A MNEMONIC FOR THE GRADED-CASE GOLOD-SHAFAREVICH INEQUALITY

DAVID ANICK AND WARREN DICKS*

Abstract. We draw attention to an easy-to-remember explanation for the graded-case inequality of Golod and Shafarevich. We review, unify, and simplify some of the classic material on this inequality, thereby offering a new, concise exposition for it.

Let K be a field, and $B = \bigoplus_{n\in\mathbb{Z}} B_n = K\langle X \mid R\rangle$ be a $\mathbb{Z}$-graded, associative K-algebra that is presented with a generating set X and a relating set R, both of which are positively graded. For each $n \in \mathbb{Z}$, set $\mathbf{b}_n := \dim_K(B_n)$. One form of the *graded-case inequality of Golod and Shafarevich* is

$$(\forall n \in \mathbb{Z}) \qquad \sum_{x\in X} \mathbf{b}_{n-\deg(x)} \leqslant (\sum_{r\in R} \mathbf{b}_{n-\deg(r)}) + \mathbf{b}_n. \tag{1}$$

Many important applications of this inequality can be found on its Wikipedia page [9].

The following is an easy-to-remember explanation for this inequality. The Koszul resolution

$$0 \to \operatorname{Ker}\partial \to \bigoplus_{r\in R} B \xrightarrow{\partial} \bigoplus_{x\in X} B \to B \to K \to 0 \tag{2}$$

respects $\mathbb{Z}$-gradings; hence, for each $n \in \mathbb{Z}$, the nth component of (2) is an exact sequence of K-modules

$$0_n \to (\operatorname{Ker}\partial)_n \to \bigoplus_{r\in R} B_{n-\deg(r)} \xrightarrow{\partial_n} \bigoplus_{x\in X} B_{n-\deg(x)} \to B_n \to K_n \to 0_n. \tag{3}$$

The middle part is an exact sequence of K-modules

$$\bigoplus_{r\in R} B_{n-\deg(r)} \xrightarrow{\partial_n} \bigoplus_{x\in X} B_{n-\deg(x)} \to B_n, \tag{4}$$

and, because K is a field, the K-dimension of the inner term is at most the sum of the K-dimensions of the two outer terms. Hence, (1) holds, and there are not even any cardinality restrictions.

This concludes the main point of this note, but perhaps lengthy explanatory remarks are in order. As far as we know, the preceding argument, which we chanced upon in 1982, has not appeared in print before now, and we would be interested to hear from anyone who knows that it has. There exist proofs of (1) in the literature which go via (3) without mentioning (2) or (4); see, for example, the original source [4], or Theorem 8.1.1.1 of [5]. There is a proof of a special case of (1), Theorem 2.3.4(i) of [2], that goes via (2), bypassing (3) and (4). There also exist proofs in the literature which have (2), (3), and (4) in the background; see, for example, Section 3.5 of [7].

In what follows, with an eye toward how the material might be presented or taught as a coherent unit, we provide a digest of five topics: the Golod-Shafarevich p-group theorem; the construction of the Koszul resolution for augmented algebras (2); its graded version (3); the Hilbert series form of (1); and, the group-algebra analogue of the Koszul resolution.

We are grateful to Andrei Jaikin, Clas Löfwall, Dmitry Piontkovskii, Jan-Erik Roos, and John Wilson for their expert advice concerning the literature.

1. The Golod-Shafarevitch p-group theorem

As Bourbaki intended, we let $\mathbb{N}$ denote the set of finite cardinals, $\{0, 1, 2, 3, \dots, \}$.

The following evolved through work of Golod, Shafarevich, Gaschütz, Vinberg, and Serre; it may not have been expressed in this form before.

2010 *Mathematics Subject Classification.* Primary: 16W50; Secondary: 16E05, 20D15, 20F05.

Key words and phrases. Golod-Shafarevich inequality, Koszul resolution, Golod-Shafarevich p-group theorem.

*Corresponding author; research supported by MINECO (Spain) through project number MTM2014-53644-P.

1.1. **Theorem.** *Let K be a field, $B = K \oplus \mathfrak{b}$ be an augmented K-algebra, X be a generating set for $\mathfrak{b}$ as left B-module, and R be a relating set for $\mathfrak{b}$ when generated by X. If $1 \leqslant |X| = \dim_K(\mathfrak{b}/\mathfrak{b}^2) < \aleph_0$ and $|R| \leqslant \frac{1}{4}|X|^2$, then $\dim_K(B) \geqslant \aleph_0$.*

Proof (after Serre [6]). We have an exact left-B-module sequence

$$\bigoplus_R B \xrightarrow{\partial} \bigoplus_X B \xrightarrow{\pi} \mathfrak{b} \to 0, \tag{5}$$

where $\bigoplus_R B$ denotes the direct sum of copies of B indexed by R.

Let n range over $\mathbb{N}$.

We first use the hypotheses that $|X| = \dim_K(\mathfrak{b}/\mathfrak{b}^2) < \aleph_0$. Since K is a field, the surjective map

$$\bigoplus_X (B/\mathfrak{b}) \xrightarrow{(B/\mathfrak{b})\otimes_B \pi} \mathfrak{b}/\mathfrak{b}^2$$

is injective. Hence, $\operatorname{Ker}\pi \subseteq \bigoplus_X \mathfrak{b}$. By the exactness of (5), $\partial(\bigoplus_R B) = \operatorname{Ker}\pi \subseteq \bigoplus_X \mathfrak{b}$. By the left B-linearity of ∂, $\partial(\bigoplus_R \mathfrak{b}^{n-1}) \subseteq \bigoplus_X \mathfrak{b}^n$, where we define $\mathfrak{b}^{-1} := \mathfrak{b}^0 := B$ and $\mathfrak{b}^{n+1} := \mathfrak{b}^n\cdot\mathfrak{b}$. On applying $(B/\mathfrak{b}^n)\otimes_B -$ to (5), we obtain an exact left-B-module sequence

$$\bigoplus_R (B/\mathfrak{b}^n) \xrightarrow{\overline{\partial}} \bigoplus_X (B/\mathfrak{b}^n) \to \mathfrak{b}/\mathfrak{b}^{n+1} \to 0$$

such that $\overline{\partial}\big(\bigoplus_R(\mathfrak{b}^{n-1}/\mathfrak{b}^n)\big) = \{0\}$. There is then induced an exact left-B-module sequence

$$\bigoplus_R (B/\mathfrak{b}^{n-1}) \to \bigoplus_X (B/\mathfrak{b}^n) \to \mathfrak{b}/\mathfrak{b}^{n+1} \to 0.$$

Set $\mathbf{a}_{n-2} := \dim_K(B/\mathfrak{b}^{n-1})$. Since K is a field, $\mathbf{a}_n - 1 \leqslant |X|\cdot\mathbf{a}_{n-1} \leqslant |R|\cdot\mathbf{a}_{n-2} + \mathbf{a}_n - 1$. By induction on n, $\mathbf{a}_n \leqslant \sum_{i=0}^{n} |X|^i < \aleph_0$.

We next use the hypothesis that $|X|^2 - 4|R| \geqslant 0$. Set $\lambda := \frac{|X| - \sqrt{|X|^2 - 4|R|}}{2}$ and $\mu := \frac{|X| + \sqrt{|X|^2 - 4|R|}}{2}$. Then $0 \leqslant \lambda \leqslant \mu$. Set $\mathbf{b}_{n-1} := \mathbf{a}_{n-1} - \lambda\cdot\mathbf{a}_{n-2} \in \mathbb{R}$. Then

$$\mathbf{b}_{-1} = \mathbf{a}_{-1} - \lambda\cdot\mathbf{a}_{-2} = 0 \quad \text{and} \quad \mathbf{b}_n - \mu\cdot\mathbf{b}_{n-1} = \mathbf{a}_n - (\lambda+\mu)\cdot\mathbf{a}_{n-1} + \mu\cdot\lambda\cdot\mathbf{a}_{n-2} = \mathbf{a}_n - |X|\cdot\mathbf{a}_{n-1} + |R|\cdot\mathbf{a}_{n-2} \geqslant 1,$$

by the previous paragraph. By induction on n, $\mathbf{b}_n \geqslant \sum_{i=0}^{n} \mu^i$. Now, $\sum_{i=0}^{n} \mu^i \leqslant \mathbf{b}_n = \mathbf{a}_n - \lambda\cdot\mathbf{a}_{n-1} \leqslant \mathbf{a}_n$. See also Remark 4.1 below.

It remains to use the hypothesis that $|X| \geqslant 1$. Here, $\mu \geqslant 1$, for if $|X| < 2$, then $|X| = 1$, hence $0 = \lfloor \frac{1}{4}|X|^2 \rfloor \geqslant |R|$, and, hence, $\mu = 1$. Now $n+1 \leqslant \sum_{i=0}^{n} \mu^i \leqslant \mathbf{a}_n \leqslant \dim_K(B)$ and, hence, $\dim_K(B) \geqslant \aleph_0$. □

1.2. **Remark.** In the foregoing proof, $\sum_{i=0}^{n} |X|^i \geqslant \mathbf{a}_n \geqslant \sum_{i=0}^{n} \mu^i$. If $(|X|, |R|)$ is neither $(1,0)$ nor $(2,1)$, then $\mu > 1$ and, hence, the growth rate of $\mathbf{a}_n$ is exponential.

1.3. **Historical remarks.** For details about the following, see [4] and [6].

Let p be a prime number, and G be a nontrivial, finite p-group. Set $K := \mathbb{Z}/p\mathbb{Z}$ and $B := KG$, the group algebra. Let $\mathfrak{b}$ denote the kernel of the K-algebra homomorphism $B \to K$ which carries G to $\{1\}$. For each $n \in \mathbb{N}$, set $\mathbf{d}_n := \dim_K\big(\mathrm{H}_n(G, K)\big)$. Recall that $\mathrm{H}_1(G, K) = \mathfrak{b}/\mathfrak{b}^2$. From the theory of minimal resolutions, it is known that there exist exact left-B-module sequences of the form $\cdots \to B^{\mathbf{d}_3} \to B^{\mathbf{d}_2} \to B^{\mathbf{d}_1} \to \mathfrak{b} \to 0$. By Theorem 1.1, $\mathbf{d}_2 > \frac{1}{4}\mathbf{d}_1^2$.

It is known that $\mathbf{d}_1$ equals the minimum number of elements it takes to generate G as a pro-p group, and that for any generating set of $\mathbf{d}_1$ elements, $\mathbf{d}_2$ equals the minimum number of relations it takes to present G as a pro-p group. (By the Burnside basis theorem, $\mathbf{d}_1$ equals the minimum number of elements it takes to generate G as a group. For any generating set of $\mathbf{d}_1$ elements, the minimum number of relations it takes to present G as a group is at least $\mathbf{d}_2$, but it is not known if equality holds.)

The main objective of Golod and Shafarevich in [4], and the reason for which (1) was first developed, was to prove that $\mathbf{d}_2 > \frac{1}{4}(\mathbf{d}_1 - 1)^2$. It followed from this, together with an earlier result of Shafarevich, that the class-field-tower problem had a negative solution, that is, there do exist infinite class-field towers. Gaschütz and Vinberg [8] independently refined the inequality to $\mathbf{d}_2 > \frac{1}{4}\mathbf{d}_1^2$. Serre [6] gave the above proof of this refined inequality. Nevertheless, there still remain many applications of (1) which have not been superseded.

2. The Koszul resolution for an augmented algebra

2.1. **Notation.** Let K be a field, X be a set, F be the free associative K-algebra on X, and $\mathfrak{f}$ be the two-sided ideal of F generated by X. We write $F = K\langle X\rangle$.

Let R be a family of elements of $\mathfrak{f}$, possibly with repetitions, and $\mathfrak{r}$ denote the two-sided ideal of F generated by the elements of R. Set $B := F/\mathfrak{r}$ and $\mathfrak{b} := \mathfrak{f}/\mathfrak{r}$. In summary, $B = K \oplus \mathfrak{b}$ is an augmented associative K-algebra presented with generating set X and relating set R. We write $B = K\langle X \mid R\rangle$.

Set $K^{(X)} := \bigoplus_{x\in X} Kx$, $F^{(X)} := F \otimes_K K^{(X)}$, and $B^{(X)} := B \otimes_K K^{(X)}$; these are the free left modules on X over K, F, and B, respectively. Similar notation will apply with R in place of X. At one stage, we shall use the natural K-centralizing K-bimodule structure of $K^{(R)}$.

2.2. **Definitions.** Each element f of $\mathfrak{f}$ has a unique expression as a left F-linear combination of the elements of X, and we shall write this as $f = \sum_{x\in X} \frac{\partial f}{\partial x}\cdot x$.

We have an isomorphism of left F-modules

$$\mathfrak{f} \xrightarrow{\sim} F^{(X)}, \qquad f = \sum_{x\in X} \frac{\partial f}{\partial x}\cdot x \mapsto \sum_{x\in X} \frac{\partial f}{\partial x} \otimes x.$$

On applying $(F/\mathfrak{r}) \otimes_F -$, we obtain an isomorphism of left $F/\mathfrak{r}$-modules

$$\mathfrak{f}/\mathfrak{r}\mathfrak{f} \xrightarrow{\sim} B^{(X)}, \qquad f + \mathfrak{r}\mathfrak{f} \mapsto \sum_{x\in X} (\tfrac{\partial f}{\partial x} + \mathfrak{r}) \otimes x.$$

We have also a surjection of F-bimodules

$$F \otimes_K K^{(R)} \otimes_K F \twoheadrightarrow \mathfrak{r}, \qquad f_1 \otimes r \otimes f_2 \mapsto f_1\cdot r\cdot f_2.$$

On applying $(F/\mathfrak{r}) \otimes_F - \otimes_F (F/\mathfrak{f})$, we obtain a surjection of left $F/\mathfrak{r}$-modules

$$B^{(R)} \twoheadrightarrow \mathfrak{r}/\mathfrak{r}\mathfrak{f}, \qquad (f + \mathfrak{r}) \otimes r \mapsto f\cdot r + \mathfrak{r}\mathfrak{f}.$$

The cokernel of the composite $B^{(R)} \twoheadrightarrow \mathfrak{r}/\mathfrak{r}\mathfrak{f} \hookrightarrow \mathfrak{f}/\mathfrak{r}\mathfrak{f} \xrightarrow{\sim} B^{(X)}$ is isomorphic to $\mathfrak{f}/\mathfrak{r}$, which is $\mathfrak{b}$. We then have an exact left-B-module sequence

$$B^{(R)} \xrightarrow{b\otimes r\mapsto \sum_{x\in X} b\cdot(\frac{\partial r}{\partial x}+\mathfrak{r})\otimes x} B^{(X)} \xrightarrow{b\otimes x\mapsto b\cdot(x+\mathfrak{r})} \mathfrak{b} \to 0. \tag{6}$$

On splicing (6) and $0 \to \mathfrak{b} \to B \to B/\mathfrak{b} \to 0$, we obtain what we call *the Koszul resolution*

$$0 \to \operatorname{Ker}\partial \to B^{(R)} \xrightarrow{\partial: b\otimes r\mapsto \sum_{x\in X} b\cdot(\frac{\partial r}{\partial x}+\mathfrak{r})\otimes x} B^{(X)} \xrightarrow{b\otimes x\mapsto b\cdot(x+\mathfrak{r})} B \to B/\mathfrak{b} \to 0.$$

The part that interests us is

$$B^{(R)} \xrightarrow{b\otimes r\mapsto \sum_{x\in X} b\cdot(\frac{\partial r}{\partial x}+\mathfrak{r})\otimes x} B^{(X)} \xrightarrow{b\otimes x\mapsto b\cdot(x+\mathfrak{r})} B. \tag{7}$$

2.3. **Remark.** Suppose that the induced map $(B/\mathfrak{b})^{(X)} \to \mathfrak{b}/\mathfrak{b}^2$ is bijective or, equivalently, that each element of R lies in $\mathfrak{f}^2$. If X is a finite, nonempty set and $|R| \leqslant \frac{1}{4}|X|^2$, then applying Theorem 1.1 to (6) shows that $\dim_K(B) = \aleph_0$.

3. The graded case of the Koszul resolution

Continuing with the notation developed in Section 2, we now hypothesize a $\mathbb{Z}$-graded K-algebra structure for B, as follows.

Let $\deg : X \to \mathbb{N}-\{0\}$, $x \mapsto \deg(x)$, be any map; there is then an induced $\mathbb{Z}$-graded K-algebra structure $F = \bigoplus_{n\in\mathbb{Z}} F_n$ with $\bigoplus_{n\in\mathbb{Z}-\mathbb{N}} F_n = \{0\}$, $F_0 = K$, $\bigoplus_{n\in\mathbb{N}-\{0\}} F_n = \mathfrak{f}$, and $x \in F_{\deg(x)}$ for each $x \in X$.

We henceforth restrict to the case where each element of R lies in $\bigcup_{n\in\mathbb{N}-\{0\}} F_n$. There is then an induced $\mathbb{Z}$-graded K-algebra structure $B = \bigoplus_{n\in\mathbb{Z}} B_n$ with $\bigoplus_{n\in\mathbb{Z}-\mathbb{N}} B_n = \{0\}$, $B_0 = K$, $\bigoplus_{n\in\mathbb{N}-\{0\}} B_n = \mathfrak{b}$, and $x + \mathfrak{r} \in B_{\deg(x)}$ for each $x \in X$. We choose a map $\deg : R \to \mathbb{N}-\{0\}$, $r \mapsto \deg(r)$, such that $r \in F_{\deg(r)}$; thus, as in [4], each occurrence of 0 in R has some positive finite degree. Notice that $\frac{\partial r}{\partial x} \in F_{\deg(r)-\deg(x)}$.

Let n range over $\mathbb{Z}$. Set $\mathbf{b}_n := \dim_K(B_n)$. Now (7) gives an exact sequence of degree-n K-modules

$$\bigoplus_{r\in R}(B_{n-\deg(r)}\otimes_K Kr) \xrightarrow{b\otimes r\mapsto \sum\limits_{x\in X} b\cdot(\frac{\partial r}{\partial x}+\mathfrak{r})\otimes x} \bigoplus_{x\in X}(B_{n-\deg(x)}\otimes_K Kx) \xrightarrow{b\otimes x\mapsto b\cdot(x+\mathfrak{r})} B_n.$$

Since K is a field, we have one form of the Golod-Shafarevich inequality:

$$\sum_{x\in X}\mathbf{b}_{n-\deg(x)} \leqslant (\sum_{r\in R}\mathbf{b}_{n-\deg(r)}) + \mathbf{b}_n. \tag{8}$$

Set $X_n := \{x \in X : \deg(x) = n\}$ and $\mathbf{x}_n := |X_n|$. Then

$$\sum_{x\in X}\mathbf{b}_{n-\deg(x)} = \sum_{i\in\mathbb{Z}}\sum_{x\in X_i}\mathbf{b}_{n-\deg(x)} = \sum_{i\in\mathbb{Z}}\mathbf{x}_i{\cdot}\mathbf{b}_{n-i}.$$

Similarly, set $R_n := \{r \in R : \deg(r) = n\}$ and $\mathbf{r}_n := |R_n|$; then $\sum\limits_{r\in R}\mathbf{b}_{n-\deg(r)} = \sum\limits_{i\in\mathbb{Z}}\mathbf{r}_i{\cdot}\mathbf{b}_{n-i}$. Now (8) becomes

$$\mathbf{b}_n + \sum_{i\in\mathbb{Z}}\mathbf{r}_i{\cdot}\mathbf{b}_{n-i} \geqslant \sum_{i\in\mathbb{Z}}\mathbf{x}_i{\cdot}\mathbf{b}_{n-i}. \tag{9}$$

4. Hilbert series

Let t be a new variable. We shall express elements of the power-series ring $\mathbb{R}[[t]]$ in the form $\sum\limits_{n\in\mathbb{Z}} a_n t^n$, and understand that $a_n = 0$ if $n \leqslant -1$. Set

$$P := \{\sum_{n\in\mathbb{Z}} a_n t^n \in \mathbb{R}[[t]] : a_n \geqslant 0 \text{ for all } n \in \mathbb{Z}\}.$$

Then P is both an additive submonoid and a multiplicative submonoid in $\mathbb{R}[[t]]$. Let $\succeq$ be the relation on $\mathbb{R}[[t]]$ such that $\alpha \succeq \beta$ if and only if $\alpha - \beta \in P$.

4.1. **Remark.** In terms of this relation, the penultimate paragraph of the proof of Theorem 1.1 says, since

$$(1-\lambda t){\cdot}(1-\mu t){\cdot}(\sum_{n\in\mathbb{Z}}\mathbf{a}_n t^n) \;=\; (1-|X|{\cdot}t+|R|{\cdot}t^2){\cdot}(\sum_{n\in\mathbb{Z}}\mathbf{a}_n t^n) \;\succeq\; (1-t)^{-1},$$

$$\sum_{n\in\mathbb{Z}}\mathbf{a}_n t^n \;\succeq\; (1-\mu t)^{-1}{\cdot}(1-\lambda t)^{-1}{\cdot}(1-t)^{-1} \;\succeq\; (1-\mu t)^{-1}{\cdot}(1-t)^{-1}.$$

Continuing with the notation developed in Section 3, we henceforth restrict to the case where $\mathbf{x}_n$, $\mathbf{r}_n \in \mathbb{N}$; as in the proof of Theorem 1.1, $\mathbf{b}_n \in \mathbb{N}$. We define the *Hilbert series* of B, X, and R, to be the elements of $\mathbb{R}[[t]]$ given by $\mathrm{H}(B) := \sum\limits_{n\in\mathbb{Z}}\mathbf{b}_n t^n$, $\mathrm{h}(X) := \sum\limits_{n\in\mathbb{Z}}\mathbf{x}_n t^n$, and $\mathrm{h}(R) := \sum\limits_{n\in\mathbb{Z}}\mathbf{r}_n t^n$, respectively. Notice that the constant terms are 1, 0, and 0, respectively. Now (9) says that $\mathrm{H}(B) + \mathrm{h}(R){\cdot}\mathrm{H}(B) \succeq \mathrm{h}(X){\cdot}\mathrm{H}(B)$. Hence, $\big(1-\mathrm{h}(X)+\mathrm{h}(R)\big){\cdot}\mathrm{H}(B) \succeq 0$. By considering the constant terms, we see that

$$\big(1-\mathrm{h}(X)+\mathrm{h}(R)\big){\cdot}\mathrm{H}(B) \succeq 1; \tag{10}$$

this is esentially Lemma 2 of [4]. In fact, one can read directly from (2) that

$$\big(1-\mathrm{h}(X)+\mathrm{h}(R)\big){\cdot}\mathrm{H}(B) - \mathrm{H}(\operatorname{Ker}\partial) \;=\; \mathrm{H}(K) \;=\; 1.$$

4.2. **Key points.** Consider any $\gamma \in t{\cdot}\mathbb{R}[[t]]$.

If $\gamma \succeq \mathrm{h}(R)$, then $\big(1-\mathrm{h}(X)+\gamma\big){\cdot}\mathrm{H}(B) \;\succeq\; \big(1-\mathrm{h}(X)+\mathrm{h}(R)\big){\cdot}\mathrm{H}(B) \;\succeq\; 1$.

If it is also the case that $\big(1-\mathrm{h}(X)+\gamma\big)^{-1} \succeq 0$, then $\mathrm{H}(B) \;\succeq\; \big(1-\mathrm{h}(X)+\gamma\big)^{-1} \;\succeq\; 0$.

If it is further the case that X is finite and $\gamma \neq \mathrm{h}(X)$, or, more generally, that $\big(1-\mathrm{h}(X)+\gamma\big)^{-1} \notin \mathbb{R}[t]$, then $\mathrm{H}(B)$ has infinitely many nonzero coefficients, and, hence, $\dim_K(B) = \aleph_0$.

Finally, we restrict to the case where X is concentrated in degree 1.

4.3. **Corollary** (Golod)**.** *Let K be a field, X be a finite, nonempty set, and ε be an element of $\left[0, \frac{|X|}{2}\right]$. For each integer $n \geqslant 2$, let R_n be a family of X-homogenous elements in $K\langle X\rangle$ of X-degree n such that $|R_n| \leqslant \varepsilon^2(|X|-2\varepsilon)^{n-2}$. $\big($When $\varepsilon = \frac{|X|-1}{2}$, this says $|R_n| \leqslant (\frac{|X|-1}{2})^2$.$\big)$ Then $\dim_K\big(K\langle X \mid \bigcup\limits_{n\geqslant 2} R_n\rangle\big) = \aleph_0$.*

Proof. Set $\gamma := \sum_{n\geqslant 2} \big(\varepsilon^2(|X|-2\varepsilon)^{n-2}t^n\big) = \sum_{m\geqslant 0} \big(\varepsilon^2(|X|-2\varepsilon)^m t^m t^2\big) = \frac{\varepsilon^2 t^2}{1-(|X|-2\varepsilon)t}$.

Set $\alpha := 1-(|X|-\varepsilon)t$ and $\beta := \varepsilon t$. Then $\alpha-\beta = 1-|X|t$, $\alpha+\beta = 1-(|X|-2\varepsilon)t$, $\gamma = \frac{\beta^2}{\alpha+\beta}$,

$$1-|X|t+\gamma = (\alpha-\beta)+\tfrac{\beta^2}{\alpha+\beta} = \tfrac{\alpha^2}{\alpha+\beta}, \quad \text{and} \quad \big(1-|X|t+\gamma\big)^{-1} = \tfrac{1}{\alpha}+\tfrac{\beta}{\alpha^2}.$$

The result now follows by 4.2, since $|X| > \varepsilon \geqslant 0$. □

4.4. **Historical remarks.** Golod [3] then used Corollary 4.3 to create new phenomena: finitely generated, non-nilpotent, nil algebras and infinite, residually finite, finitely generated p-groups, for each prime p. These were the first infinite, finitely generated torsion groups.

5. The Fox resolution for group algebras

We now recall the group-algebra analogue of the Koszul resolution.

5.1. **Notation.** Let G be a group. Let $\mathrm{d}(G)$ denote the smallest of those cardinals κ such that G can be generated by κ elements. Let G' denote the derived subgroup of G, and set $G^{\mathrm{ab}} := G/G'$, the abelianization of G.

Let $\langle X \mid R\rangle$ be a presentation for G. Clearly, $\mathrm{d}(G^{\mathrm{ab}}) \leqslant \mathrm{d}(G) \leqslant |X|$.

Let K be a field and set $B := KG$, the group algebra. Let $\mathfrak{b}$ denote the kernel of the K-algebra homomorphism $B \to K$ which carries G to $\{1\}$. Let F be the group algebra over K for the free group on X, $\mathfrak{f}$ be the two-sided ideal of F generated by $\{x-1 \mid x\in X\}$, and $\mathfrak{r}$ be the two-sided ideal of F generated by $\{r-1 \mid r\in R\}$. Then $B = F/\mathfrak{r}$ and $\mathfrak{b} = \mathfrak{f}/\mathfrak{r}$.

Set $K^{(X)} := \bigoplus_{x\in X} Kx$, $F^{(X)} := F\otimes_K K^{(X)}$, and $B^{(X)} := B\otimes_K K^{(X)}$, and similarly with R in place of X.

5.2. **Definitions.** It is not difficult to see that the left ideal of F generated by $\{x-1 \mid x\in X\}$ is closed under right multiplication by the elements of $X\cup X^{-1}$, and, hence, is the whole of $\mathfrak{f}$. We have a left-F-module map $F^{(X)} \to \mathfrak{f}$ which sends each $1\otimes x$ to $x-1$; to construct an inverse, we shall define a left-F-module map $\mathfrak{f} \to F^{(X)}$ which sends each $x-1$ to $1\otimes x$.

We view $F^{(X)}$ as an (F,K)-bimodule, and form the bimodule-algebra over K suggestively written in matrix form as $\left(\begin{smallmatrix} F & F^{(X)} \\ 0 & K\end{smallmatrix}\right)$. There then exists a unique K-algebra homomorphism $\left(\begin{smallmatrix} \phi_{1,1} & \phi_{1,2} \\ 0 & \phi_{2,2}\end{smallmatrix}\right) : F \to \left(\begin{smallmatrix} F & F^{(X)} \\ 0 & K\end{smallmatrix}\right)$ which sends each $x\in X$ to the invertible element $\left(\begin{smallmatrix} x & 1\otimes x \\ 0 & 1\end{smallmatrix}\right)$. Thus, the $\phi_{i,j}$ are K-module maps, $\phi_{1,1}(1)=1$, $\phi_{1,2}(1)=0$, $\phi_{2,2}(1)=1$, and, for all f, $g\in F$,

$$\phi_{1,1}(f\cdot g) = \phi_{1,1}(f)\cdot\phi_{1,1}(g),\ \ \phi_{1,2}(f\cdot g) = \phi_{1,1}(f)\cdot\phi_{1,2}(g)+\phi_{1,2}(f)\cdot\phi_{2,2}(g),\ \text{and}\ \phi_{2,2}(f\cdot g) = \phi_{2,2}(f)\cdot\phi_{2,2}(g).$$

In particular, $\phi_{1,1}$ and $\phi_{2,2}$ are K-algebra homomorphisms. Also, for all $x\in X$,

$$\phi_{1,1}(x) = x,\ \ \phi_{1,2}(x) = 1\otimes x,\ \text{and}\ \phi_{2,2}(x) = 1.$$

In particular, $\phi_{1,1}$ is the identity map on F, and $\phi_{2,2}(\mathfrak{f}) = \{0\}$. Hence, $\phi_{1,2}$ restricted to $\mathfrak{f}$ is a left F-module map $\mathfrak{f} \to F^{(X)}$ which sends each $x-1$ to $1\otimes x$, as desired. Now each element f of $\mathfrak{f}$ has a unique expression as a left F-linear combination of the elements of $\{x-1 \mid x\in X\}$, which we write as $f = \sum_{x\in X} \frac{\partial f}{\partial(x-1)}\cdot(x-1)$.

We have an isomorphism of left F-modules

$$\mathfrak{f} \xrightarrow{\sim} F^{(X)}, \qquad f = \sum_{x\in X} \tfrac{\partial f}{\partial(x-1)}\cdot(x-1) \mapsto \sum_{x\in X} \tfrac{\partial f}{\partial(x-1)}\otimes x.$$

On applying $(F/\mathfrak{r})\otimes_F -$, we obtain an isomorphism of left $F/\mathfrak{r}$-modules

$$\mathfrak{f}/\mathfrak{r}\mathfrak{f} \xrightarrow{\sim} B^{(X)}, \qquad f+\mathfrak{r}\mathfrak{f} \mapsto \sum_{x\in X} (\tfrac{\partial f}{\partial(x-1)}+\mathfrak{r})\otimes x.$$

We have also a surjection of F-bimodules

$$F\otimes_K K^{(R)}\otimes_K F \twoheadrightarrow \mathfrak{r}. \qquad f_1\otimes r\otimes f_2 \mapsto f_1\cdot(r-1)\cdot f_2.$$

On applying $(F/\mathfrak{r})\otimes_F - \otimes_F (F/\mathfrak{f})$, we obtain a surjection of left $F/\mathfrak{r}$-modules

$$B^{(R)} \twoheadrightarrow \mathfrak{r}/\mathfrak{r}\mathfrak{f}, \qquad (f+\mathfrak{r})\otimes r \mapsto f\cdot(r-1)+\mathfrak{r}\mathfrak{f}.$$

The cokernel of the composite $B^{(R)} \twoheadrightarrow \mathfrak{r}/\mathfrak{r}\mathfrak{f} \hookrightarrow \mathfrak{f}/\mathfrak{r}\mathfrak{f} \xrightarrow{\sim} B^{(X)}$ is isomorphic to $\mathfrak{f}/\mathfrak{r}$, which is $\mathfrak{b}$. We then have an exact left-B-module sequence

(11) $$B^{(R)} \xrightarrow{\partial\cdot b\otimes r\mapsto \sum_{x\in X} b\cdot(\frac{\partial(r-1)}{\partial(x-1)}+\mathfrak{r})\otimes x} B^{(X)} \xrightarrow{b\otimes x\mapsto b\cdot(x-1+\mathfrak{r})} \mathfrak{b}\to 0.$$

5.3. **Theorem.** *Let $\langle X \mid R\rangle$ be a presentation of a nontrivial, finite group G. If $|X| = \mathrm{d}(G^{\mathrm{ab}})$, then $|R| > \frac{1}{4}|X|^2$; equivalently, if $\mathrm{d}(G) = \mathrm{d}(G^{\mathrm{ab}}) = |X|$, then $|R| > \frac{1}{4}(\mathrm{d}(G))^2$.*

Proof. Since G^{ab} is a finite abelian group, $G^{\mathrm{ab}} \simeq \bigoplus_{i=1}^{d}(\mathbb{Z}/I_i)$ for some finite chain $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_d$ of proper ideals of $\mathbb{Z}$. Let p be a prime number such that $I_d \subseteq p\mathbb{Z}$, and set $K := \mathbb{Z}/p\mathbb{Z}$. Then $K \otimes_{\mathbb{Z}} G^{\mathrm{ab}} \simeq K^d$, and $d = \dim_K(K \otimes_{\mathbb{Z}} G^{\mathrm{ab}}) \leqslant \mathrm{d}(G^{\mathrm{ab}}) \leqslant d$. Thus, $\dim_K(K \otimes_{\mathbb{Z}} G^{\mathrm{ab}}) = \mathrm{d}(G^{\mathrm{ab}}) = |X|$.

It is well known and straightforward to prove that $\mathfrak{b}/\mathfrak{b}^2 \simeq K \otimes_{\mathbb{Z}} G^{\mathrm{ab}}$ with $(g-1) + \mathfrak{b}^2 \leftrightarrow 1 \otimes gG'$. Then $\dim_K \mathfrak{b}/\mathfrak{b}^2 = \dim_K(K \otimes_{\mathbb{Z}} G^{\mathrm{ab}}) = |X|$. The result now follows from Theorem 1.1 applied to (11). □

5.4. **Corollary** (Golod-Shafarevich)**.** *Let p be a prime number, and $\langle X \mid R\rangle$ be a presentation of a nontrivial, finite p-group G. If $|X| = \mathrm{d}(G)$, then $|R| > \frac{1}{4}(\mathrm{d}(G))^2$.*

Proof. By the Burnside basis theorem, $\mathrm{d}(G) = \dim_K(K \otimes_{\mathbb{Z}} G^{\mathrm{ab}})$ for $K = \mathbb{Z}/p\mathbb{Z}$. Hence, $\mathrm{d}(G) = \mathrm{d}(G^{\mathrm{ab}})$, and the result follows from the second part of Theorem 5.3. □

5.5. **Definitions** (continued)**.** For $f \in F-\{0\}$, we set $\deg(f) := \max\{i \in \mathbb{N} : f \in \mathfrak{f}^i\}$. For each $r \in R$, we have then defined $\deg(r-1) \in \mathbb{N} - \{0\}$, unless $r = 1$ in F, in which case we shall choose some value $\deg(r-1) \in \mathbb{N} - \{0\}$. Then $\frac{\partial(r-1)}{\partial(x-1)} \in \mathfrak{f}^{\deg(r-1)-1}$, for each $x \in X$.

Let n range over $\mathbb{Z}$. Define $\mathfrak{b}^n$ to be B if $n \leqslant 0$, and, as usual, to be $\mathfrak{b}^{n-1}\cdot\mathfrak{b}$ if $n \geqslant 1$. In (11), we find, for each $r \in R$,

$$\partial(\mathfrak{b}^n \otimes_K Kr) \subseteq \bigoplus_{x\in X}(\mathfrak{b}^{n+\deg(r-1)-1} \otimes_K Kx), \text{ and, hence, } \partial(\mathfrak{b}^{n-\deg(r-1)+1} \otimes_K Kr) \subseteq \bigoplus_{x\in X}(\mathfrak{b}^n \otimes_K Kx).$$

On applying $(B/\mathfrak{b}^n) \otimes_B -$ to (11), we get an exact left-B-module sequence

$$\bigoplus_{r\in R}(B/\mathfrak{b}^{n-\deg(r-1)+1}) \otimes_K Kr \to \bigoplus_{x\in X}(B/\mathfrak{b}^n) \otimes_K Kx \to (\mathfrak{b} + \mathfrak{b}^{n+1})/\mathfrak{b}^{n+1}.$$

Set $\mathbf{a}_n := \dim_K(B/\mathfrak{b}^{n+1})$, and define δ_n to be 0 if $n \leqslant -1$ and to be 1 if $n \geqslant 0$. Since K is a field, $|X|\mathbf{a}_{n-1} \leqslant (\sum_{r\in R} \mathbf{a}_{n-\deg(r-1)}) + (\mathbf{a}_n - \delta_n)$.

We set $R_n := \{r \in R : \deg(r-1) = n\}$ and $\mathbf{r}_n := |R_n|$. We henceforth restrict to the case where $|X|$, $\mathbf{r}_n \in \mathbb{N}$; as in the proof of Theorem 1.1, $\mathbf{a}_n \in \mathbb{N}$. We define $\mathrm{h}(R) := \sum_{n\in\mathbb{Z}} \mathbf{r}_n t^n \in \mathbb{R}[[t]]$. We define $\mathrm{h}(X)$ similarly, and find $\mathrm{h}(X) = |X|t$. Set $\mathbf{b}_n := \dim_K(\mathfrak{b}^n/\mathfrak{b}^{n+1}) = \mathbf{a}_n - \mathbf{a}_{n-1}$ and $\mathrm{H}(B) := \sum_{n\in\mathbb{Z}} \mathbf{b}_n t^n$. Notice that $(1-t)\cdot\sum_{n\in\mathbb{Z}} \mathbf{a}_n t^n = \mathrm{H}(B)$ and $\sum_{n\in\mathbb{Z}} \delta_n t^n = (1-t)^{-1}$. Now

$$\big(1 - \mathrm{h}(X) + \mathrm{h}(R)\big)\cdot(1-t)^{-1}\cdot\mathrm{H}(B) \succeq (1-t)^{-1}.$$

This is the form of Vinberg's inequality for filtered algebras [8]. The method of proof outlined here is based on the proof of Theorem 1.1 above, which, in turn, may have been suggested by Vinberg's work.

Set $\alpha := 1 - \mathrm{h}(X) + \mathrm{h}(R) \in \mathbb{R}[[t]]$. We claim that if there exists some $\varepsilon \in [0,1]$ such that the real series resulting from replacing the ts in α with εs converges to a value $\alpha(\varepsilon) \in \,]-\infty, 0]$, then $\mathrm{H}(B) \notin \mathbb{R}[t]$, and, in particular, $\dim_K(B) = \aleph_0$. This is clear if $\varepsilon \neq 1$. If $\varepsilon = 1$, then $\alpha \in \mathbb{Z}[t]$; here, if $\alpha(1) \neq 0$, we may replace ε with a value slightly smaller than 1 to pass to the preceding case, while if $\alpha(1) = 0$, then $\alpha\cdot(1-t)^{-1} \in \mathbb{Z}[t]$, and the desired conclusion holds. Thus the claim is proved. If $|X| \geqslant 1$, $\mathbf{r}_1 = 0$, and $|R| \leqslant \frac{1}{4}|X|^2$, then we may take ε to be $\min\{\frac{2}{|X|}, 1\}$ to recover Theorem 5.3. Notice that $\mathbf{r}_1 = 0$ if and only if $\mathbf{b}_1 = |X|$, and if $|X| = 1$, then $\lfloor \frac{1}{4}|X|^2 \rfloor = 0$.

5.6. **Historical remarks.** Suppose that $G = \langle X \mid R\rangle$ is a group presentation such that $\mathrm{d}(G^{\mathrm{ab}}) = |X| < \aleph_0$. Theorem 5.3 says that if $|R| \leqslant \frac{1}{4}|X|^2$, then either G is trivial $\big($where $(|X|, |R|) = (0,0)\big)$ or G is infinite. Wilson [10] showed that if $|R| < \frac{1}{4}|X|^2$, then either G is infinite cyclic $\big($where $(|X|, |R|) = (1,0)\big)$ or G maps onto a residually finite, infinite p-group, for some prime p. His proof is based on Vinberg's inequality and the methods of Golod [3]. A recent introduction to related results can be found in [1].

References

[1] Mikhail Ershov, *Golod-Shafarevich groups: a survey. Internat. J. Algebra Comput.* **22** (2012), no. 5, 1230001, 68 pp.

[2] Pavel Etingof and Ching-Hwa Eu, *Koszulity and the Hilbert series of preprojective algebras. Math. Res. Lett.* **14** (2007), 589–596.

[3] E. S. Golod, *On nil-algebras and finitely approximable p-groups.* (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **28** (1964), 273–276.

[4] E. S. Golod and I. R. Šafarevič, *On the class field tower.* (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **28** (1964), 261–272.

[5] I. N. Herstein, *Noncommutative rings.* The Carus Mathematical Monographs, No. **15**. Published by The Mathematical Association of America; distributed by John Wiley & Sons, Inc., New York, 1968. xi+199 pp.

[6] Jean-Pierre Serre, *Galois cohomology. Translated from the French by Patrick Ion and revised by the author.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1997. x+210 pp.

[7] V. A. Ufnarovskij, *Combinatorial and asymptotic methods in algebra.* pp. 1–196 in *Algebra VI. Combinatorial and asymptotic methods of algebra. Non-associative structures.* (A. I. Kostrikin and I. R. Shafarevich, editors) Encyclopaedia Math. Sci. **57**, Springer, Berlin, 1995. iii+287 pp.

[8] È. B. Vinberg, *On the theorem concerning the infinite-dimensionality of an associative algebra.* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 209–214.

[9] Wikipedia, *Golod-Shafarevich theorem.* `http://en.wikipedia.org/wiki/Golod%E2%80%93Shafarevich_theorem`

[10] John S. Wilson, *Finitely presented soluble groups.* pp. 296–316 in *Geometry and cohomology in group theory* (*Durham, 1994*). (Peter H. Kropholler, Graham A. Niblo, and Ralph Stöhr, editors) London Math. Soc. Lecture Note Ser., **252**. Cambridge Univ. Press, Cambridge, 1998. xii + 316 pp.

Laboratory for Water and Surface Studies, Department of Chemistry, Tufts University, 62 Pearson Rd., Medford MA02155, USA

E-mail address: `david.anick@rcn.com`

Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), SPAIN

E-mail address: `dicks@mat.uab.cat`