

KAZHDAN QUOTIENTS OF GOLOD-SHAFAREVICH GROUPS

MIKHAIL ERSHOV

with an appendix on

Uniform non-amenability of Golod-Shafarevich groups

by MIKHAIL ERSHOV and ANDREI JAIKIN-ZAPIRAIN

and an appendix on

Subgroup growth of Golod-Shafarevich groups

by ANDREI JAIKIN-ZAPIRAIN

ABSTRACT. The main goal of this paper is to prove that every Golod-Shafarevich group has an infinite quotient with Kazhdan's property (T) . In particular, this gives an affirmative answer to the well-known question about non-amenability of Golod-Shafarevich groups.

1. INTRODUCTION

1.1. Golod-Shafarevich groups. In early 60's, Golod and Shafarevich [GS] found a sufficient condition for a group given by generators and relators to be infinite. The original condition from [GS] admits several generalizations, but the following one is usually taken as the definition of Golod-Shafarevich groups.

Definition. Fix a prime number p .

- (a) Consider a group presentation $\langle X|R \rangle$, where X is finite. For each $i \in \mathbb{N}$ let r_i be the number of elements of R which have degree i with respect to the Zassenhaus p -filtration. The presentation $\langle X|R \rangle$ is said to satisfy the *Golod-Shafarevich* (GS) condition with respect to p , if there exists a real number $t \in (0, 1)$ such that

$$1 - H_X(t) + H_R(t) < 0 \text{ where } H_X(t) = |X|t \text{ and } H_R(t) = \sum_{i=1}^{\infty} r_i t^i.$$

- (b) A group Γ is called *Golod-Shafarevich* if it has a presentation satisfying the Golod-Shafarevich condition.

Remark: The reference to prime p will usually be omitted, as we will never consider GS groups with respect to different primes at the same time.

The Golod-Shafarevich condition can also be defined for pro- p groups (using essentially the same definition) and certain kinds of associative algebras (with suitable notion of degree) – see Section 2 for details. In fact, to prove that any (abstract)¹ Golod-Shafarevich group is infinite, one reduces the problem to the corresponding question for

2000 *Mathematics Subject Classification.* Primary 20F05, 20F69, Secondary 20E18, 20F65, 22D10, 43A07.

Key words and phrases. Golod-Shafarevich, property (T) , amenability.

¹We shall often refer to ordinary groups as ‘abstract groups’ to distinguish them from pro- p groups

algebras defined as quotients of $K\langle\langle u_1, \dots, u_n \rangle\rangle$, where $K\langle\langle u_1, \dots, u_n \rangle\rangle$ is the algebra of non-commutative power series in u_1, \dots, u_n over a field K . More specifically, given a finitely generated group Γ , one considers the pro- p completion $\Gamma_{\hat{p}}$ of Γ and its completed group algebra $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$. Every presentation $\langle X|R \rangle$ for Γ yields the corresponding presentation for $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$ as a quotient of $\mathbb{F}_p\langle\langle u_1, \dots, u_n \rangle\rangle$ where $n = |X|$, and a clever dimension-counting argument shows that $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$ is infinite provided $\langle X|R \rangle$ satisfies the GS condition. If the algebra $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$ is infinite, the groups $\Gamma_{\hat{p}}$ and Γ must also be infinite.

Besides being infinite, Golod-Shafarevich (GS) groups are known to be “large” in various senses: for instance, any Golod-Shafarevich group Γ has an infinite torsion quotient (see [Go]), the pro- p completion of Γ contains a non-abelian free pro- p group (see [Ze]), and if $\{\omega_n\Gamma\}$ is the Zassenhaus p -series of Γ , the sequence $\{\log_p |\Gamma/\omega_n\Gamma|_{n=1}^\infty\}$ has exponential growth (see discussion in Section 2).

1.2. The main theorem and a naive approach to it. In [Er], the author constructed the first examples of Golod-Shafarevich groups with Kazhdan’s property (T) . Following this discovery, Lubotzky proposed a related conjecture that every Golod-Shafarevich group should have an infinite quotient with property (T) . The main goal of this paper is to prove Lubotzky’s conjecture:

Theorem 1.1. *Every Golod-Shafarevich group has an infinite quotient with Kazhdan’s property (T) .*

The result of [Er] was quite surprising and seemed to go against the general theme of “largeness” of Golod-Shafarevich groups. On the other hand, Theorem 1.1 can be considered a natural addition to the list of “largeness” properties of Golod-Shafarevich groups. Another important consequence of this theorem is an affirmative answer to a well-known question about non-amenability of Golod-Shafarevich groups (see, e.g., [H, Open Problem 5.2]), since quotients of amenable groups are amenable, while an infinite discrete group cannot be both amenable and Kazhdan.

Corollary 1.2. *Golod-Shafarevich groups cannot be amenable.*

In fact, in Appendix A to this paper it will be shown that Golod-Shafarevich groups satisfy a stronger form of non-amenability, called *uniform non-amenability*.

Before discussing the proof of Theorem 1.1, we observe that the analogous result is known to be true for (non-elementary) hyperbolic groups, that is, every hyperbolic group has an infinite quotient with property (T) . The latter is a combination of the following two deep results and the fact that property (T) is preserved by quotients:

- (a) There exist hyperbolic groups with property (T) .
- (b) Any two hyperbolic groups have a common infinite quotient.

Since Golod-Shafarevich groups with (T) exist by [Er], a “naive” approach to Theorem 1.1 would be to try to prove the analogue of (b) for Golod-Shafarevich groups (with respect to a fixed prime p), but such an assertion is almost definitely false.

In fact, the following would be sufficient to prove Theorem 1.1. Suppose that given a Golod-Shafarevich group Γ and a presentation $\langle X|R \rangle$ for Γ with $1 - H_X(t) + H_R(t) < 0$ for some $t = t(\Gamma) \in (0, 1)$, we can construct a group Γ' with property (T) which has a presentation $\langle X|R' \rangle$ (with the same generating set X as Γ) such that $H_{R'}(t) < -(1 - H_X(t) + H_R(t))$ for $t = t(\Gamma)$ (so, in particular, Γ_1 is also a Golod-Shafarevich group). Then the group $\Gamma'' = \langle X \mid R \cup R' \rangle$ is a common quotient of Γ and Γ' , and Γ'' is Golod-Shafarevich (hence infinite) since $1 - H_X(t) + H_{R \cup R'}(t) \leq 1 - H_X(t) + H_R(t) + H_{R'}(t) < 0$. Thus, Γ has an infinite quotient Γ'' , which has property (T) being a quotient of Γ' .

The supply of Golod-Shafarevich groups with property (T) given by [Er] was clearly too short for the above strategy to work for an arbitrary Golod-Shafarevich group Γ . In [EJ] a larger collection of Golod-Shafarevich groups with (T) was constructed within the class of *Kac-Moody-Steinberg* groups, also defined in [EJ]. This collection turns out to be sufficient for the proof of Theorem 1.1, but the approach outlined in the previous paragraph cannot be applied directly.

1.3. Outline of the proof of the main theorem. A key ingredient in the proof of Theorem 1.1 is analysis of presentations for finite index subgroups of Golod-Shafarevich pro- p groups. It is easy to see that if G is a GS group (abstract or pro- p), then standard presentations for finite index subgroups of G (obtained via Schreier rewriting process) do not necessarily satisfy the GS condition. However, as we will show in Section 3, finite index subgroups of Golod-Shafarevich pro- p groups always satisfy the generalized Golod-Shafarevich (GGS) condition. Similarly to the usual GS condition, the generalized Golod-Shafarevich condition for a presentation $\langle X|R \rangle$ is defined by inequality of the form $1 - H_{D,X}(t) + H_{D,R}(t) < 0$, where $H_{D,X}(t)$ and $H_{D,R}(t)$ generalize the previously defined series $H_X(t)$ and $H_R(t)$. This time we start with a degree function $D : X \rightarrow \mathbb{N}$ on the generating set X , then extend D in a canonical way to the free pro- p group on X and put $H_{D,X}(t) = \sum_{x \in X} t^{D(x)}$ and $H_{D,R}(t) = \sum_{r \in R} t^{D(r)}$ (the usual GS inequality corresponds to the case $D(x) = 1$ for all $x \in X$).

The key result of Section 3 (see Theorem 3.14) asserts that if G is any GGS pro- p group, then for any real number $M > 0$ there exists a finite index subgroup K of G , a (pro- p) presentation $\langle X|R \rangle$ of K , a degree function D on $F(X)$, and a real number $t_0 \in (0, 1)$ such that $1 - H_{D,X}(t_0) + H_{D,R}(t_0) < -M$.

Now given a GGS (abstract) group Γ , one can prove the existence of an infinite quotient with (T) for Γ using the following three-step algorithm:

- (1) Let $\Gamma_{\hat{p}}$ be the pro- p completion of Γ (then $\Gamma_{\hat{p}}$ is a GGS pro- p group). Find a finite index subgroup K of $\Gamma_{\hat{p}}$ which has a presentation $\langle X|R \rangle$ satisfying

$$1 - H_{D,X}(t_0) + H_{D,R}(t_0) < -6 \cdot 10^5$$

for some $t_0 \in (0, 1)$ and degree function D .

- (2) (see Theorem 4.3 and a remark after it) Let Δ be a dense finitely generated subgroup of K . Show that there is a Kac-Moody-Steinberg group Λ with property (T) such that Δ and Λ have a common infinite quotient Ω . Note that we can take $\Delta = K \cap \iota(\Gamma)$ (where $\iota : \Gamma \rightarrow \Gamma_{\hat{p}}$ is the canonical map), and thus a finite index subgroup of Γ (namely $\iota^{-1}(K) \cap \Gamma$) has an infinite quotient with (T) .
- (3) Deduce from Step (2) that the entire group Γ has an infinite quotient with (T) .

Of course, Step (1) can be accomplished by Theorem 3.14, and Step (3) is possible thanks to the following general result due to Andrei Jaikin-Zapirain (see Theorem 4.5): If Γ is any finitely generated group such that some finite index subgroup of Γ has an infinite quotient with (T) , then Γ itself has an infinite quotient with (T) .

We now comment on how to construct the group Ω in Step (2) and show that it is infinite. First we construct another presentation $\langle X'|R' \rangle$ of the group K (from Step (1)) and a new degree function D' such that $H_{D',X'}(t'_0) = 12$, $H_{D',R'}(t'_0) < \frac{1}{1000}$ for some $t'_0 \in (0, 1)$, and an extra technical condition holds (this is possible by Theorem 3.15). If $p \geq 67$, we construct Ω by adapting the “naive” approach to Theorem 1.1 discussed earlier, and we are able to make Ω a GGS group (and hence infinite).

If $p < 67$, this technique does not apply directly; in fact, we do not know any examples of groups with (T) which are GGS with respect to $p < 67$. In this case we prove that Ω is infinite by showing that there exist a finite field \mathbb{F} of characteristic p and an \mathbb{F} -algebra A which is GGS as an algebra over \mathbb{F} (hence infinite), such that the completed group algebra $\mathbb{F}_p[[\Omega_{\hat{p}}]]$ can be mapped onto a finite codimension subalgebra of A . Note that neither $\mathbb{F}_p[[\Omega_{\hat{p}}]]$ nor A will be GGS as algebras over \mathbb{F}_p . What makes such phenomenon possible is that many relations needed to define A as an algebra over \mathbb{F}_p become redundant if the ground field is changed from \mathbb{F}_p to \mathbb{F} .

1.4. Organization of the paper. In Section 2 we set up basic notations and discuss various forms of the Golod-Shafarevich condition for abstract groups, pro- p groups and associative algebras. In Section 3 we analyze presentations for subgroups of generalized Golod-Shafarevich groups. Finally, Section 4 contains the proof of Theorem 1.1.

The paper concludes with two appendices. Appendix A by the author and Andrei Jaikin-Zapirain contains a proof of uniform non-amenability of Golod-Shafarevich groups. Basic background on Kazhdan's property (T) and amenability is given at the beginning of Appendix A. In Appendix B by Andrei Jaikin-Zapirain a new lower bound for the subgroup growth of Golod-Shafarevich groups is established.

Acknowledgements. I am extremely grateful to Andrei Jaikin-Zapirain whose generous contributions to this manuscript yielded significant improvement of the main results and helped clarify the exposition. In particular, the proofs of Theorem 4.5 and Proposition 3.19 and generalization of many results in Section 3 are all due to him. I would also like to thank Efim Zelmanov for introducing me to Golod-Shafarevich groups several years ago and Martin Kassabov and Alex Lubotzky for useful discussions at various stages of this project.

2. PRO- p PRESENTATIONS AND GOLOD-SHAFAREVICH THEOREM

While the main result of this paper is about abstract groups, we will work primarily with presentations of finitely generated pro- p groups. For this reason, the simplest terminology and notations will be reserved for pro- p groups and their presentations.

Throughout the paper p will be a fixed prime number. Given a finite set X , by $F(X)$ we denote the free pro- p group on X . If F is a finitely generated free pro- p group and X is a free generating set for F , we shall canonically identify F with $F(X)$.

A *presentation* is a pair (X, R) , where X is a finite set and R is a countable subset of $F(X)$. By $\text{Gr}(X, R)$ we will denote the pro- p group defined by the presentation (X, R) , that is,

$$\text{Gr}(X, R) = F(X) / \langle R \rangle^{F(X)},$$

where $\langle R \rangle^{F(X)}$ is the (closed) normal subgroup of $F(X)$ generated by R .

The free abstract group on X will be denoted by $F_{\text{abs}}(X)$, and ordinary presentations of abstract groups will be referred to as *abstract presentations*.

Given a finite set $U = \{u_1, \dots, u_n\}$ and a field K , we put $K\langle\langle U \rangle\rangle = K\langle\langle u_1, \dots, u_n \rangle\rangle$, the algebra of non-commutative power series in u_1, \dots, u_n over K . Presentations of pro- p groups will be studied via the Magnus embedding of a free pro- p group $F(\{x_1, \dots, x_n\})$ into $(\mathbb{F}_p\langle\langle u_1, \dots, u_n \rangle\rangle)^*$ given by $x_i \mapsto 1 + u_i$ for $1 \leq i \leq n$. Such use of letters X and U will be kept throughout the paper.

2.1. Degree and weight functions. We start by defining the notions of degree and weight functions on algebras of the form $K\langle\langle U \rangle\rangle$.

Definition. Let $U = \{u_1, \dots, u_n\}$ be a finite set and let K be a field.

- (a) A function $d : K\langle\langle U \rangle\rangle \rightarrow \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is called a *degree function* on $K\langle\langle U \rangle\rangle$ if
- (i) $d(f) = \infty$ if and only if $f = 0$
 - (ii) $d(f) = 0$ if and only if f is invertible in $K\langle\langle U \rangle\rangle$
 - (iii) $d(fg) = d(f) + d(g)$ for any $f, g \in K\langle\langle U \rangle\rangle$
 - (iv) If $f = \sum_{\alpha} c_{\alpha} m_{\alpha}$ where each $c_{\alpha} \in K$, each m_{α} is of the form $u_{i_1} \dots u_{i_k}$ and all m_{α} are distinct, then $d(f) = \min\{d(m_{\alpha}) : c_{\alpha} \neq 0\}$.
- (b) A function $w : K\langle\langle U \rangle\rangle \rightarrow [0, 1]$ is called a *weight function* on $K\langle\langle U \rangle\rangle$ if
- (i) $w(f) = 0$ if and only if $f = 0$
 - (ii) $w(f) = 1$ if and only if f is invertible in $K\langle\langle U \rangle\rangle$
 - (iii) $w(fg) = w(f)w(g)$ for any $f, g \in K\langle\langle U \rangle\rangle$
 - (iv) If $f = \sum_{\alpha} c_{\alpha} m_{\alpha}$ where each $c_{\alpha} \in K$, each m_{α} is of the form $u_{i_1} \dots u_{i_k}$ and all m_{α} are distinct, then $w(f) = \max\{w(m_{\alpha}) : c_{\alpha} \neq 0\}$.

Remark: Every degree function d on $K\langle\langle U \rangle\rangle$ satisfies the condition

$$(iv)' \quad d(f+g) \geq \min\{d(f), d(g)\} \text{ for any } f, g \in K\langle\langle U \rangle\rangle$$

(this follows from (iv)), so d is a ring valuation on $K\langle\langle U \rangle\rangle$. Similarly, for any weight function w on $K\langle\langle U \rangle\rangle$ we have

$$(iv)'' \quad w(f+g) \leq \max\{w(f), w(g)\} \text{ for any } f, g \in K\langle\langle U \rangle\rangle.$$

Next we define the corresponding notions for free pro- p groups. Let F be a finitely generated free pro- p group and $X = \{x_1, \dots, x_n\}$ a free generating set for F . Consider another set with n elements $U = \{u_1, \dots, u_n\}$ and embed F into $\mathbb{F}_p\langle\langle U \rangle\rangle^*$ by sending x_i to $1 + u_i$.

Definition. Let F , X and U be as above.

- (a) Two functions $\Phi : F \rightarrow \mathbb{R}$ and $\varphi : \mathbb{F}_p\langle\langle U \rangle\rangle \rightarrow \mathbb{R}$ will be called *X -compatible* if

$$\Phi(f) = \varphi(f - 1) \text{ for any } f \in F$$

- (b) A function $D : F \rightarrow \mathbb{N} \cup \{\infty\}$ is called a *degree function* on (F, X) if it is X -compatible with some degree function d on $\mathbb{F}_p\langle\langle U \rangle\rangle$.
- (c) The unique degree function D on (F, X) such that $D(x) = 1$ for each $x \in X$ is called the *standard degree function* on (F, X) .
- (d) A function $W : F \rightarrow \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is called a *weight function* on (F, X) if it is X -compatible with some weight function w on $\mathbb{F}_p\langle\langle U \rangle\rangle$.

Clearly, there is a bijection between degree functions on (F, X) and $\mathbb{F}_p\langle\langle U \rangle\rangle$, and the same is true for weight functions. From now on we shall be mostly concerned with degree and weight functions on (F, X) .

It is easy to see that any weight function on (F, X) is of the following form: take any $t_1, \dots, t_n \in (0, 1)$, set $W(x_i) = t_i$, and then extend W uniquely to F using (iii) and (iv) in the definition of a weight function on $\mathbb{F}_p\langle\langle U \rangle\rangle$. Any weight function can be approximated (in the suitable sense) by integral weight functions defined below.

Definition. Let D be a degree function on (F, X) and $t \in (0, 1)$. The function $W : F \rightarrow (0, 1)$ defined by

$$W(f) = t^{D(f)}$$

will be called the (D, t) -*weight function* on (F, X) . Weight functions of this form will be called *integral*.

The following definition will be convenient for the study of Golod-Shafarevich groups.

Definition. Let X be a finite set and S a countable subset of $F = F(X)$.

(a) If W is a weight function on (F, X) , we set

$$W(S) = \sum_{f \in S} W(f) \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

(b) Let D be a degree function on (F, X) , and assume that for each $n \in \mathbb{N}$ the set $\{s \in S : D(s) = n\}$ is finite. Then the power series

$$H_{D,S}(t) = \sum_{f \in S} t^{D(f)}$$

will be called the *Hilbert series* of S (with respect to D). We will write H_S instead of $H_{D,S}$ whenever D is clear from context. Note that for any $t_0 \in [0, 1)$ we have $H_{D,S}(t_0) = W(S)$ where W is the (D, t_0) -weight function on (F, X) .

Similarly, one defines weights and Hilbert series for subsets of power series algebras $K\langle\langle U \rangle\rangle$.

Next we define a class of filtrations on finitely generated pro- p groups which come from degree functions.

Definition. Let G be a finitely generated pro- p group, X a finite set and $\pi : F(X) \rightarrow G$ a surjective homomorphism. Let D be a degree function on $(F(X), X)$. We define the filtration $\{G_n\}$ of G by

$$G_n = \{g \in G : g = \pi(f) \text{ for some } f \in F(X) \text{ with } D(f) \geq n.\}$$

We will call $\{G_n\}$ the D -filtration of G .

It is easy to see that each G_n is an open subgroup of G and $\cap \{G_n\} = \{1\}$. Furthermore, $\{G_n\}$ is a p -filtration of G , that is, $[G_i, G_j] \subseteq G_{i+j}$ and $(G_i)^p \subseteq G_{pi}$. If D is the standard degree function on $(F(X), X)$, the filtration $\{G_n\}$ is known as the *Zassenhaus p -filtration* of G .

Similarly, if d is a degree function on a power series algebra $K\langle\langle U \rangle\rangle$, one defines the d -filtration on any algebra which is a quotient of $K\langle\langle U \rangle\rangle$.

Finally, we recall the classical definition of a Hilbert series which can be associated to any algebra with a suitable filtration.

Definition. Let A be an associative algebra over some field K , and let $A = A_0 \supseteq A_1 \supseteq A_2 \supseteq \dots$ be a descending chain of K -subalgebras in A such that $A_i \cdot A_j \subseteq A_{i+j}$ and $\dim_K(A_i/A_{i+1}) < \infty$. The series

$$\text{Hilb}_A(t) = \sum_{n=0}^{\infty} \dim_K(A_n/A_{n+1})t^n$$

will be called the *classical Hilbert series* of A with respect to $\{A_n\}$.

2.2. Golod-Shafarevich inequality. Although the Golod-Shafarevich inequality is usually introduced as a tool for proving that some groups are infinite, it is actually a result about algebras. The following theorem which we call the *generalized Golod-Shafarevich inequality* (GGs inequality) is proved in Koch's Appendix to [Hab] (see formula (2.11) on p.105 of [Hab] and a remark at the end of this section).

Theorem 2.1 (GGs inequality). *Let U be a finite set, K a field and d a degree function on $K\langle\langle U \rangle\rangle$. Let $A = K\langle\langle U \rangle\rangle/I$ for some ideal I , and let S be a generating set for I such that $\{s \in S : d(s) = n\}$ is finite for each $n \in \mathbb{N}$.² Let $Hilb_A(t)$ be the classical Hilbert series of A with respect to the d -filtration. Then*

$$\frac{(1 - H_{d,U}(t) + H_{d,S}(t)) \cdot Hilb_A(t)}{1 - t} \geq \frac{1}{1 - t},$$

where for power series $\sum \alpha_i t^i$ and $\sum \beta_i t^i$, inequality $\sum \alpha_i t^i \geq \sum \beta_i t^i$ means that $\alpha_i \geq \beta_i$ for each i .

The important consequence of the GGS inequality is that if $1 - H_{d,U}(t_0) + H_{d,S}(t_0) < 0$ for some $0 < t_0 < 1$, the series $Hilb_A(t_0)$ must diverge, so in particular A is infinite-dimensional. We now explain how GGS inequality applies to (pro- p) groups.

Let G be a pro- p group given by a presentation (X, R) , where $X = \{x_1, \dots, x_n\}$. As before, consider the standard embedding of $F(X)$ into $\mathbb{F}_p\langle\langle U \rangle\rangle$ where $U = \{u_1, \dots, u_n\}$, and let I be the ideal of $\mathbb{F}_p\langle\langle U \rangle\rangle$ generated by the set $S = \{r - 1 : r \in R\}$. It is well known [Ko, Theorem 7.17] that there is a (canonical) isomorphism between the completed group algebra $F_p[[G]]$ and $\mathbb{F}_p\langle\langle U \rangle\rangle/I$ making the following diagram commutative:

$$(2.1) \quad \begin{array}{ccc} F(X) & \xrightarrow{\quad\quad\quad} & \mathbb{F}_p\langle\langle U \rangle\rangle \\ \downarrow & & \downarrow \\ G & \xrightarrow{\quad\quad\quad} & \mathbb{F}_p[[G]] \longrightarrow \mathbb{F}_p\langle\langle U \rangle\rangle/I \end{array}$$

(where all other maps are defined in the obvious way).

Now let D and d be X -compatible degree functions on $(F(X), X)$ and $\mathbb{F}_p\langle\langle U \rangle\rangle$, respectively, so that $H_{D,X}(t) = H_{d,U}(t)$ and $H_{D,R}(t) = H_{d,S}(t)$. Define the D -filtration on $\mathbb{F}_p[[G]]$ to be the image of the d -filtration on $\mathbb{F}_p\langle\langle U \rangle\rangle/I$ under the isomorphism $\mathbb{F}_p\langle\langle U \rangle\rangle/I \rightarrow \mathbb{F}_p[[G]]$ in (2.1) and let $Hilb_{D,\mathbb{F}_p[[G]]}(t)$ be the associated (classical) Hilbert series. Then the inequality of power series in Theorem 2.1 can be rewritten as follows:

$$(2.2) \quad \frac{(1 - H_{D,X}(t) + H_{D,R}(t)) \cdot Hilb_{D,\mathbb{F}_p[[G]]}(t)}{1 - t} \geq \frac{1}{1 - t}$$

Thus, if

$$1 - H_{D,X}(t_0) + H_{D,R}(t_0) < 0 \text{ for some } t_0 \in (0, 1), \quad (GGS)$$

then $Hilb_{D,\mathbb{F}_p[[G]]}(t_0)$ diverges, so the algebra $\mathbb{F}_p[[G]]$ (and hence the group G) is infinite.

Definition. (a) A presentation (X, R) is said to satisfy the *generalized Golod-Shafarevich (GGS) condition* if there exists a degree function D on $(F(X), X)$ and $t_0 \in (0, 1)$ such that $1 - H_{D,X}(t_0) + H_{D,R}(t_0) < 0$.

(b) A pro- p (resp. an abstract) group is called a *generalized Golod-Shafarevich group* if it has a pro- p (resp. abstract) presentation satisfying the GGS condition.

(c) A pro- p (resp. an abstract) group is called a *Golod-Shafarevich group* if it has a pro- p (resp. abstract) presentation satisfying the GGS condition with respect to the standard degree function.

The following result summarizes the above discussion.

²It is easy to see that any ideal of $K\langle\langle U \rangle\rangle$ has a generating set with this property.

Corollary 2.2. *Let (X, R) be a presentation satisfying the GGS condition for some degree function D on $(F(X), X)$ and $t_0 \in (0, 1)$. Let $G = \text{Gr}(X, R)$ be the pro- p group defined by this presentation. Then G is infinite, and moreover, the series $\text{Hilb}_{D, \mathbb{F}_p[[G]]}(t_0)$ diverges.*

The next result shows that the Hilbert series $\text{Hilb}_{D, \mathbb{F}_p[[G]]}(t)$ can be expressed directly in terms of the D -filtration of G .

Proposition 2.3. *Let G be a pro- p group with a finite generating set X , and let D be some degree function on $(F(X), X)$. Let $\{G_n\}$ and $\{A_n\}$ be the D -filtrations on G and $A = \mathbb{F}_p[[G]]$, respectively. Let $a_n(G) = \dim_{\mathbb{F}_p}(A_n/A_{n+1})$ and $c_n(G) = \dim_{\mathbb{F}_p}(G_n/G_{n+1})$ for $n \in \mathbb{N}$. Then*

$$(2.3) \quad \sum_{n=0}^{\infty} a_n(G) t^n = \prod_{n=1}^{\infty} \left(\frac{1 - t^{np}}{1 - t^n} \right)^{c_n(G)}.$$

Proof. In the special case when D is the standard degree function (and hence $\{G_n\}$ is the Zassenhaus filtration on G), Proposition 2.3 is an easy consequence of Quillen's theorem [Qu] (see [DDMS, Chapter 12] for a complete proof). Even though this special case is sufficient for the proof of Theorem 1.1, we give a proof of Proposition 2.3 for an arbitrary degree function D .

Let $L = \bigoplus_{n \geq 1} G_n/G_{n+1}$ and $B = \text{gr}(A) = \bigoplus_{n \geq 0} B_n$ where $B_n = A_n/A_{n+1}$. Then L is a graded p -Lie algebra with respect to operations $[gG_{n+1}, hG_{m+1}] = [g, h]G_{n+m+1}$ and $(gG_{n+1})^p = g^p G_{pn+1}$, and similarly B is a graded associative algebra with $B_n^p \subseteq B_{np}$. We can think of G as a subset of A ; it follows from the definition of D -filtrations on G and A that $G_n \subseteq 1 + A_n$ for all n , and hence we have a natural map $\iota : L \rightarrow B$ such that $\iota(gG_{n+1}) = (g - 1) + A_{n+1}$. It is easy to see that ι is a homomorphism of graded p -Lie algebras and thus uniquely extends to a homomorphism of graded associative algebras $\iota_* : \mathcal{U}(L) \rightarrow B$ where $\mathcal{U}(L)$ is the universal p -envelope of L . First we show that ι_* is surjective.

Identify A with a quotient of $\mathbb{F}_p\langle\langle U \rangle\rangle$ via (2.1). Let $\pi : \mathbb{F}_p\langle\langle U \rangle\rangle \rightarrow A$ be the corresponding surjection and let d be the degree function on $\mathbb{F}_p\langle\langle U \rangle\rangle$ X -compatible with D . Then $A_n = \pi((\mathbb{F}_p\langle\langle U \rangle\rangle)_n)$ where

$$(\mathbb{F}_p\langle\langle U \rangle\rangle)_n = \{f \in \mathbb{F}_p\langle\langle U \rangle\rangle : d(f) \geq n\}.$$

Let $d_j = d(u_j) = D(x_j)$. For $n \in \mathbb{N}$ let θ_n be the composite map $(\mathbb{F}_p\langle\langle U \rangle\rangle)_n \rightarrow A_n \rightarrow B_n$. Since d is a degree function, B_n is spanned by elements of the form $\theta_n(u_{i_1} \dots u_{i_s})$ with $\sum_{k=1}^s d_{i_k} = n$. But for each such element we have equality $\theta_n(u_{i_1} \dots u_{i_s}) = \prod_{k=1}^s \theta_{d_{i_k}}(u_{i_k})$ in B . Finally, it is easy to see that $\theta_{d_j}(u_j) = \iota(x_j G_{d_j+1}) \in \iota(L)$ for any $j \in \mathbb{N}$. Thus, ι_* is indeed surjective.

Now observe that the left-hand side of (2.3) is just the series $\text{Hilb}_B(t)$ while the right-hand side of (2.3) is equal to $\text{Hilb}_{\mathcal{U}(L)}(t)$ by the Poincare-Birkhoff-Witt theorem for graded p -Lie algebras. Since $\iota_* : \mathcal{U}(L) \rightarrow B$ is surjective and grading-preserving, (2.3) holds if and only if ι_* is injective.

First consider the case of finite G . If $|G| = p^k$, then $\dim_{\mathbb{F}_p} B = |G| = p^k$, $\dim_{\mathbb{F}_p} L = k$ and thus (again by the Poincare-Birkhoff-Witt theorem) $\dim_{\mathbb{F}_p} \mathcal{U}(L) = p^k$. Since $\mathcal{U}(L)$ and B have the same dimension, the surjective homomorphism ι_* must also be injective.

Finally, in the general case for any $N \in \mathbb{N}$ the group G/G_N is finite, and it is easy to see that $a_n(G) = a_n(G/G_N)$ and $c_n(G) = c_n(G/G_N)$ for $n < N$. Since we already established (2.3) for G/G_N and N is arbitrary, it follows that (2.3) holds for G as well. \square

Finally, we state a simple test for the GGS condition.

Lemma 2.4. *A presentation (X, R) satisfies the GGS condition if and only if there exists a weight function W on $(F(X), X)$ such that $1 - W(X) + W(R) < 0$.*

Proof. Let $F = F(X)$. The forward direction is clear since if D is a degree function on (F, X) and W is the (D, t) -weight function for some t , then $1 - W(X) + W(R) = 1 - H_{D,X}(t) + H_{D,R}(t)$.

The converse follows from the fact that any weight function can be approximated by integral weight functions. More precisely, let W be any weight function on (F, X) . For each $t \in (0, 1)$ let D_t be the unique degree function on (F, X) such that $D_t(x) = \lfloor \frac{\log W(x)}{\log t} \rfloor$ for each $x \in X$, and let W_t be the (D_t, t) -weight function on (F, X) . Then for any $f \in F$ we have $W_t(f) \leq W(f)$ and $W_t(f) \rightarrow W(f)$ as $t \rightarrow 1$. Since $|X| < \infty$ and $1 - W(X) + W(R) < 0$, we must have $1 - W_t(X) + W_t(R) < 0$ for some $t \in (0, 1)$. \square

Remark: Unfortunately, we are not aware of any reference where Theorem 2.1 appears as stated in this paper. Here we explain in detail how Theorem 2.1 follows from Koch's Appendix to [Hab]. We keep all the notations from the statement of Theorem 2.1. The numerical inequality (2.11) in [Hab, p.105] is easily seen to be equivalent to the inequality of powers series in Theorem 2.1, but the setup in [Hab] is somewhat restrictive as it is assumed that

- (i) The set S of defining relators of A is finite;
- (ii) $A = \mathbb{F}_p[[G]]$ for some finitely generated pro- p group G , and the presentation of A under consideration is obtained from some finite presentation of G in a standard way (as described in this section).

Fortunately, condition (ii) is not used in the proof of (2.11) at all. We claim that the restriction given by (i) is also inessential. Indeed, for $n \in \mathbb{N}$ let $S_n = \{s \in S : d(s) < n\}$ (this set is finite if S satisfies the hypotheses of Theorem 2.1), let I_n be the ideal of $K\langle\langle U \rangle\rangle$ generated by I_n and $A_n = K\langle\langle U \rangle\rangle / I_n$. It is clear that $H_{S_n}(t) \equiv H_S(t) \pmod{t^n}$ and $Hilb_{A_n}(t) \equiv Hilb_A(t) \pmod{t^n}$. Thus if for each n Theorem 2.1 holds for the pair (A_n, S_n) (which is proved in [Hab]), it automatically holds for (A, S) .

3. PRESENTATIONS FOR SUBGROUPS OF GOLOD-SHAFAREVICH GROUPS.

For a finitely generated pro- p group G we let $\Phi(G) = [G, G]G^p$ be the Frattini subgroup of G . The following basic facts will be frequently used (see [DDMS, Chapter 1]):

Claim 3.1. *The following hold:*

- (1) *A subset X generates G (topologically) if and only if X generates G modulo $\Phi(G)$.*
- (2) *If F is a free pro- p group of rank d , then $F/\Phi(F) \cong (\mathbb{Z}/p\mathbb{Z})^d$, and $X \subset F$ is a free generating set if and only if $X \bmod \Phi(F)$ is a basis for $F/\Phi(F)$.*

3.1. Some properties of weight functions. Let X be a finite set and $F = F(X)$. In this subsection we shall prove that the restriction of a weight function on (F, X) to an open subgroup F' of F is a weight function on (F', X') for some free generating set X' of F' (see Corollary 3.6). In fact, we shall give an algorithm for constructing such a set X' , and we will use this algorithm later on.

Lemma 3.2. *Let X be a finite set, $F = F(X)$ and W a weight function on (F, X) . The following hold:*

- (a) $W(fg) \leq \max\{W(f), W(g)\}$ for any $f, g \in F$
- (b) $W(x_i x_j^k) = \max\{W(x_i), W(x_j)\}$ if x_i, x_j are distinct elements of X and $p \nmid k$

$$(c) \ W([f, g]) \leq W(f)W(g).$$

Proof. (a) and (b) follow from the identity $fg - 1 = (f - 1) + (g - 1) + (f - 1)(g - 1)$ (and the definition of weight functions on groups). Similarly (c) follows from the identity $[f, g] - 1 = f^{-1}g^{-1}((f - 1)(g - 1) - (g - 1)(f - 1))$. \square

Lemma 3.3. *Let $X = \{x_1, \dots, x_d\}$ be a finite set, $F = F(X)$ and W a weight function on (F, X) . Let F' be a subgroup of index p in F . Then there exists a (free) generating set $\tilde{X} = \{\tilde{x}_1, \dots, \tilde{x}_d\}$ of F such that $W(\tilde{x}_i) = W(x_i)$ for $1 \leq i \leq d$ and $F' \supset \tilde{X} \setminus \{\tilde{x}_j\}$ for some j .*

Proof. Let $I = \{i : x_i \notin F'\}$, and choose $j \in I$ for which $W(x_j)$ is minimal possible. Define the set $\tilde{X} = \{\tilde{x}_1, \dots, \tilde{x}_d\}$ as follows:

If $i \in \{1, \dots, d\} \setminus I$ or $i = j$, we set $\tilde{x}_i = x_i$. If $i \in I \setminus \{j\}$, we let \tilde{x}_i be the (unique) element of the form $x_i x_j^k$, with $0 \leq k \leq p - 1$, which lies in F' (such k exists since $[F : F'] = p$ and $x_j \notin F'$).

Since \tilde{X} is obtained from X by a sequence of Nielsen transformations, it is a free generating set, and F' contains $\tilde{X} \setminus \{\tilde{x}_j\}$ by construction. Finally, $W(\tilde{x}_i) = W(x_i)$ by Lemma 3.2(b). \square

Lemma 3.4. *Let X be a finite set, $F = F(X)$, fix some $x \in X$ and let $Y = X \setminus \{x\}$. Let $F' = \langle Y \rangle \Phi(F)$. The following hold:*

- (a) F' is the unique subgroup of F of index p containing Y .
- (b) F' is freely generated by the set

$$(3.1) \quad X' = \cup_{y \in Y} \{y, [y, x], [y, x, x], \dots, [y, \underbrace{x, \dots, x}_{p-1 \text{ times}}]\} \cup \{x^p\};$$

Proof. (a) is obvious. To prove (b) note that $\{1, x, \dots, x^{p-1}\}$ is a transversal for F' in F , and the Schreier rewriting process yields that the set $\tilde{X} = \cup_{y \in Y} \{y, y^x, y^{x^2}, \dots, y^{x^{p-1}}\} \cup \{x^p\}$ generates F' . It is easy to prove by induction that for any $z \in F$ and $k \in \mathbb{N}$ we have

$$[z, \underbrace{x, \dots, x}_{k \text{ times}}] = f_k(x, z) \cdot z^{x^k} \text{ where } f_k(x, z) \in \langle z^{x^i} : 0 \leq i < k \rangle.$$

Thus, $\langle X' \rangle = \langle \tilde{X} \rangle = F'$. Finally, since $|X'| = p(|X| - 1) + 1$, by Schreier formula X' is a free generating set for F' . \square

Lemma 3.5. *Let $X = \{x_1, \dots, x_d\}$ be a finite set, $F = F(X)$ and W a weight function on (F, X) . Let F' be a finitely generated (pro- p) subgroup of F , let $X' = \{x'_1, \dots, x'_e\}$ be a free generating set of F' , and let W' be the unique weight function on (F', X') such that $W'(x') = W(x')$ for any $x' \in X'$. The following hold:*

- (a) For any $f \in F'$ we have $W'(f) \geq W(f)$.
- (b) Assume that one of the following holds:
 - (i) $F' = F$ (so that $d = e$) and $W(x_i) = W(x'_i)$ for $1 \leq i \leq d$.
 - (ii) F' is the unique subgroup of F of index p containing $X \setminus \{x\}$ for some $x \in X$, and X' is given by (3.1).

Then $W'(f) = W(f)$ for any $f \in F'$, and thus the restriction of W to F' is a weight function on (F', X') .

Proof. The following notations will be used in all parts of the proof. Let $U = \{u_1, \dots, u_d\}$, and let $\iota : F \rightarrow \mathbb{F}_p\langle\langle U \rangle\rangle$ be the unique homomorphism such that $\iota(x_i) = 1 + u_i$. By definition there is a weight function w on $\mathbb{F}_p\langle\langle U \rangle\rangle$ such that

$$W(f) = w(\iota(f) - 1) \text{ for any } f \in F.$$

Similarly, let $U' = \{u'_1, \dots, u'_e\}$ and $\iota' : F' \rightarrow \mathbb{F}_p\langle\langle U' \rangle\rangle$ the unique homomorphism such that $\iota'(x'_i) = 1 + u'_i$. Then there exists a weight function w' on $\mathbb{F}_p\langle\langle U' \rangle\rangle$ such that

$$W'(f) = w'(\iota'(f) - 1) \text{ for any } f \in F'. \quad (***)$$

(a) Let $\varphi : \mathbb{F}_p\langle\langle U' \rangle\rangle \rightarrow \mathbb{F}_p\langle\langle U \rangle\rangle$ be the unique homomorphism making the following diagram commutative:

$$\begin{array}{ccc} F' & & \\ \iota' \downarrow & \searrow \iota|_{F'} & \\ \mathbb{F}_p\langle\langle U' \rangle\rangle & \xrightarrow{\varphi} & \mathbb{F}_p\langle\langle U \rangle\rangle \end{array}$$

In other words, φ is defined by $\varphi(u'_i) = \iota(x'_i) - 1$ (such φ exists since for any $f \in F$ the element $\iota(f) - 1$ lies in the ideal of $\mathbb{F}_p\langle\langle U \rangle\rangle$ generated by U).

For any $f \in F'$ we have $W(f) = w(\iota(f) - 1) = w(\varphi(\iota'(f)) - 1) = w(\varphi(\iota'(f) - 1))$. Thus, in view of $(***)$, to prove (a) it suffices to show that

$$w'(h) \geq w(\varphi(h)) \text{ for any } h \in \mathbb{F}_p\langle\langle U' \rangle\rangle.$$

First note that for any $u'_i \in U'$ we have $w'(u'_i) = W'(x'_i) = W(x'_i) = w(\varphi(u'_i))$. Since w', w and φ preserve multiplication, we have $w'(h) = w(\varphi(h))$ whenever h is a U' -monomial (that is, a monomial in U').

Now take any $h \in \mathbb{F}_p\langle\langle U' \rangle\rangle$, and write $h = \sum c_\alpha m'_\alpha$ where $c_\alpha \in \mathbb{F}_p$ and $\{m'_\alpha\}$ are distinct monic U' -monomials. Then $w'(h) = \max\{w'(m'_\alpha) : c_\alpha \neq 0\}$ by (iv) in the definition of a weight function, while (iv)" yields

$$w(\varphi(h)) = w(\sum c_\alpha \varphi(m'_\alpha)) \leq \max\{w(\varphi(m'_\alpha)) : c_\alpha \neq 0\} = w'(h).$$

(b)(i) Define the isomorphisms $\theta : F \rightarrow F$ and $\psi : \mathbb{F}_p\langle\langle U \rangle\rangle \rightarrow \mathbb{F}_p\langle\langle U' \rangle\rangle$ by $\theta(x_i) = x'_i$ and $\psi(u_i) = u'_i$ for $1 \leq i \leq d$. Then we have a commutative diagram slightly different from the one in the proof of (a):

$$(3.2) \quad \begin{array}{ccc} F & \xrightarrow{\theta} & F \\ \iota \downarrow & & \downarrow \iota' \\ \mathbb{F}_p\langle\langle U \rangle\rangle & \xrightarrow{\psi} & \mathbb{F}_p\langle\langle U' \rangle\rangle \end{array}$$

We claim that

$$(3.3) \quad W(f) = W'(\theta(f)) \text{ for any } f \in F.$$

Indeed, in view of (3.2), it is enough to prove that

$$w(h) = w'(\psi(h)) \text{ for any } h \in \mathbb{F}_p\langle\langle U \rangle\rangle.$$

Since w (resp. w') is a weight function on $\mathbb{F}_p\langle\langle U \rangle\rangle$ (resp. $\mathbb{F}_p\langle\langle U' \rangle\rangle$) and $\psi : \mathbb{F}_p\langle\langle U \rangle\rangle \rightarrow \mathbb{F}_p\langle\langle U' \rangle\rangle$ is a ring isomorphism which sends U -monomials to U' -monomials, we are reduced to showing that $w(u_i) = w'(\psi(u_i))$ for $1 \leq i \leq d$. The latter is proved by the following chain of equalities

$$w(u_i) = W(x_i) = W(x'_i) = W'(x'_i) = w'(u'_i) = w'(\psi(u_i)),$$

where the first and fourth equalities hold by the definition of w and w' and the second and third ones hold by the hypotheses of Lemma 3.5.

Now given $\delta > 0$, let $F_\delta = \{f \in F : W(f) < \delta\}$ and $F'_\delta = \{f \in F : W'(f) < \delta\}$. Then F_δ and F'_δ are both open subgroups of F , and by Lemma 3.5(a) we have $F'_\delta \subseteq F_\delta$. On the other hand, (3.3) implies that $[F : F'_\delta] = [F : F_\delta]$. Combining these two facts, we conclude that $F'_\delta = F_\delta$ for any $\delta > 0$, which is equivalent to the assertion of (b)(i).

(b)(ii) Without loss of generality we can assume that $x = x_1$. Choose a total order on the set U such that u_1 is the smallest element. Consider the following order on the set of monic U -monomials $M = \{u_{i_1} \dots u_{i_k} : u_{i_j} \in U\}$:

$m < \tilde{m}$ if either $w(m) < w(\tilde{m})$, or $w(m) = w(\tilde{m})$ and $m < \tilde{m}$ lexicographically.

Given nonzero $f \in \mathbb{F}_p\langle\langle U \rangle\rangle$, we define the leading term of f , denoted $\text{LT}(f)$, to be the largest monic U -monomial which appears in f with nonzero coefficient. We also set $\text{LT}(0) = 0$. Note that

- (1) $\text{LT}(fh) = \text{LT}(f)\text{LT}(h)$ for any $f, h \in \mathbb{F}_p\langle\langle U \rangle\rangle$
- (2) $w(f) = w(\text{LT}(f))$ for any $f \in \mathbb{F}_p\langle\langle U \rangle\rangle$
- (3) If $\{f_\alpha\}$ is a collection of elements of $\mathbb{F}_p\langle\langle U \rangle\rangle$ with distinct leading terms and $c_\alpha \in \mathbb{F}_p$, then $\text{LT}(\sum c_\alpha f_\alpha)$ is the largest element of the set $\{\text{LT}(f_\alpha) : c_\alpha \neq 0\}$.

From (2), (3) and the proof of (a) it is clear that to prove (b)(ii) it suffices to show that the elements of the set $Z = \{\varphi(m'_\alpha) : m'_\alpha \text{ is a monic } U'\text{-monomial}\}$ have distinct leading terms.

Any element of Z can be uniquely written as a product of elements of the set $\{\varphi(u') : u' \in U'\} = \{\iota(x') - 1 : x' \in X'\}$. Recall that

$$X' = \cup_{y \in X \setminus \{x_1\}} \{[y, \underbrace{x_1 \dots x_1}_{k \text{ times}}] : 0 \leq k \leq p-1\} \cup \{x_1^p\}.$$

Clearly $\text{LT}(\iota(\underbrace{[x_j, x_1, \dots, x_1]}_{k \text{ times}}) - 1) = u_j u_1^k$ for any $j \geq 2$ and $\text{LT}(\iota(x_1^p) - 1) = u_1^p$. Property

(1) above easily implies that the elements of Z have distinct leading terms. \square

Corollary 3.6. *Let X be a finite set, $F = F(X)$ and W a weight function on (F, X) . Let F' be an open subgroup of F . Then there exists a free generating set X' of F' such that the restriction of W to F' is a weight function on (F', X') .*

Proof. It is enough to consider the case when F' has index p in F . In this case the result follows directly from Lemma 3.5(b) and Lemma 3.3. \square

We finish this section with the converse of Lemma 3.5(b)(ii). While this result is not critical for our purposes, it will help clarify the exposition.

Proposition 3.7. *Let X be a finite set, $F = F(X)$ and W a weight function on (F, X) . Let X' be another free generating set for F such that W is a weight function on (F, X') . Then there exists a bijection $\sigma : X' \rightarrow X$ such that $W(\sigma(x)) = W(x)$ for any $x \in X$. In particular, $W(X') = W(X)$.*

Before proving Proposition 3.7 we introduce an important definition and establish an auxiliary result.

Definition. Let X be a finite set, $F = F(X)$, and fix some $x \in X$.

- (a) An element $f \in F$ will be called *X -linear in x* if $f \notin \langle X \setminus \{x\} \rangle \Phi(F)$. We will say '*linear in x* ' instead of ' *X -linear in x* ' when X is clear from the context.

- (b) We will say that $f \in F$ is *linear* if $f \notin \Phi(F)$. Clearly, $f \in F$ is linear if and only if f is X -linear in some $x \in X$.

The following straightforward claim explains this terminology:

Claim 3.8. *Let $X = \{x_1, \dots, x_d\}$, $U = \{u_1, \dots, u_d\}$, and embed $F = F(X)$ in $\mathbb{F}_p\langle\langle U \rangle\rangle$ by $x_i \mapsto 1 + u_i$. Then $f \in F$ is X -linear in x_k if and only if the expansion of f as a power series in $\{u_1, \dots, u_d\}$ contains a term cu_k where c is a nonzero element of \mathbb{F}_p .*

Lemma 3.9. *Let $X = \{x_1, \dots, x_d\}$ be a finite set, $F = F(X)$ and W a weight function on (F, X) . Take $f \in F$, and write $f = f_L f_Q$ where $f_L = x_1^{k_1} \dots x_d^{k_d}$ with $0 \leq k_i \leq p-1$ and $f_Q \in \Phi(F)$ (such factorization is unique). The following hold:*

- (a) f is X -linear in x_i if and only if $k_i \neq 0$.
- (b) $W(f_L) = \max\{W(x_i) : f \text{ is } X\text{-linear in } x_i\}$.
- (c) $W(f) = \max\{W(f_L), W(f_Q)\}$.
- (d) If f is X -linear in x_i , then $W(f) \geq W(x_i)$.

Proof. (a) is obvious. Let $U = \{u_1, \dots, u_d\}$, embed $F = F(X)$ in $\mathbb{F}_p\langle\langle U \rangle\rangle$ by $x_i \mapsto 1 + u_i$. By definition, there exists a weight function w on $\mathbb{F}_p\langle\langle U \rangle\rangle$ such that $W(h) = w(h-1)$ for each $h \in F$.

Note that $f_L - 1 = \sum_{i=1}^d k_i u_i + r$ where each term in r has degree ≥ 2 and involves only u_i 's with $k_i \neq 0$. Thus,

$$W(f_L) = w(f_L - 1) = \max\{w(u_i) : k_i \neq 0\} = \max\{W(x_i) : k_i \neq 0\}.$$

In view of (a), this proves (b).

Since the expansion of f_Q in u_1, \dots, u_d has no linear terms, the terms of maximal w -weight in the expansions of $f_L - 1$ and $f_Q - 1$ are distinct. Since $f - 1 = f_L f_Q - 1 = (f_L - 1) + (f_Q - 1) + (f_L - 1)(f_Q - 1)$, (c) follows. Finally, (d) is a direct consequence of (b) and (c). \square

Proof of Proposition 3.7. Let $\overline{F} = F/\Phi(F)$, and for each $f \in F$ let \bar{f} be the image of f in \overline{F} . Define the function $\overline{W} : \overline{F} \rightarrow [0, 1)$ by setting

$$\overline{W}(\bar{f}) = \inf\{W(h) : \bar{h} = \bar{f}\}$$

We claim that

$$\overline{W}(\bar{f}) = \max\{W(x_i) : f \text{ is } X\text{-linear in } x_i\} \quad (***)$$

Indeed, define f_L as in Lemma 3.9. Then $\overline{f_L} = \bar{f}$, whence $\overline{W}(\bar{f}) \leq W(f_L)$. On the other hand, for any $h \in F$ with $\bar{h} = \bar{f}$ we have $h_L = f_L$, so by Lemma 3.9(c) $W(h) \geq W(h_L) = W(f_L)$. Thus, $\overline{W}(\bar{f}) = W(f_L)$, and $W(f_L) = \max\{W(x_i) : f \text{ is } X\text{-linear in } x_i\}$ by Lemma 3.9(b).

Now given $w \in [0, 1)$, let $n_w = |\{x \in X : W(x) \leq w\}|$. It follows from (***) that

$$|\bar{f} \in \overline{F} : \overline{W}(\bar{f}) \leq w| = p^{n_w}.$$

Thus, n_w is uniquely determined by W (not by X), and the assertion of Proposition 3.7 easily follows. \square

3.2. Transformations of presentations. Recall that if (X, R) is a presentation, $\text{Gr}(X, R)$ denotes the pro- p group defined by this presentation, that is, $\text{Gr}(X, R) = F(X)/\langle R \rangle^{F(X)}$.

Definition. Let (X, R) be a presentation, $F = F(X)$ and $\pi : F \rightarrow \text{Gr}(X, R)$ the natural surjection. Another presentation (X', R') will be called a *subpresentation* of (X, R) if

- (a) X' is a finite subset of F .
- (b) The (closed) subgroup $F' = \langle X' \rangle$ of F is freely generated by X'
- (c) If π' is the restriction of π to F' , then R' generates $\text{Ker } \pi'$ as a (closed) normal subgroup of F' .

Note that the group $\text{Gr}(X', R')$ can be canonically identified with the subgroup $H = \text{Im } \pi'$ of $G = \text{Gr}(X, R)$.

Definition. A *transformation* T is an “operation” of replacing a presentation (X, R) by its subpresentation (X', R') . We shall symbolically write $T : (X, R) \rightarrow (X', R')$.

We now describe four types of transformations which we call elementary.

1. p -descent. Let (X, R) be a presentation, $F = F(X)$, $G = \text{Gr}(X, R)$ and $\pi : F \rightarrow G$ the natural surjection. Let $x \in X$ be such that

$$(3.4) \quad \text{no relator from } R \text{ is } X\text{-linear in } x.$$

Let $Y = X \setminus \{x\}$, and define $X', R' \subset F(X)$ by setting

$$\begin{aligned} X' &= \cup_{y \in Y} \{y, [y, x], [y, x, x], \dots, [y, \underbrace{x, \dots, x}_{p-1 \text{ times}}]\} \cup \{x^p\}; \\ R' &= \{[\underbrace{r, x, \dots, x}_k : r \in R, 0 \leq k \leq p-1]\}. \end{aligned}$$

We claim that (X', R') is a subpresentation of (X, R) and $\text{Gr}(X', R')$ is a subgroup of index p in $G = \text{Gr}(X, R)$. The transformation $(X, R) \rightarrow (X', R')$ will be called the *p -descent at x* .

Indeed, by Lemma 3.4 X' is a free generating set for the subgroup $F' = \langle Y \rangle \Phi(F)$ which has index p in F . Since $R \subset F'$ by (3.4) (and hence $\text{Ker } \pi \subseteq F'$), the subgroup $H = \pi(F')$ has index p in G . Furthermore, $\{1, x, \dots, x^{p-1}\}$ is a transversal for F' in F , so the subset $R'' = \{r^{x^k} : r \in R, 0 \leq k \leq p-1\}$ generates $\text{Ker } \pi$ as a normal subgroup of F' . Thus, (X', R'') is a subpresentation of (X, R) and $\text{Gr}(X', R'') = H$. The same argument as in the proof of Lemma 3.4 shows that $\langle R'' \rangle = \langle R' \rangle$. Thus, (X', R') is also a subpresentation of (X, R) with $\text{Gr}(X', R') = \text{Gr}(X', R'') = H$.

The next three types of transformations do not change the group, that is, they replace a presentation (X, R) of a group G by another presentation (X', R') of G .

2. Change of generators. A *change of generators* is a transformation of the form $(X, R) \rightarrow (X', R)$ where X' is a free generating set of $F(X)$. Clearly, $\text{Gr}(X, R) = \text{Gr}(X', R)$. A basic change of generators is given by a Nielsen transformation, that is, we set $X' = X \setminus \{x\} \cup \{xx_1^{\pm 1}\}$ where x and x_1 are distinct elements of X .

3. Change of relators. A *change of relators* is a transformation of the form $(X, R) \rightarrow (X, R')$ (where R and R' generate the same normal subgroup of $F(X)$). By definition $\text{Gr}(X, R) = \text{Gr}(X, R')$. A basic change of relators is given by setting $R' = R \setminus \{r\} \cup \{rr_1^{\pm 1}\}$ where r and r_1 are distinct elements of R .

4. Cleanup. Suppose that (X, R) is a presentation of such that $X \cap R \neq \emptyset$, and let $C \subseteq X \cap R$. We can construct a new presentation (X', R') of the same group by eliminating

C both from sets of generators and relators and replacing each element of C by 1 in the remaining relators.

Formally, we set $X' = X \setminus C$ and $R' = \varphi(R \setminus C)$ where $\varphi : F(X) \rightarrow F(X')$ is the homomorphism which acts as identity on $F(X')$ and sends all elements of C to 1. It is an easy exercise to check that $\text{Gr}(X', R') = \text{Gr}(X, R)$. The transformation $(X, R) \rightarrow (X', R')$ will be called *the cleanup of the set C* .

Finally, we introduce one more type of transformations (which will not be considered elementary). It composes a change of generators with a cleanup and shows the usefulness of cleanups.

Pair elimination. Let (X, R) be a presentation. Suppose that some $r \in R$ is X -linear in some $x \in X$. By Claim 3.1(b) $\tilde{X} = X \setminus \{x\} \cup \{r\}$ is a free generating set for $F(X)$. Let (X', R') be the presentation obtained from (X, R) by first making the change of generators $(X, R) \rightarrow (\tilde{X}, R)$ followed by the cleanup of the singleton set $\{r\}$ (so that $X' = X \setminus \{x\}$). The transformation $(X, R) \rightarrow (X', R')$ will be called *the elimination of the pair (x, r)* .

3.3. Weight-preserving transformations.

Definition. A *weighted presentation* is a triple (X, R, W) where (X, R) is a presentation and W is a weight function on $(F(X), X)$.

Suppose that (X, R, W) is a weighted presentation and $(X, R) \rightarrow (X', R')$ is an elementary transformation. We shall be concerned with the following questions.

- (a) Let W' be the restriction of W to $F(X')$. Is W' a weight function on $(F(X'), X')$? In other words, is (X', R', W') a weighted presentation?
- (b) What is the relationship between $W(R)$ and $W(R')$?

Affirmative answer to question (a) for most elementary transformations was obtained in Subsection 3.1. We now analyze question (b).

Definition. Let (X, R, W) be a weighted presentation and $F = F(X)$.

- (a) A change of generators $(X, R) \rightarrow (X', R)$ will be called *W -good* if W is a weight function on (F, X') .
- (b) A change of relators $(X, R) \rightarrow (X, R')$ will be called *W -good* if $W(R') \leq W(R)$.
- (c) A *W -good elementary transformation* is any p -descent, any cleanup or a W -good change of generators or relators.
- (d) A transformation $(X, R) \rightarrow (X', R')$ will be called *W -good* if it is obtained by a sequence of W -good elementary transformations.

If D is a degree function on (X, R) , a transformation $(X, R) \rightarrow (X', R')$ will be called *D -good* if it is W -good where W is the (D, t) -weight function on $(F(X), X)$ for some $t \in (0, 1)$ (clearly, the value of t is not essential).

Lemma 3.10. *Let (X, R, W) be a weighted presentation and $T : (X, R) \rightarrow (X', R')$ a W -good elementary transformation.*

- (a) *The restriction of W to $F(X')$ is a weight function on $(F(X'), X')$.*
- (b) *Suppose that T is the p -descent at some $x \in X$, let $\tau = W(x)$ and $c = \frac{1-\tau^p}{1-\tau}$. Then*

$$W(R') \leq cW(R) \text{ and } W(X') - 1 = c(W(X) - 1).$$

- (c) *Suppose that T is a change of generators or relators. Then $W(R') \leq W(R)$ and $W(X') = W(X)$.*
- (d) *Suppose T is the cleanup of a set $C \subseteq X \cap R$. Then $W(X') = W(X) - W(C)$ and $W(R') \leq W(R) - W(C)$.*

Now assume that the transformation $T : (X, R) \rightarrow (X', R')$ is D -good for some degree function D on $(F(X), X)$.

(e) If T is a p -descent at some $x \in X$ and $n = D(x)$, then

$$\frac{1 - H_{X'}(t) + H_{R'}(t)}{1 - t} \leq \frac{1 - H_X(t) + H_R(t)}{1 - t} \cdot \frac{1 - t^{pn}}{1 - t^n}$$

as power series (where all Hilbert series are with respect to D).

(f) If T is a change of generators or relators or a cleanup, then

$$\frac{1 - H_{X'}(t) + H_{R'}(t)}{1 - t} \leq \frac{1 - H_X(t) + H_R(t)}{1 - t}$$

as power series.

(g) Let $G = \text{Gr}(X, R)$, $G' = \text{Gr}(X', R')$, and let $\{G_n\}$ (resp. $\{G'_n\}$) be the D -filtration of G (resp. G'). Then $G'_n = G' \cap G_n$.

Proof. (a) If T is a change of generators, the assertion holds by definition. If T is a change of relators or a cleanup, the assertion is obvious. If T is a p -descent, the assertion holds by Lemma 3.5(b)(ii).

(b) For any $f, g \in F(X)$ we have $W([f, g]) \leq W(f)W(g)$, and thus

$$W(\underbrace{[r, x, \dots, x]}_{k \text{ times}}) \leq W(r)W(x)^k = W(r)\tau^k,$$

which implies that $W(R') \leq cW(R)$. Furthermore, computation of leading terms in the proof of Lemma 3.5(b)(ii) shows that $W(\underbrace{[y, x, \dots, x]}_{k \text{ times}}) = W(y)\tau^k$ for any $y \in X \setminus \{x\}$ and

$W(x^p) = \tau^p$. Thus, $W(X') = (W(X) - \tau)c + \tau^p$, so $W(X') - 1 = cW(X) - c\tau + \tau^p - 1 = c(W(X) - 1)$.

(c) If T is a change of generators, the equality $W(X') = W(X)$ holds by Proposition 3.7. The other assertions are obvious.

(d) It is clear that $W(X') = W(X \setminus C) = W(X) - W(C)$ and $W(R \setminus C) = W(R) - W(C)$. Recall that $R' = \varphi(R \setminus C)$ where $\varphi : F(X) \rightarrow F(X')$ is the homomorphism which acts as identity on X' and sends C to $\{1\}$. Clearly, application of φ cannot increase the weights, so $W(R') \leq W(R \setminus C) = W(R) - W(C)$.

(e) First note that (b) applied to the (D, t) -weight function implies that (e) holds as a numerical inequality for any $t \in (0, 1)$.

The proof of (b) also shows that $H_{X'}(t) - 1 = (H_X(t) - 1) \cdot \frac{1 - t^{pn}}{1 - t^n}$ as power series. Since $H_{\sqcup_{i=1}^\infty S_i}(t) = \sum_{i=1}^\infty H_{S_i}(t)$ as power series for pairwise disjoint sets $\{S_i\}$, to finish the proof of (e) it suffices to check that

$$(3.5) \quad \frac{H_{R'}(t)}{1 - t} \leq \frac{H_R(t)}{1 - t} \cdot \frac{1 - t^{pn}}{1 - t^n} \text{ as power series when } |R| = 1.$$

If $R = \{r\}$ with $D(r) = m$, then $H_R(t) = t^m$, and the proof of (b) shows that $H_{R'}(t) = \sum_{i=0}^{p-1} t^{m+in+\varepsilon_i}$ with $\varepsilon_i \in \mathbb{Z}_{\geq 0} \cup \infty$ (where we set $t^\infty = 0$). Thus,

$$H_R(t) \cdot \frac{1 - t^{pn}}{1 - t^n} - H_{R'}(t) = \sum_{i=0}^{p-1} t^{m+in} (1 - t^{\varepsilon_i}).$$

This implies (3.5) since the power series $\frac{\sum_{i=0}^{p-1} t^{m+in}(1 - t^{\varepsilon_i})}{1 - t}$ clearly has non-negative coefficients.

(f) follows from the proof of (d) similarly to how (e) was deduced from the proof of (b).

(g) The assertion appears to be automatic, but it is not. Let $F = F(X)$, $F' = F(X')$ and $\pi : F \rightarrow G$ the natural surjection. Given $g \in G'$ and $f \in F$ with $\pi(f) = g$, we need to show that there exists $f' \in F'$ with $\pi(f') = g$ and $D(f') \geq D(f)$.

If T is a change of generators or relators, then $F' = F$ and there is nothing to prove. If T is a p -descent, the assertion is still clear since by construction $F' = \pi^{-1}(G')$. Finally, if $T : (X, R) \rightarrow (X', R')$ is the cleanup of a set $C \subseteq X \cap R$, we put $f' = \varphi(f)$ where $\varphi : F \rightarrow F'$ is the homomorphism which acts as identity on X' and sends C to $\{1\}$. Then $f' \in F'$, and it is clear that $\pi(f') = \pi(f) = g$ and $D(f') \geq D(f)$. \square

3.4. Finitary version of the GGS inequality. In this subsection we prove a fundamental inequality relating the Hilbert series of a presentation of a pro- p group G with that of a suitable presentation of a finite index subgroup of G . In particular, this inequality implies that the generalized GS condition is preserved under the passage to finite index subgroups.

Theorem 3.11. *Let (X, R) be a presentation and D a degree function on $(F(X), X)$. Let $G = \text{Gr}(X, R)$ and $\{G_n\}$ the D -filtration of G . Let K be a finite index subgroup of G , and for each $n \in \mathbb{N}$ define $c_n(G/K) = \log_p[KG_n : KG_{n+1}]$ (note that $c_n = 0$ for sufficiently large n). The following hold:*

(a) *There exists a D -good transformation $(X, R) \rightarrow (X', R')$ such that*

(i) $K = \text{Gr}(X', R')$

(ii) *The following inequality of power series holds:*

$$(3.6) \quad \frac{1 - H_{X'}(t) + H_{R'}(t)}{1 - t} \leq \frac{1 - H_X(t) + H_R(t)}{1 - t} \prod_{n=1}^{\infty} \left(\frac{1 - t^{pn}}{1 - t^n} \right)^{c_n(G/K)}$$

where all Hilbert series are with respect to D .

(b) *Assume in addition that $K \subseteq G_m$ for some $m \in \mathbb{N}$. Then in (a) we can require that $D(x) \geq m$ for any $x \in X'$ (and hence $D(f) \geq m$ for any $f \in F(X')$).*

Remark: The following observations were made by Jaikin-Zapirain:

(1) If K is a normal subgroup of G , Proposition 2.3 implies that

$$(3.7) \quad \prod_{n=1}^{\infty} \left(\frac{1 - t^{pn}}{1 - t^n} \right)^{c_n(G/K)} = \text{Hilb}_{\mathbb{F}_p[G/K]}(t),$$

where $\text{Hilb}_{\mathbb{F}_p[G/K]}(t)$ is the classical Hilbert series of the group algebra $\mathbb{F}_p[G/K]$ corresponding to the degree function D . In fact, there is a natural way to define $\text{Hilb}_{\mathbb{F}_p[G/K]}(t)$ even when K is not normal, and (3.7) still holds – we shall not prove this fact as it will not be used in the sequel.

(2) In view of (3.7), Theorem 3.11 can be considered as a finitary version of the GGS inequality (2.2). In fact, (2.2) can be deduced from Theorem 3.11 (this is what part (b) is useful for). Indeed, take $m \in \mathbb{N}$, and let (X', R') be the presentation satisfying the conclusion of Theorem 3.11(a)(b) with $K = G_m$. It is easy to see that $\text{Hilb}_{\mathbb{F}_p[G/G_m]}(t) \equiv \text{Hilb}_{\mathbb{F}_p[[G]]}(t) \pmod{t^m}$, and Theorem 3.11(b) ensures that $1 - H_{X'}(t) + H_{R'}(t) \equiv 1 \pmod{t^m}$. It follows that the (coefficient-wise) inequality of power series in (2.2) holds at least up to degree $m - 1$. Since m is arbitrary, we deduce (2.2).

Theorem 3.11(a) is reduced to the following lemma by straightforward induction on $[G : K]$.

Lemma 3.12. *Let $(X, R), G, D, K, \{G_n\}$ be as in Theorem 3.11. Choose a subgroup $L \subseteq G$ of index p such that $K \subseteq L$, and set $L_n = L \cap G_n$ and $c_n(L/K) = \log_p[KL_n : KL_{n+1}]$. Then there exist a D -good transformation $(X, R) \rightarrow (X', R')$ and $i \in \mathbb{N}$ with the following properties:*

- (a) $L = \text{Gr}(X', R')$;
- (b) $\{L_n\}$ is the D -filtration of L ;
- (c) $c_n(L/K) = c_n(G/K)$ for $n \neq i$ and $c_i(L/K) = c_i(G/K) - 1$.
- (d) The following inequality of power series holds:

$$\frac{1 - H_{X'}(t) + H_{R'}(t)}{1 - t} \leq \frac{1 - H_X(t) + H_R(t)}{1 - t} \cdot \frac{1 - t^{pi}}{1 - t^i}.$$

Proof. Let $F = F(X)$, $\pi : F \rightarrow G$ the natural surjection, and $F' = \pi^{-1}(L)$, so that $[F : F'] = p$. By Lemmas 3.3 and 3.5(b)(i) there exists a free generating set \tilde{X} of F such that

- (i) The change of generators $(X, R) \rightarrow (\tilde{X}, R)$ is D -good.
- (ii) $F' \supseteq \tilde{X} \setminus \{x\}$.

We claim that

- (iii) R contains no elements which are \tilde{X} -linear in x .

Indeed, by (ii) we have $F' = \langle \tilde{X} \setminus \{x\} \rangle \Phi(F)$, so (iii) simply asserts that $R \subset F'$. Suppose this is not the case. Then $\langle R \rangle F' = F$, whence $L = \pi(F') = \pi(\langle R \rangle F') = G$, a contradiction.

Let (X', R') be the presentation obtained from (\tilde{X}, R) by applying the p -descent at x (such p -descent is possible by condition (iii)). Condition (ii) implies that $\text{Gr}(X', R') = L$, so (a) holds. Note that (b) holds by Lemma 3.10(g).

Now let $i = D(x)$. Inequality in (d) holds by condition (i) and Lemma 3.10(e)(f). It remains to prove (c). For a subgroup H of G we set $H_n = H \cap G_n$ and $c_n(H) = \dim_{\mathbb{F}_p}(H_n/H_{n+1})$. Then it is easy to see that $c_n(G/K) = c_n(G) - c_n(K)$ and $c_n(L/K) = c_n(L) - c_n(K)$. Since L is a subgroup of index p in G , there exists unique $j \in \mathbb{N}$ such that $c_n(L) = c_n(G)$ for $n \neq j$ and $c_j(L) = c_j(G) - 1$, and we only need to prove that $j = i$. It will suffice to show that $L_i \neq G_i$ and $L_{i+1} = G_{i+1}$.

The first assertion is clear since $\pi(x) \in G_i \setminus L$ by construction. For the second assertion note that any $f \in F \setminus F'$ is \tilde{X} -linear in x and thus $D(f) \leq D(x) = i$ by Lemma 3.9(d). Hence $G_{i+1} = \{\pi(f) : f \in F \text{ and } D(f) \geq i+1\} \subseteq \pi(F') = L$, and so $L_{i+1} = G_{i+1} \cap L = G_{i+1}$. \square

Proof of Theorem 3.11(b). Suppose that the presentation (X', R') in Theorem 3.11(a) does not satisfy the conclusion of (b), so there exists $x \in X'$ with $D(x) < m$. We can assume that x has the smallest D -degree among all elements of X' .

Let $F' = F(X')$, $\pi : F' \rightarrow K$ the natural surjection and $N = \text{Ker } \pi$. Since $K \subseteq G_m$, there exists $f \in F'$ such that $D(f) \geq m$ and $\pi(f) = \pi(x)$ (here we use Lemma 3.10(g)). Note that f is not X' -linear in x (for otherwise $D(f) = D(x) < m$). Thus, $f^{-1}x$ is an element of N which is X' -linear in x , whence some $r \in R'$ must be X' -linear in x (as $N = \langle R' \rangle^{F'}$). Note that $D(r) = D(x)$ by the choice of x .

Let (X'', R'') be the presentation obtained from (X', R') by eliminating the pair (x, r) . Since $D(r) = D(x)$, this pair elimination is D -good by Lemma 3.5(b)(i), so by Lemma 3.10(f) inequality (3.6) holds with (X', R') replaced by (X'', R'') .

If the presentation (X'', R'') does not satisfy the conclusion of Theorem 3.11(b), we repeat the above procedure. Since each application of this procedure decreases the size of

the generating set by 1, after finitely many steps we obtain a presentation with desired property. \square

3.5. GGS condition for subgroups of GGS pro- p groups.

Definition. Let w, δ and ε be positive real numbers. We will say that a weighted presentation (X, R, W) satisfies the condition **Pres** (w, δ, ε) if

- (a) $W(X) = w$ and $W(R) < \delta$;
- (b) $\sum_{w \in \text{Im}(W)} w < \varepsilon$ where $\text{Im}(W) = \{W(f) : f \in F(X)\}$ is the image of W .

Our ultimate goal for applications in the next section is to prove the following result.

Theorem 3.13. *Let G be a GGS pro- p group, and let w, δ and ε be any positive real numbers. Then there exists an open normal subgroup H of G and a weighted presentation (X, R, W) of H which satisfies **Pres** (w, δ, ε) .*

However, the next theorem should probably be considered the key result of this section.

Theorem 3.14. *Let G be a GGS pro- p group. Then for any $M > 0$ there exists an open normal subgroup K of G and a weighted presentation (X, R, W) of K such that*

$$1 - W(X) + W(R) < -M.$$

Proof of Theorem 3.14. We are given that G has a presentation (X, R) such that $1 - H_X(t_0) + H_R(t_0) < 0$ for some $0 < t_0 < 1$ and degree function D on $(F(X), X)$ (where all Hilbert series are with respect to D). Let W be the (D, t_0) -weight function on $(F(X), X)$, that is, $W(f) = t_0^{D(f)}$.

Let $\{G_n\}$ be the D -filtration of G . For each $n \in \mathbb{N}$ let (X_n, R_n) be a presentation satisfying the conclusion of Theorem 3.11 with $K = G_{n+1}$. Let

$$\begin{aligned} \mu &= -(1 - W(X) + W(R)) = -(1 - H_X(t_0) + H_R(t_0)) \\ \mu_n &= -(1 - W(X_n) + W(R_n)) = -(1 - H_{X_n}(t_0) + H_{R_n}(t_0)) \text{ for } n \in \mathbb{N}. \end{aligned}$$

Let $c_k = \dim_{\mathbb{F}_p}(G_k/G_{k+1})$, and define $c_k(G/G_{n+1})$ as in Theorem 3.11. Then clearly

$$c_k(G/G_{n+1}) = \begin{cases} c_k & \text{if } k \leq n \\ 0 & \text{if } k > n \end{cases}$$

Thus, Theorem 3.11 yields

$$\mu_n \geq \mu \prod_{k=1}^n \left(\frac{1 - t_0^{pk}}{1 - t_0^k} \right)^{c_k}.$$

Proposition 2.3 and Corollary 2.2 imply that the infinite product $\prod_{k=1}^{\infty} \left(\frac{1 - t_0^{pk}}{1 - t_0^k} \right)^{c_k}$ diverges.

Since $\mu > 0$, we have $\mu_n \rightarrow \infty$. The proof is complete. \square

Our next result combined with Theorem 3.14 implies Theorem 3.13.

Theorem 3.15. *Let w, δ, ε and M be positive real numbers, with $\delta, \varepsilon < 1$. Let K be a pro- p group which has a weighted presentation (X, R, W) such that $1 - W(X) + W(R) < -M$.*

- (a) *Assume that $M > \max\{\frac{w^2}{\delta}, \frac{w}{\varepsilon}\}$. Then K has a weighted presentation (X', R', W') such that $W'(X') = w$, $W'(R') < \delta$ and $\max\{W'(x) : x \in X'\} < \varepsilon$.*
- (b) *Assume that $M > 4 \max\{\frac{w^2}{\delta}, \frac{w}{\varepsilon}\}$. Then K has a weighted presentation satisfying **Pres** (w, δ, ε) .*

The desired weighted presentation of K in Theorem 3.15(b) will be obtained from the original weighted presentation (X, R, W) using transformations and contractions of weight functions as defined below.

Definition. Let X be a finite set, $F = F(X)$ and W a weight function on (F, X) . Let $c \geq 1$ be a real number. The c -contraction of W is the unique weight function W' on (F, X) such that $W'(x) = \frac{W(x)}{c}$ for any $x \in X$.

Lemma 3.16. Let $X = \{x_1, \dots, x_n\}$ be a finite set, $F = F(X)$, W a weight function on (F, X) and W' the c -contraction of W for some $c \geq 1$. The following hold:

- (a) $W'(f) \leq \frac{W(f)}{c}$ for any $f \in F$;
- (b) If f is not linear, that is, $f \in \Phi(F)$, then $W'(f) \leq \frac{W(f)}{c^2}$.

Now take $f \in F \setminus \Phi(F)$, and write $f = f_L f_Q$ where $f_L = x_1^{k_1} \dots x_n^{k_n}$ with $0 \leq k_i \leq p-1$ and $f_Q \in \Phi(F)$. Let $\lambda = W(f)/W(f_L)$. We have

- (c) $W'(f_L) = \frac{W(f_L)}{c}$
- (d) If $c \leq \lambda$, then $W'(f) \leq \frac{W(f)}{c^2}$
- (e) If $c \geq \lambda$, then $W'(f) = W'(f_L)$

Proof. (a) is obvious, (b) is clear in view of Claim 3.8, (c) follows directly from Lemma 3.9(b), and (d) and (e) follow easily from (a), (b), (c) and Lemma 3.9(c). \square

Proof of Theorem 3.15. (a) We start with a special case when R contains no linear relators (we warn the reader that this special case occurs very rarely in practice). In this case we set $(X', R') = (X, R)$ and define the new weight function W' to be the c -contraction of W where $c = W(X)/w$.

Clearly, $W'(X) = W(X)/c$. Since $1 - W(X) + W(R) < -M$, we have $W(X) > M$ and $W(X) > W(R)$. Thus, Lemma 3.16(b) yields

$$W'(R) \leq \frac{W(R)}{c^2} = \frac{W(R) \cdot w^2}{W(X)^2} < \frac{w^2}{W(X)} < \frac{w^2}{M} < \delta.$$

Since $c > \frac{M}{w}$, for any $x \in X$ we have

$$W'(x) = \frac{W(x)}{c} < \frac{1}{c} < \frac{w}{M} < \varepsilon.$$

General case: We need to find a way to treat linear relators efficiently. Here is the idea. Let us say that a subset C of R is W -regular if there exists a free generating set \tilde{X} of $F = F(X)$ such that $C \subseteq \tilde{X}$ and W is a weight function on (F, \tilde{X}) (note that W -regular subsets may only contain linear relators, but for a linear relator r it may happen that the singleton set $\{r\}$ is already not W -regular). If C is a W -regular set of relators, we can remove it using a W -good transformation by first making the W -good change of generators $(X, R) \rightarrow (\tilde{X}, R)$ followed by the cleanup of C . To make use of this observation, we proceed as follows.

First, we “optimize” the set of relators in a suitable way (using a change of relators). Then we apply the c -contraction to W for suitable c and show that the set R_{bad} of all linear relators in R can be divided into two disjoint subsets R_{bad}^1 and R_{bad}^2 with the following properties: the set R_{bad}^1 is regular for the contracted weight function, while the weights of elements of R_{bad}^2 contract at least by the factor of c^2 (so they behave as non-linear relators during the contraction). Then we remove all relators from R_{bad}^1 as described above, and after an additional contraction we obtain a weighted presentation with desired properties.

We now begin the formal proof.

Step 1: *Optimizing the set of relators.*

First note that R contains only finitely many linear relators since $W(R) < \infty$ while for any linear relator $r \in R$ we have $W(r) \geq \min\{W(x) : x \in X\}$.

Let $R_{bad} = \{r_1, \dots, r_b\}$ be the set of all linear relators in R . Without loss of generality, we can assume that $W(r_1) \leq W(r_i)$ for any $i \geq 2$. Let $x_1 \in X$ be the generator such that r_1 is linear in x_1 and $W(x_1)$ is largest among all generators with this property. After applying a change of relators of the form $r_i \rightarrow r_i r_1^{k_i}$ for $2 \leq i \leq b$, we can assume that all relators besides r_1 are NOT linear in x_1 . By the choice of r_1 , this change of relators is W -good, so the total weight of relators cannot increase. Let $\tilde{R}_{bad} = \{\tilde{r}_2, \dots, \tilde{r}_{\tilde{b}}\}$ be the new set of linear relators excluding r_1 (note that \tilde{b} may be smaller than b since some of the relators $r_i r_1^{k_i}$ may not be linear).

Now apply the same procedure to the set \tilde{R}_{bad} : after reordering, assume that $W(\tilde{r}_2) \leq W(\tilde{r}_i)$ for $i \geq 3$, choose $x_2 \in X$ such that \tilde{r}_2 is linear in x_2 and $W(x_2)$ is largest possible etc.. Note that x_2 is different from x_1 .

After repeating this procedure at most b times and changing the notations, we can assume that the set $R_{bad} = \{r_1, \dots, r_m\}$ of linear relators in R has the following property: there exist distinct elements $x_1, \dots, x_m \in X$ such that for $1 \leq i \leq m$ we have

- (i) r_i is linear in x_i and not linear in x_j for $j < i$;
- (ii) if r_i is linear in x for some other $x \in X$, then $W(x) \leq W(x_i)$.

Note that (i),(ii) and Lemma 3.9(b) imply that

- (iii) $W((r_i)_L) = W(x_i)$ for $1 \leq i \leq m$ (using the notations of Lemmas 3.9 and 3.16)

Step 2: *Initial contraction.*

For $1 \leq i \leq m$, let $\lambda_i = \frac{W(r_i)}{W(x_i)} = \frac{W(r_i)}{W((r_i)_L)}$. Note that $\lambda_i \geq 1$ by Lemma 3.9(c). Define $X_{bad} = \{x_1, \dots, x_m\}$, $X_{good} = X \setminus X_{bad}$ and $R_{good} = R \setminus R_{bad}$. We have

$$1 - W(X_{bad}) - W(X_{good}) + W(R_{bad}) + W(R_{good}) < -M.$$

Since $W(r_i) \geq W(x_i)$, we have $W(R_{bad}) \geq W(X_{bad})$, whence $W(X_{good}) > M$.

Let $c = W(X_{good})/w$. Note that $c > \frac{M}{w}$.

Define W_a to be the c -contraction of W on $(F(X), X)$. Note that

$$W_a(X_{good}) = \frac{W(X_{good})}{c} = w \quad \text{and} \quad W_a(X) = w + \frac{W(X_{bad})}{c} \geq w.$$

Now divide R_{bad} into two disjoint subsets:

$$R_{bad}^1 = \{r_i \in R_{bad} : \lambda_i \leq c\} \text{ and } R_{bad}^2 = R_{bad} \setminus R_{bad}^1,$$

and let X_{bad}^j be the subset of X_{bad} corresponding to R_{bad}^j for $j = 1, 2$, that is, $X_{bad}^j = \{x_i \in X_{bad} : r_i \in R_{bad}^j\}$.

Claim 3.17. *The following hold:*

- (1) If $r \in R_{good} \sqcup R_{bad}^2$, then $W_a(r) \leq W(r)/c^2$
- (2) If $r_i \in R_{bad}^1$, then $W_a(r_i) = W_a(x_i)$.

Proof. (1) holds by Lemma 3.16(b)(d), and (2) follows directly from condition (iii) above and Lemma 3.16(c)(e). \square

Step 3: *Cleanup and the second contraction.*

Now let $\tilde{X} = X \setminus X_{bad}^1 \cup R_{bad}^1$. We claim that \tilde{X} is a free generating set of $F = F(X)$. Indeed, condition (i) ensures that the matrix $A \in GL_m(\mathbb{F}_p)$ expressing \tilde{X} in terms of X mod $\Phi(F)$ is upper-triangular with nonzero diagonal entries, hence invertible. Thus, \tilde{X} generates F modulo $\Phi(F)$ and hence generates F . Finally, \tilde{X} is a free generating set of F since $|\tilde{X}| = |X|$.

Since $W_a(r_i) = W_a(x_i)$ for every $r_i \in R_{bad}^1$, the change of generators $(X, R) \rightarrow (\tilde{X}, R)$ is W_a -good by Lemma 3.5(b)(ii). We now apply this change of generators followed by the cleanup of the set R_{bad}^1 , and let (X', R') be the obtained presentation (note that $X' = X \setminus X_{bad}^1$).

By Lemma 3.10(d) we have $W_a(X') = W_a(X) - W_a(X_{bad}^1)$ and $W_a(R') \leq W_a(R) - W_a(R_{bad}^1)$. We shall now estimate the quotient $W_a(R')/W_a(X')$. Using Claim 3.17, we have

$$\begin{aligned} \frac{W_a(R')}{W_a(X')} &\leq \frac{W_a(R_{good}) + W_a(R_{bad}^2)}{W_a(X_{good}) + W_a(X_{bad}^2)} \leq \frac{1}{c} \cdot \frac{W(R_{good}) + W(R_{bad}^2)}{W(X_{good}) + W(X_{bad}^2)} = \\ &\quad \frac{1}{c} \cdot \frac{W(R) - W(R_{bad}^1)}{W(X) - W(X_{bad}^1)} \leq \frac{1}{c} \cdot \frac{W(R)}{W(X)} < \frac{1}{c}, \end{aligned}$$

where the last two steps hold since $W(R) < W(X)$ and $W(R_{bad}^1) \geq W(X_{bad}^1)$.

Now let $c_1 = W_a(X')/w$. Note that $c_1 \geq 1$ since $W_a(X') \geq W_a(X_{good}) = w$. Let W' be the weight function on $(F(X'), X')$ obtained by the c_1 -contraction of W . Then we have $W'(X') = W_a(X')/c_1 = w$ and $W'(R') \leq W_a(R')/c_1$, so

$$W'(R') \leq W'(X') \cdot \frac{W_a(R')}{W_a(X')} < \frac{w}{c} < \frac{w^2}{M} < \delta.$$

Finally, for any $x \in X'$ we have $W'(x) \leq W_a(x) \leq \frac{W(x)}{c} < \frac{1}{c} \leq \frac{w}{M} < \varepsilon$.

(b) By (a) we can find a weighted presentation (X', R', W') of K such that $W'(X') = 2w$, $W'(R') < \delta$ and $\max\{W'(x) : x \in X'\} < \frac{\varepsilon}{2}$.

Let D be the unique degree function on $(F(X'), X')$ such that $\frac{W'(x)}{2} < (\frac{1}{2})^{D(x)} \leq W'(x)$ for any $x \in X'$. Let W'' be the $(D, \frac{1}{2})$ -weight function on $(F(X'), X')$, so that $\frac{W'(x)}{2} < W''(x) \leq W'(x)$ for any $x \in X'$.

By construction $w = \frac{W'(X')}{2} < W''(X')$ and $W''(R') \leq W'(R') < \delta$. Furthermore, if $N = \min\{D(x) : x \in X'\}$, then $\sum_{z \in \text{Im}(W'')} W''(z) \leq \sum_{n=N}^{\infty} (\frac{1}{2})^n = (\frac{1}{2})^{N-1}$ and

$$\left(\frac{1}{2}\right)^N = \max\{W''(x) : x \in X'\} \leq \max\{W'(x) : x \in X'\} < \frac{\varepsilon}{2},$$

so $\sum_{z \in \text{Im}(W'')} W''(z) < \varepsilon$.

Finally, let $c' = \frac{W''(X')}{w}$ and \widetilde{W} be the c' -contraction of W'' . It is clear that the weighted presentation (X', R', \widetilde{W}) satisfies **Pres**(w, δ, ε). \square

Remark: In the final step of the proof of Theorem 3.15(b) we can also take \widetilde{W} to be the (D, τ) -weight function on $(F(X'), X')$, where $\tau \in (0, \frac{1}{2})$ is such that $\widetilde{W}(X') = w$ (such τ exists by continuity). Thus, \widetilde{W} can be chosen an integral weight function.

3.6. From pro- p to abstract groups. We finish with the analogue of Theorem 3.13 dealing with abstract groups (see Proposition 3.19 below). The proofs of the following two results were suggested by Jaikin-Zapirain.

Lemma 3.18. *Let (X, R) be a presentation, $G = \text{Gr}(X, R)$ and Γ and Λ two dense countable subgroups of G . Let W be a weight function on $(F(X), X)$. Then for any $\delta_1 > 0$ there exists a subset R_{aux} of $F(X)$ such that $W(R_{aux}) < \delta_1$ and the images of Γ and Λ in $\text{Gr}(X, R \cup R_{aux})$ coincide.*

Proof. Let $\pi : F(X) \rightarrow G$ be the natural surjection. Choose countable subsets $Y = \{y_1, y_2, \dots\}$ and $Z = \{z_1, z_2, \dots\}$ of $F(X)$ such that $\pi(Y) = \Gamma$ and $\pi(Z) = \Lambda$.

Since Λ is dense in G and π maps open sets to open sets, we conclude that $\pi^{-1}(\Lambda)$ is dense in $F(X)$. Thus, for each $i \in \mathbb{N}$ we can find $y'_i \in F(X)$ such that $\pi(y'_i) \in \Lambda$ and $W(y_i^{-1}y'_i) < \frac{\delta_1}{2^{i+1}}$. Similarly, for each $i \in \mathbb{N}$ we can find $z'_i \in F(X)$ such that $\pi(z'_i) \in \Gamma$ and $W(z_i^{-1}z'_i) < \frac{\delta_1}{2^{i+1}}$.

Let $R_{aux} = \{y_i^{-1}y'_i, z_i^{-1}z'_i : i \in \mathbb{N}\}$; clearly $W(R_{aux}) < \delta_1$. If Γ' (resp. Δ') is the image of Γ (resp. Λ) in $\text{Gr}(X, R \cup R_{aux})$, then by construction $\Gamma' \subseteq \Delta'$ and $\Delta' \subseteq \Gamma'$, so $\Gamma' = \Delta'$. \square

Notation: If (X, R) is a presentation, by $\text{Gr}_{abs}(X, R)$ we will denote the abstract group generated by (the image of) X in $\text{Gr}(X, R)$.

Remark: If R happens to be a subset of $F_{abs}(X)$ (the free abstract group on X), the abstract group $\langle X|R \rangle$ need not coincide with $\text{Gr}_{abs}(X, R)$; what is true is that $\text{Gr}_{abs}(X, R)$ is isomorphic to the image of $\langle X|R \rangle$ in its pro- p completion.

Proposition 3.19. *Let G be a GGS pro- p group, and let w, δ and ε be positive real numbers. Let Γ be a dense finitely generated subgroup of G . Then there exists a weighted presentation (X, R, W) such that*

- (i) (X, R, W) satisfies the condition **Pres**(w, δ, ε).
- (ii) The abstract group $\text{Gr}_{abs}(X, R)$ is a quotient of a finite index subgroup of Γ .

Proof. By Theorem 3.13 we can find an open subgroup \tilde{H} of G , a presentation (X, \tilde{R}) of \tilde{H} and a weight function W on $(F(X), X)$ such that (X, \tilde{R}, W) satisfies **Pres**(w, δ, ε). Note that $\tilde{H} \cap \Gamma$ and $\text{Gr}_{abs}(X, \tilde{R})$ are both dense countable subgroups of $\text{Gr}(X, \tilde{R})$.

By Lemma 3.18, given $\delta_1 > 0$ we can find $R_{aux} \subset F(X)$ with $W(R_{aux}) < \delta_1$ such that if $R = \tilde{R} \cup R_{aux}$, then the images of $\tilde{H} \cap \Gamma$ and $\text{Gr}_{abs}(X, \tilde{R})$ in $\text{Gr}(X, R)$ coincide. By definition the image of $\text{Gr}_{abs}(X, \tilde{R})$ in $\text{Gr}(X, R)$ is $\text{Gr}_{abs}(X, R)$, so (ii) holds. Since $W(R) < W(\tilde{R}) + \delta_1$, the weighted presentation (X, R, W) satisfies **Pres**(w, δ, ε) for sufficiently small δ_1 , so (i) also holds. \square

Remark: The reader might wonder how the class

$$\mathfrak{GGS} = \{\text{GGS abstract groups}\}$$

$$\mathfrak{GGS}' = \{\text{dense finitely generated subgroups of GGS pro-}p\text{ groups}\}$$

to which Proposition 3.19 applies. Clearly, any group in \mathfrak{GGS}' is residually- p , which is not necessarily true for groups in \mathfrak{GGS} ; however, for any $\Gamma \in \mathfrak{GGS}$ the image of Γ in its pro- p completion belongs to \mathfrak{GGS}' . Thus, to prove Theorem 1.1 it will suffice to prove the corresponding result for groups in \mathfrak{GGS}' , and in fact this is what we will do (see Theorem 4.6). On the other hand, we do not know whether any group in \mathfrak{GGS}' also belongs to \mathfrak{GGS} .

4. CONSTRUCTING KAZHDAN QUOTIENTS

In this section we prove the main result of this paper (Theorem 1.1). The first three subsections will be devoted to the proof of Theorem 1.1 “up to finite index”:

Theorem 4.1. *Let Γ be a Golod-Shafarevich group. Then Γ has a subgroup of finite index which surjects onto an infinite group with property (T).*

In the last subsection we will deduce Theorem 1.1 from Theorem 4.1 using a very general argument.

4.1. Golod-Shafarevich groups with property (T). In [EJ], a large class of groups, called *Kac-Moody-Steinberg groups*, was introduced, and many of these groups were shown to have property (T). We shall not recall the general construction, but concentrate on its special case that will be used in this paper.

Let \mathbb{F} be a finite field, let n_1, \dots, n_k be a collection of positive integers, and for each $1 \leq i \leq k$ let $M_i = \mathbb{F}^{n_i}$ (the standard n_i -dimensional \mathbb{F} -vector space). For $1 \leq i \leq k$ let \mathcal{U}_i be the group whose elements are formal symbols $\{x_i(a) : a \in M_i\}$ subject to relations

$$x_i(a) \cdot x_i(b) = x_i(a + b),$$

so that $\mathcal{U}_i \cong (M_i, +)$. Define $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ to be the group generated by $\{\mathcal{U}_1, \dots, \mathcal{U}_k\}$ subject to the following relations:

$$(R1) \ [x_i(a), x_j(b), x_i(c)] = 1 \text{ for } 1 \leq i \neq j \leq k, \ a, c \in M_i \text{ and } b \in M_j$$

$$(R2) \ [x_i(\lambda a), x_j(b)] = [x_i(a), x_j(\lambda b)] \text{ for } 1 \leq i \neq j \leq k, \ a \in M_i, \ b \in M_j \text{ and } \lambda \in \mathbb{F}.$$

In other words, $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ is the largest quotient of the free product $\mathcal{U}_1 * \dots * \mathcal{U}_k$ in which each pair $\{\mathcal{U}_i, \mathcal{U}_j\}$ generates a nilpotent subgroup of class two and the commutator map $\mathcal{U}_i \times \mathcal{U}_j \rightarrow [\mathcal{U}_i, \mathcal{U}_j]$ is \mathbb{F} -linear. Note that relations (R2) hold automatically if $|\mathbb{F}|$ is prime.

The group $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ will be called a *Kac-Moody-Steinberg group* and its subgroups $\mathcal{U}_1, \dots, \mathcal{U}_k$ will be called *the root subgroups* (for the explanation of this terminology see [EJ, Section 7]).

The following result is a special case of [EJ, Corollary 7.2]:

Theorem 4.2. *The group $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ has property (T) whenever $|\mathbb{F}| > (k - 1)^2$.*

It is easy to see that “most” of the groups $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ are Golod-Shafarevich if $k \geq 9$ and $|\mathbb{F}|$ is prime. Thus, for $p \geq 67$, Theorem 4.2 produces a large supply of Golod-Shafarevich groups with property (T) which provides a starting point for the proof of Theorem 4.1. However, if $|\mathbb{F}|$ is not prime, the group $\text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ will never be Golod-Shafarevich because of relations (R2). In this case, instead of working with the group $\Gamma = \text{KMS}(\mathbb{F}; n_1, \dots, n_k)$ we will consider the completed group algebra $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$ and show that $\mathbb{F}_p[[\Gamma_{\hat{p}}]]$ maps onto a finite codimension subalgebra of an \mathbb{F} -algebra satisfying the GS condition.

In the next two subsections we shall prove the following theorem.

Theorem 4.3. *There exist positive real numbers w, δ and ε with the following property: if (X, R, W) is a triple satisfying the condition $\mathbf{Pres}(w, \delta, \varepsilon)$, then the group $\text{Gr}_{\text{abs}}(X, R)$ has an infinite quotient with property (T). In fact, one can take $w = \frac{3}{2}$, $\delta = \frac{1}{50}$ and $\varepsilon = \frac{1}{100}$ when $p \geq 67$, and $w = 12$, $\delta = \frac{1}{100}$ and $\varepsilon = \frac{1}{1000}$ for arbitrary p .*

It is clear that Theorem 4.1 follows from Theorem 4.3 and Proposition 3.19.

Remark: Suppose K is a pro- p group with a weighted presentation (X', R', W') such that $1 - W'(X') + W'(R') < -6 \cdot 10^5$. Then by Theorem 3.15(b) K also has a presentation (X, R, W) satisfying $\mathbf{Pres}(w, \delta, \varepsilon)$ where w, δ and ε are as in Theorem 4.3. This observation justifies Step 2 of the algorithm presented in the introduction.

4.2. Proof of Theorem 4.3 for $p \geq 67$. Fix arbitrary positive real numbers w, δ and ε , and let (X, R, W) be a triple satisfying $\mathbf{Pres}(w, \delta, \varepsilon)$. In fact, in our case $p \geq 67$ we will not use the full power of the assumption (b) in $\mathbf{Pres}(w, \delta, \varepsilon)$ – it will be enough to assume that $\max\{W(x) : x \in X\} < \varepsilon$.

Divide X into 9 (disjoint) subsets X^1, \dots, X^9 such that $|W(X^i) - W(X^j)| < \varepsilon$ for each i, j . For each $1 \leq i \leq 9$ we will denote the elements of X^i by $x_{i,1}, \dots, x_{i,n_i}$. We can now identify X with a generating set for the group $\mathrm{KMS}(\mathbb{F}_p; n_1, \dots, n_9)$, which is given by the abstract presentation

$$\langle X \mid R_{KM} \rangle = \langle x_{i,k}, 1 \leq i \leq 9, 1 \leq k \leq n_i \mid x_{i,k}^p, [x_{i,k}, x_{i,l}], [x_{i,k}, x_{j,l}, x_{i,m}] \rangle.$$

Since $p > (9 - 1)^2$, the group $\mathrm{KMS}(\mathbb{F}_p; n_1, \dots, n_9) = \langle X \mid R_{KM} \rangle$ has property (T) by Theorem 4.2. Hence the group $\mathrm{Gr}_{abs}(X, R_{KM})$ also has (T) being a quotient of $\langle X \mid R_{KM} \rangle$.³

Now consider the presentation $(X, R \cup R_{KM})$. For $1 \leq i \leq 9$ let $w_i = W(X^i)$, so that $\sum_{i=1}^9 w_i = W(X) = w$. Note that

$$W(\{[x_{i,k}, x_{i,l}] : 1 \leq i \leq 9, 1 \leq k < l \leq n_i\}) = \sum_{i=1}^9 \sum_{1 \leq k < l \leq n_i} W(x_{i,k})W(x_{i,l}) < \frac{1}{2} \sum_{i=1}^9 \left(\sum_{k=1}^{n_i} W(x_{i,k}) \right)^2 = \frac{1}{2} \sum_{i=1}^9 w_i^2.$$

Similarly,

$$W(\{[x_{i,k}, x_{j,l}, x_{i,m}] : 1 \leq i \neq j \leq 9, 1 \leq k, m \leq n_i, 1 \leq l \leq n_j\}) \leq \sum_{1 \leq i \neq j \leq 9} w_i^2 w_j.$$

Therefore,

$$W(R_{KM}) < \sum_{i=1}^9 \frac{w_i^2}{2} + \sum_{1 \leq i \neq j \leq 9} (w_i^2 w_j) + \varepsilon^{p-1} w.$$

(since $\max\{W(x) : x \in X\} < \varepsilon$). Recall that we have defined subsets X^1, \dots, X^9 so that $|w_i - w_j| < \varepsilon$ for each i, j , whence $|w_i - \frac{w}{9}| < \varepsilon$ for each i . An easy computation shows that

$$\sum_{i=1}^9 w_i^2 \leq \frac{w^2}{9} + 9\varepsilon^2 \text{ and } \sum_{1 \leq i \neq j \leq 9} (w_i^2 w_j) \leq \frac{8}{81} w^3 + 9w\varepsilon^2.$$

Therefore

$$1 - W(X) + W(R \cup R_{KM}) \leq 1 - w + \frac{w^2}{18} + \frac{8w^3}{81} + 9(w + \frac{1}{2})\varepsilon^2 + w\varepsilon^{p-1} + \delta.$$

Setting $w = \frac{3}{2}$, $\delta = \frac{1}{50}$ and $\varepsilon = \frac{1}{100}$, we get $1 - W(X) + W(R \cup R_{KM}) \leq 1 - \frac{3}{2} + \frac{1}{8} + \frac{1}{3} + \frac{36}{10^4} + \frac{3}{200} + \frac{1}{50} < 0$. Thus, we proved the following claim:

³Recall that by definition $\mathrm{Gr}_{abs}(X, R_{KM})$ is the image of $\langle X \mid R_{KM} \rangle$ in its pro- p completion. We do not know whether the group $\langle X \mid R_{KM} \rangle$ is always residually- p (which is equivalent to the injectivity of the map $\langle X \mid R_{KM} \rangle \rightarrow \mathrm{Gr}_{abs}(X, R_{KM})$).

Claim 4.4. Assume that $w = \frac{3}{2}$, $\delta = \frac{1}{50}$ and $\varepsilon = \frac{1}{100}$. Then

$$(4.1) \quad 1 - W(X) + W(R \cup R_{KM}) < 0.$$

We are ready to finish the proof of Theorem 4.3 (so far for $p \geq 67$). By Claim 4.4 the presentation $(X, R \cup R_{KM})$ satisfies the GGS condition, hence the group $\text{Gr}_{abs}(X, R \cup R_{KM})$ is infinite. By definition, $\text{Gr}_{abs}(X, R \cup R_{KM})$ is a quotient of $\text{Gr}_{abs}(X, R)$. On the other hand, $\text{Gr}_{abs}(X, R \cup R_{KM})$ is also a quotient of $\text{Gr}_{abs}(X, R_{KM})$ and hence has property (T).

4.3. Proof of Theorem 4.3 in the general case. When p is small, we do not have any examples of groups of the form $\text{KMS}(\mathbb{F}_p; n_1, \dots, n_k)$ which are both Kazhdan and Golod-Shafarevich. Instead, we will consider the groups $\text{KMS}(\mathbb{F}_q; n_1, \dots, n_9)$ with $q = p^8$, which have property (T) by Theorem 4.2 since $p^8 > 8^2$. Let $\mathcal{B} = \{\lambda_1 = 1, \dots, \lambda_8\}$ be a basis for \mathbb{F}_q over \mathbb{F}_p . For any $n_1, \dots, n_9 \in \mathbb{N}$ the group $\text{KMS}(\mathbb{F}_q; n_1, \dots, n_k)$ is given by the abstract presentation $\langle X_{KM} | R_{KM} \rangle$ where

$$(4.2) \quad X_{KM} = \{x_{i,k}(\lambda_j), 1 \leq i \leq 9, 1 \leq k \leq n_i, 1 \leq j \leq 8\} \text{ and}$$

$$(4.3) \quad R_{KM} = R_{nilp} \cup R_{field} \text{ where}$$

$$(4.4) \quad R_{nilp} = \{x_{i,k}(\lambda)^p, [x_{i,k}(\lambda), x_{i,l}(\mu)], [x_{i,k}(\lambda), x_{j,l}(\mu), x_{i,m}(\nu)]\},$$

$$(4.5) \quad R_{field} = \{[x_{i,k}(\lambda), x_{j,l}(\mu)][x_{i,k}(\lambda\mu), x_{j,l}(1)]^{-1}\},$$

and λ, μ, ν are arbitrary elements of the basis \mathcal{B} .

Here $x_{i,k}(\alpha)$ for $\alpha \in \mathbb{F}_q \setminus \mathcal{B}$ is defined in the obvious way: write $\alpha = \sum_{j=1}^8 c_j \lambda_j$ with $0 \leq c_j \leq p-1$ and set $x_{i,k}(\alpha) = \prod_{j=1}^8 x_{i,k}(\lambda_j)^{c_j}$.

Proof of Theorem 4.3. Let $w = 12$, fix $\delta > 0$ and $\varepsilon > 0$ (to be specified at the very end), and let (X, R, W) be a triple satisfying **Pres**(12, $\delta/2, \varepsilon'$) with $\varepsilon' = \min\{\varepsilon, \delta/14\}$.

In order to find a common infinite quotient for $\text{Gr}_{abs}(X, R)$ and the above KMS groups we need to impose an additional condition on the triple (X, R, W) :

(CC) For every $z \in \mathbb{R}$, the cardinality of the set $\{x \in X : W(x) = z\}$ is divisible by 8.

This can be achieved by adding artificial generators y_1, \dots, y_m and relations $y_1 = \dots = y_m = 1$ (that is, by replacing the presentation (X, R) with $(X \sqcup Y, R \sqcup Y)$ for some finite set Y), and extending W to a weight function on $(F(X \cup Y), X \cup Y)$ by assigning appropriate weights to elements of Y . Clearly, for any $z \in \text{Im}(W)$ we need to add at most 7 generators y with $W(y) = z$ to achieve (CC). Thus, we can assume that $12 \leq W(X \cup Y) < 12 + 7s$ and $W(R \cup Y) \leq W(R) + 7s$ where $s = \sum_{z \in \text{Im}(W)} z$. By assumption $W(R) < \delta/2$ and $s < \varepsilon' \leq \delta/14$, whence $W(R) + 7s < \delta$. Now contract the weight function W on $(F(X \cup Y), X \cup Y)$ by the factor $\frac{W(X \cup Y)}{12} \geq 1$. After redenoting $X \cup Y$ by X , $R \cup Y$ by R and the contracted weight function by W we get that (X, R, W) satisfies **Pres**(12, δ, ε) and condition (CC) holds.

Since $\max\{W(x) : x \in X\} < \varepsilon$, we can divide the set X into 9 subsets X^1, \dots, X^9 , such that

- (a) $|W(X^i) - W(X^j)| < 8\varepsilon$ for all $1 \leq i, j \leq 9$
- (b) For any $1 \leq i \leq 9$ and $z \in \mathbb{R}$, the cardinality of the set $\{x \in X^i : W(x) = z\}$ is divisible by 8.

For each $1 \leq i \leq 9$ let $n_i = |X^i|/8$ (this is surely an integer by (b)), and denote elements of X^i by symbols $\{x_{i,k}(\lambda_j)\}$ with $1 \leq k \leq n_i$ and $1 \leq \lambda_j \leq 8$, such that for each i, k the weight $W(x_{i,k}(\lambda_j))$ is the same for all j . Now we can identify X with the generating

set X_{KM} from the presentation for $\text{KMS}(\mathbb{F}_q; n_1, \dots, n_9)$ given by (4.2)-(4.5), and consider the presentation $(X, R_{KM} \cup R)$. If we show that the pro- p group $\text{Gr}(X, R_{KM} \cup R)$ is infinite, we will be done by the same argument as in the case $p \geq 67$. Note that unlike the latter case the group $G = \text{Gr}(X, R_{KM} \cup R)$ is not GGS because of the relations R_{field} . To prove that it is infinite, we will map the completed group algebra $\mathbb{F}_p[[G]]$ onto a finite codimension subalgebra of certain \mathbb{F}_q -algebra, which, in turn, will be Golod-Shafarevich (as \mathbb{F}_q -algebra) and hence infinite.

Let $U = \{u_{i,k}(\lambda_j) : 1 \leq i \leq 9, 1 \leq k \leq n_i, 1 \leq j \leq 8\}$, and consider the standard embedding of $F(X)$ into $\mathbb{F}_p\langle\langle U \rangle\rangle$ given by $x_{i,k}(\lambda_j) \mapsto 1 + u_{i,k}(\lambda_j)$. Let $S = \{r - 1 : r \in R_{KM} \cup R\}$ and I the ideal of $\mathbb{F}_p\langle\langle U \rangle\rangle$ generated by S . Recall from Section 2 that the algebra $\mathbb{F}_p[[G]]$ is isomorphic to $\mathbb{F}_p\langle\langle U \rangle\rangle/I$.

Now let \tilde{U} be the set of formal symbols $\{u_{i,k} : 1 \leq i \leq 9, 1 \leq k \leq n_i\}$, and consider the unique continuous \mathbb{F}_p -algebra homomorphism $\varphi : \mathbb{F}_p\langle\langle U \rangle\rangle \rightarrow \mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle$ given by

$$\varphi(u_{i,k}(\lambda_j)) = \lambda_j u_{i,k}.$$

Clearly, $\varphi(\mathbb{F}_p\langle\langle U \rangle\rangle)$ contains the ideal of $\mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle$ generated by \tilde{U} , so $\text{Im } \varphi$ is of finite codimension. Let $S_1 = \{\varphi(r - 1) : r \in R\}$, $S_2 = \{\varphi(r - 1) : r \in R_{nilp}\}$, and $S_3 = \{a^2b, aba, ba^2 : a, b \in \tilde{U}\}$. Let \tilde{I} be the ideal of $\mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle$ generated by $S_1 \cup S_2 \cup S_3$. We will prove that

- (a) $\varphi(S) \subseteq \tilde{I}$, whence there is a homomorphism $\mathbb{F}_p\langle\langle U \rangle\rangle/I \rightarrow \mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle/\tilde{I}$ with image of finite codimension;
- (b) The algebra $\mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle/\tilde{I}$ satisfies the GGS condition and hence it is infinite.

By earlier discussion, (a) and (b) would imply Theorem 4.1.

For (a), we only need to show that $\varphi(r - 1) \in I$ for any $r \in R_{field}$ (for $r \in R_{nilp} \cup R$ we have $\varphi(r - 1) \in I$ by construction). Any $r \in R_{field}$ is of the form $r = [x_{i,k}(\lambda), x_{j,l}(\mu)][x_{i,k}(\lambda\mu), x_{j,l}(1)]^{-1}$. Let $a = \varphi(u_{i,k})$ and $b = \varphi(u_{j,l})$. Then

$$\begin{aligned} \varphi([x_{i,k}(\lambda), x_{j,l}(\mu)]) &= [1 + \lambda a, 1 + \mu b] = (1 + \lambda a)^{-1}(1 + \mu b)^{-1}(1 + \lambda a)(1 + \mu b) = \\ &= 1 + (1 + \lambda a)^{-1}(1 + \mu b)^{-1}\lambda\mu(ab - ba) \equiv 1 + \lambda\mu(ab - ba) \pmod{\tilde{I}}. \end{aligned}$$

Similarly, we have $\varphi([x_{i,k}(\lambda\mu), x_{j,l}(1)]) \equiv 1 + \lambda\mu(ab - ba) \pmod{\tilde{I}}$, whence

$$\varphi(r - 1) = \varphi([x_{i,k}(\lambda), x_{j,l}(\mu)] - [x_{i,k}(\lambda\mu), x_{j,l}(1)])\varphi([x_{i,k}(\lambda\mu), x_{j,l}(1)])^{-1} \in \tilde{I}.$$

Thus, we proved (a).

Now let w be the weight function on $\mathbb{F}_p\langle\langle U \rangle\rangle$ which is X -compatible with the weight function W on $(F(X), X)$. Let \tilde{w} be the unique weight function on $\mathbb{F}_q\langle\langle \tilde{U} \rangle\rangle$ such that $\tilde{w}(\varphi(u)) = w(u)$ for any $u \in U$ – note that such \tilde{w} exists since $w(u_{i,k}(\lambda_j)) = W(x_{i,k}(\lambda_j))$ is independent of j .

By an obvious analogue of Lemma 2.4 for algebras, to prove (b) it suffices to show that

$$1 - \tilde{w}(\tilde{U}) + \tilde{w}(S_1) + \tilde{w}(S_2) + \tilde{w}(S_3) < 0.$$

Arguing as in the proof of Lemma 3.5(a), we conclude that $\tilde{w}(\varphi(f)) \leq w(f)$ for any $f \in \mathbb{F}_p\langle\langle U \rangle\rangle$. In particular, $\tilde{w}(S_1) \leq w(\{r - 1 : r \in R\}) = W(R) < \delta$.

For $1 \leq i \leq 9$ let $U^i = \{u_{i,k} : 1 \leq k \leq n_i\}$ and $\tilde{w}_i = \tilde{w}(U^i)$, and set $\tilde{w} = \sum \tilde{w}_i$. By construction, $\tilde{w} = \tilde{w}(\tilde{U}) = w(U)/8 = 3/2$. Also note that $|\tilde{w}_i - \tilde{w}_j| = \frac{|W(X^i) - W(X^j)|}{8} < \frac{8\varepsilon}{8} = \varepsilon$ for any i, j , and thus $|\tilde{w}_i - \frac{\tilde{w}}{9}| < \varepsilon$.

Since $\max\{w(u) : u \in U\} = \max\{W(X) : x \in X\} < \varepsilon$, we have $\tilde{w}(S_3) \leq 3\varepsilon\tilde{w}^2$, and $\tilde{w}(S_2)$ can be estimated as in the case $p \geq 67$. We get

$$1 - \tilde{w}(\tilde{U}) + \tilde{w}(S_1) + \tilde{w}(S_2) + \tilde{w}(S_3) \leq 1 - \tilde{w} + \sum_{i=1}^9 \tilde{w}_i^2/2 + \sum_{i \neq j} \tilde{w}_i^2 \tilde{w}_j + \varepsilon^{p-1} \tilde{w} + 3\varepsilon \tilde{w}^2 + \delta.$$

Using the same estimates as in the case $p \geq 67$, we get

$$1 - \tilde{w}(\tilde{U}) + \tilde{w}(S_1) + \tilde{w}(S_2) + \tilde{w}(S_3) \leq 1 - \tilde{w} + \tilde{w}^2/18 + 8\tilde{w}^3/81 + 9(\tilde{w} + \frac{1}{2})\varepsilon^2 + (\tilde{w} + 3\tilde{w}^2)\varepsilon + \delta < -\frac{1}{24} + 36\varepsilon^2 + 9\varepsilon + \delta.$$

Setting $\delta = \frac{1}{50}$ and $\varepsilon = \frac{1}{1000}$, we get $1 - \tilde{w}(\tilde{U}) + \tilde{w}(S_1) + \tilde{w}(S_2) + \tilde{w}(S_3) < 0$, which finishes the proof. \square

4.4. Proof of Theorem 1.1. In this subsection we establish the following result due to Jaikin-Zapirain:

Proposition 4.5. *Let Γ be a finitely generated group and H a finite index subgroup Γ . If H maps onto an infinite group with property (T), then Γ also maps onto an infinite group with property (T).*

Proof. If H maps onto an infinite group with property (T), then so does any finite index subgroup of H . Thus, we may assume that H is normal in Γ . Let N be a normal subgroup of H such that H/N has property (T). By [Gr2], any finitely generated abstract group has a just-infinite quotient. Thus, replacing N by a larger normal subgroup (if necessary), we may assume that H/N is just-infinite.

Let S_0 be a transversal of H in Γ , and set $L = \bigcap_{s \in S_0} N^s$. Then L is a normal subgroup of Γ . We shall show that Γ/L has property (T). Since H/L is of finite index in Γ/L , it is enough to show that H/L has property (T).

Note that H/L naturally embeds in $\prod_{s \in S_0} H/N^s$. Choose a subset $S \subseteq S_0$ such that the composite map

$$\pi : H/L \rightarrow \prod_{s \in S_0} H/N^s \rightarrow \prod_{s \in S} H/N^s$$

is injective and S is minimal with this property. For each $s \in S$ let $H_s = \pi(H/L) \cap (H/N^s)$. Since $\pi(H/L)$ surjects onto each factor in $\prod_{s \in S} H/N^s$, the group H_s is normal in H/N^s . Since $H/N^s \cong H/N$ is just-infinite, H_s is either trivial or of finite index in H/N^s . But if H_a is trivial for some $a \in S$, then H/L injects in $\prod_{s \in S \setminus \{a\}} H/N^s$, contrary to the minimality of S . Thus, H_s is of finite index in H/N^s for each $s \in S$. Therefore, $\pi(H/L) \supseteq \prod H_s$ is of finite index in $\prod_{s \in S_0} H/N^s$.

Since H/N has property (T), so do $\prod_{s \in S_0} H/N^s$ and its finite index subgroups. Therefore, $H/L \cong \pi(H/L)$ also has property (T). \square

We completed the proof of Theorem 1.1 as it clearly follows from Theorem 4.1 and Proposition 4.5. Recall that Theorem 4.1 was, in turn, a consequence of Proposition 3.19 and Theorem 4.3. Thus, we actually established the following generalization of Theorem 1.1:

Theorem 4.6. *Let G be a generalized Golod-Shafarevich pro- p group and Γ a dense countable subgroup of G . Then Γ has an infinite quotient with property (T).*

It is also natural to ask for a generalization (or rather strengthening) of Theorem 1.1 of a different kind:

Question 4.7. *Let Γ be a GGS group with respect to some prime p . Does Γ always have an infinite Kazhdan quotient which is also a GGS group?*

Note that for $p \geq 67$ the existence of such quotient for some finite index subgroup of Γ follows directly from the proof of Theorem 4.1. Unfortunately, the GGS condition is likely not satisfied by the Kazhdan quotient for the entire Γ constructed in the proof of Theorem 4.5.

APPENDIX A.

UNIFORM NON-AMENABILITY OF GOLOD-SHAFAREVICH GROUPS

by MIKHAIL ERSHOV and ANDREI JAIKIN-ZAPIRAIN ⁴

The main goal of this section is to prove that generalized Golod-Shafarevich groups satisfy a strong form of non-amenability, called uniform non-amenability.

Notation. In the previous sections by the standard abuse of notation we identified the generating set X of a (pro- p) presentation (X, R) with its (canonical) image in the pro- p group $\text{Gr}(X, R)$ or abstract group $\text{Gr}_{\text{abs}}(X, R)$. In this section such abuse of notation could lead to confusion. For this reason, if (X, R) is a presentation and $\Gamma = \text{Gr}_{\text{abs}}(X, R)$, we will denote the image of X in Γ by X_Γ .

A.1. Basic definitions.

Definition. Let Γ be a finitely generated (abstract) group. Given a finite generating set S of Γ , the *Følner constant* ⁵ $h(\Gamma, S)$ is defined by

$$h(\Gamma, S) = \inf_{A \subset \text{Cay}(\Gamma, S)} \frac{|\partial A|}{|A|}$$

where infimum is taken over all finite subsets A of the Cayley graph of Γ with respect to S and

$$\partial A = \{g \in A : gs \notin A \text{ for some } s \in S \cup S^{-1}\}.$$

The group Γ is called *amenable* if $h(\Gamma, S) = 0$ for some (hence any) finite generating set S of Γ .

Definition. Let Γ be a finitely generated group.

- (a) Given a unitary representation V of Γ and a generating set S of Γ , we define $\kappa(\Gamma, S; V)$ to be the largest $\varepsilon \geq 0$ such that for any $v \in V$ there exists $s \in S$ with $\|sv - v\| \geq \varepsilon\|v\|$.
- (b) Given a generating set S of Γ , the *Kazhdan constant* $\kappa(\Gamma, S)$ is defined to be the infimum of the set $\{\kappa(\Gamma, S; V)\}$ where V runs over all unitary representations of Γ without nonzero invariant vectors.
- (c) The group Γ is called a *Kazhdan group* (equivalently Γ is said to have *Kazhdan's property (T)*) if $\kappa(\Gamma, S) > 0$ for some (hence any) finite generating set S of Γ .

Recall that to deduce Corollary 1.2 from Theorem 1.1 we used the fact that a group which is both amenable and Kazhdan must be finite. This well-known fact follows, for instance, from the following characterization of amenability:

⁴The author is supported by the Spanish Ministry of Science and Innovation, grant MTM2008-06680.

⁵We thank Goulmira Arzhantseva for correcting our terminology here

Theorem A.1 ([Hu]). *Let Γ be a group generated by a finite set S . Then Γ is amenable if and only if $\kappa(\Gamma, S; L^2(\Gamma)) = 0$.*

The constant $\kappa(\Gamma, S; L^2(\Gamma))$, which appears frequently in subsequent discussion, will be denoted by $\alpha(\Gamma, S)$ and called *Kazhdan L^2 -constant*. Note that if Γ is infinite, $L^2(\Gamma)$ has no nonzero invariant vectors whence

$$\alpha(\Gamma, S) \geq \kappa(\Gamma, S).$$

We finish this subsection with several inequalities involving Følner and Kazhdan constants.

Lemma A.2. *Let Γ be a group generated by a finite set S . If $\pi : \Gamma \rightarrow \Gamma'$ is a surjective homomorphism and $S' = \pi(S)$, then*

- (a) $h(\Gamma, S) \geq h(\Gamma', S')$
- (b) $\alpha(\Gamma, S) \geq \alpha(\Gamma', S')$ and $\kappa(\Gamma, S) \leq \kappa(\Gamma', S')$

Proof. (a) appears as Theorem 4.1 in [A+] and (b) is Lemma 3.4 in [Os2]. □

Definition ([Os3]). Let Γ be a group generated by a finite set S . Given a finite subset Y of Γ , define $\text{depth}_S(Y)$ to be the minimal $L \in \mathbb{Z}_{\geq 0}$ such that every element of Y can be expressed by a word of length $\leq L$ in $S \cup S^{-1}$.

Lemma A.3. *Let Γ be a group generated by a finite set S and Δ a subgroup of Γ generated by a finite set Y . Then for any unitary representation V of Γ we have*

$$\kappa(\Gamma, S; V) \geq \frac{\kappa(\Delta, Y; V)}{\text{depth}_S(Y)}$$

Proof. This is very well known, but since the proof is very short we give it here. Let $L = \text{depth}_S(Y)$ and take any $\kappa > \kappa(\Gamma, S; V)$. Then there is $v \in V$ with $\|sv - v\| < \kappa\|v\|$ for all $s \in S \cup S^{-1}$ (since $\|sv - v\| = \|s^{-1}v - v\|$ by unitarity). Now given $y \in Y$, write $y = s_1 \dots s_n$ with $s_i \in S \cup S^{-1}$ and $n \leq L$. We have

$$\begin{aligned} \|s_1 \dots s_n v - v\| &\leq \|s_1 \dots s_n v - s_1 \dots s_{n-1} v\| + \dots + \|s_1 s_2 v - s_1 v\| + \|s_1 v - v\| \\ &= \sum_{i=1}^n \|s_i v - v\| < L\kappa\|v\|, \end{aligned}$$

and thus $\kappa(\Delta, Y; V) < L\kappa$. □

Lemma A.4. *Let Γ be a group generated by a finite set S and Δ a subgroup of Γ generated by a finite set Y . Let $L = \text{depth}_S(Y)$. Then*

$$(a) \ h(\Gamma, S) \geq \frac{1}{|Y|L+1} h(\Delta, Y) \quad (b) \ \alpha(\Gamma, S) \geq \frac{\alpha(\Delta, Y)}{\sqrt{|Y|}L}$$

Proof. (a) appears as Theorem 7.1 in [A+], and (b) is the corrected version of Lemma 2.9 in [Os3]. For completeness, we shall sketch the proof of (b). We start with a general claim:

Claim A.5. *Let $\{V_i\}_{i=1}^\infty$ be unitary representations of a group G generated by a finite set X and $\kappa = \inf_{i \geq 1} \{\kappa(G, X; V_i)\}$. Let $V = \widehat{\bigoplus_{i=1}^\infty V_i}$ (where $\widehat{\bigoplus}$ denotes completed direct sum). Then*

$$\kappa(G, X; V) \geq \frac{\kappa}{\sqrt{|X|}}.$$

Proof of the Claim. Take any $v \in V$ and write $v = \sum_{i=1}^{\infty} v_i$ with $v_i \in V_i$. By definition of κ there exists a function $s : \mathbb{N} \rightarrow X$ such that for each $i \in \mathbb{N}$ we have $\|s(i)v_i - v_i\| \geq \kappa\|v_i\|$. Thus we have

$$\begin{aligned} \sum_{x \in X} \|xv - v\|^2 &= \sum_{x \in X} \sum_{i=1}^{\infty} \|xv_i - v_i\|^2 \geq \\ &\sum_{i=1}^{\infty} \|s(i)v_i - v_i\|^2 \geq \sum_{i=1}^{\infty} \kappa^2 \|v_i\|^2 = \kappa^2 \|v\|^2. \end{aligned}$$

Thus, for some $x \in X$ we must have $\|xv - v\|^2 \geq \frac{\kappa^2}{|X|} \|v\|^2$. \square

We proceed with the proof Lemma A.4(b). Let $V = L^2(\Gamma)$. It is easy to see that as a Δ -module $V = \widehat{\bigoplus}_{\gamma \in \Gamma/\Delta} V_\gamma$ with each $V_\gamma \cong L^2(\Delta)$. Then $\kappa(\Delta, Y; V_\gamma) = \alpha(\Delta, Y)$ for each γ and hence $\kappa(\Delta, Y; L^2(\Gamma)) \geq \frac{\alpha(\Delta, Y)}{\sqrt{|\Gamma|}}$ by the above claim. By Lemma A.3 we have $\alpha(\Gamma, S) = \kappa(\Gamma, S; L^2(\Gamma)) \geq \frac{\kappa(\Delta, Y; L^2(\Gamma))}{L}$ which finishes the proof. \square

Remark: The above argument essentially follows the proof of Lemma 2.9 in [Os3]. In fact, [Os3, Lemma 2.9] asserts the stronger inequality $\alpha(\Gamma, S) \geq \frac{\alpha(\Delta, Y)}{L}$; however, the proof implicitly contains unjustified claim that the function $s : \mathbb{N} \rightarrow X$ defined above may be chosen constant. We note that this correction does not affect the validity of any other results in [Os3].

A.2. Uniform non-amenability. In view of Theorem A.1, given a non-amenable group Γ , one can measure the “extent” of its non-amenability using either Følner constants or Kazhdan L^2 -constants. This suggests two possible definitions of *uniformly non-amenable groups*. Define the *uniform Følner constant* $h(\Gamma)$ and the *uniform Kazhdan L^2 -constant* $\alpha(\Gamma)$ by

$$h(\Gamma) = \inf_S h(\Gamma, S) \quad \text{and} \quad \alpha(\Gamma) = \inf_S \alpha(\Gamma, S),$$

where infimum is taken over all finite generating sets of Γ .

In [Os3], a group Γ is called uniformly non-amenable if $\alpha(\Gamma) > 0$, and in [A+] a group Γ is called uniformly non-amenable if $h(\Gamma) > 0$. Lemma A.6 below shows that uniform non-amenability in the sense of [Os3] implies uniform non-amenability in the sense of [A+]. To the best of our knowledge, it is an open question whether the converse implication holds. We shall talk about uniform non-amenability in the (stronger) sense of [Os3].

Lemma A.6. (see e.g. [A+, Proposition 2.4])⁶ *Let Γ be a group. Then for any finite generating set S of Γ we have $h(\Gamma, S) \geq \frac{1}{2}\alpha(\Gamma, S)^2$. In particular, $h(\Gamma) \geq \frac{1}{2}\alpha(\Gamma)^2$.*

The first examples of non-amenable but not uniformly non-amenable groups were constructed by Osin in [Os1] – this gave an answer to a question of Shalom. In recent years many important classes of groups were shown to be uniformly non-amenable: these include finitely generated linear groups with non-abelian free subgroups [BG], (non-elementary) hyperbolic groups and free Burnside groups of sufficiently large odd exponent [Os3].

Lemmas A.4(b) and A.2(b) easily imply that a finitely generated group Γ is uniformly non-amenable whenever the following condition is satisfied: there exists a non-amenable group Λ and a fixed finite generating set T of Λ such that given any finite generating

⁶Formally, [A+, Proposition 2.4] only states the inequality between the uniform Følner and Kazhdan L^2 -constants, but the proof actually establishes the full statement of Lemma A.6

set S of Γ there exists an embedding $\iota_S : \Lambda \rightarrow \Gamma$ with $\text{depth}_S(\iota_S(T)) \leq N$ for some N independent of S . In fact, this is a typical way to prove non-uniform amenability.

Our proof of uniform non-amenability of Golod-Shafarevich groups will implicitly use a statement of this kind, but we will not be able to satisfy the above condition for a fixed pair (Λ, T) . The main ingredient in our proof will be the following quantitative version of Theorem 4.3.

Theorem A.7. *Let (X, R, W) be a weight presentation satisfying **Pres**(12, 1/100, 1/1000) and $\Gamma = \text{Gr}_{\text{abs}}(X, R)$. Then there exists an infinite group Ω with (T) and a surjective homomorphism $\pi : \Gamma \rightarrow \Omega$ such that*

$$(A.1) \quad \kappa(\Omega, \pi(X_\Gamma)) \geq \frac{1}{25|X|}.$$

In particular, $\alpha(\Gamma, X_\Gamma) \geq \frac{1}{25|X|}$.

Finally, we remark that an important consequence of uniform non-amenability is uniform exponential growth, and for many classes of groups the easiest way to establish uniform exponential growth is by proving their uniform non-amenability. However, this is not the case for the class of Golod-Shafarevich groups where uniform exponential growth has already been known – this result was pointed out by Bartholdi and Grigorchuk [BaGr], with the proof based on [Gr, Lemma 8]

A.3. Proof of uniform non-amenability for GGS groups.

Proof of Theorem A.7. First note that the desired lower bound on $\alpha(\Gamma, X_\Gamma)$ indeed follows from (A.1) since

$$\alpha(\Gamma, X_\Gamma) \geq \alpha(\Omega, \pi(X_\Gamma)) \geq \kappa(\Omega, \pi(X_\Gamma)).$$

Here the first inequality holds by Lemma A.2(b), and the second one holds since Ω is infinite.

Inequality (A.1) follows easily from the analysis of the proof of Theorem 4.3. For simplicity, we shall only discuss the case $p \geq 67$; the case $p < 67$ is similar.

Recall that any group of the form $\text{KMS}(\mathbb{F}_p, \{n_1, \dots, n_9\})$ has a generating set $S = \sqcup_{i=1}^9 S_i$ where each S_i consists of n_i pairwise-commuting elements of order p and $\langle S_i \rangle = \mathcal{U}_i$, the i^{th} root subgroup of $\text{KMS}(\mathbb{F}_p, \{n_1, \dots, n_9\})$.

In the proof of Theorem 4.3 we established the following: there exist a group $\Lambda = \text{KMS}(\mathbb{F}_p, \{n_1, \dots, n_9\})$ with $\sum n_i = |X|$, an infinite group Ω and surjective homomorphisms $\pi : \Gamma \rightarrow \Omega$ and $\theta : \Lambda \rightarrow \Omega$ such that $\pi(X_\Gamma) = \theta(S)$ where S is a generating set of Λ of the above form.

Since $p > (9-1)^2$, [EJ, Corollary 7.2] implies that Λ has property (T) and yields the following bound for the Kazhdan constant:

$$\kappa(\Lambda, \cup \mathcal{U}_i) \geq \sqrt{\frac{2}{9} \left(1 - \frac{9-1}{\sqrt{p}}\right)} > \frac{1}{25}.$$

By the assumption on S we have $\text{depth}_S(\cup \mathcal{U}_i) \leq p \cdot \max\{n_i\} < p|X|$, and thus Lemma A.3 yields

$$(A.2) \quad \kappa(\Lambda, S) \geq \frac{\kappa(\Lambda, \cup \mathcal{U}_i)}{p|X|} \geq \frac{1}{25p|X|}.$$

Since $\kappa(\Omega, \pi(X_\Gamma)) = \kappa(\Omega, \theta(S)) \geq \kappa(\Lambda, S)$ by Lemma A.2(b), the proof is complete. \square

Definition. Let G be a pro- p group and S a generating set for G . Let (X, R) be a presentation for G and $\pi : F(X) \rightarrow G$ the natural surjection. We will say that (X, R) is a *presentation for the pair* (G, S) if $\pi(X) = S$ (note that we do not require that π is injective on X).

Recall that for a finitely generated pro- p group G we denote by $\Phi(G)$ the Frattini subgroup of G .

Lemma A.8. *Let H be a finitely generated pro- p group and $S = \{s_1, \dots, s_d\}$ a generating set for H . Let (X, R) be a presentation for the pair (H, S) , and let W be a weight function on $(F(X), X)$. Then there exists an open normal subgroup K of H with the following property: if $\tilde{S} = \{\tilde{s}_1, \dots, \tilde{s}_d\}$ is a subset of H such that*

$$\tilde{s}_i \equiv s_i \pmod{K},$$

then \tilde{S} generates H and there is a presentation (X, \tilde{R}) for the pair (H, \tilde{S}) such that $W(\tilde{R}) = W(R)$.

Proof. Let $F = F(X)$ and $\pi : F \rightarrow H$ the natural surjection. Let $\delta = \min\{W(x) : x \in X\}$,

$$O = \{f \in F : W(f) < \delta\} \text{ and } K = \pi(O)$$

Clearly, O is an open normal subgroup of F , and hence K is open and normal in H . We will show that K has the required property.

We are given that $\pi(X) = S$. Let $e = |X|$ and $\{x_1, \dots, x_e\}$ the elements of X . By assumption on \tilde{S} we can construct a subset $\tilde{X} = \{\tilde{x}_1, \dots, \tilde{x}_e\} \subset F$ such that $\tilde{x}_i \equiv x_i \pmod{O}$ for $1 \leq i \leq e$ and $\pi(\tilde{X}) = \tilde{S}$.

It is clear that $O \subseteq \Phi(F)$. Thus, \tilde{X} generates F modulo $\Phi(F)$, and so \tilde{X} is a generating set for F ; furthermore, \tilde{X} is a free generating set for F since $|\tilde{X}| = |X|$. It follows that there exists an isomorphism $\theta : F \rightarrow F$ such that $\theta(x_i) = \tilde{x}_i$ for $1 \leq i \leq e$. It is easy to show that

$$(A.3) \quad W(\theta(f)) = W(f) \text{ for any } f \in F$$

(this is ensured by the definition of O and the choice of \tilde{X}).

Now consider the map $\tilde{\pi} = \pi\theta : F \rightarrow H$. Note that $\tilde{\pi}(X) = \pi(\tilde{X}) = \tilde{S}$ (so \tilde{S} generates H) and the set $\tilde{R} = \theta^{-1}(R)$ generates $\text{Ker } \tilde{\pi}$ as a (closed) normal subgroup of F . Thus, (X, \tilde{R}) is a presentation for (H, \tilde{S}) , and the equality $W(\tilde{R}) = W(R)$ holds by (A.3). \square

We are now ready to prove uniform non-amenability of GGS groups. As in Theorem 4.6, we prove a slightly more general result.

Theorem A.9. *Let G be a generalized Golod-Shafarevich pro- p group. Then any dense finitely generated subgroup of G is uniformly non-amenable.*

Proof. By Theorem 3.13 some open subgroup H of G has a weighted presentation (X, R, W) satisfying **Pres**(12, 1/100, 1/1000). Let K be an open normal subgroup of H satisfying the conclusion of Lemma A.8.

Let S be the image of X in H , so that (X, R) is a presentation for (H, S) . Let $\{s_1, \dots, s_d\}$ be the elements of S .

Now let Γ be any dense finitely generated subgroup of G and Q any finite generating set of Γ . By density of Γ there exists a subset $\tilde{S} = \{\tilde{s}_1, \dots, \tilde{s}_d\}$ of Γ such that

- (i) $\tilde{s}_i \equiv s_i \pmod{K}$ for $1 \leq i \leq d$ (so in particular $\tilde{S} \subset H$).
- (ii) $\text{depth}_Q(\tilde{S}) \leq [G : K]$.

Condition (i) and Lemma A.8 imply that there is a presentation (X, \tilde{R}) for the pair (H, \tilde{S}) with $W(\tilde{R}) = W(R)$. In particular, the triple (X, \tilde{R}, W) satisfies **Pres**(12, 1/100, 1/1000) since (X, R, W) has this property.

Let Λ be the abstract subgroup of Γ generated by \tilde{S} . By the choice of \tilde{R} we can identify Λ with the group $\text{Gr}_{abs}(X, \tilde{R})$ so that $X_\Lambda = \tilde{S}$. Then by Theorem A.7 we have

$$\alpha(\Lambda, \tilde{S}) \geq \frac{1}{25p|X|}.$$

On the other hand, by condition (ii) and Lemma A.4(b) we have

$$\alpha(\Gamma, Q) \geq \frac{\alpha(\Lambda, \tilde{S})}{[G : K]\sqrt{|\tilde{S}|}} \geq \frac{\alpha(\Lambda, \tilde{S})}{[G : K]\sqrt{|X|}}.$$

Thus, $\alpha(\Gamma, Q)$ is bounded below by $\frac{1}{25p|X|^{3/2}[G:K]}$, the quantity which does not depend on Q (in fact, it does not even depend on Γ). This inequality finishes the proof. \square

A.4. Explicit bound for uniform Kazhdan L^2 -constants in GS groups. Let (X, R) be a presentation satisfying the GS condition.⁷ Let $G = \text{Gr}(X, R)$ and Γ a dense finitely generated subgroup of G . In this subsection we shall obtain an explicit lower bound for the Kazhdan L^2 -constant $\alpha(\Gamma)$ which depends only on the Hilbert series $H_X(t)$ and $H_R(t)$ (with respect to the standard degree function on $F(X)$). In the course of the proof we estimate from above the (finite) index of a subgroup Δ of Γ to which the proof of Theorem 4.1 applies. Such estimate is of independent interest, especially in view of the remark following Question 4.7.

Theorem A.10. *Let (X, R) be a presentation satisfying the GS condition and $G = \text{Gr}(X, R)$. Choose $0 < t_1 < t_0 < 1$ such that $1 - H_X(t_i) + H_R(t_i) < 0$ for $i = 0, 1$, and set*

$$\mu = -(1 - H_X(t_0) + H_R(t_0)), \quad \rho = \frac{t_0}{t_1}.$$

Fix $M \geq 6 \cdot 10^5$, and set

$$k_0 = \left\lceil \frac{\log(M) - \log(\mu)}{\log(\rho)} \right\rceil \quad \text{and} \quad N = \sum_{i=0}^{k_0} |X|^i.$$

The following hold:

- (a) *There exists a subgroup of G of index at most p^N which has a weighted presentation (X', R', W) with $1 - W(X') + W(R') < -M$.*
- (b) *Let Γ be a dense finitely generated subgroup of G . Then there exists $C > 0$ which depends only on $p, N, |X|$ (and can be computed explicitly) such that $\alpha(\Gamma) \geq C$.*

Proof. To prove (a) we analyze the proof of Theorem 3.14. Let W be the (D, t_0) -weight function on $(F(X), X)$, where D is the standard degree function. For $k \in \mathbb{N}$ let (X_k, R_k) be defined as in the proof of Theorem 3.13 and $\mu_k = -(1 - W(X_k) + W(R_k))$.

Claim A.11. $\mu_{k_0} \geq M$.

⁷Here we indeed restrict our considerations from GGS groups to GS groups. This restriction does not seem to be essential, but it does simplify some arguments

Proof of Claim A.11. Let $A = \mathbb{F}_p[[G]]$, let $\{A_n\}$ be the D -filtration of A and $a_n = \dim A_n/A_{n+1}$. Let $\{c_n\}$ be as in the proof of Theorem 3.13. Then

$$\mu_k \geq \mu \prod_{i=1}^k \left(\frac{1-t_0^{pi}}{1-t_0^i} \right)^{c_i} \geq \mu \sum_{i=0}^k a_i t_0^i.$$

where the second inequality holds by Proposition 2.3. Since $1 - H_X(t_1) + H_R(t_1) < 0$, by the Golod-Shafarevich theorem the series $\sum_{i=0}^{\infty} a_i t_1^i$ diverges and thus

$$(A.4) \quad \limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \frac{1}{t_1}.$$

Since $\{A_n\}$ is the filtration of A corresponding to the standard degree function, for any $n, m \in \mathbb{N}$ we have $(A_n/A_{n+1}) \cdot (A_m/A_{m+1}) = A_{n+m}/A_{n+m+1}$ in $gr(A)$, and thus $a_{n+m} \leq a_n a_m$. This observation and (A.4) easily imply that $a_n \geq (1/t_1)^n$ for any n . Therefore,

$$\mu_{k_0} \geq \mu \sum_{i=0}^{k_0} a_i t_0^i \geq \mu (t_0/t_1)^{k_0} = \mu \rho^{k_0} \geq M \geq 6 \cdot 10^5.$$

□

By Claim A.11 the presentation $(X', R') = (X_{k_0}, R_{k_0})$ satisfies the inequality $1 - W(X') + W(R') < -M$. By construction $\text{Gr}(X_k, R_k)$ is the $(k+1)^{\text{st}}$ term of the Zassenhaus p -series of $\text{Gr}(X, R)$, whence $\log_p[\text{Gr}(X, R) : \text{Gr}(X_k, R_k)] \leq \sum_{i=0}^k |X|^i$ for any k . In particular, $\log_p[\text{Gr}(X, R) : \text{Gr}(X', R')] \leq \sum_{i=0}^{k_0} |X|^i = N$, which proves (a).

(b) We start by analyzing the proof of Theorem A.9. By Theorem 4.3 the group $\text{Gr}(X', R')$ has a weighted presentation (X'', R'', W'') satisfying **Pres**(12, 1/100, 1/1000). Thus, in the proof of Theorem A.9 we can set $H = \text{Gr}(X', R')$.

Now let K be an open normal subgroup of H which satisfies the conclusion of Lemma A.8 applied to the weighted presentation (X'', R'', W'') . We choose K so that the index $[G : K]$ is minimal possible. In the proof of Theorem A.9 we showed that $\alpha(\Gamma) \geq \frac{1}{25p|X''|^{3/2}[G:K]}$. Thus, it remains to find upper bounds for $|X''|$ and $[G : K]$ in terms of p , $|X|$ and N . Since we already have such a bound for $[G : H]$, it is enough to find a bound for $[H : K]$.

Given a finite set \tilde{X} and a weight function \tilde{W} on $(F(\tilde{X}), \tilde{X})$, we set $w_{\min}(\tilde{X}, \tilde{W}) = \min\{\tilde{W}(x) : x \in \tilde{X}\}$, $w_{\max}(\tilde{X}, \tilde{W}) = \max\{\tilde{W}(x) : x \in \tilde{X}\}$ and $\theta(\tilde{X}, \tilde{W}) = \frac{\log(w_{\max})}{\log(w_{\min})}$. Using the proof of Lemma A.8 it is easy to show that

$$\log_p[H : K] \leq \sum_{i=0}^m |X''|^m \text{ where } m = \lceil \theta(X'', W'') \rceil.$$

Thus, it remains to find upper bounds for $|X''|$ and $\theta(X'', W'')$.

Recall that the weighted presentation (X'', R'', W'') is obtained from (X, R, W) using a sequence of p -descents, good changes of generators and relators, cleanups and weight contractions. It is easy to see that

- (i) If $(\tilde{X}, \tilde{R}, \tilde{W}) \rightarrow (\tilde{X}', \tilde{R}', \tilde{W})$ is a \tilde{W} -good change of generators or relators or a cleanup, then $|\tilde{X}'| \leq |\tilde{X}|$ and $\theta(\tilde{X}', \tilde{W}') \leq \theta(\tilde{X}, \tilde{W})$.
- (ii) If $(\tilde{X}, \tilde{R}, \tilde{W}) \rightarrow (\tilde{X}', \tilde{R}', \tilde{W})$ is a p -descent, then $|\tilde{X}'| = p(|\tilde{X}| - 1) + 1 < p|\tilde{X}|$ and $\theta(\tilde{X}', \tilde{W}') \leq p\theta(\tilde{X}, \tilde{W})$.
- (iii) If $(\tilde{X}, \tilde{R}, \tilde{W}) \rightarrow (\tilde{X}, \tilde{R}, \tilde{W}')$ is a weight contraction, then $\theta(\tilde{X}, \tilde{W}') \leq \theta(\tilde{X}, \tilde{W})$.

Since a p -descent replaces the group defined by a presentation by a subgroup of index p and all other operations listed above do not modify the group, the number of p -descents in the sequence is equal to $\log_p[\mathrm{Gr}(X, R) : \mathrm{Gr}(X'', R'')] = \log_p[G : H] \leq N$. Also observe that $\theta(X, W) = 1$ (since by $W(x) = t_0$ for all $x \in X$ by construction). Using (i), (ii) and (iii), we conclude that $|X''| < p^N |X|$ and $\theta(X'', W'') \leq p^N$. These are desired bounds for $|X''|$ and $\theta(X'', W'')$. The proof is complete. \square

APPENDIX B. ON SUBGROUP GROWTH OF GOLOD-SHAFAREVICH GROUPS

by ANDREI JAIKIN-ZAPIRAIN ⁸

As mentioned in the introduction, there are many results indicating that Golod-Shafarevich groups are “large”. One of the examples is the main theorem of this paper. We add one more property to this list: we will show that Golod-Shafarevich groups have many subgroups of finite index.

If G is an abstract group, we denote by $a_m(G)$ the number of subgroups of index m in G . Similarly, if G is a profinite group, $a_m(G)$ denotes the number of open subgroups ⁹ of index m in G . If G is finitely generated (abstract or profinite) then $a_m(G)$ is finite for all m . Note that the number of subgroups of index m in an abstract group G coincides with the number of open subgroups of index m in the profinite completion \hat{G} of G : $a_m(G) = a_m(\hat{G})$.

Proofs of the following results on subgroup growth can be found in [LS, Chapters 1-3]. For a free group F_d on $d \geq 2$ generators we have $(m!)^{d-1} \leq a_m(F_d) \leq m(m!)^{d-1}$, and $a_m(G) \leq a_m(F_d)$ for any d -generated abstract group G . The same holds for profinite groups: If \hat{F}_d denotes a free profinite group on d generators, then $a_m(\hat{F}_d) \leq (m!)^d$ and $a_m(G) \leq a_m(\hat{F}_d)$ for any d -generated profinite group G . For a free pro- p group $(F_d)_{\hat{p}}$ on d generators we have $a_m((F_d)_{\hat{p}}) \leq 2^{md}$. From this we may deduce that for any finitely generated pro- p group G we have $\log_2(\log_2(a_m(G))) \leq \log_2(m) + C$ for some constant C (of course $a_m(G) = 0$ if m is not a power of p). It is also known that there is a constant C' such that $\log_2(\log_2(a_m((F_d)_{\hat{p}}))) \geq \log_2(m) + C'$ if $m = p^k$ for some k .

A well-known characterization of p -adic analytic pro- p groups says that a pro- p group G is p -adic analytic if and only if there is a constant C such that

$$\log_2(\log_2(a_m(G))) \leq \log_2(\log_2 m) + C \text{ for all } m.$$

By a result of A. Shalev [Sh], if a pro- p group G is not p -adic analytic, there is a constant C such that

$$\log_2(\log_2(a_m(G))) \geq 2 \cdot \log_2(\log_2 m) + C \text{ for infinitely many values of } m.$$

In an unpublished work A. Shalev (see [LS, Theorem 4.6.4]) showed that if G is a Golod-Shafarevich pro- p group, then for any $\varepsilon > 0$ we have

$$\log_2(\log_2(a_m(G))) \geq (3 - \varepsilon) \log_2(\log_2 m) \text{ for infinitely many values of } m.$$

The following result significantly improves this bound:

Theorem B.1. *Let G be a generalized Golod-Shafarevich pro- p group. Then there exists a constant $\beta = \beta(G) > 0$ such that*

$$\log_2(\log_2(a_m(G))) \geq (\log_2 m)^\beta$$

for infinitely many values of m .

⁸The author is supported by the Spanish Ministry of Science and Innovation, grant MTM2008-06680.

⁹If G is a finitely generated profinite group, by an important result of Nikolov and Segal [NS] every subgroup of finite index in G is automatically open

Proof. Let (X, R) be a presentation of G satisfying the GGS condition for some degree function D on $(F(X), X)$ and $t_0 \in (0, 1)$. Let $\{G_n\}$ be the D -filtration of G and put

$$c_n = \dim_{\mathbb{F}_p} G_n/G_{n+1} = \log_p [G_n : G_{n+1}].$$

Let $t_1 = \frac{1}{\limsup \sqrt[n]{c_n}}$. Since t_1 coincides with the radius of convergence of $\text{Hilb}_{D, \mathbb{F}_p[[G]]}(t)$ (for instance, this follows easily from Proposition 2.3), we have $t_0 > t_1$.

Let $\varepsilon = \min\{\frac{t_1}{2}, \frac{t_0 - t_1}{2}\}$. For infinitely many n we have $c_{n-1} \geq (\frac{1}{t_1 + \varepsilon})^{n-1}$ and $|G/G_n| \leq p^{(\frac{1}{t_1 - \varepsilon})^{n-1}}$. We fix n satisfying these two conditions.

Let (X', R') be the presentation of $K = G_n$ obtained in Theorem 3.11. Observe first that $d(G_n) \geq H_{X'}(t_0) - H_{R'}(t_0)$, where $d(G_n)$ is the minimal number of generators of G_n . Indeed, let $m = d(G_n)$ and let $\{y_1, \dots, y_m\} \subseteq X'$ be a generating set for G_n . Put $Y = \{y_1, \dots, y_{m-1}\}$ and consider the group $H = \text{Gr}(X', R' \cup Y)$. This group is (pro)cyclic and in particular it is not a generalized Golod-Shafarevich group. Hence

$$1 - H_{X'}(t_0) + H_{R'}(t_0) + H_Y(t_0) \geq 0.$$

On the other hand, $H_Y(t_0) \leq (m-1)t_0 \leq m-1 = d(G_n) - 1$, whence $d(G_n) \geq H_{X'}(t_0) - H_{R'}(t_0)$.

If we substitute t_0 in the expression (3.6) of Theorem 3.11, we obtain that

$$d(G_n) - 1 \geq H_{X'}(t_0) - H_{R'}(t_0) - 1 \geq (H_X(t_0) - H_R(t_0) - 1) \prod_{i=1}^{n-1} \left(\frac{1 - t_0^{pi}}{1 - t_0^i} \right)^{c_i}.$$

Note that $\prod_{i=1}^{n-1} \left(\frac{1 - t_0^{pi}}{1 - t_0^i} \right)^{c_i} \geq (1 + t_0^{n-1})^{c_{n-1}}$, and $c_{n-1} \geq (\frac{1}{t_1 + \varepsilon})^{n-1}$ by our choice of n . Therefore,

$$\prod_{i=1}^{n-1} \left(\frac{1 - t_0^{pi}}{1 - t_0^i} \right)^{c_i} \geq [(1 + t_0^{n-1})^{\frac{1}{t_0^{n-1}}}]^{(\frac{t_0}{t_1 + \varepsilon})^{n-1}} \geq 2^{(\frac{t_0}{t_1 + \varepsilon})^{n-1}}.$$

Since $(H_X(t_0) - H_R(t_0) - 1) > 0$ and $\frac{t_0}{t_1 + \varepsilon} > 1$, there exists $\alpha > 0$ (independent of n) such that

$$\begin{aligned} d(G_n) - 1 &\geq (H_X(t_0) - H_R(t_0) - 1) 2^{(\frac{t_0}{t_1 + \varepsilon})^{n-1}} \\ &\geq \alpha \cdot 2^{(\frac{1}{t_1 - \varepsilon})^{\alpha(n-1)}} \geq \alpha \cdot 2^{(\log_p |G/G_n|)^{\alpha}}. \end{aligned}$$

Let $m = p|G/G_n|$. Note that any pro- p group K contains $p^{d(K)} - 1$ subgroups of index p since $K/[K, K]K^p \cong (\mathbb{Z}/p\mathbb{Z})^{d(K)}$. Therefore, if we count the number of subgroups of index p in G_n , we obtain that $a_m(G) \geq p^{\alpha \cdot 2^{(\log_p m - 1)^{\alpha}}}$. This finishes the proof. \square

REFERENCES

- [A+] G. N. Arzhantseva, J. Burillo, M. Lustig, L. Reeves, H. Short, E. Ventura, *Uniform non-amenability*, Adv. Math. 197 (2005), no. 2, 499–522.
- [BaGr] L. Bartholdi and R. Grigorchuk, *Lie methods in growth of groups and groups of finite width*. Computational and geometric aspects of modern algebra (Edinburgh, 1998), 1–27, London Math. Soc. Lecture Note Ser., 275, Cambridge Univ. Press, Cambridge, 2000.
- [BG] E. Breuillard and T. Gelander, *Uniform independence in linear groups*. Invent. Math. 173 (2008), no. 2, 225–263.
- [DDMS] J. D. Dixon, M. P. F. du Sautoy, A. Mann and D. Segal, *Analytic pro- p groups*. Second edition. Cambridge Studies in Advanced Mathematics, 61. Cambridge University Press, Cambridge, 1999.
- [EJ] M. Ershov, A. Jaikin-Zapirain, *Property (T) for noncommutative universal lattices*, preprint, 2008, arXiv: 0809.4095

- [Er] M. Ershov, *Golod-Shafarevich groups with property (T) and Kac-Moody groups*. Duke Math. J., 145 (2008), no. 2, 309–339
- [GS] E. Golod and I. Shafarevich, *On the class field tower. (Russian)* Izv. Akad. Nauk SSSR Ser. Mat. 28 (1964), 261–272
- [Go] E. Golod, *On nil algebras and finitely approximable groups. (Russian)* Izv. Akad. Nauk SSSR Ser. Mat. 28 (1964), 273–276
- [Gr] R. I. Grigorchuk, *On the Hilbert-Poincaré series of graded algebras that are associated with groups. (Russian)* Mat. Sb. 180 (1989), no. 2, 207–225, 304; translation in Math. USSR-Sb. 66 (1990), no. 1, 211–229
- [Gr2] R. I. Grigorchuk, *Just-infinite branch groups*, New horizons in pro- p groups, 125–185, Progr. Math., 184, Birkhäuser Boston, Boston, MA, 2000.
- [Hab] K. Haberland, *Galois cohomology of algebraic number fields*. With two appendices by Helmut Koch and Thomas Zink. VEB Deutscher Verlag der h: Wissenschaften, Berlin, 1978.
- [H] P. de la Harpe, *Uniform growth in groups of exponential growth*. Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part II (Haifa, 2000). Geom. Dedicata 95 (2002), 1–17
- [Hu] A. Hulanicki, *Means and Følner condition on locally compact groups*. Studia Math. 27 (1966) 87–104.
- [Ko] H. Koch, *Galois theory of p -extensions*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2002. xiv+190 pp.
- [LS] A. Lubotzky and D. Segal, *Subgroup growth*. Progress in Mathematics, 212. Birkhäuser Verlag, Basel, 2003.
- [NS] N. Nikolov and D. Segal, *On finitely generated profinite groups I,II*. Ann. of Math. (2) 165 (2007), no. 1, 171–238 and 239–273
- [Os1] D. Osin, *Weakly amenable groups*. Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001), 105–113, Contemp. Math., 298, Amer. Math. Soc., Providence, RI, 2002.
- [Os2] D. Osin, *Kazhdan constants of hyperbolic groups*. (Russian) Funktsional. Anal. i Prilozhen. 36 (2002), no. 4, 46–54; translation in Funct. Anal. Appl. 36 (2002), no. 4, 290–297
- [Os3] D. Osin, *Uniform non-amenability of free Burnside groups*. Arch. Math. (Basel) 88 (2007), no. 5, 403–412.
- [Qu] D. Quillen, *On the associated graded ring of a group ring*. J. Algebra 10 (1968) 411–418.
- [Sh] A. Shalev, *Growth functions, p -adic analytic groups, and groups of finite coclass*. J. London Math. Soc. (2) 46 (1992), no. 1, 111–122
- [Wi] J. Wilson, *Profinite groups*, London Mathematical Society Monographs. New Series, 19. The Clarendon Press, Oxford University Press, New York, 1998
- [Ze] E. Zelmanov, *On groups satisfying the Golod-Shafarevich condition*. New horizons in pro- p groups, 223–232, Progr. Math., 184, Birkhäuser Boston, Boston, MA, 2000

UNIVERSITY OF VIRGINIA

E-mail address: ershov@virginia.edu

DEPARTAMENTO DE MATEMATICAS UAM; INSTITUTO DE CIENCIAS MATEMATICAS CSIC-UAM-UC3M-UCM

E-mail address: andrei.jaikin@uam.es