Chapter 6
The Burnside Problem

In this chapter we discuss some instances of the Burnside Problem. There are three versions of this problem, the first one being the

General Burnside Problem: *Is it true that if a group G is finitely generated and torsion, then it is finite?*

We discuss the General Burnside problem for locally finite groups (Section 6.2), for polycyclic-by-finite and solvable groups (Section 6.3), as well as its bounded version for linear groups (Section 6.4). Finally, in Section 6.5 we discuss the Kurosh–Levitzky problem (on nil algebras) and explain the construction of Golod and Shafarevich yielding a negative answer to the Kurosh–Levitzky problem and therefore to the General Burnside Problem.

6.1 Formulation of the Burnside Problems

The *General Burnside Problem*, posed by William Burnside in 1902 [46] – one of the oldest and most influential questions in group theory – asks whether or not a finitely generated group in which every element has finite order is necessarily finite.

Problem 6.1 (General Burnside Problem). Is it true that if a group G is finitely generated and torsion, then it is finite?

Sometimes, the word *periodic* is used instead of "torsion". In order to approach the study of the General Burnside Problem, we introduce the following useful notion.

Definition 6.2 (Burnside property). A class $\mathscr{C}$ of groups satisfies the *Burnside property* if for every torsion group G in $\mathscr{C}$ the following holds: every finitely generated subgroup of G is finite.

A weaker formulation of the General Burnside Problem is the following. First recall that a group G is said to be *periodic with bounded exponent*, or just a group

T. Ceccherini-Silberstein and M. D'Adderio, *Topics in Groups and Geometry*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-88109-2_6

with *bounded exponent*, if there exists an integer $n \geq 1$ such that $g^n = 1_G$ for all $g \in G$.

Problem 6.3 (Bounded Burnside Problem). Is it true that if a group G is finitely generated and of bounded exponent, then G is finite?

There is also a restricted version of the Burnside problem.

Problem 6.4 (Restricted Burnside Problem). Is it true that for every $m, n \in \mathbb{N}$ there are finitely many (up to isomorphism) finite groups G with m generators and of bounded exponent n?

6.2 Locally Finite Groups and the General Burnside Problem

Definition 6.5. Let $\mathscr{P}$ be a property of groups (e.g., being finite). We say that a group G is *locally* $\mathscr{P}$ if every finitely generated subgroup of G satisfies $\mathscr{P}$.

Example 6.6. Every abelian torsion group is locally finite. This immediately follows from the structure theorem of finitely generated abelian groups (see Corollary 1.30).

We can rephrase the General Burnside Problem in the following way.

Problem 6.7 (Reformulation of the General Burnside Problem). Is every torsion group locally finite?

Notice that Lemma 2.34 says that the General Burnside Problem has a positive solution for nilpotent groups, equivalently, the class of nilpotent groups has the Burnside property.

The class of locally finite groups is clearly closed under taking subgroups, homomorphic images and finite direct products (**exercise**). Next we show that it is also closed under extensions.

Lemma 6.8. *Let G be a group. Let $H \leq G$ be a normal subgroup and suppose that both H and G/H are locally finite. Then G is locally finite.*

Proof. Let $K \leq G$ be a finitely generated subgroup of G. The image of K in G/H is $KH/H \cong K/(K \cap H)$, which is finite by assumption. Hence $[K : K \cap H] < \infty$, which implies that $K \cap H$ is finitely generated (cf. Corollary 1.11). Hence $K \cap H \leq H$ is finite by the assumption on H. We deduce that K is also finite. □

In fact, we can show that every group contains a largest locally finite subgroup. Here, "largest" means that it contains all the other locally finite subgroups: this is stronger than "maximal". In order to do so, the following two propositions will be useful.

Proposition 6.9. *Let $K, L \trianglelefteq G$, and let both K and L be locally finite. Then KL is locally finite.*

Proof. $KL/K \cong L/(K \cap L)$ is locally finite, hence KL is an extension of locally finite groups (K and KL/K). Therefore, by the previous lemma, KL itself is locally finite. □

Let I be a *directed set*, that is, a set equipped with an order $\preceq$ such that for every $i_1, i_2 \in I$ there exists an $i \in I$ with $i_1 \preceq i$ and $i_2 \preceq i$. Let G be a group. A family $(H_i)_{i \in I}$ of subgroups of G is said to be *increasing* if $H_i \leq H_j$ for all $i, j \in I$ such that $i \preceq j$. If in addition we have $\bigcup_{i \in I} H_i = G$, then we say that the family $(H_i)_{i \in I}$ *exhausts* G.

Proposition 6.10. *Let G be a group. Let $(H_i)_{i \in I}$ be an increasing and exhausting family of subgroups of G. Suppose that H_i is locally finite for every $i \in I$. Then G itself is locally finite.*

Proof. Let $X \subseteq G$ be a finite subset. Since $(H_i)_{i \in I}$ is increasing and exhausting, we can find $i = i(X)$ such that $X \subseteq H_i$. Since H_i is locally finite, it follows that the subgroup generated by X is finite. This shows that G is locally finite. □

Corollary 6.11. *Every group G contains a largest locally finite normal subgroup $L(G)$ such that $G/L(G)$ does not contain nontrivial locally finite normal subgroups.*

Proof. Let I denote the set of all locally finite normal subgroups of G. Equip I with the order given by inclusion and observe that, by Proposition 6.9, it is a directed set. It follows from Proposition 6.10 that the subgroup $L(G) := \bigcup_{H \in I} H$ is locally finite. Moreover, since conjugation by elements in G preserves local-finiteness of subgroups, we have that $L(G) \trianglelefteq G$. It is clear from the construction that every locally finite normal subgroup of G is contained in $L(G)$.

On the other hand, if $H/L(G) \leq G/L(G)$ is a locally finite normal subgroup, then $H \leq G$ is locally finite by Lemma 6.8. But we have just seen that H must be contained in $L(G)$, hence $H/L(G)$ is the trivial subgroup. □

6.3 The General Burnside Problem for Polycyclic-by-Finite and Solvable Groups

The following theorem gives a positive solution to the General Burnside Problem for polycyclic-by-finite and solvable groups.

Theorem 6.12 (General Burnside Problem for polycyclic-by-finite and solvable groups). *Let G be a torsion group. Then*

(1) *if G is solvable, then it is locally finite;*

(2) *if G is polycyclic-by-finite (e.g polycyclic), then it is finite.*

Proof. If G is solvable, then consider its derived series. The quotients of this series are abelian and torsion, and therefore locally finite (cf. Example 6.6). Then G is locally finite by recursively applying Lemma 6.8.

If G is polycyclic-by-finite, consider any finite subnormal series with cyclic quotients. Since G is torsion, these quotients are necessarily finite. Since extensions of finite groups by finite groups are finite, it follows that G is finite. □

6.4 The Bounded Burnside Problem for Linear Groups

This section is devoted to the proof of the following theorem, which is due to Burnside.

Theorem 6.13 (Bounded Burnside Problem for linear groups). *Let G be a subgroup of* $\mathrm{GL}(n,\mathbb{K})$*, where $\mathbb{K}$ is a field. Suppose that G is finitely generated and of bounded exponent. Then G is finite.*

Proof. Notice that if $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$, then $G \leq \mathrm{GL}(n,\mathbb{K}) \leq \mathrm{GL}(n,\overline{\mathbb{K}})$. Hence, without loss of generality, we can assume that $\mathbb{K}$ is algebraically closed.

We can also assume that G acts irreducibly on $V = \mathbb{K}^n$. Indeed, we can always find a chain of subspaces

$$\{0\} \leq V_1 \leq V_2 \leq \cdots \leq V_s = V,$$

such that G acts irreducibly on each V_i/V_{i-1}, $i = 1,2,\ldots,s$ (cf. the proof of the first Claim in Section 4.5). Taking a basis for each factor, consider the basis of V obtained by taking the union of these bases. Then, in this basis, G will be in block upper triangular form:

$$g = \begin{pmatrix} M_1(g) & * & \cdots & * \\ 0 & M_2(g) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_s(g) \end{pmatrix},$$

for every $g \in G$, where $M_i(g) \in \mathrm{GL}(V_i/V_{i-1})$, $i = 1,2,\ldots,s$. Consider now the homomorphism

$$g \mapsto \varphi(g) := \begin{pmatrix} M_1(g) & 0 & \cdots & 0 \\ 0 & M_2(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_s(g) \end{pmatrix}.$$

The kernel of this map consists of matrices of the form

$$\begin{pmatrix} I_{n_1} & * & \cdots & * \\ 0 & I_{n_2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_{n_s} \end{pmatrix}$$

where $I_{n_i} \in \mathrm{GL}(n_i,\mathbb{K})$ denotes the identity matrix and $n_i = \dim_{\mathbb{K}}(V_i/V_{i-1})$, $i = 1,2,\ldots,s$. Thus $\ker(\varphi)$, being a subgroup of $\mathrm{UT}(n,\mathbb{K})$, is a nilpotent group (cf. Example 2.5.(b)). Since every nilpotent group is solvable, we deduce from Theorem 6.12 and the hypothesis that G is torsion, that $\ker(\varphi)$ is locally finite. On the other hand, the image of φ is a finite direct product of linear groups acting irreducibly, by construction. Hence if we know the result for G acting irreducibly, then $\varphi(G)$, being the finite direct product of locally finite groups, is locally finite by Lemma 6.8. We then deduce that G, being an extension of locally finite groups, is locally finite by applying once more Lemma 6.8. Since G is finitely generated, it is in fact finite.

To finish the proof, we need some more Wedderburn theory (cf. Section 4.4).

Let A be an algebra, $A \subseteq \mathrm{End}_{\mathbb{K}}(V)$ with $n := \dim_{\mathbb{K}} V < \infty$ and suppose that A acts irreducibly. Then A acts completely reducibly on V and therefore, by Proposition 4.29, its radical vanishes: $N(A) = \{0\}$. Then, since $\mathbb{K}$ is algebraically closed, Wedderburn's Theorem (Theorem 4.20) guarantees that $A \cong A/N(A) \cong M_{n_1}(\mathbb{K}) \oplus \cdots \oplus M_{n_r}(\mathbb{K})$. Let us show that A is in fact simple, that is, that $r = 1$. Suppose, by contradiction, that $r \geq 2$. Let e_1 denote the identity of $M_{n_1}(\mathbb{K}) \subseteq A$ (here we are identifying $M_{n_1}(\mathbb{K})$ with the corresponding ideal of A). Then $e_1 V$ is a proper A-submodule of V, contradicting the irreducibility of the action of A on V. Thus $A \cong M_s(\mathbb{K}) \subseteq \mathrm{End}_{\mathbb{K}}(V) \cong M_n(\mathbb{K})$. In particular, $s \leq n$. Let us show that, in fact, $s = n$, so that $A = \mathrm{End}_{\mathbb{K}}(V) \cong M_n(\mathbb{K})$.

Choosing a suitable basis of V, we can assume that $E_{1,1}$, the matrix that has 1 in the $(1,1)$ position and zero elsewhere, belongs to A. Since $E_{1,1}V \neq \{0\}$, we can find a vector $v \in V$ such that $E_{1,1}v \neq 0$. Now, $AE_{11}v$ is an A-submodule of V, therefore $AE_{1,1}v = V$, since V is A-irreducible. Moreover, $AE_{1,1}$ is the set of matrices with possibly nonzero entries only in the first column, i.e. of the form

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ * & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & 0 \end{pmatrix} \in M_s(\mathbb{K}).$$

Now $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} Ae_{11}v \leq s$, so that $n \leq s$. Hence $n = s$, and $A = \mathrm{End}_{\mathbb{K}}(V)$. We just proved:

Theorem 6.14 (Wedderburn). *Let $\mathbb{K}$ be an algebraically closed field and let V be a finite-dimensional vector space over $\mathbb{K}$. Suppose that $A \subseteq \mathrm{End}_{\mathbb{K}}(V)$ acts irreducibly on V. Then $A = \mathrm{End}_{\mathbb{K}}(V)$.* □

Corollary 6.15 (Burnside). *Let $\mathbb{K}$ be an algebraically closed field and let V be a finite-dimensional vector space over $\mathbb{K}$. If $G \subseteq \mathrm{GL}(V)$ and G acts irreducibly on V, then $\mathrm{span}_{\mathbb{K}}(G) = \mathrm{End}_{\mathbb{K}}(V)$.* □

We now finish the proof of the Burnside Theorem for linear groups.

Recall that $G \leq \mathrm{GL}(n, \mathbb{K})$, G is finitely generated and there exists a positive integer d such that $g^d = 1$ for all $g \in G$. Also, by the preceding arguments, we may assume that G acts irreducibly. Then, by Corollary 6.15, G spans $M_n(\mathbb{K})$. Hence there exist $g_1, g_2, \ldots, g_{n^2} \in G$ which form a basis for $M_n(\mathbb{K})$. Let $g \in G$. Observe that since $\mathbb{K}$ is algebraically closed g is triangularizable, that is, g is similar to a matrix of the form

$$\begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \alpha_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \alpha_n \end{pmatrix}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$. Since $g^d = 1$, we have $\alpha_i^d = 1$ for all i. It follows that we have at most d possible values for each α_i. This implies that, as g varies in G, the distinct values of the traces $\mathrm{tr}(g)$ are at most d^n. Hence the number of distinct n^2-tuples $(\mathrm{tr}(g_1 g), \mathrm{tr}(g_2 g), \ldots, \mathrm{tr}(g_{n^2} g))$ as g varies in G is at most $(d^n)^{n^2} = d^{n^3}$.

Claim. *Given $g \in G$, the tuple $(\mathrm{tr}(g_1 g), \mathrm{tr}(g_2 g), \ldots, \mathrm{tr}(g_{n^2} g))$ determines g uniquely.*

We first observe that the bilinear form $f \colon M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to \mathbb{K}$ defined by $f(a,b) := \mathrm{tr}(ab)$ for all $a, b \in M_n(\mathbb{K})$ satisfies the condition

$$f(ab, c) = f(a, bc)$$

for all $a, b, c \in M_n(\mathbb{K})$. It follows that the set $I = \{a \in M_n(\mathbb{K}) : f(a,b) = 0 \text{ for all } b \in M_n(\mathbb{K})\}$ is an ideal of $M_n(\mathbb{K})$. As the algebra $M_n(\mathbb{K})$ is simple (**exercise**), we have that $I = \{0\}$, equivalently, f is non-degenerate.

Let then $g', g'' \in G$ and suppose that $\mathrm{tr}(g_i g') = \mathrm{tr}(g_i g'')$ for all $i = 1, 2, \ldots, n^2$. By subtracting we get $\mathrm{tr}(g_i(g' - g'')) = 0$ for $i = 1, \ldots, n^2$. Since the g_i's span $M_n(\mathbb{K})$, this implies $f(a, (g' - g'')) = \mathrm{tr}(a(g' - g'')) = 0$ for all $a \in M_n(\mathbb{K})$. It follows that $g' - g'' = 0$ by the non-degeneracy of f. The claim follows.

All this shows that there are finitely many possibilities for $g \in G$. Hence G is finite, and this finishes the proof of the Burnside theorem for linear groups. □

6.5 The Golod–Shafarevich Construction

In this section we describe the negative solution to the General Burnside Problem provided by Golod and Shafarevich.

It turns out that their solution goes through the negative solution to a problem in associative algebras strictly connected to the General Burnside Problem.

Let $\mathbb{K}$ be a field and let A be an associative algebra over $\mathbb{K}$. For two subsets S and T of such an algebra we set $ST := \mathrm{span}_{\mathbb{K}}\{st \mid s \in S, t \in T\}$. For a positive integer n, we denote by S^n the product $SS \cdots S$ of S with itself n times: by associativity, this is well defined.

Definition 6.16. An element $a \in A$ is called *nilpotent* if there exists an integer $n \geq 1$ such that $a^n = 0$. We say that A is a *nil algebra* if every $a \in A$ is nilpotent. We say that A is *nilpotent* if there exists an integer $n \geq 1$ such that $A^n = \{0\}$.

Note that every nilpotent algebra is nil. Conversely, we have the following theorem, whose proof is left as an **exercise**.

Theorem 6.17 (Wedderburn). *A finite-dimensional nil algebra is nilpotent.*

Remark 6.18. In fact, Wedderburn proved even more, namely that a finite-dimensional algebra which admits a linear basis consisting of nilpotent elements is nilpotent.

The following problem is related to the General Burnside Problem.

Problem 6.19 (Kurosh–Levitzky). Let A be a finitely generated nil algebra. Does this imply that A is nilpotent (and hence finite-dimensional)?

An algebra in which any finitely-generated subalgebra is nilpotent is called *locally nilpotent*. The following proposition is somewhat similar to Lemma 6.8. Its proof is left as an **exercise**.

Proposition 6.20. *Let A be a algebra. Let $I \leq A$ be an ideal and suppose that both A/I and I are locally nilpotent. Then A is locally nilpotent.*

So for any associative algebra A there exists a largest locally nilpotent ideal which is called the *Levitzky radical*.

Remark 6.21. In analogy with the reformulation of the General Burnside Problem, the Kurosh–Levitzky Problem asks whether nil and local nilpotence are equivalent conditions.

The answer to the Kurosh–Levitzky Problem is negative. To see this, we fix some notation.

Let $\mathbb{K}$ be a field. Denote by $\mathbb{K}\langle x_1, x_2, \ldots, x_n\rangle = \mathbb{K}\langle X\rangle$ the free associative algebra with coefficients in $\mathbb{K}$ freely generated by $X = \{x_1, x_2, \ldots, x_n\}$. We simply call it the *free algebra generated by X*.

Let $R \subseteq \mathbb{K}\langle X\rangle$ be any subset. We denote by (R) the (two-sided) ideal generated by R, i.e. the set consisting of all finite sums $\sum_i a_i r_i b_i$ where $r_i \in R$ and $a_i, b_i \in \mathbb{K}\langle X\rangle$. We then say that the algebra $\mathbb{K}\langle X\rangle/(R)$ has the *presentation* $\langle X \mid R\rangle$ and that the elements of X (resp. R) are the corresponding *generators* (resp. *relators*).

A unital algebra A is said to be *graded* if it has a direct sum decomposition into $\mathbb{K}$-subspaces

$$A = A_0 \oplus A_1 \oplus A_2 \oplus \cdots = \bigoplus_{i \in \mathbb{N}} A_i \tag{6.1}$$

where $A_0 := \mathbb{K}1_A$ and $A_i A_j \subseteq A_{i+j}$ for all $i, j = 0, 1, \ldots$ We say that the elements of A_i are the *homogeneous elements of degree i*. An ideal I of a graded algebra A is said to be *homogeneous* provided that for every element $a \in I$, the homogeneous parts of a are also contained in I. If I is a homogeneous ideal of a graded algebra A, then A/I is also a graded algebra, and it has decomposition

$$A/I = \bigoplus_{i \in \mathbb{N}} (A_i + I)/I.$$

Example 6.22. (1) Let X be a set. Then the free algebra $\mathbb{K}\langle X\rangle$ generated by X is graded. Indeed, the homogeneous elements of degree i are the homogeneous (non-commutative) polynomials of degree i together with the 0 polynomial.

(2) The algebra $A := \mathbb{K}[x_1, \ldots, x_n]$ of (commutative) polynomials with coefficients in $\mathbb{K}$ is also graded. Here, the homogeneous elements of degree i are the homogeneous (commutative) polynomials of degree i together with the 0 polynomial.

Let A be a graded algebra as in (6.1) and suppose that $\dim_{\mathbb{K}} A_i < \infty$ for all i. Then the associated *Hilbert series* is the formal power series

$$H_A(t) := \sum_{i\geq 0} \dim_{\mathbb{K}}(A_i)t^i.$$

Given two formal powers series $\sum_{i\geq 0} a_i t^i$ and $\sum_{i\geq 0} b_i t^i$, we write

$$\sum_{i\geq 0} a_i t^i \preceq \sum_{i\geq 0} b_i t^i$$

provided that $a_i \leq b_i$ for all $i \geq 0$, and define their product as

$$\sum_{k\geq 0} c_k t^k := \left(\sum_{i\geq 0} a_i t^i\right)\left(\sum_{j\geq 0} b_j t^i\right)$$

where

$$c_k := a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$$

for all $k \geq 0$.

Suppose that $R \subseteq \mathbb{K}\langle X\rangle$ is a subset consisting of homogeneous linearly independent elements of degree ≥ 2, and let r_i denote the number of elements of degree i in R for all $i \geq 2$. We set

$$H_R(t) := r_2 t^2 + r_3 t^3 + \cdots .$$

Then the ideal (R) generated by R is a graded ideal and the algebra $A = \langle X \mid R\rangle := \mathbb{K}\langle X\rangle/(R)$ is a graded algebra.

The following theorem constitutes the key ingredient of the Golod–Shafarevich construction.

Theorem 6.23 (Golod–Shafarevich). *With the above notation we have*

$$H_A(t)(1 - nt + H_R(t)) \succeq 1. \tag{6.2}$$

Before proving the theorem, let us show how we can derive from it a negative answer to the Kurosh–Levitzky Problem.

Suppose that we manage to find a real number $0 < t_0 < 1$ such that

(1) $H_R(t)$ converges at t_0 and
(2) $1 - nt_0 + H_R(t_0) < 0$.

Then $H_A(t)$ does not converge at t_0. In fact if it converges, then necessarily $H_A(t_0) \geq 0$, since $t_0 > 0$, which, together with (2), contradicts (6.2). This implies that A is infinite-dimensional: in fact for a finite-dimensional algebra A, the power series $H_A(t)$ is a polynomial, which converges everywhere.

Remark 6.24. This argument can be used in several ways to conclude that an algebra with a given presentation is infinite-dimensional.

For example, suppose that we are given a finite subset $R \subseteq \mathbb{K}\langle X\rangle_2$ of quadratic relators such that $r := |R| < n^2/4$ (recall that $n = |X|$). Then for $t = 2/n$, one has $1 - nt + H_R(t) = 1 - nt + rt^2 < 0$, and therefore the algebra $A = \langle X \mid R\rangle$ is infinite-dimensional.

Let $\mathbb{K}$ be countable, and observe that $\mathbb{K}\langle X\rangle$ is also countable. Denote by $\mathbb{K}_0\langle X\rangle$ the ideal of $\mathbb{K}\langle X\rangle$ consisting of all elements with 0 constant term and let $a_1, a_2, \ldots$ be an enumeration of all the elements of $\mathbb{K}_0\langle X\rangle$. Finally, recursively define $R \subseteq \mathbb{K}_0\langle X\rangle$ as follows. Let $R_1 \subseteq \mathbb{K}_0\langle X\rangle$ be the set of all homogeneous components of $(a_1)^2$ and let $n_2 \in \mathbb{N}$ be greater than the degrees of all elements in R_1.

Example 6.25. If $a_1 = x_1 + x_2x_3$, then $a_1^2 = x_1^2 + x_1x_2x_3 + x_2x_3x_1 + x_2x_3x_2x_3$, so $R_1 = \{x_1^2, x_1x_2x_3 + x_2x_3x_1, x_2x_3x_2x_3\}$. Here we can take $n_2 \geq 5$.

Suppose we have defined $R_i \subseteq \mathbb{K}_0\langle X\rangle$ and $n_{i+1} \in \mathbb{N}$. Then we define $R_{i+1} \subseteq \mathbb{K}_0\langle X\rangle$ as the set of all homogeneous components of $(a_1)^2, (a_2)^{n_2}, \ldots, (a_{i+1})^{n_{i+1}}$ and we choose $n_{i+2} \in \mathbb{N}$ greater than the degrees of all elements in R_{i+1}. Note that $R_i \subseteq R_{i+1}$. We then set $R := \bigcup_{i\geq 1} R_i$.

Remark 6.26. The choice of starting with a_1^2 is made in order to ensure that all the elements of R have degree at least 2 (recall that the a_i's have zero constant term).

Set

$$B := \mathbb{K}_0\langle X\rangle/(R).$$

We first notice that B is nil. Indeed, R contains all the homogeneous components of $(a_i)^{n_i}$ for all $i \geq 1$, so that every element of B is nilpotent. Note that B is clearly finitely generated, so in order to prove that B is a counterexample to the Kurosh–Levitzky problem, we only need to show that it is infinite-dimensional.

Consider now the (graded) algebra $A := \mathbb{K}\langle X\rangle/(R)$ and observe that $A \cong \mathbb{K}1_A \oplus B$ as vector spaces, so that B is infinite-dimensional if and only if A is. Now, to prove that $\dim_{\mathbb{K}} A$ is infinite, it will be enough to show that there exists a t_0 such that $1 - nt_0 + \sum_{i\geq 2} r_i t_0^i < 0$. Recall that r_i is the number of homogeneous elements of degree i in R. By construction, the r_i's are either 0 or 1, hence we can assume that $r_i = 1$ for all i, since $\sum_{i\geq 2} r_i t^i \preceq \sum_{i\geq 2} t^i$.

Now $1 - nt + \sum_{i\geq 2} t^i$ converges for all $0 < t < 1$ and for these values of t we have

$$1 - nt + \sum_{i\geq 2} t^i = 1 - nt + \frac{t^2}{1-t} = \frac{(n+1)t^2 - (n+1)t + 1}{1-t}.$$

Consider the inequality

$$\frac{(n+1)t^2 - (n+1)t + 1}{1-t} < 0. \tag{6.3}$$

The discriminant of the quadratic polynomial at the numerator of (6.3) is positive for $n \geq 4$ and, in this case, the corresponding roots are

$$\alpha^{\pm} := \frac{n+1 \pm \sqrt{(n+1)^2 - 4(n+1)}}{2(n+1)}.$$

Note that $0 < \alpha^- < \alpha^+ < 1$ and that (6.3) is satisfied for every $\alpha^- < t < \alpha^+$. It follows that we can find $0 < t_0 < 1$ such that $1 - nt_0 + H_R(t_0) < 0$.

This completes the proof that A is infinite-dimensional over $\mathbb{K}$, so that B is a counterexample to the Kurosh–Levitzky problem.

Proof of the Golod–Shafarevich theorem (Theorem 6.23). It is straightforward to check the inequalities implicit in (6.2) for the constant term and the coefficient of t. Indeed for the constant term it reduces to $1 \geq 1$, while for the coefficient of t, it reduces to $\dim_{\mathbb{K}} A_1 - n = n - n \geq 0$.

To check it for the other coefficients, we proceed as follows. Let $\mathbb{K}\langle X\rangle = \bigoplus_{i\geq 0} \mathbb{K}\langle X\rangle_i$ be the decomposition in homogeneous components, so that we have $\dim_{\mathbb{K}} \mathbb{K}\langle X\rangle_i = n^i$. Set $I := (R)$, and denote by $I = \bigoplus_{i\geq 2} I_i$ the decomposition of the graded ideal I into homogeneous components (note that $I_i = I \cap \mathbb{K}\langle X\rangle_i$ for all i). Also set $a_i := \dim_{\mathbb{K}} A_i$ and observe that we have $\dim_{\mathbb{K}} I_i = n^i - a_i$. Moreover, set $R_i := R \cap \mathbb{K}\langle X\rangle_i$ and $r_i = |R_i|$ for all i. For every i we choose a subspace $\widetilde{A}_i$ such that $\mathbb{K}\langle X\rangle_i = I_i \oplus \widetilde{A}_i$, so that $\dim_{\mathbb{K}} \widetilde{A}_i = a_i$.

We clearly have $I = \mathbb{K}\langle X\rangle R + IX$. It follows that

$$I_s = \sum_{i=2}^{s} \mathbb{K}\langle X\rangle_{s-i} R_i + I_{s-1} X$$

for all $s \geq 2$. In fact, we have

$$I_s = \sum_{i=2}^{s} \widetilde{A}_{s-i} R_i + I_{s-1} X$$

for $s \geq 2$. To see this, it is enough to observe that $I_{s-i} R_i \subseteq I_{s-1} X$ for $i \geq 2$, so that

$$\begin{aligned} I_s &= \sum_{i=2}^{s} \mathbb{K}\langle X\rangle_{s-i} R_i + I_{s-1} X = \sum_{i=2}^{s} (I_{s-i} \oplus \widetilde{A}_{s-i}) R_i + I_{s-1} X \\ &= \sum_{i=2}^{s} \widetilde{A}_{s-i} R_i + \sum_{i=2}^{s} I_{s-i} R_i + I_{s-1} X = \sum_{i=2}^{s} \widetilde{A}_{s-i} R_i + I_{s-1} X. \end{aligned}$$

Hence, looking at the dimension over $\mathbb{K}$ of these subspaces, we deduce that

$$n^s - a_s \leq \sum_{i=2}^{s} a_{s-i} r_i + (n^{s-1} - a_{s-1}) n,$$

that is,

$$a_s + \sum_{i=2}^{s} a_{s-i} r_i - n a_{s-1} \geq 0.$$

It remains only to notice that, for $s \geq 1$, the coefficient of t^s in $H_A(t)(1 - nt + H_R(t))$ is exactly $a_s - n a_{s-1} + \sum_{i=2}^{s} a_{s-i} r_i$. □

We are now going to use our counterexample to the Kurosh–Levitzky problem to produce a counterexample to the General Burnside Problem.

Let $\mathbb{K}$ be a countable or finite field of characteristic $ch(\mathbb{K}) = p > 0$. Consider the algebra A that we just constructed. We have $A = \mathbb{K}1 \oplus B$ as vector spaces, where $1 = 1_A$ is the unit of the algebra A, and $B = \mathbb{K}_0\langle X\rangle/(R)$ is an infinite dimensional nil algebra. Consider the set $A^\times$ of all invertible elements of A.

Now for every $b \in B$, the element $1+b \in A$ is invertible, and in fact has finite order. Indeed, since B is nil, there exists an $\ell \geq 1$ such that $b^{p^\ell} = 0$; hence $(1+b)^{p^\ell} = 1 + b^{p^\ell} = 1$ (here we are using $ch(\mathbb{K}) = p > 0$). It follows that $1 + B \subseteq A^\times$.

Let us denote again by x_i the cosets $x_i + (R)$ in $A \equiv \mathbb{K}\langle X \rangle/(R)$, where $x_i \in X$. Then the elements $1 + x_i$, $i = 1, 2, \ldots, n$ are invertible and torsion. Consider the subgroup $G \subseteq 1 + B \subseteq A^\times$ generated by $\{1+x_1, 1+x_2, \ldots, 1+x_n\}$. This is clearly a finitely generated torsion group.

Theorem 6.27. *G is infinite.*

Proof. Suppose not and assume $|G| = d$. Then every element g of G can be expressed as

$$g = (1+x_{i_1})(1+x_{i_2})\cdots(1+x_{i_r})$$

with $r < d$ (we don't need the inverses, since $(1+x_j)^{k_j} = 1$ for some k_j). Indeed, suppose that $g = (1+x_{j_1})\cdots(1+x_{j_\ell})$ with ℓ minimal and $\ell \geq d$. Then at least two of the $d+1$ elements

$$1, 1+x_{j_1}, (1+x_{j_1})(1+x_{j_2}), \ldots, (1+x_{j_1})(1+x_{j_2})\cdots(1+x_{j_d})$$

must be equal, say $(1+x_{j_1})(1+x_{j_2})\cdots(1+x_{j_h})$ and $(1+x_{j_1})(1+x_{j_2})\cdots(1+x_{j_k})$, where $0 \leq h < k \leq d$. But then

$$g = (1+x_{j_1})(1+x_{j_2})\cdots(1+x_{j_h})(1+x_{j_{k+1}})(1+x_{j_{k+2}})\cdots(1+x_{j_\ell})$$

is a product of $h + (\ell - k) < \ell$ generators, contradicting the minimality of ℓ. So $\ell < d$.

Let us show that the set of products $\{x_{j_1}x_{j_2}\cdots x_{j_r} : 1 \leq j_i \leq n, 1 \leq i \leq r < d\}$ spans B. To do so, it is sufficient to prove that every word $w = x_{i_1}x_{i_2}\cdots x_{i_d}$ of length d is a linear combination of shorter words.

We have

$$(1+x_{i_1})(1+x_{i_2})\cdots(1+x_{i_d}) = (1+x_{j_1})(1+x_{j_2})\cdots(1+x_{j_r})$$

with $r < d$ as we have just shown. Keeping the factor $x_{i_1}x_{i_2}\cdots x_{i_d}$ on the left-hand side, and bringing everything else to the right-hand side, gives the desired expression of w as a linear combination of shorter words.

It follows that $\dim_{\mathbb{K}} A = 1 + \dim_{\mathbb{K}} B \leq 1 + (1 + n + \cdots + n^{d-1})$, contradicting the fact that A is infinite-dimensional. Therefore G must be infinite. □

6.6 Notes

William Burnside [47] solved the Bounded Burnside Problem for linear groups. Issai Schur [305] proved the General Burnside Problem for linear groups. The Bounded Burnside Problem has been checked for exponent $n = 2$ (trivial: abelian groups), $n = 3$ (Burnside [46]), $n = 4$ (Ivan N. Sanov [302]) and $n = 6$ (Marshall Hall [153]).

In 1964 Evgenii S. Golod and Igor R. Shafarevich [125] constructed a 2-generated infinite p-group, thus providing a counterexample to the General Burnside Problem.

In 1980 Rostislav I. Grigorchuk [129] constructed his renowned group of intermediate growth which, among other most important properties, provides a negative solution to the General Burnside Problem. See the Notes to Chapter 7 for more on the Grigorchuk group.

In 1968 Pëtr S. Novikov and Sergei I. Adyan [258] found a counterexample to the Bounded Burnside Problem for all odd exponents $n \geq 4381$.

In 1992 both Sergei V. Ivanov and Igor Lysënok announced a counterexample to the Bounded Burnside Problem for all but finitely many exponents: Ivanov [188] for $n \geq 2^{48}$ and Lysënok [220] for $n \geq 8000$.

In 1980 Alexander Yu. Olshanskii [260] constructed the so-called Tarski monsters. A *Tarski monster* is an infinite group G such that every proper subgroup H of G, other than the identity subgroup, is a cyclic group of order a fixed prime number p. Such a group G is necessarily finitely generated. In fact it is clearly generated by every two non-commuting elements. Then Olshanskii showed that there is a Tarski p-group for every prime $p > 10^{75}$.

In 1991 Efim I. Zelmanov [361, 362] gave a positive solution to the Restricted Burnside Problem.

The Kurosh–Levitzky problem goes back to Alexander G. Kurosh [206] and Jakob Levitzky [214] in the early 1940s.

For a comprehensive relatively recent account on the Burnside problem, we also refer to Adyan's survey [3].

6.7 Exercises

Exercise 6.1. Show that the class of locally finite groups is closed under taking subgroups, homomorphic images, and finite direct products.

Exercise 6.2. Let $\mathbb{K}$ be a field. Show that the algebra $M_n(\mathbb{K})$ is simple.

Exercise 6.3. Show that a finitely-dimensional algebra which admits a linear basis consisting of nilpotent elements is nilpotent. This proves Wedderburn theorem (Theorem 6.17).