Journal of Algebra and Its Applications
Vol. 10, No. 4 (2011) 675–686
© World Scientific Publishing Company
DOI: 10.1142/S0219498811004690

World Scientific
www.worldscientific.com

CHARACTERS AND SOLUTIONS TO EQUATIONS IN FINITE GROUPS

ALON AMIT[*,‡] and UZI VISHNE[†,§]

[*]*Facebook Inc., Palo Alto, CA, USA*

[†]*Department of Mathematics, Bar-Ilan University*
Ramat-Gan 52900, Israel
[‡]*alon.amit@facebook.com*
[§]*vishne@math.biu.ac.il*

Received 31 March 2010
Accepted 15 July 2010

Communicated by I. M. Isaacs

The number of ways an element of a finite group can be expressed as a square, a commutator, or more generally in the form $w(x_1, \dots, x_r)$, where w is a word in the free group, defines a natural class function. We investigate some properties of these class functions, in particular their tendency to be characters or virtual characters of the underlying group. Generalizing classical results of Frobenius and others, we prove that generalized commutators yield characters in this manner, and use this to exhibit a criterion for nilpotency based on a certain equation associated with the irreducible characters.

Keywords: Equations in groups; characters; repeated commutators; nilpotency conditions.

Mathematics Subject Classification: Primary: 20D60; Secondary: 20F69, 20F18, 60B15

1. Introduction

Let G be a finite group. For every $h \in G$, let $N_{x^2}(h)$ denote the number of solutions to $x^2 = h$ in the group. This is obviously a class function, and a celebrated theorem of Frobenius and Schur asserts that N_{x^2} is a difference of characters; moreover the inner product of N_{x^2} with any irreducible character is 0, 1 or -1 [7, Chap. 4]. This result was extended in various directions, e.g. to n-powers (see below), and recently to norms with respect to an automorphism, in [3].

More generally, let w be a word in the free group $\mathbb{F}_r$ on r letters. Substitution $(g_1, \dots, g_r) \mapsto w(g_1, \dots, g_r)$ defines a natural function from r-tuples in G to G itself. We let $N_{w,G}(h)$ denote the number of solutions to $w(x_1, \dots, x_r) = h$ for $x_1, \dots, x_r \in G$ (if G is clear from the context, we sometimes omit it and write $N_w(h)$). Any automorphism $\sigma \in \mathrm{Aut}(G)$ carries the set of solutions of the equation $w(x_1, \dots, x_r) = h$ onto the set of solutions to $w(x_1, \dots, x_r) = \sigma(h)$, showing that

N_w is invariant under σ. In particular N_w is a class function, and so can be expressed as a linear combination of the irreducible characters of G. It is natural to investigate the coefficients in those expansions.

It was proved by Frobenius [5, Article 78, pp. 394–403 (1908)] that N_{x^m} is always a virtual character (see also [7, Problem 4.7]). Frobenius also proved (see [5, pp. 1–37]) that the function $N_{[x,y]}(h)$ is a character. This was generalized in [13], where $N_{x_1x_2\dots x_nx_1^{-1}x_2^{-1}\dots x_n^{-1}}$ is computed explicitly, and shown to always be a character; the author uses this computation to give a new proof for Itô's theorem. Serre shows in [10, Chap. 7] how counting solutions of similar type (restricted to certain conjugacy classes) can be applied to the inverse Galois problem. Various additional results of this kind, especially for the symmetric groups, can be found in [12, Examples 7.67–7.69].

Recall that a group whose character table entries are rational integers, is called rational. Following Corollary 3.4 below, we call a finite group G *semi-rational* if $N_{w,G}$ is a virtual character for every word w, and analogously say that w is semi-rational, if $N_{w,G}$ is a virtual character for any finite group G. After presenting some general reduction techniques and disposing of the abelian case in Sec. 2, we give in Sec. 3 various sufficient conditions for a group to be semi-rational, and point out that $\mathrm{SL}_2(7)$ is not semi-rational.

In Sec. 4 we change perspective, and show that the generalized commutators $\gamma_r = [x_1, x_2, \dots, x_r]$ are semi-rational (this was independently proved in [1]). From this one can conclude, for example, that in an odd p-group G the number of solutions for $[[[x_1, x_2], x_3], x_4] = h^2$ is equal to the number of solutions for $[[[x_1, x_2], x_3], x_4] = h$, for any element $h \in G$ (see Corollary 4.14). We observe that the limiting behavior of the corresponding counting functions is independent of G (so that all finite groups are "probabilistically nilpotent"), and present a criterion for nilpotency involving properties of the matrix $A_{\chi,\psi} = \frac{\langle \chi\psi, \chi\rangle}{\psi(1)}$ constructed from the character table of the group (Corollary 4.12).

Another way to view N_w is as a distribution function: if $X_1, \dots, X_r$ are uniformly and independently distributed random variables with values in G, then $|G|^{-r}N_w$ is the distribution function of the random variable $w(X_1, \dots, X_r)$. The fact that $|G|^{-r}N_{\gamma_r}$ tends to the delta distribution concentrated on the identity element, when $r \to \infty$, can be used to show (Theorem 4.10) that the matrix A has spectral radius 1.

We fix notation, as follows. By $\hat{G}$ we denote the set of irreducible characters of G, and $\widehat{\mathbb{Z}}$ is the $\mathbb{Z}$-module generated by all the entries in the character table of G. Letting ω denote a primitive root of unity of order $|G|$, we always have that $\widehat{\mathbb{Z}} \subseteq \mathbb{Z}[\omega]$. The standard inner product of class functions is denoted by

$$\langle a, b\rangle = \frac{1}{|G|} \sum a(g)\bar{b(g)}, \tag{1}$$

so if $f = \sum \alpha_\chi \chi$ is a class function (summing over the irreducible characters of G), the coefficients are $\alpha_\chi = \langle f, \chi\rangle$. In particular, for a word $w \in \mathbb{F}_r$, $N_w = \sum \alpha_\chi \chi$

where

$$\alpha_\chi = \frac{1}{|G|} \sum_{(g_1,\ldots,g_r)\in G^r} \bar{\chi}(w(g_1,\ldots,g_r)). \tag{2}$$

We say that N_w is integral if all the coefficients α_χ are in $\mathbb{Z}$.

2. Reduction of Words and Groups

Consider a word w which is a product $w'w''$ with disjoint sets of variables. In this case we write $w = w' * w''$. Since $w'w'' = h$ iff $w' = k$ and $w'' = k^{-1}h$ for some $k \in G$, and since w' and w'' are independent by assumption, we have the convolution formula $N_{w'*w''} = N_{w'} * N_{w''}$, namely

$$N_{w'*w''}(h) = \sum_{k\in G} N_{w'}(k) N_{w''}(k^{-1}h). \tag{3}$$

By the generalized orthogonality relation for characters [7, Theorem 2.13], the convolution $\chi * \chi'$ of irreducible characters is zero if $\chi' \neq \chi$ and is equal to $\frac{|G|}{\chi(1)}\chi$ (an integral multiple of χ) if $\chi' = \chi$.

Corollary 2.1. *If $N_{w'}$ and $N_{w''}$ are integral, then so is $N_{w'*w''}$.*

Occasionally, N_w can be analyzed more easily by bringing w to a more convenient form. If w and w' belong to the same orbit under the action of $\mathrm{Aut}(\mathbb{F}_r)$, then $N_{w,G} = N_{w',G}$ for every finite group G. For example, $x_1x_2^{-1}x_1x_2$ can be transformed to $x_1^2x_2^2$ by sending x_1 to x_1x_2, and so $N_{x_1x_2^{-1}x_1x_2} = N_{x_1^2} * N_{x_1^2}$.

Question 2.2. Suppose $N_{w,G} = N_{w',G}$ for every finite group G. Does it follow that w' is mapped to w by some automorphism of $\mathbb{F}_r$?

This question can be refined for a fixed group G. Let the rank $r \geq 1$ also be fixed. The *group of identities* of G, denoted here by $K(G)$, is the set of elements $w \in \mathbb{F}_r$ which are mapped to 1 by every homomorphism $\phi : \mathbb{F}_r \to G$. It is thus the intersection of all kernels of such homomorphisms. Since G is finite, there are only finitely many homomorphisms $\mathbb{F}_r \to G$, so the group of identities is a finite index characteristic subgroup of $\mathbb{F}_r$. Finally if w and w' fall into the same class in the quotient $\mathbb{F}_r/K(G)$, then they represent the same element in every substitution, and in particular $N_{w,G} = N_{w',G}$.

Question 2.3. Suppose $N_{w,G} = N_{w',G}$ for a fixed group G. Does it follow that w' can be mapped by an automorphism of $\mathbb{F}_r$ to some $w'' \in \mathbb{F}_r$ which is equivalent to w modulo $K(G)$?

Allowing for the automorphism here is not redundant: If x_1 and x_2 are two of the generators of $\mathbb{F}_r$, then $N_{x_1} = N_{x_2}$ for every group G, but $x_1^{-1}x_2 \notin K(G)$ unless G is trivial.

When G is abelian of exponent e, it is easy to verify that $K(G)$ is the subgroup of $\mathbb{F}_r$ generated by commutators and e powers; therefore $\mathbb{F}_r/K(G) \cong (\mathbb{Z}/e\mathbb{Z})^r$. Rewriting a word w modulo $K(G)$, it is equivalent to a word of the form $x_1^{\alpha_1} \cdots x_r^{\alpha_r}$ where $0 \leq \alpha_1, \ldots, \alpha_r < e$. To complete the analysis in the abelian case, we only need to compute the words N_{x^α}, and then apply Eq. (2).

Proposition 2.4. *When G is abelian, N_{x^α} is a sum of distinct irreducible characters.*

Proof. The exponentiation map $g \mapsto g^\alpha$ is a homomorphism; we denote the image by G^α and the kernel by G_α. Then $G/G_\alpha \cong G^\alpha$, so

$$\langle N_{x^\alpha}, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \psi(\bar{g}^{\alpha}) = \frac{1}{|G^\alpha|} \sum_{h \in G^\alpha} \psi(\bar{h}) = \langle 1, \psi_{G^\alpha} \rangle$$

which is either zero or one. □

3. Integrality of the Coefficients

Let $w \in \mathbb{F}_r$ and let N_w be the induced counting function, as defined in the introduction. The following property of the functions N_w is due to Solomon (see also [6]).

Proposition 3.1 ([11, Theorem 1]). *Assume $r > 1$. For every $g \in G$, the value $N_w(g)$ is divisible by $|\mathrm{C}_G(g)| = |G|/|[g]|$.*

Let χ be an irreducible character of G. Writing N_w as a linear combination of the characters, the coefficient $\langle N_w, \chi \rangle$ of χ (see Eq. (1)) can be written as

$$\frac{1}{|G|} \sum_{(g_1,\ldots,g_r) \in G^r} \bar{\chi}(w(g_1, \ldots, g_r)).$$

The following argument, building on Proposition 3.1, is due to Stanley ([12, Exerise 7.69.j]).

Proposition 3.2. *Assume $r > 1$. For every $w \in \mathbb{F}_r$, the coefficients of N_w are all in $\widehat{\mathbb{Z}}$.*

Proof. For a conjugacy class C, let χ_C denote the characteristic function of C, namely $\chi_C(g) = 1$ if $g \in C$ and $\chi_C(g) = 0$ otherwise. Obviously N_w is an integral combination of the χ_C. In fact, by Proposition 3.1, N_w is an integral combination of the functions $\frac{|G|}{|C|}\chi_C$.

Direct computation with (1) gives $\frac{|G|}{|C|}\chi_C = \sum_\psi \psi(C)\bar{\psi}$ (where $\psi(C)$ stands for the value $\psi(g)$ for some $g \in C$), and therefore N_w is a $\widehat{\mathbb{Z}}$-linear combination of the irreducible characters. □

When all the coefficients $\langle N_w, \chi \rangle$ are rational integers, N_w is a virtual character of G. From the proposition we know that these coefficients are in $\mathbb{Z}[\omega]$ (where ω

denotes a primitive root of unity of degree $|G|$). Therefore, they are rational integers if and only if they are invariant under the Galois group of $\mathbb{Q}[\omega]/\mathbb{Q}$.

If $\sigma \in \mathrm{Gal}(\mathbb{Q}[\omega]/\mathbb{Q})$, then $\sigma(\omega) = \omega^s$ for some s prime to $|G|$ (and every such s defines a corresponding automorphism). Let $m_s : G \to G$ be the exponentiation by s, which is one-to-one and onto. Also let s' be the inverse of s modulo $|G|$. For every $h \in G$, $\rho(h)$ is diagonalizable with eigenvalues that are powers of ω, and so we have that $\sigma(\chi(h)) = \chi(h^s)$. Therefore

$$\begin{aligned}\sigma(\langle N_w, \chi\rangle) &= \langle \sigma \circ N_w, \sigma \circ \chi\rangle \\ &= \langle N_w, \sigma \circ \chi\rangle \\ &= \frac{1}{|G|}\sum_{h\in G} N_w(h)\bar{\chi}(h^s) \\ &= \frac{1}{|G|}\sum_{h'\in G} N_w(m_{s'}(h'))\bar{\chi}(h') \\ &= \langle N_w \circ m_{s'}, \chi\rangle. \end{aligned} \tag{4}$$

Corollary 3.3. *N_w is a virtual character iff $N_w(h) = N_w(h')$ whenever h and h' generate the same subgroup of G.*

If such h and h' are always conjugate in G (the most well known example being S_n), then the character table consists of integers, so that in our notation $\widehat{\mathbb{Z}} = \mathbb{Z}$. Such groups are called rational, and from Proposition 3.2 we have the following corollary.

Corollary 3.4. *Every* (*finite*) *rational group is semi-rational.*

We emphasize that N_w is invariant under any automorphism of G. This implies that the coefficient of χ in N_w is equal to the coefficient of any $\chi \circ \sigma$ for $\sigma \in \mathrm{Aut}(G)$. Moreover, by Corollary 3.3, we get the following proposition.

Proposition 3.5. *If h and h' lie in the same orbit of* $\mathrm{Aut}(G)$ *whenever they generate the same cyclic subgroup, then G is semi-rational.*

Proposition 3.5 is strictly stronger than the claim involving rational groups; for example, the alternating groups satisfy this condition but are not rational. Other irrational examples can be detected using the following observation.

Corollary 3.6. *In particular, if every two elements of the same order are conjugate in G, then G is semi-rational.*

Example 3.7. In $G = \mathrm{SL}_2(5)$ there are 7 orbits under the conjugation action of $\mathrm{GL}_2(5)$ (I, $-I$, and the five non-scalar orbits, characterized by trace), each orbit with its own orbit of elements: $1, 2, 3, 4, 5, 6$ and 10. This proves that $\mathrm{SL}_2(5)$ is semi-rational.

On the other hand, direct enumeration of the values of $N_{[x,[x,y]]}$ for the group $\mathrm{SL}_2(7)$, shows that it is not a virtual character, so this group is not semi-rational.

Abelian groups are almost never rational (except for the elementary abelian groups of exponent 2). However, we still have the following example.

Example 3.8. All abelian groups are semi-rational.

This is because exponentiation by s prime to $|G|$ is an automorphism when the group is abelian. Also see Proposition 2.4.

We do not know if the converse to Proposition 3.5 holds.

Question 3.9. Let G be a semi-rational group. Does it follow that every two elements h, h' generating the same subgroup of G are in the same orbit of $\mathrm{Aut}(G)$?

4. Generalized Commutators

Words in the commutator subgroup of $\mathbb{F}_r$ are of special interest, in light of the classical result that any finitely generated verbal quotient of the free group can be presented with (at most) one relation of the form $x_1^d = 1$, and all others in the commutator subgroup [9, Theorem 2.3].

Define the iterated commutators by $\gamma_1 = x_1$ and $\gamma_k = [x_k, \gamma_{k-1}]$ for $k \geq 2$, where $x_1, x_2, \dots$ are generators of a free group. We first observe that every finite group is "probabilistically nilpotent".

Remark 4.1. Let $X_1, X_2, \dots$ be independent random variables over G, such that $\Pr\{X_k = 1\}$ is bounded away from zero for k large enough. Let $W_k = \gamma_k(X_1, \dots, X_k) = [W_{k-1}, X_k]$ be the iterated commutators as before. Then $\Pr\{W_k = 1\} \to 1$ as $k \to \infty$.

Indeed, $W_k \neq 1$ implies $X_1, \dots, X_k \neq 1$, with exponentially declining probability.

This observation will later be translated to a statement on the norm of a certain matrix computed from character values. But first we need to compute the distribution of a commutator explicitly.

Let $\phi : \mathbb{C}[G] \to \mathbb{C}$ be the augmentation map defined by $\phi(g) = 1$ for every $g \in G$. It is convenient to work with "distribution elements" $a \in \mathbb{C}[G]$, which satisfy $\phi(a) = 1$. If $a = \sum a_g g$ and the coefficients a_g are all positive real numbers, then a defines a distribution on the group by $\Pr\{X = g\} = a_g$. One advantage of this notation is that the product of two distribution elements in the group algebra is again a distribution element, representing the convolution of the two distributions. In particular, for a conjugacy class C, we let $\hat{C} = \frac{1}{|C|}\sum_{g \in C} g$, the uniform distribution on C. Distributions which are uniform on classes are represented by central distribution elements, those of the form $\sum p_C \hat{C}$ with $\sum p_C = 1$; if X has this distribution, then $\Pr\{X \in C\} = p_C$, while $\Pr\{X = x\} = \frac{p_{[x]}}{|[x]|}$. In particular, the uniform distribution on the group is represented by the element $\sum \frac{|C|}{|G|}\hat{C}$.

Suppose that X has a central distribution, so the probability $x \mapsto \Pr\{X = x\}$ is a class function. We can then write it (uniquely) as a linear combination of the irreducible characters, so for suitable coefficients q_ψ, we have

$$\Pr\{X = x\} = \frac{1}{|G|} \sum_{\psi} \frac{q_\psi}{\psi(1)} \psi(x), \tag{5}$$

where $\psi(1)$ is the dimension of the representation associated to ψ. Lemma 4.6 explains why it is reasonable to write the coefficients in this form. Recall that $\hat{G}$ denotes the set of irreducible representations of G.

Definition 4.2. Let $q : \hat{G} \to \mathbb{C}$ be a function. The distribution defined by Eq. (5) (if this is indeed a distribution, namely the entries are positive and sum to 1), is denoted by $P(q)$.

Notice that $q_1 = 1$ is a necessary condition, as seen by summing over all $x \in G$.

Example 4.3. Let $\delta_1 : \hat{G} \to \mathbb{R}$ be the delta function, $\delta_1(1) = 1$ and $\delta_1(\psi) = 0$ for $\psi \neq 1$. For this vector equation (5) becomes $\Pr\{X = x\} = 1/|G|$, so $P(\delta_1)$ is the uniform distribution on G.

Let $A : \hat{G} \times \hat{G} \to \mathbb{C}$ denote the matrix

$$A_{\chi,\psi} = \frac{\langle \chi\psi, \chi \rangle}{\psi(1)}. \tag{6}$$

We use the indexing by $\hat{G}$ to multiply A by vectors such as the above q (namely $(A \cdot q)_\chi = \sum_{\psi \in \hat{G}} A_{\chi\psi} q_\psi$). Notice that A is a real (in fact rational) matrix, as $\langle \chi\psi, \chi \rangle$ counts the components isomorphic to χ in the tensor product of the representations associated to χ and ψ.

Example 4.4. (1) If χ is linear, the row associated with χ in A is $A_{\chi,\psi} = \delta_{\psi,1}$. In particular the first row of A is always $1, 0, \dots, 0$, and for abelian groups $A_{\chi,\psi} = \delta_{\psi,1}$ so A has rank 1 and $A^2 = A$.

(2) Let $n = 2k+1$ be odd. The dihedral group $G = D_n$ has 2 linear representations, 1 and ϵ, and k two-dimensional ones, with characters satisfying $\chi^2 = 1 + \epsilon + \chi$. The matrix A is thus $(2+k) \times (2+k)$, with a 2×2 block $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, a $2 \times k$ block of zeros, a $k \times 2$ block of ones, and the remaining $k \times k$ block being the identity matrix. The eigenvalues are $0, 1$ and $\frac{1}{2}$ (the latter with multiplicity k).

(3) In Theorem 4.10 we show that the spectral radius of A is 1. To put this in perspective, we exhibit the spectrum in some explicit cases. For $G = S_4$, the eigenvalues of A are $0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1$. For the alternating group $G = A_5$, the characteristic polynomial is $\frac{1}{120}(\lambda - 1)(\lambda - \frac{1}{3})(120\lambda^3 - 118\lambda^2 + 10\lambda - 4)$, with a real zero around $\lambda = 0.932$ and two complex ones (with absolute values around 0.189). For $G = S_5$, the characteristic polynomial of A is $\frac{1}{120}\lambda^3(\lambda - 1)(120\lambda^3 - 98\lambda^2 - 13\lambda - 2)$, with one real root around 0.949 and two complex ones with absolute values around 0.132.

Proposition 4.5. *Let $u^* : \hat{G} \to \mathbb{R}$ be defined by $u^*_\psi = \psi(1)^2$. Then $Au^* = u^*$.*

Proof. Compute

$$\begin{aligned}(A \cdot u^*)_\chi &= \sum_\psi A_{\chi,\psi} u_\psi \\ &= \sum_\psi \langle \chi\psi, \chi\rangle \psi(1) \\ &= \sum_\psi \langle \psi, \bar{\chi}\chi\rangle \psi(1) \\ &= \dim(\bar{\chi}\chi) = \chi(1)^2 = u^*_\chi.\end{aligned}$$ □

Notice that taking u^* as the vector of coefficients, Eq. (5) becomes

$$\begin{aligned}\Pr\{X = x\} &= \frac{1}{|G|} \sum_\psi \psi(1)\psi(x), \\ &= \frac{1}{|G|} \sum_\psi \psi(x)\psi(\bar{1}) = \delta_{1,x},\end{aligned}$$

so $P(u^*)$ is the delta distribution, concentrated on the identity. Obviously, if X has this distribution, then so does $[U, X]$. The next lemma shows how this fact implies Proposition 4.5.

Lemma 4.6. *Let $q : \hat{G} \to \mathbb{C}$ be such that $P(q)$ is a distribution. If X has this distribution and U is uniform, then the commutator $[X, U] = XUX^{-1}U^{-1}$ has the distribution $P(A \cdot q)$.*

Proof. Recall the formula for multiplication of conjugacy classes,

$$\hat{C} \cdot \hat{C}' = \sum_{C''} \frac{|C''|}{|G|} \sum_{\chi \in \hat{G}} \frac{\chi(\hat{C})\chi(\hat{C}')\bar{\chi}(\hat{C}'')}{\chi(1)} \hat{C}'',$$

which can be restated as

$$\Pr\{XY \in C'' \mid X \in C, Y \in C'\} = \frac{|C''|}{|G|} \sum_{\chi \in \hat{G}} \frac{\chi(\hat{C})\chi(\hat{C}')\bar{\chi}(\hat{C}'')}{\chi(1)}. \tag{7}$$

Taking C as the conjugacy class of x in Eq. (5), our assumption on X can be written as

$$p_C = \Pr\{X \in C\} = \frac{|C|}{|G|} \sum_\psi \frac{q_\psi}{\psi(1)} \psi(\hat{C}).$$

Since U is uniform, the element $UX^{-1}U^{-1}$ is uniform on the conjugacy class of X^{-1}. Thus X and $UX^{-1}U^{-1}$ are independent given the conjugacy class of X. Now

$$\begin{aligned}
\Pr\{[X,U] \in C''\} &= \sum_C \Pr\{[X,U] \in C'' \,|\, X \in C\} \Pr\{X \in C\} \\
&= \sum_C \Pr\{XY^{-1} \in C'' \,|\, X \in C,\, Y \in C\} p_C \\
&= \sum_C \Pr\{XY \in C'' \,|\, X \in C,\, Y \in C^{-1}\} p_C \\
&= \frac{|C''|}{|G|} \sum_C \sum_{\chi \in \hat{G}} \frac{\chi(\hat{C})\bar{\chi}(\hat{C})\bar{\chi}(\hat{C}'')}{\chi(1)} p_C \\
&= \frac{|C''|}{|G|} \sum_{\chi \in \hat{G}} \sum_C (\chi(\hat{C})\bar{\chi}(\hat{C}) p_C) \frac{\bar{\chi}(\hat{C}'')}{\chi(1)}.
\end{aligned}$$

Letting $r : \hat{G} \to \mathbb{R}$ be the vector inducing the distribution of $[X,U]$, we get from the last computation that

$$\begin{aligned}
r_\chi &= \sum_C (\chi(\hat{C})\bar{\chi}(\hat{C}) p_C) \\
&= \sum_C \left(\chi(\hat{C})\bar{\chi}(\hat{C}) \frac{|C|}{|G|} \sum_\psi \frac{q_\psi}{\psi(1)} \psi(\hat{C}) \right) \\
&= \sum_\psi \frac{q_\psi}{\psi(1)} \frac{1}{|G|} \sum_C (|C| \chi(\hat{C})\bar{\chi}(\hat{C})\psi(\hat{C})) \\
&= s \sum_\psi \frac{\langle \chi\psi, \chi \rangle}{\psi(1)} q_\psi = (A \cdot q)_\chi.
\end{aligned}$$

□

By induction, we have the following corollary.

Corollary 4.7. *Suppose X_1 has the distribution $P(q)$, and $X_2, \ldots$ are uniform. Then the generalized commutator $W_k = [W_{k-1}, X_k]$ has distribution $P(A^{k-1} \cdot q)$.*

Remark 4.8. If X and Y are independent with central distributions, and the distribution of Y is symmetric with respect to inversion, then $[X,Y]$ and $[Y,X]$ have the same distribution. In particular in Lemma 4.6, $[U,X]$ has the same distribution as $[X,U]$.

Indeed, $[Y,X] = [YXY^{-1}, Y^{-1}]$ is conjugate to $[X, Y^{-1}]$, which has the same distribution as $[X,Y]$.

Let $\mathcal{C}$ denote the set of conjugacy classes of G. We define $T : \mathcal{C} \times \hat{G} \to \mathbb{C}$ by

$$T_{C,\psi} = \frac{|C|}{|G|} \frac{\psi(\hat{C})}{\psi(1)}.$$

Then Eq. (5) implies $\Pr\{X \in C\} = (T \cdot q)_C$, in other words, the distribution $P(q)$ is equal to $T \cdot q$. Notice that although A is a real matrix, T in general is not. Let $D : \mathcal{C} \times \mathcal{C} \to \mathbb{R}$ be the diagonal matrix defined by $D_{C,C} = |C|$, and let $E : \hat{G} \times \hat{G} \to \mathbb{R}$ be the diagonal matrix defined by $E_{\psi\psi} = \psi(1)$. The character table $\Xi_{C\psi} = \psi(\hat{C})$ is known to be invertible and $T = \frac{1}{|G|} D \Xi E^{-1}$, so T is invertible.

Proposition 4.9. *$TAT^{-1} : \mathcal{C} \times \mathcal{C} \to \mathbb{C}$ is a real matrix.*

Proof. Put $p = T \cdot q$ in Lemma 4.6: if $p : \mathcal{C} \to \mathbb{R}$ is the distribution vector of X, then $TAT^{-1} \cdot p$ is the distribution vector of $[U, X]$ (where U is uniform).

It follows that the vectors in the standard basis of $\mathbb{R}^{\mathcal{C}}$ (with one entry equal to 1, and all the other entries being 0) are all mapped to distribution vectors, which are real. □

Theorem 4.10. *The spectral radius of A is 1. Moreover, the eigenspace of the eigenvalue 1 is one-dimensional, and every other eigenvalue has absolute value <1.*

Proof. We prove the claim for TAT^{-1}.

Let $p : \mathcal{C} \to \mathbb{R}$ be a distribution vector. Let the variable X_1 have this distribution, and let $X_2, \ldots$ be uniform; then $W_k = \gamma_k(X_1, \ldots, X_k)$ has the distribution $(TAT^{-1})^{k-1}p$ by Corollary 4.7, and it converges to the delta distribution Tu^* by Remark 4.1.

Multiplying by an arbitrary constant, we conclude that for every positive real vector p, $(TAT^{-1})^k p$ converges to a multiple of Tu^*. But every real vector can be written as the difference of two positive vectors; therefore the claim holds for all real vectors.

Now let v denote any eigenvector of TAT^{-1} over $\mathbb{C}$, with eigenvalue λ. Since TAT^{-1} is real, the complex conjugate $\bar{v}$ is also an eigenvector, with an eigenvalue $\bar{\lambda}$. But $v + \bar{v}$ is real, so $(TAT^{-1})^k(v + \bar{v}) = \lambda^k v + \bar{\lambda}^k \bar{v}$ must converge. This proves $|\lambda| \leq 1$, and moreover if $|\lambda| = 1$ then necessarily $\lambda = 1$. But if v is an eigenvector with eigenvalue $\lambda = 1$ then it must be real, and so $v = (TAT^{-1})^k v$ converges to a multiple of Tu^*. □

We now get to the final result of this section, namely the explicit computation of N_{γ_k}.

Theorem 4.11. *The counting functions N_{γ_k} are characters. In particular, the generalized commutators γ_k are semi-rational.*

Proof. The counting function of X_1, namely $N_{x_1}(g) = 1$, is of course $|G|$ times the uniform distribution, so by Example 4.3 we can write $N_{x_1} = P(|G|\delta_1)$.

In general, if the counting function of $W = w(x_1, \ldots, x_{k-1})$ is $N_w = P(q)$ for a suitable $q : \hat{G} \to \mathbb{C}$, then the counting function of $[U, W]$ is $P(|G|A{\cdot}q)$ by Lemma 4.6, since $[U, W]$ involves one new (uniform) variable. Therefore, by induction on k, the counting function of γ_k is $|G|^2 \cdot P((|G|A)^{k-2} \cdot A\delta_1)$. Now, $A\delta_1$ is an integral positive vector. Indeed, let $J : \hat{G} \to \mathbb{R}$ be the constant vector defined by $J_\psi = 1$ for every $\psi \in \hat{G}$. Then $A \cdot \delta_1 = J$ since

$$\begin{aligned}(A \cdot \delta_1)_\chi &= \sum_\psi A_{\chi,\psi}(\delta_1)_\psi \\ &= A_{\chi,1} = \frac{\langle \chi, \chi \rangle}{\dim(1)} = 1.\end{aligned}$$

Note that $|G|A$ is an integral matrix with positive entries, as $\psi(1)$ always divides $|G|$, and the coefficients $\langle \chi\psi, \chi \rangle$ are positive integers. Therefore $(|G|A)^{k-2} \cdot J$ is again integral and positive. Finally by the definition in Eq. (5),

$$N_{\gamma_k}(x) = \sum_\psi \frac{|G|}{\psi(1)}((|G|A)^{k-2} \cdot J)_\psi \psi(x), \tag{8}$$

which is a character since the coefficients $|G|/\psi(1)$ are integral. □

Taking $x = 1$ in (8), we obtain

$$N_{\gamma_k}(1) = |G|^{k-1} J^{\mathrm{t}} A^{k-2} J. \tag{9}$$

By definition, G is nilpotent of class at most k iff $\gamma_{k+1}(G) = 1$, iff $N_{\gamma_{k+1}}(1) = |G|^{k+1}$. Equation (9) then provides a character condition for nilpotency, where $J_\psi = 1$ as above.

Corollary 4.12. *The group G is nilpotent of class at most k iff*

$$J^{\mathrm{t}} A^{k-1} J = |G|.$$

Example 4.13. (a) The case $k = 1$: G is abelian iff $|\hat{G}| = |G|$.
(b) The case $k = 2$: $[[G, G], G] = 1$ iff

$$\sum_{\chi,\psi \in \hat{G}} \frac{\langle \chi\psi, \chi \rangle}{\psi(1)} = |G|.$$

This can be restated as a condition on two class functions: $[[G, G], G] = 1$ iff

$$\left\langle \sum_{\chi \in \hat{G}} \chi\bar{\chi}, \sum_{\psi \in \hat{G}} \frac{\psi}{\psi(1)} \right\rangle = |\hat{G}|.$$

(c) The case $k = 3$: $[[[G, G], G], G] = 1$ iff

$$\sum_{\chi,\varphi,\psi \in \hat{G}} \frac{\langle \chi\varphi, \chi \rangle \langle \varphi\psi, \varphi \rangle}{\varphi(1)\psi(1)} = |G|.$$

In light of Corollary 3.3, the integrality of N_{γ_k} implies the following corollary.

Corollary 4.14. *Let G be a finite group and s an integer prime to $|G|$. Fix $k \geq 1$. For every $g \in G$, the number of solutions to $g = \gamma_k(x_1, \ldots, x_k)$ is equal to the number of solutions to $g^s = \gamma_k(x_1, \ldots, x_k)$.*

It would be interesting to have a bijective proof of this fact.

We conclude this section by noting that precise estimate of $N_{w,S_n}(1)$, namely the number of solutions to $w(x_1, \dots, x_t) = 1$ in the symmetric group, has applications to subgroup growth of the one relator group

$$\langle a_1, \dots, a_t \,|\, w(a_1, \dots, a_t) = 1 \rangle,$$

see [8, Sec. 14.4] for details. For example, a proof that $N_{\gamma_3,S_n}(1) = \sum_{\chi,\psi \in \widehat{S_n}} \frac{\langle \chi\psi, \chi \rangle}{\psi(1)}$ is of the same order of magnitude as the partition number $p(n) = |\widehat{S_n}|$ (as we expect), will show that $G = \langle a, b, c \,|\, [[a, b], c] = 1 \rangle$ has subgroup growth of a new type.

References

[1] A. M. A. Alghamdi and F. G. Russo, A generalization of the probability that the commutator of two group elements is equal to a given element, preprint (2010), arXiv:1004.0934v1.

[2] J. L. Alperin and R. Bell, *Groups and Representations*, Graduate Texts in Mathematics, Vol. 162 (Springer-Verlag, 1995).

[3] D. Bump and D. Ginzburg, Generalized Frobenius–Schur numbers, *J. Algebra* **278** (2004) 294–313.

[4] C. J. Colbourn and J. H. Dinitz (eds.), *Handbook of Combinatorial Designs* (CRC Press, 1996).

[5] G. Frobenius, *Gesammelte Abhandlungen Band III*, ed. J. P. Serre (Springer-Verlag, 1968).

[6] I. M. Isaacs, Systems of equations and generalized characters in groups, *Canadian. J. Math* **22** (1970) 1040–1046.

[7] I. M. Isaacs, *Character Theory of Finite Groups*, Pure and Applied Mathematics Series (Academic Press, 1976).

[8] A. Lubotzky and D. Segal, *Subgroup Growth*, Progress in Mathematics, Vol. 212 (Birkhäuser, 2003).

[9] W. Magnus, A. Karrass and D. Solitar, *Combinatorial Group Theory* (Interscience, 1966).

[10] J. P. Serre, *Topics in Galois Theory*, Notes written by Henri Darmon, Jones and Bartlett, 1992.

[11] L. Solomon, The solution of equations in groups, *Arch. Math.* **20** (1969) 241–247.

[12] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2 (Cambridge University Press, 1999).

[13] T. Tambour, The number of solutions of some equations in finite groups and a new proof of Itô's theorem, *Commun. Algebra* **28**(11) (2000) 5353–5362.