Reducing (i)-(iv) modulo p^{k+2} and dividing through by p^{k+1}, we have:

(i") $a(w_{k+1} + z_{k+1}) + 2bx_{k+1} + 2cy_{k+1}$
$+f_1(\sum_{i=0}^{k} w_i p^i, \sum_{i=0}^{k} x_i p^i, \sum_{i=0}^{k} y_i p^i, \sum_{i=0}^{k} z_i p^i)/p^{k+1} \equiv 0,$

(ii") $ay_{k+1} + 2bz_{k+1} + f_2(\sum_{i=0}^{k} w_i p^i, \sum_{i=0}^{k} x_i p^i, \sum_{i=0}^{k} y_i p^i, \sum_{i=0}^{k} z_i p^i)/p^{k+1} \equiv 0,$

(iii") $-ax_{k+1} + 2cy_{k+1} + f_3(\sum_{i=0}^{k} w_i p^i, \sum_{i=0}^{k} x_i p^i, \sum_{i=0}^{k} y_i p^i, \sum_{i=0}^{k} z_i p^i)/p^{k+1} \equiv 0,$

(iv") $w_{k+1} + z_{k+1} + f_4(\sum_{i=0}^{k} w_i p^i, \sum_{i=0}^{k} x_i p^i, \sum_{i=0}^{k} y_i p^i, \sum_{i=0}^{k} z_i p^i)/p^{k+1} \equiv 0$

(with all these equivalences being modulo p). As before, there are at most p possibilities for $(w_{k+1}, x_{k+1}, y_{k+1}, z_{k+1})$. In total then, there are at most p^{n-m-1} choices for (w, x, y, z), as required. □

5.2.3 ℓ^2-flattening

We can combine Proposition 5.2.7 and the comments immediately thereafter with Proposition 2.2.12 to reduce expansion to ℓ^2-flattening. Recall that by the ℓ^2-flattening lemma, it will suffice to check that the random walk does not concentrate on a small approximate subgroup of $G_n = \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$. Theorem 5.1.1 will follow from Theorem 2.2.18 (the ℓ^2-flattening lemma for approximate subgroups) and the following result, which tells us that if H_n is an approximate subgroup of G_n in which the random walk concentrates, then for some r (suitably bounded), $H_n^{(r)}$ contains a large congruence kernel of G_n. In particular, H_n must already be too large to satisfy conclusion (ii) of Theorem 2.2.18. Henceforth let $S_n = \pi_{p^n}(S)$.

Theorem 5.2.11. *There exist $C_0(S) > 0$ and $f_S : (0,1) \to \mathbb{N}$ such that the following holds. For all $\gamma \in (0,1)$ sufficiently small, there exists $\delta \in (0,1)$ such that if $n \in \mathbb{N}$ is sufficiently large, and $H_n \subseteq G_n$ is an $|G_n|^{3\delta}$-approximate subgroup satisfying $\mu_{S_n}^{(2l)}(H_n) \geq |G_n|^{-2\delta}$ for some $C_0 \log|G_n| \leq l$, then $K_{\gamma n} \subseteq H_n^{(f_S(\gamma))}$.*

Note that if C_0, δ satisfy the conditions of Theorem 5.2.11, then so do all larger C_0 and smaller δ.

Proof of Theorem 5.1.1. Let $C > 0$ be as in Theorem 2.2.18. Let $l \geq C_0 \log|G_n|$ be as in Theorem 5.2.11, and suppose (for a contradiction) that:

$$\|\mu_{S_n}^{(2l)}\|_2 > |G_n|^{-\delta}\|\mu_{S_n}^{(l)}\|_2$$

(for $\delta > 0$ sufficiently small, yet to be specified). By Theorem 2.2.18, applied to G_n, with $\nu = \mu_{S_n}^{(l)}$, G_n has a $C|G_n|^{C\delta}$-approximate subgroup $\tilde{H}_n$ such that:

$$|\tilde{H}_n| \leq C|G_n|^{C\delta}/\|\mu_{S_n}^{(l)}\|_2^2 \tag{5.2}$$

and for some $g \in G_n$, $\mu_{S_n}^{(l)}(g\tilde{H}_n) \geq C|G_n|^{-C\delta}$. If $\|\mu_{S_n}^{(l)}\|_2 < |G_n|^{-\frac{1}{2}+\epsilon}$ there is nothing to prove, so suppose otherwise. Then by (5.2),

$$|\tilde{H}_n| \leq C|G_n|^{1+C\delta-\epsilon}.$$

For δ sufficiently small and n sufficiently large we may assume $\tilde{H}_n$ is a $|G_n|^{\delta}$-approximate subgroup satisfying $\mu_{S_n}^{(l)}(g\tilde{H}_n) \geq |G_n|^{-\delta}$. By Remark 2.2.19 we may apply Theorem 5.2.11 to $H_n = \tilde{H}_n \cdot \tilde{H}_n$, so that $K_{\gamma n} \subseteq \tilde{H}_n^{(2f_S(\gamma))}$. By Remark 2.2.14 (ii),

$$\begin{aligned}
|K_{\gamma n}| &\leq |\tilde{H}_n^{(2f_S(\gamma))}| \\
&\leq |\tilde{H}_n|(|G_n|^{\delta})^{2f_S(\gamma)-1} \\
&\leq C|G_n|^{1+(2f_S(\gamma)-1+C)\delta-\epsilon}.
\end{aligned}$$

But $|K_{\gamma n}| \gg_p |G_n|^{1-\gamma}$, so:

$$1 \ll_\gamma |G_n|^{(2f_S(\gamma)-1+C)\delta-(\epsilon-\gamma)}.$$

Choosing $\gamma < \epsilon$ and $\delta < \frac{\epsilon-\gamma}{2f_S(\gamma)-1+C}$, and for n sufficiently large, this is a contradiction. It follows that:

$$\|\mu_{S_n}^{(2l)}\|_2 \leq |G_n|^{-\delta}\|\mu_{S_n}^{(l)}\|_2.$$

Therefore, letting $\alpha > 1$ be such that

$$\|\mu_{S_n}^{(l)}\|_2 \leq |G_n|^{-\frac{1}{2}+\alpha\epsilon}$$

we can achieve the ℓ^2-flattening inequality by iterating this argument at most m times, for any $m > \frac{(\alpha-1)\epsilon}{\delta}$. That is, for any:

$$C(\epsilon) > 2^{\frac{(\alpha-1)\epsilon}{\delta}} C_0.$$

$\|\mu_{S_n}^{(l)}\|_2 < |G_n|^{-\frac{1}{2}+\epsilon}$ for some $l \leq C(\epsilon)\log|G_n|$. Since $(|G_n : \pi_n(\Gamma)|)_n$ is bounded, $(\pi_n(\Gamma))_n$ also satisfies the ℓ^2-flattening condition. The result now follows from Proposition 2.2.12. □

Remark 5.2.12. *For any $H \subseteq G_n$ and $l \in \mathbb{N}$, if $\mu_{S_n}^{(l)}(H) \geq K$, then for any $0 \leq l' \leq l$,*

$$K \leq \mu_{S_n}^{(l)}(H) = \sum_{x \in G_n} \mu_{S_n}^{(l-l')}(x^{-1})\mu_{S_n}^{(l')}(xH)$$

so that, for some $g \in G_n$, $\mu_{S_n}^{(l')}(gH) \geq K$. In particular, taking $H = g\tilde{H}_n$ as in the proof of Theorem 5.1.1 above, and by Remark 2.2.19, we may hypothesize further in Theorem 5.2.11 that $\mu_{S_n}^{(2l')}(H_n) \geq |G_n|^{-2\delta}$ for all $0 \leq l' \leq l$.

The remainder of our work will be devoted to proving Theorem 5.2.11.

5.3 Non-Concentration in Subvarieties

In this section we prove exponentially fast decay of $\mu_S^{(l)}$ on proper subvarieties. Such non-concentration shall be a consequence of the expansion results for linear algebraic group over $\mathbb{Q}$ emerging from [69]. We will apply these decay estimates (1) to produce elements whose centralisers intersect significantly with approximate subgroups and (2) to produce linearly independent matrices arising as conjugates by elements in approximate subgroups of a fixed matrix.

5.3.1 Diophantine properties of p-adic numbers

We begin by noting a key property of p-adic numbers which are algebraic over $\mathbb{Q}$. It is this property which shall allow us to upgrade results on non-concentration of the random walk on subvarieties to results giving non-concentration on sublevel sets.

Definition 5.3.1. *Let $C > 0$. A subset $A \subseteq \mathbb{Q}_p$ is* C-strongly Diophantine *if, whenever $r \in \mathbb{N}$, $f \in \mathbb{Z}[X_1, \ldots, X_r]$ is a homogeneous polynomial of degree $m \geq 1$, expressible as a sum of at most k signed monic monomials, $a_1, \ldots, a_r \in A$, and $i > C(m + \log k)$, then:*

$$f(a_1, \ldots, a_r) \notin p^i\mathbb{Z}_p \setminus \{0\}.$$

We say A is strongly Diophantine *if there exists $C > 0$ such that A is C-strongly Diophantine.*

Informally, a set A is strongly Diophantine if one must apply the ring operations in $\mathbb{Q}_p$ a large number of times to A to produce a non-zero element of small p-adic norm. Our motivation for making this definition is the following:

Proposition 5.3.2. *Let $A = \{a_1, \ldots, a_r\} \subseteq \mathbb{Q}_p$. Suppose that for every $1 \leq i \leq r$, a_i is a root of some polynomial over $\mathbb{Q}$. Then A is strongly Diophantine.*

Indeed, the strong Diophantine property shall be the *only* property of algebraic numbers which is used in the proof of Theorem 5.1.1. We start with some elementary observations:

Lemma 5.3.3. *Let $B \subseteq A \subseteq \mathbb{Q}_p$. For every $C > 0$, if A is C-strongly Diophantine, then so is B.*

Lemma 5.3.4. *For any $r \in \mathbb{N}$ and any $C > 0$, the set:*

$$\{(a_1, \ldots, a_r) : \{a_1, \ldots, a_r\} \subseteq \mathbb{Q}_p \text{ is } C\text{-strongly Diophantine}\}$$

is closed in $\mathbb{Q}_p^r$.

Lemma 5.3.5. *Let $\lambda \in \mathbb{Q}_p \backslash \{0\}$. Let $n \in \mathbb{Z}$ be such that $\|\lambda\|_p = p^{-n}$. For every $C > 0$ and every $A \subseteq \mathbb{Q}_p$, if A is C-strongly Diophantine, then $\lambda \cdot A$ is $\max(C, C+n)$-strongly Diophantine.*

We now turn to examples. First, any finite set of integers is strongly Diophantine. For a non-zero integer of small p-adic norm must be divisible in $\mathbb{Z}$ by a large power of p, and hence be of large Euclidean norm.

Example 5.3.6. *If $A \subseteq \mathbb{Z} \cap [-N, N]$, then any product of elements from A of length at most m lies in $[-N^m, N^m]$, so a sum of at most k such products lies in $[-N^m k, N^m k]$. In particular, $A \subseteq \mathbb{Q}_p$ is $\frac{1}{\log p} \max(1, \log N)$-strongly Diophantine.*

This observation extends easily to finite sets of rational numbers: a non-zero rational number of small p-adic norm may also be of small Euclidean norm, but only when it has large denominator.

Example 5.3.7. *Let $A = \{a_1, \ldots, a_r\} \subseteq \mathbb{Q}$. For every $1 \leq i \leq r$, let $p_i, q_i \in \mathbb{Z}$, with $q_i \neq 0$ be such that $a_i = p_i/q_i$. Let $L = \operatorname{lcm}(|q_1|, \ldots, |q_r|)$. Let $p_i' \in \mathbb{Z}$ be such that $a_i = p_i'/L$. We have $|p_i'| \leq L|p_i|$.*

Let f, k, m be as in Definition 5.3.1, and suppose $\|f(a_1, \ldots, a_r)\|_p \leq p^{-i}$. We have:

$$f(a_1, \ldots, a_r) = f(p'_1, \ldots, p'_r)/L^m$$

so $\|f(p'_1, \ldots, p'_r)\|_p \leq p^{-i}$. It now follows from Example 5.3.6 that A is C-strongly Diophantine, with:

$$C = \tfrac{1}{\log p} \max(1, \log L + \log(\max_{1 \leq i \leq r} |p'_i|)).$$

We are now ready to prove Proposition 5.3.2. Essentially we do this by restricting scalars to replace algebraic numbers by matrices over $\mathbb{Q}$, and thereby reducing the problem to Example 5.3.7.

Proof of Proposition 5.3.2. Let K be the field generated over $\mathbb{Q}$ by A. Let $\mathcal{B} = \{b_1, \ldots, b_r\}$ be a basis of the (finite) field extension $(K : \mathbb{Q})$; we may take $\mathcal{B} \subseteq \mathbb{Z}_p$ and $b_1 = 1$. By restriction of scalars (with respect to $\mathcal{B}$), there is a $\mathbb{Q}$-algebra monomorphism:

$$\Phi : K \hookrightarrow \mathbb{M}_d(\mathbb{Q})$$

satisfying $\Phi(q) = qI_d$ for $q \in \mathbb{Q}$. Let $n \in \mathbb{N}$ be minimal such that $\Phi(p^n a_i) \in \mathbb{M}_d(\mathbb{Q} \cap \mathbb{Z}_p)$ for all $1 \leq i \leq r$. According to Lemma 5.3.5, it suffices to prove that $p^n \cdot A$ is strongly Diophantine. Let f, k, m be as in Definition 5.3.1, and suppose $p^{mn} f(a_1, \ldots, a_r) = f(p^n a_1, \ldots, p^n a_r) \in p^i \mathbb{Z}_p \setminus \{0\}$. Let $M = \Phi(f(p^n a_1, \ldots, p^n a_r)) = f(\Phi(p^n a_1), \ldots, \Phi(p^n a_r))$ and let $\mathrm{Adj}(M) \in \mathbb{M}_d(\mathbb{Q} \cap \mathbb{Z}_p)$ be its adjugate matrix. Then:

$$\begin{aligned}
0 \neq \det(M) =& (M \cdot \mathrm{Adj}(M))_{1,1} \\
&= \sum_{i=1}^{d} (M \cdot \mathrm{Adj}(M))_{i,1} b_i \\
&= \sum_{i=1}^{d} (\sum_{j=1}^{d} M_{i,j} \, \mathrm{Adj}(M)_{j,1}) b_i \\
&= \sum_{j=1}^{d} (\sum_{i=1}^{d} M_{i,j} b_i) \, \mathrm{Adj}(M)_{j,1} \\
&= f(p^n a_1, \ldots, p^n a_r) \sum_{j=1}^{d} \mathrm{Adj}(M)_{j,1} b_j \\
&\in p^i \mathbb{Z}_p \setminus \{0\}.
\end{aligned}$$

Now let $A' \subseteq \mathbb{Q}$ be the set of all entries of all $\Phi(p^n a_i)$. By Example 5.3.7, there exists $C > 0$ such that A' is strongly Diophantine. Each entry of M is a homogeneous polynomial in A': it has degree m and consists of at most $d^m k$ signed monic monomials. $\det(M)$ is a homogeneous polynomial in the entries of M of degree d, consisting of $d!$ signed monic monomials. Thus $\det(M)$ is a homogeneous polynomial in A' of degree dm, and consisting of at most $d!(d^m k)^d$ signed monic monomials. Hence:

$$i \leq C(\log(d!(d^m k)^d) + dm) = C \log d! + Cd \log k + d(1 + \log d)m$$

and the result follows. □

5.3.2 A sieving argument

Theorem 5.3.8 (Theorem 7.2 from [17])**.** *Let K be a field of characteristic zero, $X \subseteq \mathrm{GL}_d(K)$ a finite symmetric set such that $\Gamma := \langle X \rangle$ is not virtually soluble. Let $\mathbb{G}$ be the Zariski closure of Γ and $\mathcal{R}$ be its soluble radical. Suppose $\mathcal{V}$ is an algebraic subvariety in GL_d such that $\dim(\mathcal{R}(\mathcal{V} \cap \mathbb{G})) < \dim(\mathbb{G})$. Then there exist $C_1(\mathcal{V}), C_2(X) > 0$ such that:*

$$\mu_X^{(l)}(\mathcal{V}) \leq C_1 e^{-C_2 l}.$$

Henceforth let S be as in Theorem 5.1.1. Applying Theorem 5.3.8 to the connected algebraic group $\mathbb{G} = \mathrm{SL}_d^r$, and taking $X = S^r$, we deduce:

Proposition 5.3.9. *Let $f : \mathbb{M}_2(\mathbb{Z}_p)^r \to \mathbb{Z}_p$ be a polynomial over $\mathbb{Z}$ which does not vanish on $\mathrm{SL}_2(\mathbb{Z}_p)^r$. Then there exist $C_1(f), C_2(S) > 0$ such that, letting $V(f) \subseteq \mathbb{M}_2(\mathbb{Z}_p)^r$ be the algebraic variety defined by the vanishing of f,*

$$(\times_{i=1}^r \mu_S^{(l)})(V(f)) \leq C_1 e^{-C_2 l}.$$

Proof. In view of Theorem 5.3.8, it suffices to check that $\langle X \rangle$ is Zariski-dense in $\mathrm{SL}_2(\mathbb{Q}_p)^r$. Indeed, since the vanishing set of a non-trivial polynomial on $\mathrm{SL}_2(\mathbb{Z}_p)^r$ has measure zero, it suffices to check that the closure $\overline{\langle X \rangle}$ of $\langle X \rangle$ in $\mathrm{SL}_2(\mathbb{Z}_p)^r$ is open in $\mathrm{SL}_2(\mathbb{Z}_p)^r$.

$I_2 \in S^{(2)}$, as S is symmetric. Hence $\langle S^{(2)} \rangle^r \subseteq \langle X^{(2)} \rangle$. $F_2 \hookrightarrow \langle S \rangle$, so $F_2 \hookrightarrow \langle S^{(2)} \rangle$. Applying Proposition 5.1.2, $\overline{\langle S^{(2)} \rangle} \leq_o \mathrm{SL}_2(\mathbb{Z}_p)$, so $\overline{\langle S^{(2)} \rangle^r} \leq_o \mathrm{SL}_2(\mathbb{Z}_p)^r$, as required. □

We now employ the strong Diophantine hypothesis on the coefficients of the entries of S. In the case of homogeneous f this will allow us to upgrade the conclusion of Proposition 5.3.9: not only will the random walk fail to concentrate in the set of zeroes of f, but it will also fail to concentrate in an appropriate *sublevel set* of f.

Corollary 5.3.10. *Let $f : \mathbb{M}_2(\mathbb{Z}_p)^r \to \mathbb{Z}_p$ be a homogeneous polynomial of degree $d \geq 1$ over $\mathbb{Z}$ which does not vanish on $\mathrm{SL}_2(\mathbb{Z}_p)^r$. There exists $C_3 > 0$ such that, for $0 < l \leq C_3 m$,*

$$(\times_{i=1}^r \mu_S^{(l)})(\{\underline{X} \in \mathbb{M}_2(\mathbb{Z}_p)^r : p^m \mid f(\underline{X})\}) \leq C_1 e^{-C_2 l}.$$

Proof. Let $C > 0$ be such that the set A of entries occurring in elements of S is C-strongly Diophantine. Suppose f is a sum of at most e signed monic monomials. Let $g_1, \ldots, g_r \in \mathrm{SL}_2(\mathbb{Z}_p)$ be words of length l in S. Then the entries of the g_i are all expressible as sums of 2^l monic monomials of degree l in the elements of A, so $f(g_1, \ldots, g_r)$ is a sum of $2^{dl}e$ signed monic monomials of degree dl in the elements of A. Then for

$$m > C(\log(2^{dl}e) + dl) = Cd(1 + \log 2)l + C \log e,$$

if $f(g_1, \ldots, g_r) \equiv 0 \mod p^m$ then $f(g_1, \ldots, g_r) = 0$. The result now follows from Proposition 5.3.9, for any $C_3 < C^{-1}(d(1 + \log 2) + \log e)^{-1}$. □

Let H_n be as in Theorem 5.2.11. Since $\mu_{S_n}^{(2l)}$ concentrates on H_n and does not concentrate on sublevel sets of polynomials not vanishing on G_n, (as witnessed by Corollary 5.3.10), it follows that $\mu_{S_n}^{(2l)}$ concentrates somewhat on the complements of such sublevel sets in H_n. In particular we have:

Corollary 5.3.11. *Let f, C_3 be as in Corollary 5.3.10. There exist $C_4, C_5 > 0$ and $\delta'(\delta) > 0$, with $\lim_{\delta \to 0} \delta' = 0$ such that, for $\delta' n \leq 2l \leq C_3 m$ (and for δ sufficiently small),*

$$(\times_{i=1}^r \mu_{S_n}^{(2l)})(\{\underline{X} \in H_n^r : f(\underline{X}) \not\equiv 0 \mod p^m\}) \geq C_4 e^{-C_5 \delta l}.$$

Proof. By definition of H_n, and Remark 5.2.12,

$$(\times_{i=1}^r \mu_{S_n}^{(2l)})(H_n^r) \geq |G_n|^{-2r\delta}$$

so by Corollary 5.3.10, a lower bound for $(\times_{i=1}^r \mu_{S_n}^{(2l)})(\{\underline{X} \in H_n^r : f(\underline{X}) \not\equiv 0 \mod p^m\})$ is given by:

$$|G_n|^{-2r\delta} - C_1 e^{-2C_2 l} \geq C p^{-6r\delta n} - C_1 e^{-2C_2 l} \geq C e^{\frac{-6r\delta l \log p}{\delta'}} - C_1 e^{-2C_2 l}$$

(for some $C > 0$). Choosing:

$$\delta < \frac{C_2 \delta'}{3r \log p}$$

we have the required result, for any $C_4 < C, C_5 \leq 6r \log p / \delta'$. □

Corollary 5.3.12. *Under the conditions of Corollary 5.3.11, there exists $D(S) > 0$ such that, if $\delta' n \leq 2l \leq \min(C_3 m, Dn)$, then there exist $C_6, C_7 > 0$ such that:*

$$|\{\underline{X} \in (H_n \cap B_{S_n}(2l))^r : f(\underline{X}) \not\equiv 0 \mod p^m\}| \geq C_6 e^{C_7 l}.$$

Note that, since δ' could be chosen arbitrarily small in Corollary 5.3.11 (by choosing δ sufficiently small), we may in particular set $\delta' < \min(C_3 \frac{m}{n}, D)$, so that l satisfying the conditions of Corollary 5.3.12 exists.

Proof. Recall from Kesten's Theorem that there is a constant $C_S > 0$ such that for any $g \in \langle S \rangle$,

$$\mu_S^{(2l)}(g) \ll_S e^{-C_S l}.$$

Since the set of entries of elements of S is strongly Diophantine, and since all entries of elements of $B_S(2l)$ are expressible as a sum of at most 4^l monic monomials of total degree l in the entries of elements of S, there exists $D(S) > 0$ such that for $2l < Dn$, the restriction of $\pi_{p^n} : \mathrm{SL}_2(\mathbb{Z}_p) \to G_n$ to $B_S(2l)$ is injective.

Hence for any $g \in G_n$,

$$\mu_{S_n}^{(2l)}(g) \ll_S e^{-C_S l}.$$

Note also that, for any $X \subseteq G_n$, $\mu_{S_n}^{(2l)}(X) \leq |X| \max_{g \in X}(\mu_{S_n}^{(2l)}(g))$. It follows from Corollary 5.3.11 that:

$$|\{\underline{X} \in (H_n \cap B_{S_n}(2l))^r : f(\underline{X}) \not\equiv 0 \mod p^m\}| \gg_S e^{(C_S - C_5\delta)l}.$$

Choosing δ sufficiently small, we have the required result. □

The range of values of l for which Corollaries 5.3.10-5.3.12 may be applied directly will prove to be too restrictive for our purposes. However for certain special polynomials, it will be possible to generalise the conclusion of Corollary 5.3.10 to much larger values of l, using the following elementary observation:

Lemma 5.3.13. *Let $V \subseteq G_n^r$. Suppose that for some $M > 0$, $l_0 \in \mathbb{N}$,*

$$(\times_{i=1}^r \mu_{S_n}^{(2l_0)})(\underline{h}V) \leq M$$

for all $\underline{h} \in G_n^r$. Then for all $l \geq l_0$,

$$(\times_{i=1}^r \mu_{S_n}^{(2l)})(V) \leq M.$$

Proof. For any $g \in G_n$,

$$\mu_{S_n}^{(2l)}(g) = \sum_{h \in G_n} \mu_{S_n}^{(2l-2l_0)}(h)\mu_{S_n}^{(2l_0)}(h^{-1}g).$$

Hence:

$$\begin{aligned}(\times_{i=1}^r \mu_{S_n}^{(2l)})(V) &= \sum_{\underline{h} \in G_n^r} (\times_{i=1}^r \mu_{S_n}^{(2l-2l_0)})(\underline{h}) \sum_{\underline{g} \in V} (\times_{i=1}^r \mu_{S_n}^{(2l_0)})(\underline{h}^{-1}\underline{g}) \\ &= \sum_{\underline{h} \in G_n^r} (\times_{i=1}^r \mu_{S_n}^{(2l-2l_0)})(\underline{h})(\times_{i=1}^r \mu_{S_n}^{(2l_0)})(\underline{h}^{-1}V) \\ &\leq \max_{\underline{h} \in G_n^r} (\times_{i=1}^r \mu_{S_n}^{(2l_0)})(\underline{h}^{-1}V) \leq M.\end{aligned}$$

□

5.3.3 Examples

We illustrate the results of the previous section by applying them to some specific polynomials. At first glance, our choice of examples may seem somewhat arbitrary and unmotivated, but, as the reader will note later, non-vanishing of these polynomials will be what allows us to produce a large set of simultaneously diagonalisable elements in $H_n \cap B_{S_n}(2l)$, and to guarantee that by taking brackets of matrices conjugated by these elements we may generate most of G_n.

Example 5.3.14. *Define, for W, X, Y, Z 2×2 matrices,*

$$\mathrm{Arr}(W, X, Y, Z) = \begin{pmatrix} W_{1,1} & W_{2,1} & W_{1,2} & W_{2,2} \\ X_{1,1} & X_{2,1} & X_{1,2} & X_{2,2} \\ Y_{1,1} & Y_{2,1} & Y_{1,2} & Y_{2,2} \\ Z_{1,1} & Z_{2,1} & Z_{1,2} & Z_{2,2} \end{pmatrix}.$$

Letting $g_i = \begin{pmatrix} w_i & x_i \\ y_i & z_i \end{pmatrix}$, consider the following homogeneous polynomials over $\mathbb{Z}$:

(0) $f^{(0)}(g_1) = w_1x_1y_1z_1$,

(i) $f^{(i)}(g_1) = \mathrm{tr}(g_1)$,

(ii) $f^{(ii)}(g_1, g_2, g_3) = \det(\mathrm{Arr}(I_2, g_1, g_2, g_3))$,

(iii) $f^{(iii)}(g_1, g_2, g_3, g_4, g_5) = \det(\mathrm{Arr}(g_2 - g_1, g_3 - g_1, g_4 - g_1, g_5 - g_1))$,

(iv) $f^{(iv)}(g_1,\dots,g_9)=\det(M(g_1,\dots,g_9))$, *where* $M(g_1,\dots,g_9)\in\mathbb{M}_9(\mathbb{Z}_p)$ *is the matrix:*

$$\begin{pmatrix} 2(w_1z_1+x_1y_1) & 2y_1z_1 & -2w_1x_1 & 2x_1z_1 & z_1^2 & -x_1^2 & -2w_1y_1 & -y_1^2 & w_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2(w_9z_9+x_9y_9) & 2y_9z_9 & -2w_9x_9 & 2x_9z_9 & z_9^2 & -x_9^2 & -2w_9y_9 & -y_9^2 & w_9^2 \end{pmatrix}.$$

In each case, we produce a tuple of elements of $\mathrm{SL}_2(\mathbb{Z}_p)$ *at which the polynomial doesn't vanish. Then, by Corollaries 5.3.10-5.3.12, we conclude that, for suitable constants* $C_1,\dots,C_7>0$,

(a) If $2l\le C_3m$ *then:*

$$(\times_{i=1}^r\mu_S^{(l)})(\{\underline{X}\in\mathbb{M}_2(\mathbb{Z}_p)^r : f(\underline{X})\equiv 0 \mod p^m\})\le C_1e^{-C_2l};$$

(b) If $\delta' n\le 2l\le C_3m$ *then:*

$$(\times_{i=1}^r\mu_{S_n}^{(2l)})(\{\underline{X}\in H_n^r : f(\underline{X})\not\equiv 0 \mod p^m\})\ge C_4e^{-C_5\delta l};$$

(c) If $\delta' n\le 2l\le \min(C_3m,Dn)$ *then:*

$$|\{\underline{X}\in(H_n\cap B_{S_n}(2l))^r : f(\underline{X})\not\equiv 0 \mod p^m\}|\ge C_6e^{C_7l}.$$

The tuples of elements which witness non-vanishing are as follows.

(0) If $g_1=\begin{pmatrix}2&1\\1&1\end{pmatrix}$ *then* $f^{(0)}(g_1)=2$.

(i) If $g_1=I_2$ *then* $f^{(i)}(g_1)=2$.

(ii) If $g_1=\begin{pmatrix}1&1\\0&1\end{pmatrix}$, $g_2=\begin{pmatrix}1&0\\1&1\end{pmatrix}$, $g_3=\begin{pmatrix}2&1\\1&1\end{pmatrix}$ *then* $f^{(ii)}(g_1,g_2,g_3)=1$.

(iii) If $g_1=I_2$, $g_2=\begin{pmatrix}1&0\\1&1\end{pmatrix}$, $g_3=\begin{pmatrix}1&1\\0&1\end{pmatrix}$, $g_4=\begin{pmatrix}2&1\\1&1\end{pmatrix}$, $g_5=\begin{pmatrix}1&1\\1&2\end{pmatrix}$ *then* $f^{(iii)}(g_1,g_2,g_3,g_4,g_5)=1$.

(iv) If $g_1=I_2$, $g_2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$, $g_3=\begin{pmatrix}1&1\\0&1\end{pmatrix}$, $g_4=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$, $g_5=\begin{pmatrix}1&0\\1&1\end{pmatrix}$, $g_6=\begin{pmatrix}1&0\\-1&1\end{pmatrix}$, $g_7=\begin{pmatrix}1&1\\-1&0\end{pmatrix}$, $g_8=\begin{pmatrix}-1&1\\-1&0\end{pmatrix}$, $g_9=\begin{pmatrix}2&1\\1&1\end{pmatrix}$, *then* $f^{(iv)}(g_1,g_2,g_3,g_4,g_5,g_6,g_7,g_8,g_9)=512$.

5.3.4 Producing commuting elements

We now establish a series of non-vanishing results for polynomials defined in terms of $X \in \mathbb{M}_2(\mathbb{Z}_p) \setminus \{0\}$, under certain conditions. A key point here is that the bounds we obtain are independent of X (as will turn out to be necessary when we apply these bounds later). We therefore cannot proceed by applying Corollary 5.3.12 directly to the "obvious" polynomials defined in terms of X; rather we must first trap the subvarieties defined by the vanishing of these polynomials inside new subvarieties, whose defining polynomials do not depend on X. We then apply Corollary 5.3.12 to these new defining polynomials.

In the present section, we use this strategy to study the map $\mathrm{SL}_2(\mathbb{Z}_p) \to \mathbb{Z}_p$ given by $g \mapsto \mathrm{tr}(gX)$. The fibres of this map are not Zariski-dense, which means that the set of traces of elements in $H_n \cdot X$ is large (since the random walk concentrates on H_n). In particular, taking $X \in H_n$, we see that many traces occur among elements of $H_n^{(2)}$, so that $H_n^{(2)}$ intersects many conjugacy classes in G_n. Applying the approximate class equation (Lemma 2.2.15), we locate an element of $H_n^{(2)}$ (which may be taken to be semisimple) with large centralizer in $H_n^{(2)}$.

Proposition 5.3.15. *There exist $C_8, C_9 > 0$ such that if $X \not\equiv 0 \mod p^k$ and $\delta' n \leq 2l \leq \min(C_3 m, Dn)$ then:*

$$|\{(h_1, h_2, h_3) \in (H_n \cap B_{S_n}(2l))^3 : (\mathrm{tr}(X), \mathrm{tr}(h_1 X), \mathrm{tr}(h_2 X), \mathrm{tr}(h_3 X)) \not\equiv 0 \mod p^{m+k}\}| \geq C_8 e^{C_9 l}.$$

Proof. Let $[X] = (a, b, c, d)^T$ (the co-ordinate vector of $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with respect to the standard basis). Let $h_1, h_2, h_3 \in G_n$. Then:

$$\mathrm{Arr}(I_2, h_1, h_2, h_3)[X] = (\mathrm{tr}(X), \mathrm{tr}(h_1 X), \mathrm{tr}(h_2 X), \mathrm{tr}(h_3 X))^T. \tag{5.3}$$

If $h_1, h_2, h_3 \in H_n \cap B_{S_n}(2l)$ are such that:

$$(\mathrm{tr}(X), \mathrm{tr}(h_1 X), \mathrm{tr}(h_2 X), \mathrm{tr}(h_3 X)) \equiv 0 \mod p^{m+k}$$

then:

$$\det(\mathrm{Arr}(I_2, h_1, h_2, h_3)) \equiv 0 \mod p^m.$$

The result follows from Example 5.3.14 (ii) (using Corollary 5.3.12). □

Remark 5.3.16. *Letting h_1, h_2, h_3 be such that* $\det(\mathrm{Arr}(I_2, h_1, h_2, h_3)) \not\equiv 0 \mod p^m$, *if $X, Y \in \mathbb{M}_2(\mathbb{Z}_p)$ are such that:*

$$(\mathrm{tr}(X), \mathrm{tr}(h_1X), \mathrm{tr}(h_2X), \mathrm{tr}(h_3X)) \equiv (\mathrm{tr}(Y), \mathrm{tr}(h_1Y), \mathrm{tr}(h_2Y), \mathrm{tr}(h_3Y)) \mod p^{m+k}$$

then by (5.3),

$$X \equiv Y \mod p^k.$$

Taking $k = n - m$ in the above, we have for any $V \subseteq \mathbb{M}_2(\mathbb{Z}/p^n\mathbb{Z})$,

$$|V| \leq |\{(\mathrm{tr}(X), \mathrm{tr}(h_1X), \mathrm{tr}(h_2X), \mathrm{tr}(h_3X)) : X \in V\}| \cdot |p^{n-m}\mathbb{M}_2(\mathbb{Z}/p^n\mathbb{Z})|.$$

Noting that:

$$|\{(\mathrm{tr}(X), \mathrm{tr}(h_1X), \mathrm{tr}(h_2X), \mathrm{tr}(h_3X)) : X \in V\}| \leq |\mathrm{tr}(V)||\mathrm{tr}(h_1V)||\mathrm{tr}(h_2V)||\mathrm{tr}(h_3V)|$$

we conclude that for some $g \in \{I_2, h_1, h_2, h_3\}$,

$$|\mathrm{tr}(gV)| \geq |V|^{\frac{1}{4}} p^{-m}.$$

Proposition 5.3.17. *There exist $C_{10}, C_{11} > 0$ such that the following holds. Let $a_1, \ldots, a_r \in \mathbb{Z}_p$; $X_1, \ldots, X_r \in \mathbb{M}_2(\mathbb{Z}_p)$, with $X_1, \ldots, X_r \not\equiv 0 \mod p^k$. Then for $2l_0 \leq C_3 m$,*

$$(\mu_{S_n}^{(2l_0)})(\{g \in \mathbb{M}_2(\mathbb{Z}_p) : \mathrm{tr}(X_i g) \equiv a_i \mod p^{m+k} \text{ for some } 1 \leq i \leq r\}) \leq C_{10} e^{-C_{11} l_0}.$$

Proof. Let $U_i = \{g \in B_S(2l_0) : \mathrm{tr}(X_i g) \equiv a_i \mod p^{m+k}\}$. Then for g_1, g_2, g_3, $g_4, g_5 \in U_i$,

$$\mathrm{Arr}(g_2 - g_1, g_3 - g_1, g_4 - g_1, g_5 - g_1)[X_i] \equiv 0 \mod p^{m+k}$$

so:

$$\det(\mathrm{Arr}(g_2 - g_1, g_3 - g_1, g_4 - g_1, g_5 - g_1)) \equiv 0 \mod p^m.$$

As noted in Example 5.3.14 (iii), this polynomial doesn't vanish on $\mathrm{SL}_2(\mathbb{Z}_p)^5$. By Corollary 5.3.10, for some $C_1, C_2 > 0$,

$$(\times_{i=1}^5 \mu_{S_n}^{(2l_0)})(U_i^5) \leq C_1 e^{-C_2 l_0}$$

hence

$$\mu_{S_n}^{(2l_0)}(U_i) \leq C_1^{\frac{1}{5}} e^{\frac{-C_2 l_0}{5}}.$$

The result follows, since:

$$\mu_{S_n}^{(2l_0)}(\bigcup_{j=1}^{r} U_j) \leq r\mu_{S_n}^{(2l_0)}(U_i), \text{ for some } 1 \leq i \leq r.$$

□

Corollary 5.3.18. *For $\delta' > 0$ sufficiently small and $\delta' n \leq 2l \leq Dn$,*

$$|\{g \in H_n \cap B_{S_n}(2l) : \mathrm{tr}(X_i g) \not\equiv a_i \mod p^{m+k} \text{ for some } 1 \leq i \leq r\}| \geq C_{12} e^{C_{13} l}.$$

Proof. Let $C_3 > 0$ be as in Proposition 5.3.17. Let $\delta' n \leq 2l_0 \leq C_3 m$. Since C_{10}, C_{11} in Proposition 5.3.17 were independent of $X_1, \ldots, X_r$, and since for all $h \in G_n$, $X_i \equiv 0 \mod p^k$ iff $hX_i \equiv 0 \mod p^k$,

$$V := \{g \in \mathbb{M}_2(\mathbb{Z}_p) : \mathrm{tr}(X_i g) \equiv a_i \mod p^{m+k} \text{ for some } 1 \leq i \leq r\}$$

satisfies the hypothesis of Lemma 5.3.13 with $M = C_{10} e^{-C_{11} l_0}$. Hence for $2l_0 \leq 2l \leq Dn$,

$$\mu_{S_n}^{(2l)}(V) \leq C_{10} e^{-C_{11} l_0}.$$

Arguing as in Corollaries 5.3.11 and 5.3.12, we have the required result. □

Theorem 5.3.19. *There exist $C_{14}, C_{15}, C_{16} > 0$ such that, for $\delta' n \leq 2l, 2l' \leq Dn$, with $2l' \leq C_3 m$ there exists $g \in H_n^{(2)} \cap B_{S_n}(2(l + l'))$ such that $\mathrm{tr}(g) \not\equiv \pm 2 \mod p^m$ and:*

$$|H_n^{(2)} \cap C_{G_n}(g)| \geq C_{14} e^{C_{15} l - C_{16}(m + \delta n)}.$$

Proof. Let $h_1, h_2, h_3 \in H_n \cap B_{S_n}(2l')$ be as in Proposition 5.3.15, and set $h_4 = I_2$. Let $V = \{g \in H_n \cap B_{S_n}(2l) : \mathrm{tr}(h_i g) \not\equiv \pm 2 \mod p^m \text{ for } i = 1, 2, 3, 4\}$. Applying Corollary 5.3.18 with $r = 8$; $a_1 = \ldots = a_4 = 2$; $a_5 = \ldots = a_8 = -2$ and $X_i = X_{i+4} = h_i$ for $i = 1, 2, 3, 4$, we have, for some $C_{12}, C_{13} > 0$:

$$|V| \geq C_{12} e^{C_{13} l}.$$

As was noted in Remark 5.3.16, for some $i \in \{1, 2, 3, 4\}$,

$$|\mathrm{tr}(h_i V)| \geq C_{12}^{\frac{1}{4}} e^{\frac{C_{13} l}{4}} p^{-m}.$$

In particular, there exists $U \subseteq h_i V$ consisting of at least $C_{12}^{\frac{1}{4}} e^{\frac{C_{13} l}{4}} p^{-m}$ elements which are pairwise non-conjugate in G_n. It follows from Lemma 2.2.15 that for some $g \in U$,

$$|H_n^{(2)} \cap C_{G_n}(g)| \geq C_{12}^{\frac{1}{4}} e^{\frac{C_{13} l}{4}} p^{-m} (|G_n|^{3\delta})^{-3}.$$

This yields the required lower bound. □

5.3.5 Producing independent conjugates

In this section we study the adjoint action of $\mathrm{SL}_2(\mathbb{Z}_p)$ on the Lie ring $\mathfrak{sl}_2(\mathbb{Z}_p)$ of traceless matrices over $\mathbb{Z}_p$. This representation is irreducible, so given traceless $X \in \mathbb{M}_2(\mathbb{Z}_p) \setminus p\mathbb{M}_2(\mathbb{Z}_p)$, the orbit of X spans $\mathfrak{sl}_2(\mathbb{Z}_p)$. Using the concentration of the random walk on H_n, we deduce that there exist $g_1, g_2 \in H_n$ such that the $\mathbb{Z}_p$-span of I_2, X, X^{g_1}, X^{g_2} contains a large ball around 0 in $\mathbb{M}_2(\mathbb{Z}_p)$.

First recall the adjoint representation $\mathrm{Ad} : \mathrm{SL}_2(\mathbb{Z}_p) \to \mathrm{Aut}_{\mathbb{Z}_p}(\mathfrak{sl}_2(\mathbb{Z}_p))$ of $\mathrm{SL}_2(\mathbb{Z}_p)$, given by $\mathrm{Ad}(g)X = X^g$. Consider the $\mathbb{Z}_p$-basis $\mathcal{B} = \{D_{1,2}, E_{1,2}, E_{2,1}\}$ for $\mathfrak{sl}_2(\mathbb{Z}_p)$, given by:

$$D_{1,2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

With respect to $\mathcal{B}$, the matrix of $\mathrm{Ad}(g)$ is given by:

$$M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{Ad}(g)) = \begin{pmatrix} wz + xy & yz & -wx \\ 2xz & z^2 & -x^2 \\ -2wy & -y^2 & w^2 \end{pmatrix}, \text{ where } g = \begin{pmatrix} w & x \\ y & z \end{pmatrix}.$$

There is a natural symmetric bilinear form $\langle -, - \rangle$ defined on $\mathbb{M}_2(\mathbb{Q}_p)$ by:

$$\langle X, Y \rangle = \mathrm{tr}(X \cdot Y^T).$$

We have:

$$\langle D_{1,2}, D_{1,2} \rangle = 2,\ \langle E_{1,2}, E_{1,2} \rangle = 1,\ \langle E_{2,1}, E_{2,1} \rangle = 1,\ \langle D_{1,2}, E_{1,2} \rangle = 0,\ \langle D_{1,2}, E_{2,1} \rangle = 0,$$
$$\langle E_{1,2}, E_{2,1} \rangle = 0.$$

It follows that for $X, Y \in \mathfrak{sl}_2(\mathbb{Z}_p)$:

$$\begin{aligned} \langle X^g, Y \rangle = 2((wz + xy)X_{1,1} + yzX_{1,2} - wxX_{2,1})Y_{1,1} \\ +(2xzX_{1,1} + z^2X_{1,2} - x^2X_{2,1})Y_{1,2} \\ +(-2wyX_{1,1} - y^2X_{1,2} + w^2X_{2,1})Y_{2,1}. \end{aligned} \tag{5.4}$$

Proposition 5.3.20. *There exist $C_{19}, C_{20} > 0$ such that, if $X, Y \in \mathfrak{sl}_2(\mathbb{Z}_p)$, with $X, Y \not\equiv 0 \mod p$, then for $\delta' n \leq 2l \leq \min(C_3 m, Dn)$:*

$$|\{g \in H_n \cap B_{S_n}(2l) : \langle X^g, Y \rangle \not\equiv 0 \mod p^m\}| \geq C_{19} e^{C_{20} l}.$$

Proof. Let $U = \{g \in B_S(2l) : \langle X^g, Y\rangle \equiv 0 \mod p^m\}$. Let $g_1, \ldots, g_9 \in U$. Recall the matrix $M(g_1, \ldots, g_9) \in \mathbb{M}_9(\mathbb{Z}_p)$ defined in Example 5.3.14 (iv) to be:

$$\begin{pmatrix} 2(w_1z_1 + x_1y_1) & 2y_1z_1 & -2w_1x_1 & 2x_1z_1 & z_1^2 & -x_1^2 & -2w_1y_1 & -y_1^2 & w_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2(w_9z_9 + x_9y_9) & 2y_9z_9 & -2w_9x_9 & 2x_9z_9 & z_9^2 & -x_9^2 & -2w_9y_9 & -y_9^2 & w_9^2 \end{pmatrix}$$
$$(\text{for } g_i = \begin{pmatrix} w_i & x_i \\ y_i & z_i \end{pmatrix}).$$

Define $\underline{v}(X, Y) \in \mathbb{Z}_p^9$ to be:

$$(X_{1,1}Y_{1,1}, X_{1,2}Y_{1,1}, X_{2,1}Y_{1,1}, X_{1,1}Y_{1,2}, X_{1,2}Y_{1,2}, X_{2,1}Y_{1,2}, X_{1,1}Y_{2,1}, X_{1,2}Y_{2,1}, X_{2,1}Y_{2,1})^T.$$

It follows from (5.4) that:

$$M(g_1, \ldots, g_9)\underline{v}(X, Y) = (\langle X^{g_1}, Y\rangle, \ldots, \langle X^{g_9}, Y\rangle)^T \equiv 0 \mod p^m$$

(since $g_1, \ldots, g_9 \in U$). However $X, Y \notin p\mathbb{M}_2(\mathbb{Z}_p)$, so $\underline{v}(X, Y) \notin (p\mathbb{Z}_p)^9$. It follows that $\det(M(g_1, \ldots, g_9)) \equiv 0 \mod p^m$.

As noted in Example 5.3.14 (iv), $\det(M(g_1, \ldots, g_9))$ does not vanish on $\mathrm{SL}_2(\mathbb{Z}_p)^9$. It follows from Corollary 5.3.10 that for some $C_1, C_2 > 0$,

$$(\times_{i=1}^9 \mu_{S_n}^{(2l)})(U^9) \leq C_1 e^{-C_2 l}$$

hence

$$\mu_{S_n}^{(2l)}(U) \leq C_1^{\frac{1}{9}} e^{\frac{-C_2 l}{9}}.$$

Arguing as in Corollaries 5.3.11 and 5.3.12, the result follows. □

Corollary 5.3.21. *Let X, l, m be as in Proposition 5.3.20. Then there exist $g_1, g_2 \in H_n \cap B_{S_n}(2l)$ such that the following hold.*

(i) For any $Y \in \mathbb{M}_2(\mathbb{Z}_p)$, there exist $\kappa, \lambda, \mu, \nu \in \mathbb{Z}_p$ such that,

$$p^{m-1}Y \equiv \kappa I_2 + \lambda X + \mu X^{g_1} + \nu X^{g_2} \mod p^n.$$

(ii) For any $Z \in \mathfrak{sl}_2(\mathbb{Z}_p)$ with $Z \not\equiv 0 \mod p$, at least one of $[X, Z], [X^{g_1}, Z]$, $[X^{g_2}, Z] \not\equiv 0 \mod p^m$ (here $[X_1, X_2] = X_1X_2 - X_2X_1$ denotes the Lie bracket).

Proof. There exists $W_1 \in \mathfrak{sl}_2(\mathbb{Z}_p)$ such that $W_1 \not\equiv 0 \mod p$ and $\langle X, W_1\rangle = 0$. By Proposition 5.3.20, there exists $g_1 \in H_n \cap B_{S_n}(2l)$ such that $\langle X^{g_1}, W_1\rangle \not\equiv 0 \mod p^m$.

Now, there exists $W_2 \in \mathfrak{sl}_2(\mathbb{Z}_p)$ such that:

$$W_2 \not\equiv 0 \mod p \text{ and } \langle X, W_2\rangle = \langle X^{g_1}, W_2\rangle = 0.$$

By Proposition 5.3.20 again, there exists $g_2 \in H_n \cap B_{S_n}(2l)$ such that $\langle X^{g_2}, W_2\rangle \not\equiv 0 \mod p^m$. Note that:

$$\mathbb{Z}_p \cdot X = \{W \in \mathfrak{sl}_2(\mathbb{Z}_p) : \langle W, W_1\rangle = \langle W, W_2\rangle = 0\}.$$

We check that g_1, g_2 have the required properties.

(i) Let $Y \in \mathbb{M}_2(\mathbb{Z}_p)$. Let $\tilde{g}_1, \tilde{g}_2$ be lifts to $\mathrm{SL}_2(\mathbb{Z}_p)$ of g_1, g_2, respectively. First, $Y' = Y - \frac{\mathrm{tr}(Y)}{2} I_2 \in \mathfrak{sl}_2(\mathbb{Z}_p)$. Second,

$$Y'' = Y' - \frac{\langle Y', W_2\rangle}{\langle X^{\tilde{g}_2}, W_2\rangle} X^{\tilde{g}_2} \in \mathfrak{sl}_2(\mathbb{Q}_p)$$

satisfies $\langle Y'', W_2\rangle = 0$. Third,

$$Y''' = Y'' - \frac{\langle Y'', W_1\rangle}{\langle X^{\tilde{g}_1}, W_1\rangle} X^{\tilde{g}_1} \in \mathfrak{sl}_2(\mathbb{Q}_p)$$

satisfies $\langle Y''', W_1\rangle = \langle Y''', W_2\rangle = 0$. As noted above, it follows that for some $a \in \mathbb{Z}_p$, $Y''' = aX$. Altogether, we have:

$$\begin{aligned} p^{m-1}Y = \frac{\mathrm{tr}(Y)}{2} p^{m-1} I_2 + a p^{m-1} X \\ + \left(\frac{\langle Y', W_1\rangle}{\langle X^{\tilde{g}_1}, W_1\rangle} - \frac{\langle Y', W_2\rangle}{\langle X^{\tilde{g}_2}, W_2\rangle}\right) p^{m-1} X^{\tilde{g}_1} + \frac{\langle Y', W_2\rangle}{\langle X^{\tilde{g}_2}, W_2\rangle} p^{m-1} X^{\tilde{g}_2}. \end{aligned}$$

Since $\frac{p^{m-1}}{\langle X^{\tilde{g}_1}, W_1\rangle}, \frac{p^{m-1}}{\langle X^{\tilde{g}_2}, W_2\rangle} \in \mathbb{Z}_p$, our expression for $p^{m-1}Y$ is of the required form upon reducing modulo p^n.

(ii) Suppose $[Z, X], [Z, X^{g_1}], [Z, X^{g_2}] \equiv 0 \mod p^m$. Since $Z \not\equiv 0 \mod p$, at least one of $[Z, D_{1,2}], [Z, E_{1,2}], [Z, E_{2,1}] \not\equiv 0 \mod p$. But by (i), each of $p^{m-1}D_{1,2}$, $p^{m-1}E_{1,2}, p^{m-1}E_{2,1}$ is expressible as a $\mathbb{Z}_p$-combination of X, X^{g_1}, X^{g_2}, so $[Z, p^{m-1}D_{1,2}], [Z, p^{m-1}E_{1,2}], [Z, p^{m-1}E_{2,1}] \equiv 0 \mod p^m$, a contradiction.

□

5.4 Proof of the Main Theorem

In this section we gather together the results of Section 5.3, along with those of Section 2.6, to prove Theorem 5.2.11.

Section 5.4.1 relies only on Corollary 5.3.21. It witnesses that, given $X, Y \in \mathfrak{sl}_2(\mathbb{Z}_p)$, we can produce matrices spanning a large ball in $\mathfrak{sl}_2(\mathbb{Z}_p)$ by taking conjugates in H_n and nested brackets. In other words, our spanning matrices shall be conjugates of a fixed matrix of the form:

$$[Y^{h_s}, [Y^{h_{s-1}}, \ldots, [Y^{h_1}, X]\ldots]] \tag{5.5}$$

Section 5.4.2 combines Theorem 5.3.19 with the approximate class equation to produce a large set of simultaneously diagonalisable elements in $H_n^{(4)}$. We shall use these elements to produce a subset A of $\mathbb{Z}/p^n\mathbb{Z}$ to which the sum-product theorem is applicable. Specifically, for g, h semisimple elements lying in the same maximal torus, having eigenvalues $\lambda^{\pm 1}, \mu^{\pm 1}$ respectively, we have:

$$(\mu^2 - \mu^{-2})[g, X] = (\lambda - \lambda^{-1})(hXh^{-1} - h^{-1}Xh) \tag{5.6}$$

for any 2-by-2 matrix X. Our set A is (a slight modification of) the set of $\mu^2 - \mu^{-2}$, as h ranges over the intersection of our torus with $H_n^{(4)}$. This process is closely reminiscent of an argument used crucially in [36].

In Section 5.4.3 we combine the results of Sections 5.4.1 and 5.4.2 with the sum-product theorem. We begin with $Z \in \mathfrak{sl}_2(\mathbb{Z}_p)$ lying in a sufficiently small (but still moderately large) ball around 0. As observed in Section 5.4.3, we can write Z as a linear combination of H_n-conjugates of a nested bracket as in (5.5). Using the sum-product results of Section 2.6, we can decompose the coefficients of this linear combination as differences-of-sums-of-products of elements of A (up to some error). With this decomposition in hand, we can repeatedly apply (5.6) to replace the nested bracket by a difference-of-sums-of-conjugates of a fixed X. The conclusion is that there exist $M, N \in \mathbb{N}$ (depending only on S) such that every such Z is approximable by an element of

$$\Sigma_M X^{H_n^{(N)}} - \Sigma_M X^{H_n^{(N)}}$$

that is, of the difference set of the M-fold sumset of the set of $H_n^{(N)}$-conjugates of X.

At long last, in Section 5.4.4 we deduce Theorem 5.2.11 from this last result. Under the approximate correspondence $g \approx I_2 + p^m X$ between elements $g \in K_m$ and $p^m X \in p^m \mathfrak{sl}_2(\mathbb{Z}_p)$, products are approximated by sums and inversion is approximated by subtraction. Taking X such that $g \in K_m \cap H_n^{(2)}$, the fact that a large ball in $\mathfrak{sl}_2(\mathbb{Z}_p)$ is filled (up to small error) by $\Sigma_M X^{H_n^{(N)}} - \Sigma_M X^{H_n^{(N)}}$ translates to the set of corresponding products of elements of H_n filling a large congruence kernel in $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ (again, up to small error). Iterating the same procedure eliminates this error and concludes the proof of Theorem 5.2.11.

5.4.1 Generating traceless matrices

Let $X, Z \in \mathbb{M}_2(\mathbb{Z}_p) \setminus p\mathbb{M}_2(\mathbb{Z}_p)$ with $p \mid \mathrm{tr}(X)$ and $\mathrm{tr}(Z) = 0$. Then by Corollary 5.3.21 (ii), there exists $h \in H_n \cap B_{S_n}(2l)$ such that $[X^h, Z] \not\equiv 0 \mod p^m$. Hence, setting $h_1 = h^{-1}$ and conjugating, $[Z^{h_1}, X] \not\equiv 0 \mod p^m$. Let $m' < m$

be such that $[Z^{h_1}, X] \in p^{m'}\mathbb{M}_2(\mathbb{Z}_p) \setminus p^{m'+1}\mathbb{M}_2(\mathbb{Z}_p)$. $\mathrm{tr}([Z^{h_1}, X]) = 0$ so, applying Corollary 5.3.21 (ii) to $\frac{1}{p^{m'}}[Z^{h_1}, X]$, we obtain $h_2 \in H_n \cap B_{S_n}(2l)$ such that $[Z^{h_2}, [Z^{h_1}, X]] \not\equiv 0 \mod p^{2m}$.

Iterating this argument, for any $s \in \mathbb{N}$, there exist $h_1, \ldots, h_s \in H_n \cap B_{S_n}(2l)$ such that:

$$X' := [Z^{h_s}, [Z^{h_{s-1}}, \ldots, [Z^{h_1}, X] \ldots]] \not\equiv 0 \mod p^{sm}.$$

Let $m'' < sm$ be such that $X' \in p^{m''}\mathbb{M}_2(\mathbb{Z}_p) \setminus p^{m''+1}\mathbb{M}_2(\mathbb{Z}_p)$. Let $Y \in \mathbb{M}_2(\mathbb{Z}_p)$. Applying Corollary 5.3.21 (i) to $\frac{1}{p^{m''}}X'$, we have:

$$p^{m+m''}Y \equiv \kappa I_2 + \lambda X' + \mu(X')^{g_1} + \nu(X')^{g_2} \mod p^n \tag{5.7}$$

for some $\kappa, \lambda, \mu, \nu \in \mathbb{Z}_p$, $g_1, g_2 \in H_n \cap B_{S_n}(2l)$. Moreover if $\mathrm{tr}(Y) = 0$ we may take $\kappa = 0$.

5.4.2 Trace amplification

Recall Theorem 5.3.19:

Theorem 5.3.19. *There exist $C_{14}, C_{15}, C_{16} > 0$ such that, for $\delta' n \leq 2l, 2l' \leq Dn$, with $2l' \leq C_3 m$ there exists $g \in H_n^{(2)} \cap B_{S_n}(2(l+l'))$ such that $\mathrm{tr}(g) \not\equiv \pm 2 \mod p^m$ and:*

$$|H_n^{(2)} \cap C_{G_n}(g)| \geq C_{14} e^{C_{15} l - C_{16}(m+\delta n)}.$$

Corollary 5.4.1. *Under the conditions of Theorem 5.3.19,*

$$|H_n^{(4)} \cap C_{K_1}(g)| \geq \frac{C_{14}}{|\mathrm{SL}_2(p)|} e^{C_{15} l - C_{16}(m+\delta n)}.$$

Proof. There exists $g_0 \in G_n$ such that:

$$|(H_n^{(2)} \cap C_{G_n}(g)) \cap K_1 g_0| \geq \frac{1}{|\mathrm{SL}_2(p)|} |H_n^{(2)} \cap C_{G_n}(g)|.$$

Let $h_0 \in (H_n^{(2)} \cap C_{G_n}(g)) \cap K_1 g_0$. Then:

$$\{h h_0^{-1} : h \in (H_n^{(2)} \cap C_{G_n}(g)) \cap K_1 g_0\} \subseteq H_n^{(4)} \cap C_{K_1}(g).$$

The required result follows. □

For such g, let λ_g be a root of the characteristic polynomial $\chi_g(x) = x^2 - \mathrm{tr}(g)x + 1$ of g. Recall that $\mathbb{Z}_p[\lambda_g]$ is contained in $\mathcal{O} = \mathbb{Z}_p, \mathbb{Z}_p[\theta], \mathbb{Z}_p[\sqrt{p}]$ or $\mathbb{Z}_p[\theta\sqrt{p}]$, where θ satisfies the relation $\theta^2 = \alpha$, for some $\alpha \in \mathbb{Z}_p$ mapping to a non-square in $\mathbb{F}_p$. Embedding $G_n \hookrightarrow \mathrm{SL}_2(\mathcal{O}/(p^n))$, we have for some $A \in \mathrm{GL}_2(\mathcal{O})$:

$$g^A = \begin{pmatrix} \lambda_g & 0 \\ 0 & \lambda_g^{-1} \end{pmatrix}. \tag{5.8}$$

We note an elementary piece of linear algebra:

Lemma 5.4.2. *Let g be as above.*

(i) $\lambda_g \not\equiv \lambda_g^{-1} \mod p^{2m}$.

(ii) At least of one $g_{1,1} - g_{2,2}, g_{1,2}, g_{2,1} \not\equiv 0 \mod p^{2m}$.

(iii) If $h \in \mathrm{SL}_2(\mathcal{O})$ centralises g then h^A is congruent modulo p^{n-2m} to a diagonal matrix in $\mathrm{SL}_2(\mathcal{O})$.

Proof. (i) By Theorem 5.3.19, we have $\lambda_g + \lambda_g^{-1} \not\equiv \pm 2 \mod p^m$. First, suppose (for a contradiction) that $\lambda_g \equiv \pm 1 \mod p^m$. Then:

$$1 = \lambda_g \lambda_g^{-1} \equiv \pm \lambda_g^{-1}$$

so $\lambda_g, \lambda_g^{-1} \equiv \pm 1 \mod p^m$, which is impossible, by the above.

Now, suppose that $\lambda_g \equiv \lambda_g^{-1} \mod p^{2m}$. Then:

$$(\lambda_g + 1)(\lambda_g - 1) = \lambda_g(\lambda_g - \lambda_g^{-1}) \equiv 0 \mod p^{2m}$$

and at least one of $\lambda_g + 1, \lambda_g - 1 \equiv 0 \mod p^m$, a contradiction.

(ii) If $g_{1,1} - g_{2,2}, g_{1,2}, g_{2,1} \equiv 0 \mod p^{2m}$, then:

$$1 = \det(g) \equiv g_{1,1}^2 \equiv g_{2,2}^2 \mod p^{2m}$$

so $g_{1,1} \equiv g_{2,2} \equiv \pm 1 \mod p^{2m}$ (as p is odd), and $\mathrm{tr}(g) \equiv \pm 2 \mod p^{2m}$, a contradiction.

(iii) Consider $h^A = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$. If h centralises g, then:

$$(\lambda_g - \lambda_g^{-1})x, (\lambda_g - \lambda_g^{-1})y \equiv 0 \mod p^n$$

so $x, y \equiv 0 \mod p^{n-2m}$.

Finally, $1 = wz - xy \equiv wz \mod p^{2n-4m}$, so in particular, w, z are units in $\mathcal{O}$, and $w(w^{-1} - z) \equiv 0 \mod p^{2n-4m}$. Hence $z \equiv w^{-1} \mod p^{2n-4m}$. In summary, we have:

$$h^A \equiv \begin{pmatrix} \lambda_h & 0 \\ 0 & \lambda_h^{-1} \end{pmatrix} \mod p^{n-2m}$$

for some invertible $\lambda_h \in \mathcal{O}$.

□

Lemma 5.4.3. *Let $g, h, A, \lambda_g, \lambda_h$ be as above. For any $X \in \mathbb{M}_2(\mathcal{O})$,*

$$(\lambda_g - \lambda_g^{-1})(hXh^{-1} - h^{-1}Xh) \equiv (\lambda_h^2 - \lambda_h^{-2})[g, X] \mod p^{n-2m}.$$

Proof. Recalling that $hXh^{-1} - h^{-1}Xh = \mathrm{tr}(h)[h, X]$, and conjugating by A, it suffices to show that:

$$(\lambda_g - \lambda_g^{-1})[h^A, X^A] \equiv (\lambda_h - \lambda_h^{-1})[g^A, X^A] \mod p^{n-2m}.$$

But, letting $X^A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and for any invertible $\lambda \in \mathcal{O}$,

$$[\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}, X^A] = (\lambda - \lambda^{-1}) \begin{pmatrix} 0 & b \\ -c & 0 \end{pmatrix}$$

and the result follows. □

Lemma 5.4.4. $\lambda_h^2 - \lambda_h^{-2} \in \mathbb{Z}_p \mod p^{n-6m}$.

Proof. Let $Y = g - \frac{1}{2}\mathrm{tr}(g)I_2$. Then:

$$Y^A = \tfrac{1}{2}(\lambda_g - \lambda_g^{-1}) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In particular, $\lambda_g - \lambda_g^{-1}$ divides every entry of $[Y^A, X^A] = [g^A, X^A]$, so divides every entry of $[g, X]$. Hence $Z_X := \frac{1}{\lambda_g - \lambda_g^{-1}}[g, X] \in \mathbb{M}_2(\mathcal{O})$, and:

$$(\lambda_h^2 - \lambda_h^{-2})Z_X \equiv hXh^{-1} - h^{-1}Xh \mod p^{n-4m} \tag{5.9}$$

(since $\lambda_g - \lambda_g^{-1} \not\equiv 0 \mod p^{2m}$). Now let $g = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$. Then:

$$[g, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}] = \begin{pmatrix} -y & (w-z) \\ 0 & y \end{pmatrix}, [g, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}] = \begin{pmatrix} x & 0 \\ (z-w) & -x \end{pmatrix}.$$

Hence by Lemma 5.4.2 (ii), $Z_X \not\equiv 0 \mod p^{2m}$ for some $X \in \{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\}$.

On the other hand, by (5.9), $(\lambda_h^2 - \lambda_h^{-2})Z_X$ is congruent modulo p^{n-4m} to an element of $\mathbb{M}_2(\mathbb{Z}_p)$. In particular, $\pi_{p^{n-6m}}(\lambda_h^2 - \lambda_h^{-2}) \in \mathbb{Z}/p^{n-6m}\mathbb{Z}$. □

Lemma 5.4.5. *There exist $C'_{14}, C'_{15}, C'_{16} > 0$ such that the following holds. Let l, m be as in Theorem 5.3.19. Then:*

$$|\pi_{p^{n-6m}}(\{\lambda_h^2 - \lambda_h^{-2} : h \in H_n^{(4)} \cap C_{K_1}(g)\})| \geq C'_{14} p^{C'_{15} l - C'_{16}(m+\delta n)}.$$

Proof. Let $\mathcal{I} \subseteq 1 + p\mathcal{O}$ be a set of representatives of $\{\lambda_h : h \in H_n^{(4)} \cap C_{K_1}(g)\}$ (so that $\mathcal{I}$ contains at most one representative of each element of $\mathcal{O}/(p^n)$). For $\alpha \in \mathcal{O}$, suppose there exists $\lambda \in \mathcal{I}$ such that:

$$\lambda^2 - \lambda^{-2} \equiv \alpha \mod p^n.$$

Then, since $\lambda^2, \lambda^{-2} \equiv 1 \mod p$, $\alpha \equiv 0 \mod p$. Define:

$$F_\alpha(X) = X^4 - \alpha X^2 - 1 \in \mathcal{O}[X].$$

Then $F_\alpha(\lambda) \equiv 0 \mod p^n$, but $F'_\alpha(\lambda) = 4\lambda^3 - 2\alpha\lambda \equiv 4 \mod p$. By Hensel's Lemma, there is a unique $\lambda' \in \mathcal{O}$ such that $F_\alpha(\lambda') = 0$ and $\lambda \equiv \lambda' \mod p^n$. But $F_\alpha(X)$ has at most 4 roots in $\mathcal{O}$, so at most 4 elements $\lambda \in \mathcal{I}$ satisfy $\lambda^2 - \lambda^{-2} \equiv \alpha \mod p^n$. Setting $\mathcal{J} := \{\lambda^2 - \lambda^{-2} : \lambda \in \mathcal{I}\}$, it follows that:

$$|\mathcal{J}| \geq |\mathcal{I}|/4.$$

To summarise,

$$|\pi_{p^{n-6m}}(\mathcal{J})| \geq |\mathcal{J}|/p^{6m} \geq |\mathcal{I}|/4p^{6m} = |H_n^{(4)} \cap C_{K_1}(g)|/4p^{6m}$$

and the result follows from Corollary 5.4.1. □

Putting everything from this section together, we obtain:

Proposition 5.4.6. *There exists $C(S) > 0$ such that the following holds. Let $\gamma > 0$ and let $H_n \subseteq G_n$ be as in Theorem 5.2.11, with δ sufficiently small (depending on γ). Then there exists $A \subseteq \mathbb{Z}/p^n\mathbb{Z}$; $Y \in \mathbb{M}_2(\mathbb{Z}_p)$; $x_0 \in \mathcal{O}$ such that $x_0 \not\equiv 0 \mod p^{\frac{2}{C_3}\delta' n}$; $\frac{1}{x_0}Y \in \mathbb{M}_2(\mathcal{O})$; $\frac{1}{x_0}Y$ is not congruent modulo p to a multiple of I_2 and:*

(i) $|A| \geq (p^n)^{C-\gamma}$;

(ii) For all $X \in \mathbb{M}_2(\mathcal{O})$, $x \in A$, there exists $h \in H_n^{(4)}$ such that:

$$x[\tfrac{1}{x_0}Y, X] \equiv hXh^{-1} - h^{-1}Xh \mod p^{n(1-\gamma)}.$$

Proof. Take Y as in the proof of Lemma 5.4.4; $x_0 = \lambda_g - \lambda_g^{-1}$. As was observed there, $\frac{1}{x_0}Y \in \mathbb{M}_2(\mathcal{O})$ is similar in $\mathbb{M}_2(\mathcal{O})$ to $\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, so is not congruent modulo p to a multiple of I_2.

We take A to be a set of lifts to $\mathbb{Z}/p^n\mathbb{Z}$ of:

$$\pi_{p^{n-6m}}(\{\lambda_h^2 - \lambda_h^{-2} : h \in H_n^{(4)} \cap C_{K_1}(g)\}).$$

Such a set of lifts exists in $\mathbb{Z}/p^n\mathbb{Z}$, by Lemma 5.4.4.

We may take $l = \frac{n}{2D}$, $m = \frac{\delta' n}{C_3}(= \frac{2l'}{C_3})$ in Theorem 5.3.19, and, setting δ sufficiently small depending on γ, conditions (i), (ii) follow from Lemmas 5.4.5, 5.4.3, respectively. □

5.4.3 Applying the sum-product estimate

In this section we prove:

Theorem 5.4.7. *There exist $M, N \in \mathbb{N}$, $c, \gamma_0 \in (0,1)$ (depending only on S) such that for any $0 < \gamma < \gamma_0$, and for δ sufficiently small (depending on γ), the following holds. There exist $i_1, i_2 \in \mathbb{N}$, with $i_2 < i_1 \ll_S \gamma n$ and $i_1 - i_2 > c\gamma n$, such that for any $X, Z \in \mathbb{M}_2(\mathbb{Z}_p)$, satisfying $X \notin p\mathbb{M}_2(\mathbb{Z}_p)$, $p \mid \mathrm{tr}(X)$ and $\mathrm{tr}(Z) = 0$,*

$$p^{i_2} Z \in \Sigma_M X^{H_n^{(N)}} - \Sigma_M X^{H_n^{(N)}} \mod p^{i_1}.$$

This will follow from the construction achieved at (5.7): we take Z in place of Y in the LHS of (5.7) and approximate the RHS by an element of $\Sigma_M X^{H_n^{(N)}} - \Sigma_M X^{H_n^{(N)}}$ by iterating the construction achieved in Proposition 5.4.6. In so doing, we shall be required to approximate the coefficients λ, μ, ν appearing in (5.7) by sums-of-products of elements of A. This we can do by using the sum-product estimates from Section 2.6. However, in order to apply Corollary 2.6.4, we must apply Proposition 5.4.6 not to $H_n \subseteq G_n$ itself, but to $\pi_{p^{\beta n}}(H_n) \subseteq G_{\beta n}$, for some $\beta \in (0,1)$ (to be determined). By Remark 2.2.14 (i), $\pi_{p^{\beta n}}(H_n)$ still satisfies the hypothesis of Theorem 5.2.11, so Proposition 5.4.6 is indeed applicable.

We thus obtain $A \subseteq \mathbb{Z}/p^{\beta n}\mathbb{Z}$ with $|A| \geq (p^{\beta n})^{C-\gamma}$. Let $\tilde{A} \subseteq \mathbb{Z}/p^n\mathbb{Z}$ be a set of lifts of A. Choose $\gamma_0 < \frac{C}{2}$. We take $\alpha = \frac{C}{2}$ in Corollary 2.6.4 and obtain corresponding $\epsilon > 0$; $r, s \in \mathbb{Z}_{>0}$. Taking $\beta < \epsilon$, $\tilde{A}$ satisfies the conditions of Corollary 2.6.4 with $m = \beta n$, so that there exist $j \leq k \leq n$, with $k - j \geq \frac{\alpha\beta n}{4}$ and $k \ll_S \beta n$, such that:

$$p^j\mathbb{Z}/p^k\mathbb{Z} \subseteq \pi_{p^k}(\Sigma_r \tilde{A}^{(s)} - \Sigma_r \tilde{A}^{(s)}). \tag{5.10}$$

Let $x_1, \ldots, x_s \in \tilde{A}$; $k_1, \ldots, k_s \in H_n$; $X_1 \in \mathbb{M}_2(\mathcal{O})$. By Proposition 5.4.6 we have:

$$x_1[\frac{1}{x_0}Y, X_1] \in X_1^{H_n^{(4)}} - X_1^{H_n^{(4)}} \mod p^{\beta n(1-\gamma)}.$$

Replacing X_1 by $X_1^{k_1^{-1}}$ and conjugating by k_1 in the above, we obtain:

$$x_1[\frac{1}{x_0}Y^{k_1}, X_1] \in X_1^{H_n^{(6)}} - X_1^{H_n^{(6)}} \mod p^{\beta n(1-\gamma)}.$$

We now iteratively apply this last estimate: recursively define, for $1 \leq N \leq s$,

$$X_{N+1} = x_N[\frac{1}{x_0}Y^{k_N}, X_N]$$

and suppose by induction that there exists $W_N \in \mathbb{M}_2(\mathcal{O})$ such that:

$$X_N \equiv W_N \mod p^{\beta n(1-\gamma)}; W_N \in \Sigma_{K_N} X_1^{H_n^{(M_N)}} - \Sigma_{K_N} X_1^{H_n^{(M_N)}}.$$

By Proposition 5.4.6,

$$X_{N+1} \equiv x_N[\frac{1}{x_0}Y^{k_N}, W_N] \equiv W_N^{k_N^{-1}h_N^{-1}k_N} - W_N^{k_N^{-1}h_N k_N} \mod p^{\beta n(1-\gamma)}$$

for some $h_N \in H_n^{(4)}$, and:

$$W_N^{k_N^{-1}h_N^{-1}k_N} - W_N^{k_N^{-1}h_N k_N} \in \Sigma_{2K_N} X_1^{H_n^{(M_N+6)}} - \Sigma_{2K_N} X_1^{H_n^{(M_N+6)}}.$$

We may thus approximate X_{s+1}:

$$\frac{x_1 \cdots x_s}{x_0^s}[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X_1]]] \in \Sigma_{2^{s-1}} X_1^{H_n^{(6s)}} - \Sigma_{2^{s-1}} X_1^{H_n^{(6s)}} \mod p^{\beta n(1-\gamma)}. \tag{5.11}$$

Now consider Z: letting $\mathcal{O}$ be generated over $\mathbb{Z}_p$ by θ, there exist $Z_1, Z_2 \in \mathbb{M}_2(\mathbb{Z}_p)$ with $\mathrm{tr}(Z_1) = \mathrm{tr}(Z_2) = 0$ such that $x_0^s Z = Z_1 + \theta Z_2$. Letting Y be as in Proposition 5.4.6 and m, m'' be as in (5.7) we can write:

$$\begin{aligned} p^{m+m''} Z_i \equiv \lambda_i[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]] + \mu_i[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]]^{g_1} \\ +\nu_i[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]]^{g_2} \mod p^n \end{aligned}$$

(by (5.7)) for some $\lambda_i, \mu_i, \nu_i \in \mathbb{Z}_p$. Moreover, recall from (5.7) that $m + m'' < (s+1)m$ so that we may take $m + m'' \ll_S \delta' n$ (as m need only satisfy the conditions of Proposition 5.3.20). By (5.10), there exist $x_{j,k}, y_{j,k} : \mathbb{Z}_p \to \tilde{A}$ such that:

$$\lambda_i p^j \equiv \sum_{j=1}^{r} \Big(\prod_{k=1}^{s} x_{j,k}(\lambda_i) - \prod_{k=1}^{s} y_{j,k}(\lambda_i) \Big) \mod p^k,$$

$$\mu_i p^j \equiv \sum_{j=1}^{r} \Big(\prod_{k=1}^{s} x_{j,k}(\mu_i) - \prod_{k=1}^{s} y_{j,k}(\mu_i) \Big) \mod p^k,$$

$$\nu_i p^j \equiv \sum_{j=1}^{r} \Big(\prod_{k=1}^{s} x_{j,k}(\nu_i) - \prod_{k=1}^{s} y_{j,k}(\nu_i) \Big) \mod p^k.$$

Replacing k with $\min(k, \beta n(1-\gamma))$ in this last estimate, combining it with (5.11) and setting $X_1 = X$,

$$\frac{\lambda_i p^j}{x_0^s}[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]], \frac{\mu_i p^j}{x_0^s}[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]],$$
$$\frac{\nu_i p^j}{x_0^s}[Y^{k_s}, [Y^{k_{s-1}}, \ldots, [Y^{k_1}, X]]] \in \Sigma_{2^s r} X^{H_n^{(6s)}} - \Sigma_{2^s r} X^{H_n^{(6s)}} \mod p^k.$$

Substituting these last approximations into our expression for Z_i, there exist $W_1, W_2 \in \Sigma_{2^s 3r} X^{H_n^{(6s+1)}} - \Sigma_{2^s 3r} X^{H_n^{(6s+1)}}$ such that:

$$p^{j+m+m''} Z_i \equiv x_0^s W_i \mod p^k$$

so that:

$$x_0^s p^{j+m+m''} Z \equiv x_0^s (W_1 + \theta W_2) \mod p^k.$$

Finally let i be such that $x_0^s \equiv 0 \mod p^i$ but $x_0^s \not\equiv 0 \mod p^{i+1}$ Recall that by Proposition 5.4.6, $i \ll_S \delta' n$. Then:

$$p^{j+m+m''} Z \equiv (W_1 + \theta W_2) \mod p^{k-i}.$$

In particular, since the left-hand side lies in $\mathbb{M}_2(\mathbb{Z}_p)$, $W_2 \equiv 0 \mod p^{k-i}$. Set $i_1 = k - i, i_2 = j + m + m''$. Then:

$$i_1 - i_2 = (k - j) - (i + m + m'') \gg_S (\beta - \delta')n \gg \beta n.$$

for δ sufficiently small, and $i_1 \leq k \ll_S \beta n$. Choosing δ sufficiently small, we have completed the proof of Theorem 5.4.7, with $\gamma \asymp_S \beta, M = 2^s 3r, N = 6s + 1$.

5.4.4 Generating a large subgroup

In this section we deduce Theorem 5.2.11 from Theorem 5.4.7: We shall permit ourselves to shrink γ as required in the sequel, making the corresponding reductions in δ as we do so.

The first step is to note that $H_n^{(2)}$ contains elements of intermediate depth in the filtration of G_n by the K_i (here "intermediate" means of depth approximately γn). The proof we give here, though self-contained, should call to mind the kind of arguments already employed in Section 5.3.2.

Lemma 5.4.8. *There exists $j_0 \in \mathbb{N}$, satisfying $\gamma n \asymp_S j_0$, such that $H_n^{(2)} \cap (K_{j_0} \setminus K_{j_0+1}) \neq \emptyset$.*

Proof. As in Theorem 5.2.11, $\mu_{S_n}^{(2l)}(H_n) \geq |G_n|^{-2\delta}$ for any $l \leq C_0 \log|G_n|$. There exists $g \in G_n$ such that:

$$\mu_{S_n}^{(2l)}(H_n \cap gK_{\gamma n}) \geq |G_n|^{-2\delta}|G_{\gamma n}|^{-1}.$$

Hence $\mu_{S_n}^{(4l)}(H_n^{(2)} \cap K_{\gamma n}) \geq |G_n|^{-4\delta}|G_{\gamma n}|^{-2}$. Arguing as in the proof of Corollary 5.3.12, it follows from Kesten's Theorem and the strong Diophantine hypothesis that for $4l \leq Dn$,

$$\begin{aligned}\mu_{S_n}^{(4l)}((H_n^{(2)} \cap K_{\gamma n}) \setminus \{1\}) &\gg_S |G_n|^{-4\delta}|G_{\gamma n}|^{-2} - e^{-C_S l} \\ &\geq e^{-6(2\delta+\gamma)n\log p} - e^{-C_S l}.\end{aligned}$$

We set $l = C'\gamma n$, for $C' > 0$ sufficiently large that $e^{-6(2\delta+\gamma)n\log p} - e^{-C_S l} > 0$. For γ sufficiently small (depending on C'), $4l \leq Dn$ is satisfied.

The conclusion is that there exists $h \in H_n^{(2)} \cap B_{S_n}(4l)$ such that $h \equiv I_2 \mod p^{\gamma n}$ but $h \neq I_2$ and in particular, one of $h_{1,2}, h_{2,1}, h_{1,1} - h_{2,2} \neq 0$. Now each of $h_{1,2}, h_{2,1}$ and $h_{1,1} - h_{2,2}$ are expressible as a homogeneous polynomial, of degree at most $4l$, and consisting of at most 2^{3l+1} signed monic monomials, in the set A of entries of the elements of S, and A is a C-strongly Diophantine set for some $C > 0$. It follows that one of $h_{1,2}, h_{2,1}, h_{1,1} - h_{2,2} \not\equiv 0 \mod p^m$ whenever:

$$m > C(\log(2^{3l+1}) + 4l) = CC'(4 + 3\log 2)\gamma n + C\log 2,$$

so that $h \notin K_m$ for such m. Therefore for some $\gamma n \leq j_0 \leq CC'(4 + 3\log 2)\gamma n$, $h \in K_{j_0} \setminus K_{j_0+1}$, as required. □

Henceforth we let M, N, i_1, i_2 be as in Theorem 5.4.7. Our next Lemma exploits the correspondence between $\mathrm{SL}_2(\mathbb{Z}_p)$ and $\mathfrak{sl}_2(\mathbb{Z}_p)$: given $g \in K_i$, there exists $Z \in \mathfrak{sl}_2(\mathbb{Z}_p)$ such that $I_2 + p^i Z$ is a good approximation to g. Conversely, given $Z \in \mathfrak{sl}_2(\mathbb{Z}_p)$, $I_2 + p^i Z$ may be approximated by some $g \in K_i$. Using Theorem 5.4.7, we can quantify this last statement: under appropriate hypotheses, our approximation g to $I_2 + p^i Z$ may be chosen to lie within some small power of H_n.

Lemma 5.4.9. *Suppose that, for some $K, i \in \mathbb{N}$, with $i \geq i_1$, there exists $g \in H_n^{(K)} \cap (K_i \setminus K_{i+1})$. Then, for all $Z \in \mathbb{M}_2(\mathbb{Z}_p)$ satisfying $\mathrm{tr}(Z) = 0$,*

$$I_2 + p^{i+i_2} Z \in H_n^{(2MK+4MN)} \mod p^{i+i_1}.$$

Proof. Let $X \in \mathbb{M}_2(\mathbb{Z}_p) \setminus p\mathbb{M}_2(\mathbb{Z}_p)$ be such that $g = I_2 + p^i X$. We have:

$$1 = \det(g) = 1 + p^i \mathrm{tr}(X) + p^{2i}\det(X)$$

so that $\mathrm{tr}(X) \equiv 0 \mod p^i$.

Hence:

$$g^{-1} = I_2 - p^i X + p^i \operatorname{tr}(X) I_2 \equiv I_2 - p^i X \mod p^{2i}.$$

Apply Theorem 5.4.7 to X, Z, so that there exist $h_1, \ldots, h_M, h'_1, \ldots, h'_M \in H^{(N)}$ such that:

$$p^{i_2} Z \equiv X^{h_1} + \ldots + X^{h_M} - X^{h'_1} - \ldots - X^{h'_M} \mod p^{i_1}.$$

Then:

$$\begin{aligned} I_2 + p^{i+i_2} Z &\equiv (I_2 + p^i X)^{h_1} \cdots (I_2 + p^i X)^{h_M} (I_2 - p^i X)^{h'_1} \cdots (I_2 - p^i X)^{h'_M} \mod p^{i+i_1} \\ &\equiv g^{h_1} \cdots g^{h_M} (g^{-1})^{h'_1} \cdots (g^{-1})^{h'_M} \mod p^{2i} \end{aligned}$$

as required. □

We can now argue inductively (taking Lemma 5.4.8 as our base case and iteratively applying Lemma 5.4.9). The upshot shall be that a small power of H_n contains elements at most levels of the filtration of G_n by the K_i.

Proposition 5.4.10. *Let $j_0 \in \mathbb{N}$ be as in Lemma 5.4.8. For $r \in \mathbb{N}$, let $N_r = (2M)^r(\frac{4MN}{2M-1} + 2) - \frac{4MN}{2M-1}$. Then for any $j_0 + ri_2 \leq j \leq j_0 + r(i_1 - 1)$, there exists $g_j \in H_n^{(N_r)} \cap (K_j \setminus K_{j+1})$.*

Proof. We proceed by induction on r. The base case $r = 0$ is immediate from Lemma 5.4.8. Let $r \geq 0$ and suppose the claim holds up to r. Let:

$$j_0 + (r+1)i_2 \leq j \leq j_0 + (r+1)(i_1 - 1).$$

Let $i_2 \leq j' \leq i_1 - 1$ be such that:

$$j_0 + ri_2 \leq j - j' \leq j_0 + r(i_1 - 1).$$

Let $g_{j-j'} \in H_n^{(N_r)} \cap (K_{j-j'} \setminus K_{j-j'+1})$. Let $Z \in p^{j'-i_2}\mathbb{M}_2(\mathbb{Z}_p) \setminus p^{j'-i_2+1}\mathbb{M}_2(\mathbb{Z}_p)$ with $\operatorname{tr}(Z) = 0$. By Lemma 5.4.9, there exists $g_j \in H_n^{(2MN_r+4MN)} = H_n^{(N_{r+1})}$ such that $g_j \equiv I_2 + p^{i_2+j-j'} Z \mod p^{i_1+j-j'}$. In particular, $g_j \in K_j \setminus K_{j+1}$. □

Since $i_2 < i_1 \ll_S \gamma n$, $i_1 - i_2 > \gamma c n$, we have $i_2 < i_1(1-c)$ (replacing c by a smaller constant if required). Choose n sufficiently large (depending on γ) such that $\frac{1}{i_1} < \frac{c}{2}$. Let r_0 be sufficiently large that:

$$\frac{r_0 + 1}{r_0} < \frac{2 - c}{2 - 2c} < \frac{1 - \frac{1}{i_1}}{1 - c} \tag{5.12}$$

so that $i_2(r_0 + 1) < i_1(r_0 + 1)(1 - c) < r_0(i_1 - 1)$. Hence, for $r \geq r_0$, if j satisfies:

$$j_0 + r(i_1 - 1) \leq j \leq j_0 + (r+1)(i_1 - 1)$$

then $j_0 + (r+1)i_2 \leq j \leq j_0 + (r+1)(i_1 - 1)$ so that, by Proposition 5.4.10, $H_n^{(N_{r+1})} \cap (K_j \setminus K_{j+1}) \neq \emptyset$.

Now take $R \in \mathbb{N}$ satisfying:

$$R > \frac{2}{c\gamma} \tag{5.13}$$

so that $n \leq j_0 + R(i_1 - 1)$. Thus for any $j_0 + r_0(i_1 - 1) \leq j < n$, $H_n^{(N_R)} \cap (K_j \setminus K_{j+1}) \neq \emptyset$. Applying Lemma 5.4.9, we conclude:

Lemma 5.4.11. *Let r_0 be as in (5.12). For every $j_0 + r_0(i_1 - 1) \leq j \leq n - i_1$ and every $Z \in \mathbb{M}_2(\mathbb{Z}_p)$ satisfying $\mathrm{tr}(Z) = 0$,*

$$I_2 + p^{j+i_2} Z \in H_n^{(N_{R+1})} \mod p^{j+i_1}.$$

As a consequence, we have:

Theorem 5.4.12. *Let $i_3 = j_0 + r_0(i_1 - 1) + i_2$. Let $R' \in \mathbb{N}$ satisfy $R' \geq \frac{n - i_3}{i_1 - i_2}$. Then:*

$$K_{i_3} \subseteq H_n^{(R' N_{R+1})}$$

Proof. We shall prove by induction on $R' \in \mathbb{N}$ that, if:

$$i_3 + R'(i_1 - i_2) \leq j \leq i_3 + (R'+1)(i_1 - i_2) \tag{5.14}$$

and $Z \in \mathbb{M}_2(\mathbb{Z}_p)$ satisfies $\det(I_2 + p^{i_3} Z) \equiv 1 \mod p^j$ then:

$$I_2 + p^{i_3} Z \in H_n^{((R'+1)N_{R+1})} \mod p^j. \tag{5.15}$$

In particular, taking $j = n$, we obtain the required result for R' as in the statement.

In the base case $R' = 0$, supposing that $i_3 \leq j \leq i_3 + i_1 - i_2$, and letting Z be such that:

$$\det(I_2 + p^{i_3} Z) = 1 + p^{i_3} \mathrm{tr}(Z) + p^{2i_3} \det(Z) \equiv 1 \mod p^j,$$

$\mathrm{tr}(Z) \equiv 0 \mod p^{j - i_3}$, since $i_3 \geq i_1 - i_2$. Let $W \in \mathbb{M}_2(\mathbb{Z}_p)$ be such that $\mathrm{tr}(Z + p^{j-i_3} W) = 0$, so that:

$$I_2 + p^{i_3} Z + p^j W \in H_n^{(N_{R+1})} \mod p^{i_3 + i_1 - i_2}$$

by Lemma 5.4.11. Hence $I_2 + p^{i_3} Z \in H_n^{(N_{R+1})} \mod p^j$.

Suppose our claim (5.15) holds up to $j = j' \leq n - i_1 + i_2$, with R' as in (5.14). We show the claim holds for $j' \leq j \leq j' + i_1 - i_2$. Let $K \in \mathbb{N}$ and let $g \in H_n^{(K)}$ be such that $g \equiv I_2 + p^{i_3} Z \mod p^{j'}$.

Then for some $W \in \mathbb{M}_2(\mathbb{Z}_p)$,

$$(I_2 + p^{i_3}Z)g^{-1} = I_2 + p^{j'}W$$

so:

$$\begin{aligned} 1 + p^{j'}\operatorname{tr}(W) + p^{2j'}\det(W) &= \det(I_2 + p^{j'}W) \\ &= \det((I_2 + p^{i_3}Z)g^{-1}) \\ &= \det(I_2 + p^{i_3}Z) \\ &\equiv 1 \mod p^j. \end{aligned}$$

Hence $p^{j-j'} \mid \operatorname{tr}(W)$. Let $V \in \mathbb{M}_2(\mathbb{Z}_p)$ be such that $\operatorname{tr}(W + p^{j-j'}V) = 0$. Then by Lemma 5.4.11 again,

$$I_2 + p^{j'}(W + p^{j-j'}V) = I_2 + p^{j'}W + p^jV \in H_n^{(N_{R+1})} \mod p^{j'+i_1-i_2}$$

so that $I_2 + p^{j'}W \in H_n^{(N_{R+1})} \mod p^j$. Hence:

$$I_2 + p^{i_3}Z = (I_2 + p^{j'}W)g \in H_n^{(K+N_{R+1})} \mod p^j.$$

The required result follows by induction. □

Why does the containment $K_{i_3} \subseteq H_n^{(R'N_{R+1})}$ from Theorem 5.4.12 constitute a proof of Theorem 5.2.11? Recall from the statement of Theorem 5.4.12 that:

$$i_3 = j_0 + r_0(i_1 - 1) + i_2$$

where:

(i) i_1, i_2 are as defined in Theorem 5.4.7: they satisfy $i_1 \asymp_S \gamma n$, $i_2 \ll_S \gamma n$;

(ii) j_0 is as defined in Lemma 5.4.8: it satisfies $j_0 \asymp_S \gamma n$;

(iii) r_0 is as defined at (5.12): it need only satisfy:

$$\frac{r_0 + 1}{r_0} < \frac{2 - c}{2 - 2c}$$

where $c = c(S)$ is as in the statement of Theorem 5.4.7. In particular we may take r_0 to depend only on S.

Overall then, $i_3 \asymp_S \gamma n$, as required. Next, recall from the statement of Proposition 5.4.10 that:

$$N_r = (2M)^r\left(\frac{4MN}{2M-1} + 2\right) - \frac{4MN}{2M-1},$$

where M, N are as in Theorem 5.4.7 (they depend only on S). Moreover, R is introduced at (5.13); it need only satisfy the condition $R > 2/c\gamma$, where c, once again, is as in the statement of Theorem 5.4.7. Specifically, we can take R to depend on S and γ alone, so that the same holds for N_{R+1}. Finally, R' was introduced in the statement of Theorem 5.4.12, and subject only to the condition $R' \geq (n-i_3)/(i_1-i_2)$. As explained above, $i_3 \asymp_S \gamma n$ and, according to the statement of Theorem 5.4.7, $i_1 - i_2 > c\gamma n$, so:

$$\frac{n-i_3}{i_1-i_2} \ll_S \frac{1}{\gamma}.$$

We may thus take $R'N_{R+1}$ to be a function of γ and S alone. This concludes the proof of Theorem 5.2.11, and hence that of Theorem 5.1.1.

5.5 What's Next?

5.5.1 The higher-rank case $\mathrm{SL}_d(\mathbb{Z}_p)$

Replacing SL_2 with SL_d throughout, a very similar argument would yield the analogue of Theorem 5.1.1 for $\mathrm{SL}_d(\mathbb{Z}_p)$. To be explicit:

Theorem 5.5.1. *Let $S \subseteq \mathrm{SL}_d(\mathbb{Z}_p)$ be a finite symmetric set, generating a subgroup Γ whose closure $\overline{\Gamma}$ in $\mathrm{SL}_d(\mathbb{Z}_p)$ is open. Let $A \subseteq \mathbb{Z}_p$ be the set of entries occurring in elements of S. Suppose that A is strongly Diophantine. Then $(\pi_{p^n}(\Gamma), \pi_{p^n}(S))_n$ is a family of two-sided expanders.*

As in the SL_2 case, the result is due to Bourgain and Gamburd for generating sets supported over $\mathbb{Z}$ [7], and once again, their argument goes through without changes once we have pushed our non-concentration estimates down into the congruence quotients (using the strong Diophantine hypothesis).

In terms of establishing non-concentration in subvarieties, the only stage at which the proof in higher rank differs substantially from the case of SL_2 is when we produce a large set of simultaneously diagonalizable elements. Recall that for SL_2, this was achieved in Section 5.3.4, using the fact that elements without repeated eigenvalues are characterised by their trace. Such a characterisation clearly breaks down in higher rank, so a different approach is needed. Following [7], we note that an element $g \in \mathrm{SL}_d(\mathbb{Z}_p)$ has a repeated eigenvalue in some field extension if and only if the *Sylvester determinant* $\mathrm{Res}(\chi_g, \chi_g')$ vanishes, so such elements satisfy a non-trivial polynomial identity. In practice, we shall further require non-vanishing estimates

for the Sylvester determinant applied to translates gh of g, which are uniform in the translating element $h \in \mathrm{SL}_d(\mathbb{Z}_p)$. This is more tricky, but can still be done by considering a sufficiently large tuple of solutions $g_1, \ldots, g_r$, as was done in Sections 5.3.4 and 5.3.5. These ideas are to be explored in detail elsewhere.

5.5.2 Diophantine generating sets

In Chapter 1 we hinted at an analogy between Theorem 5.1.1 and the result of [8] on spectral gap in SU(2). Let us now pursue that analogy a little further. In [8] spectral gap was proved for the action of generating set of dense subgroups in SU(2) satisfying the following condition:

Definition 5.5.2. *A finite symmetric subset $S \subseteq \mathrm{SU}(2)$ is* Diophantine *if there exists $D(S) \in (0,1)$ such that for any $n \in \mathbb{N}$ and any $w \in B_S(n)$, either $w = \pm I_2$ or:*

$$\|w \pm I_2\| \geq D^n.$$

Here $\|\cdot\|$ is the ℓ^2-norm on $\mathbb{M}_2(\mathbb{C})$. In other words, we must apply the multiplication operation in the group many times to S to produce an element lying close to $\{\pm I_2\}$ but outside it.

Taking cosets of the congruence kernels K_n as our neighbourhoods, there is an obvious corresponding notion of a Diophantine subset of $\mathrm{SL}_2(\mathbb{Z}_p)$:

Definition 5.5.3. *A finite symmetric subset $S \subseteq \mathrm{SL}_2(\mathbb{Z}_p)$ is* Diophantine *if there exists $C > 0$ such that for any $n \in \mathbb{N}$ and any $w \in B_S(n)$, either $w = \pm I_2$ or $\pm w \notin K_{Cn}$.*

Compare this with the strong Diophantine condition introduced in Section 5.3.1: there we had a subset of $\mathbb{Z}_p$ and could not produce an element of small p-adic norm without applying the ring operations to our set many times. Considering the polynomials $f_1(g) = g_{1,2}$, $f_2(g) = g_{2,1}$, $f_3(g) = g_{1,1} - g_{2,2}$, for $g \in \mathrm{SL}_2(\mathbb{Z}_p)$, and applying Proposition 5.3.2, we deduce:

Proposition 5.5.4. *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z}_p)$ be finite symmetric. Let $A \subseteq \mathbb{Z}_p$ be the set of entries occurring in elements of S. Suppose that for every $a \in A$, there exists $f_a(X) \in \mathbb{Q}[X]$ such that $f_a(a) = 0$. Then S is Diophantine.*

Given the result of [8], it is plausible that a Diophantine generating set is all that is needed to produce spectral gap results is p-adic groups too. We therefore propose the following generalization of Theorem 5.1.1:

Conjecture 5.5.5. *Let $S \subseteq \mathrm{SL}_2(\mathbb{Z}_p)$ be a finite symmetric set, generating a subgroup Γ whose closure $\overline{\Gamma}$ in $\mathrm{SL}_2(\mathbb{Z}_p)$ is open. Suppose S is Diophantine. Then $(\pi_{p^n}(\Gamma), \pi_{p^n}(S))_n$ is an expander family.*

The obvious difficulty in applying the methods of this chapter to Conjecture 5.5.5 is the recovery of the non-concentration results of Section 5.3. There, the strong Diophantine hypothesis allowed us to take the upper bound from Theorem 5.3.8 on the return probability of the random walk to the subvariety defined by a homogeneous polynomial, and upgrade it to a bound on the return probability to a corresponding sublevel set. It may be too much to hope for that the same should be achievable only under the hypothesis of a *Diophantine* generating set for an *arbitrary* such polynomial, but we may still expect to be able to obtain bounds for the *specific* polynomials used in Sections 5.3.4 and 5.3.5.

5.5.3 Towards a positive-characteristic analogue

Taking the application of the Solovay-Kitaev procedure to expanders explored in this chapter, together with our analysis of the Solovay-Kitaev procedure for groups with an analytic structure over other pro-p domains such as $\mathbb{F}_q[[t]]$, leads one to the possibility of constructing expander congruence quotients for these groups as well. Certainly such groups are sufficiently quasirandom to make relevant the Bourgain-Gamburd machinery [43].

What about the non-concentration of the random walk in algebraic subvarieties? For the very simplest case of the group $\mathrm{SL}_2(\mathbb{F}_p[[t]])$, and for generators supported over $\mathbb{F}_p[t]$, this is provided for by our work in Chapter 4 (specifically by Theorem 4.1.1). We therefore propose:

Conjecture 5.5.6. *Let $S \subseteq \mathrm{SL}_2(\mathbb{F}_p[t])$ be a finite symmetric set, generating a non-elementary subgroup Γ. Then $(\pi_{t^n}(\Gamma), \pi_{t^n}(S))_n$ is a family of two-sided expanders, where $\pi_{t^n} : \mathrm{SL}_2(\mathbb{F}_p[[t]]) \to \mathrm{SL}_2(\mathbb{F}_p[t]/(t^n))$ is the congruence map.*

Indeed, Theorem 4.1.1 was originally proved as a tool with which to attack this conjecture. The analogue for $\mathrm{SL}_2(\mathbb{F}_p[t]/(t^n))$ of the remainder of Section 5.3 and Sections 5.4.1 and 5.4.2 then goes through without changes.

Problems arise, however, when we reach the stage of the argument at which the results on sum-product are required. To the author's knowledge, versions of Bourgain's results on $\mathbb{Z}/p^n\mathbb{Z}$ for the ring $\mathbb{F}_p[t]/(t^n)$ are not known at this time (though Bourgain has suggested [4] that $\mathbb{F}_p[t]/(t^n)$ may be susceptible to analysis by the same kind of tools employed in [3]).

In any case, the most obvious analogue of Proposition 2.6.3 for $\mathbb{F}_p[t]/(t^n)$, namely:

Not-a-Theorem 5.5.7. *For all $\alpha_1, \alpha_2 > 0$, there exists $\epsilon > 0$; $r, s \in \mathbb{Z}_{>0}$ such that, for $A \subseteq \mathbb{F}_p[t]/(t^n)$ satisfying:*

$$|\pi_{t^m}(A)| > p^{\alpha_1 m} \text{ whenever } n \geq m > \epsilon n,$$

there exists $0 \leq k < \alpha_2 n$ such that:

$$(t^k)/(t^n) \subseteq \Sigma_r A^{(s)} - \Sigma_r A^{(s)}.$$

is plainly false. For let $A = \pi_{t^n}(\mathbb{F}_p[t^2])$. Then for any $\alpha_1 \in (0, \frac{1}{2})$, $\epsilon > 0$ and $n \geq m > \epsilon n$, A satisfies:

$$|\pi_{t^m}(A)| > p^{\alpha_1 m}$$

provided n is sufficiently large. But $\Sigma_r A^{(s)} - \Sigma_r A^{(s)} = A$ for all $r, s \in \mathbb{Z}_{>0}$, as A is a subring of $\mathbb{F}_p[t]/(t^n)$, and A does not contain any non-trivial ideal $(t^k)/(t^n)$ of $\mathbb{F}_p[t]/(t^n)$ (except $(t^{n-1})/(t^n)$, when n is odd).

A stronger hypothesis on A would appear to be called for, then, perhaps requiring that the subring of $\mathbb{F}_p[t]/(t^n)$ generated by A contains a large ideal. Even if such a result on sum-product growth were to hold though, its hypothesis would not obviously apply to the set A constructed in Section 5.4.2: the constraints on A coming from Proposition 5.4.6 do not obviously preclude A from generating a smaller subring.

As a model case of Conjecture 5.5.6, it might be instructive to investigate *subring subgroups* of $\mathrm{SL}_2(\mathbb{F}_p[t]/(t^n))$, such as the congruence image of $\mathrm{SL}_2(\mathbb{F}_p[t^2])$, and aim for a proof that the random walk does not concentrate in these subgroups (assuming all reasonable claims about the sum-product phenomenon in $\mathbb{F}_p[t]/(t^n)$). The tools used in such a proof may be applicable to filling this particular gap in the proof of Conjecture 5.5.6.

Index

Bibliography

[1] L. Babai and Á. Seress, On the diameter of permutation groups. *European J. Combin.* **13** (1992), 231-243.

[2] Y. Benoist and N. de Saxcé, A spectral gap theorem in simple Lie groups. *arXiv:1405.1808* (2014).

[3] J. Bourgain, The sum-product theorem in $\mathbb{Z}_q$ with q arbitrary. *Journal d'Analyse Mathématique* **106** (2008), Issue 1, 1-93.

[4] J. Bourgain, private communication, September 2013.

[5] J. Bourgain and A. Gamburd, Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Annals of Mathematics* **167** (2008), 625-642.

[6] J. Bourgain and A. Gamburd, Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$: I. *J. Eur. Math. Soc.* **10** (2008), Issue 4, 987-1011.

[7] J. Bourgain and A. Gamburd, Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$: II. *J. Eur. Math. Soc.* **11** (2009), Issue 5, 1057-1103.

[8] J. Bourgain and A. Gamburd, On the spectral gap for finitely-generated subgroups of SU(2). *Invent. math.* **171** (2008), Issue 1, 83-121.

[9] J. Bourgain and A. Gamburd, A spectral gap theorem in $\mathrm{SU}(d)$. *J. Eur. Math. Soc.* **14** (2012), Issue 5, 1455-1511.

[10] J. Bourgain, A. Gamburd and P. Sarnak, Affine linear sieve, expanders, and sum-product. *Invent. math.* **179** (2010), Issue 3, 559-644.

[11] J. Bourgain and P. P. Varjú, Expansion in $\mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z})$, q arbitrary. *Invent. math.* **188** (2012), Issue 1, 151-173.

[12] H. Bradford, New Uniform Diameter Bounds in Pro-p Groups. *arXiv:1410.3007* (2014).

[13] H. Bradford, Expansion, Random Walks and Sieving in $\mathrm{SL}_2(\mathbb{F}_p[t])$. *arXiv:1501.03199* (2015).

[14] E. Breuillard, Heights on GL_2 and Free Subgroups. In *Geometry, Rigidity and Group Actions*, Zimmers Festschrift, B. Farb and D. Fisher eds., Chicago Univ. Press (2011).

[15] E. Breuillard, A Height Gap Theorem for Finite Subsets of $\mathrm{GL}_d(\bar{\mathbb{Q}})$ and Non Amenable Subgroups. *Ann. of Math* **174**-2 (2011), 1057-1110.

[16] E. Breuillard, A Strong Tits Alternative. *arXiv:0804.1395* (2008).

[17] E. Breuillard, Approximate subgroups and super-strong approximation. *arXiv:1407.3158* (2014).

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

[19] E. Breuillard, B. Green, R. Guralnick and T. Tao, Expansion in finite simple groups of Lie type. *arXiv:1309.1975* (2013).

[20] R. Camina, Subgroups of the Nottingham group. *J. Algebra* **196** (1997), 101-113.

[21] R. Camina, The Nottingham group. In *New horizons in pro-p groups*, Progr. Math. 184, Birkhauser, Boston (2000), 205-221.

[22] R.W. Carter, *Simple groups of Lie type*, Pure and applied mathematics 28, Wiley-Interscience, London 1972.

[23] I.S. Cohen, On the structure and ideal theory of complete local rings. *Trans. Amer. Math. Soc.* **59** (1946), 54-106.

[24] C.M. Dawson and M.A. Nielsen, The Solovay-Kitaev algorithm. *Quantum Information & Computation* **6** (2006), Issue 1, 81-95.

[25] P. Diaconis, Random walks on groups: characters and geometry. In *Groups St Andrews 2001 in Oxford*, vol. 2, London Math. Soc. Lecture Note Ser. 304-305, Cambridge University Press, Cambridge (2003), 120-142.

[26] P. Diaconis and L. Saloff-Coste, Comparison techniques for random walk on finite groups. *Ann. Probab.* **21** (1993), Issue 4, 2131-2156.

[27] O. Dinai, Expansion properties of finite simple groups. Ph. D. thesis, The Hebrew University of Jerusalem 2009.

[28] O. Dinai, Uniform poly-log diameter bounds for some families of finite groups. *Proc. Amer. Math. Soc.* **134** (2006), Issue 11, 3137-3142.

[29] O. Dinai, Diameters of Chevalley groups over local rings. *Archiv der Mathematik* **99** (2012), Issue 5, 417-424.

[30] J.D. Dixon, M.P.F. Du Sautoy, A. Mann and D. Segal, *Analytic pro-p groups*, 2nd Edition, Cambridge studies in advanced mathematics 61, Cambridge University Press, Cambridge 1999.

[31] I. Fesenko, On just infinite pro-p groups and arithmetically profinite extensions of local fields. *J. Reine Angew. Mathematik* **517** (1999), 61-80.

[32] A. Gamburd, S. Hoory, M. Shahshahani, A. Shalev and B. Virág, On the girth of random Cayley graphs. *Random Structures and Algorithms* **35** (2009), Issue 1, 100-117.

[33] A. Gamburd and M. Shahshahani, Uniform diameter bounds for some families of Cayley graphs. *Int. Math. Res. Notices* **71** (2004), 3813-3824.

[34] W.T. Gowers, Quasirandom Groups. *Combinatorics, Probability and Computing* **17** (2008), Issue 03, 363-387.

[35] L. K. Grover, A fast quantum mechanical algorithm for database search. In *28th ACM Symposium on Theory of Computation*, Association for Computing Machinery, New York (1996), 212-219.

[36] H. A. Helfgott, Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Annals of Mathematics* **167** (2008), 601-623.

[37] H. A. Helfgott, Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$. *J. Eur. Math. Soc.* **13** (2011), Issue 3, 761-851.

[38] H. A. Helfgott, Growth in groups: ideas and perspectives. *arXiv:1303.0239* (2013).

[39] H. A. Helfgott and Á. Seress, On the diameter of permutation groups. *arXiv:1109.3550* (2011).

[40] H. A. Helfgott, Á. Seress and A. Zuk, Random generators of the symmetric group: diameter, mixing time and spectral gap. *J. Algebra* **421** (2015), 349-368.

[41] S. Hoory, N. Linial and A. Wigderson, Expander Graphs and their Applications. *Bull. Amer. Math. Soc.* **43** (2006), Issue 4, 439-561.

[42] R. Howe, Kirillov theory for compact p-adic analytic groups, *Pacific J. Math.* **73** (1977), 365-381.

[43] A. Jaikin-Zapirain, On the representation growth of pro-p groups. Preprint (2000).

[44] D.L. Johnson, *Topics in the Theory of Group Presentations*, London Mathematical Society lecture note series 42, Cambridge University Press, Cambridge 1980.

[45] F. Jouve, E. Kowalski and D. Zywina, Splitting fields of characteristic polynomials of random elements in arithmetic groups. *Israel J. Math.* **193** (2013), Issue 1, 263-307.

[46] H. Kesten, Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** (1959), Issue 2, 336-354.

[47] O.H. King, The subgroup structure of finite classical groups in terms of geometric configurations. In *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser. 327, Cambridge Univ. Press, Cambridge (2005), 29-56.

[48] A. A. Kirillov, Unitary representations of nilpotent lie groups, Russian Math. Surveys **17** (1962), 57-110.

[49] B. Klopsch, Normal subgroups in substitution groups of the formal power series. *J. Algebra* **228** (2000), Issue 1, 91-106.

[50] V. Landazuri and G. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups. *J. Algebra* **32** (1974), 418-443.

[51] S. Lang and A. Weil, Number of points of varieties in finite fields. *Amer. J. Math.* **76** (1954), 819-827.

[52] M. Lazard, Groupes analytiques p-adiques. *Inst. Hautes Etudes Scientifiques, Publ. Math.* **26** (1965), 389-603.

[53] A. Lubotzky, Cayley graphs: eigenvalues, expanders and random walks. In *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser. 218, Cambridge Univ. Press, Cambridge (1995), 155-189.

[54] A. Lubotzky, Expander Graphs in Pure and Applied Mathematics. *Bull. Amer. Math. Soc.* **49** (2012), Issue 1, 113-162.

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

[56] A. Lubotzky and C. Meiri, Sieve methods in group theory II: The mapping class group. *Geom. Dedicata* **159** (2012), 327-336.

[57] A. Lubotzky and C. Meiri, Sieve methods in group theory III: $\mathrm{Aut}(F_n)$. *Int. J. Algebra Comput.* **22** (2012), Issue 7, 1250062.

[58] A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs. *Combinatorica* **8** (1988), Issue 3, 261-277.

[59] A. Lubotzky and L. Rosenzweig, The galois group of random elements of linear groups. *Amer. J. of Math.* **136**, (2014), 1347-1383.

[60] W. Magnus, A. Karrass and D. Solitar, *Combinatorial Group Theory: Presentations of Groups in Terms of Generators and Relations*, 2nd Revised Edition, Dover Publications, Mineola NY 2004.

[61] G.A. Margulis, Explicit constructions of concentrators. *Problems of Inform. Transm.* **10** (1975), 325-332.

[62] C. R. Matthews, L. N. Vaserstein and B. Weisfeiler, Congruence properties of Zariski-dense subgroups. I. *Proc. London Math. Soc.* **48** (1984), Issue 3, 514-532.

[63] N. Nikolov and L. Pyber, Product decompositions of quasirandom groups and a Jordan type theorem. *J. Eur. Math. Soc.* **13** (2011), Issue 4, 1063-1077.

[64] M. V. Nori, On subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$. *Invent. math.* **88** (1987), Issue 2, 257-275.

[65] R. Pink, Strong approximation for Zariski dense subgroups over arbitrary global fields. *Comment. Math. Helv.* **75** (2000), Issue 4, 608-643.

[66] L. Pyber and E. Szabó, Growth in finite simple groups of Lie type of bounded rank. *arXiv:1005.1858* (2010).

[67] I. Rivin, Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.* **142** (2008), Issue 2, 353-379.

[68] A. Salehi Golsefidy, Expansion properties of linear groups. Preprint (2012).

[69] A. Salehi Golsefidy and P. P. Varjú, Expansion in Perfect Groups. *Geom. Funct. Anal.* **22** (2012), Issue 6, 1832-1891.

[70] P. Sarnak and X. Xue, Bounds for multiplicities of automorphic representations. *Duke Math. J.* **64** (1991), 207-227.

[71] N de Saxcé, Trou dimensionnel dans les groupes de Lie compacts semisimples via les séries de Fourier. *Journal d'Analyse Mathématique*, **120** (2013), Issue 1, 311-331.

[72] A. Selberg, On the estimation of Fourier coefficients of modular forms. In *Theory of numbers*, Proc. Sympos. Pure Math. 8, Amer. Math. Soc., Providence, R.I. (1965), 1-15.

[73] A. Shalev, Lie methods in the theory of pro-p groups. In *New horizons in pro-p groups*, Progr. Math. 184, Birkhauser, Boston (2000), 1-54.

[74] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings, 35th Annual Symposium on Fundamentals of Computer Science*, Los Alamitos, 1994. IEEE Press.

[75] T. Tao, Product set estimates for non-commutative groups. *Combinatorica* **28** (2008), Issue 5, 547-594.

[76] T. Tao, *Expansion in finite simple groups of Lie type*, Graduate studies in mathematics 164, Amer. Math. Soc., Providence, R.I. 2015.

[77] P. P. Varjú, Expansion in $\mathrm{SL}_d(\mathcal{O}_K/I)$, I square free. *J. Eur. Math. Soc.* **14** (2012), Issue 1, 273-305.

[78] P. P. Varjú, Random walks in compact groups. *Documenta Mathematica* **18** (2013), 1137-1175.

[79] B. Weisfeiler, Strong approximation for Zariski-dense subgroups of semisimple algebraic groups. *Annals of Mathematics* **120** (1984), Issue 2, 271-315.

[80] I. York, The group of formal power series under substitution. Ph.D. thesis, Nottingham University, Nottingham 1990.