Words and Groups

Dan Segal

June 25, 2009

A *word* is an expression

$$w = w(\mathbf{x}) = \prod_{j=1}^{s} x_{i(j)}^{\varepsilon(j)}$$

$$i(1), \ldots, i(s) \in \{1, \ldots, k\}, \quad \varepsilon(j) = \pm 1 \ \forall j$$

(think of k as fixed and large)

The *verbal mapping* on a group G:

$$f_w : G^{(k)} \to G$$

$$\mathbf{g} f_w = w(\mathbf{g}) = \prod_{j=1}^{s} g_{i(j)}^{\varepsilon(j)}$$

where $\mathbf{g} = (g_1, \ldots, g_k)$.

Obviously f_w only depends on the equivalence class of w, i.e. the element represented by w in the free group F on $x_1, \ldots, x_k$:

$$\mathbf{g} f_w = w \pi_{\mathbf{g}}$$

where $\pi_{\mathbf{g}} : F \to G$ sends x_i to g_i $(i = 1, \ldots, k)$.

Given a group G, a *generalized word* over G is an expression

$$w = w(\mathbf{x}) = \prod_{l=1}^{t} x_{i_l}^{\varepsilon(l)\alpha(l)} \tag{1}$$

$i_1, \ldots, i_t \in \{1, \ldots, k\}$, $\varepsilon(l) = \pm 1$, $\alpha(l) \in \mathrm{Aut}(G)$

In this case

$$f_w(\mathbf{g}) = w(\mathbf{g}) = \prod_{l=1}^{t} g_{i_l}^{\varepsilon(l)\alpha(l)}.$$

Given a *finite* group G and a word w, there is a *positive* word w^* such that $f_w = f_{w^*}$ on $G^{(k)}$: supposing G has order m, we obtain w^* from w by replacing each occurrence of x^{-1} in w by x^{m-1}, for each variable x.

So when studying the map f_w on a given finite group, we may w.l.o.g. assume that w is positive.

Topics

1. *Fibres over finite groups*

2. *Ellipticity in profinite groups*

3. *Ellipticity in finite groups*

4. *Algebraic groups*

5. *Finite simple groups*

Notation

For a subset S of a group G and $m \in \mathbb{N}$,

$$S^{*m} = \{s_1 s_2 \dots s_m \mid s_i \in S\}.$$

For a word w in k variables,

$$G_w = \left\{w(\mathbf{g})^{\pm 1} \mid \mathbf{g} \in G^{(k)}\right\},$$

$$w(G) = \langle G_w \rangle$$

$$G_{+w} = G^{(k)} f_w = \left\{w(\mathbf{g}) \mid \mathbf{g} \in G^{(k)}\right\}.$$

The word w has *width* m in G if $w(G) = G_w^{*m}$, and *positive width* m if $w(G) = G_{+w}^{*m}$.

F denotes a free group on sufficiently many variables, sometimes infinitely many.

$\delta_l(G)$ denotes the lth term of the derived series of G

Fibres over finite groups

G *abelian* $\Rightarrow$ f_w a homomorphism $\Rightarrow$

$$\left|f_w^{-1}(g)\right| = \left|f_w^{-1}(1)\right| = \frac{|G|^k}{|w(G)|} \geq |G|^{k-1}$$

$$(g \in G_{+w})$$

Definition.

$$P(G, w = g) = \frac{\left|f_w^{-1}(g)\right|}{|G|^k} \quad (k >> 0)$$

this is the *probability that* w *represents* g.

$$P(G, w) = P(G, w = 1):$$

the probability that w represents 1. Thus

G abelian $\Rightarrow$

$$P(G, w = g) = P(G, w) \geq |G|^{-1} \qquad (2)$$

$(\forall w,\ \forall g \in G_{+w})$.

Suppose that G is *not nilpotent.* Let

$$w_n = [x_1, \ldots, x_n].$$

Then $w_n(G) \neq 1$ for each n, so for each n there exists $h_n \in G$ with $1 \neq h_n \in G_{+w_n}$. Then

$$\begin{aligned} \frac{1}{|G|^n} &\leq P(G, w_n = h_n) \\ &\leq \frac{(|G|-1)^n}{|G|^n} \longrightarrow 0 \text{ as } n \to \infty. \end{aligned}$$

Thus $P(G, w = h)$ *takes arbitrarily small positive values as* w *varies over all words, and the outer inequality in (2) is violated.*

Now suppose that G is not *soluble*. Then G has a just-non-soluble quotient Q; that is, Q contains a unique minimal normal subgroup

$$\delta_l(Q) = M = S_1 \times \cdots \times S_r,$$

$S_1, \ldots, S_r$ isomorphic non-abelian simple groups.

Lemma 1 *For each* $n \in \mathbb{N}$ *there is a word* w_n *in* n *variables such that for*
$\mathbf{g} = (g_1, \ldots, g_n) \in Q^{(n)}$,

$$w_n(\mathbf{g}) = 1 \Rightarrow \langle g_1, \ldots, g_n \rangle \neq Q. \qquad (3)$$

Proof to come.

Now, Let $P(Q, n)$ denote *the probability that a random* n*-tuple in* Q *generates* Q. Then:

$$P(Q, n) \geq 1 - m2^{-n} \qquad (4)$$

where m is the number of maximal subgroups of Q. To see this, put

$$Y = \left\{ \mathbf{g} \in Q^{(n)} \mid \langle g_1, \ldots, g_n \rangle \neq Q \right\}$$

and observe that

$$|Y| = \left| \bigcup_{L <_{\max} Q} L^{(n)} \right| \leq m \left(\frac{|Q|}{2} \right)^n .$$

Recall now that Q is a quotient of our group G; combining (3) and (4) we see that for $w = w_n$,

$$P(G, w) \leq P(Q, w) \leq 1 - P(Q, n) \leq m2^{-n}.$$

Thus *if* G *is not soluble,* $P(G, w)$ *takes arbitrarily small values, and (2) is violated.*

Proof of Lemma 1. W.l.o.g. $n \geq \mathrm{d}(G)$. Let F be the free group on $\{x_1, \ldots, x_n\}$, let K be the intersection of the kernels of all epimorphisms from F onto Q, and set $E = K\delta_l(F)$. If $\pi : F \to Q$ is an epimorphism then $E\pi = \delta_l(Q) = M$; it follows that $E/K := H$ is a subdirect product in a direct product P of copies of $M = S_1 \times \cdots \times S_r$.

Such a subdirect product takes the form

$$H = \Delta_1 \times \cdots \times \Delta_r,$$

where each Δ_j is a diagonal subgroup in some sub-product of the simple factors of P. It follows that H contains an element whose projection to each simple factor in P is non-trivial, and hence lies in no proper normal subgroup of H.

Thus there exists $w \in E$ such that $\left\langle w^E \right\rangle K = E$.

Now suppose that $Q = \langle g_1, \ldots, g_n \rangle$. Define $\pi : F \to Q$ by $x_i\pi = g_i$ $(i = 1, \ldots, n)$. Then $w(\mathbf{g}) = w\pi$, and so

$$\left\langle w(\mathbf{g})^M \right\rangle = \left\langle w^E \right\rangle \pi = E\pi = M.$$

Hence $w(\mathbf{g}) \neq 1$ and the lemma follows.

Theorem 1 (Abért, Nikolov/Segal) *Let G be a finite group, and put $\varepsilon(G) = p^{-|G|}$ where p is the largest prime divisor of $|G|$.*

(i) *The following are equivalent:*

(a) *G is soluble,*

(b) $\inf_w P(G, w) > 0$,

(c) $\inf_w P(G, w) > \varepsilon(G)$.

(ii) *The following are equivalent:*

(a) *G is nilpotent,*

(b) $\inf_{w,g} \left\{ P(G, w = g) \mid g \in G_{+w} \right\} > 0$,

(c) $\inf_{w,g} \left\{ P(G, w = g) \mid g \in G_{+w} \right\} > \varepsilon(G)$.

It remains to prove (a)⇒(c) in both cases.

Case 1: Where $|G| = p^h = m$.
Fix a 'basis' $\mathbf{b} = (b_1, \ldots, b_h)$ for G, so that

$$1 < \langle b_1 \rangle < \langle b_1, b_2 \rangle < \ldots < \langle b_1, b_2, \ldots, b_h \rangle = G$$

is a central series with cyclic factors of order p. Then each element of G is uniquely of the form

$$g = b_1^{x_1} \cdots b_h^{x_h} = \mathbf{b}^{\mathbf{x}}$$

with $x_1, \ldots, x_h \in \mathbf{P} = \{0, 1, 2, , \ldots, p-1\}$.

Identify G with a subgroup of $\mathrm{GL}_m(\mathbb{F}_p)$ by taking the regular representation.

Set V_s = linear span of $(G-1)^s$ in $\mathrm{M}_m(\mathbb{F}_p)$ $(s \geq 1)$

$V_0 = \{1\}$.

G unipotent $\Rightarrow V_n = 0$ for all $n \geq m$.

For $\mathbf{j} = (j_1, j_2, \ldots)$ we set $|\mathbf{j}| = j_1 + j_2 + \ldots$.

Lemma 2 *There exist matrices* $B_{\mathbf{j}} = B_{\mathbf{j}}(\mathbf{b}) \in V_{|\mathbf{j}|}$ *and polynomials* $F_{\mathbf{j}} \in \mathbb{F}_p[X_1, \ldots, X_h]$ *such that*

$$\mathbf{b}^{\mathbf{x}} = \sum_{\mathbf{j} \in \mathbf{P}^{(h)}} F_{\mathbf{j}}(x_1, \ldots, x_h) B_{\mathbf{j}} \qquad (\mathbf{x} \in \mathbf{P}^{(h)});$$

each $F_{\mathbf{j}}$ *has total degree at most* $|\mathbf{j}|$.

Proof. Put $a_i = b_i - 1$ for each i. Then for $0 \leq x \leq p-1$ we have

$$b_i^x = \sum_{j=0}^{x} \binom{x}{j} a_i^j = 1 + \sum_{j=1}^{x} c(j) x(x-1)\ldots(x-j+1) a_i^j$$
$$= \sum_{j=0}^{p-1} F_j(x) a_i^j,$$

where

$$F_0(X) = 1$$
$$F_j(X) = c(j) X(X-1)\ldots(X-j+1) \qquad (j > 1)$$

The lemma follows on setting

$$F_{\mathbf{j}}(X_1, \ldots, X_h) = F_{j_1}(X_1) \ldots F_{j_h}(X_h),$$
$$B_{\mathbf{j}} = a_1^{j_1} a_2^{j_2} \ldots a_h^{j_h}.$$

Next, let w be a positive generalized word over G.

Lemma 3 *There exist matrices* $B(w)_{\mathbf{j}} \in V_{|\mathbf{j}|}$ *and polynomials* $F(w)_{\mathbf{j}} \in \mathbb{F}_p[X_{11}, \ldots, X_{nh}]$ *for* $\mathbf{j} \in \mathbf{P}^{(ht)}$ *such that for* $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{P}^{(h)}$ *we have*

$$w(\mathbf{b}^{\mathbf{x}_1}, \ldots, \mathbf{b}^{\mathbf{x}_k}) = \sum_{\mathbf{j} \in \mathbf{P}^{(ht)}} F(w)_{\mathbf{j}}(\mathbf{x}_1, \ldots, \mathbf{x}_k) B(w)_{\mathbf{j}};$$

each $F(w)_{\mathbf{j}}$ *has total degree at most* $|\mathbf{j}|$.

Proof. For each l, the tuple $\mathbf{b}^{\alpha(l)} = (b_1^{\alpha(l)}, \ldots, b_h^{a(l)})$ is again a basis for G, and for $\mathbf{j} \in \mathbf{P}^{(h)}$ we put $B(l)_{\mathbf{j}} = B_{\mathbf{j}}(\mathbf{b}^{\alpha(l)})$. Then for $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{P}^{(h)}$ we have

$$\begin{aligned} w(\mathbf{b}^{\mathbf{x}_1}, \ldots, \mathbf{b}^{\mathbf{x}_k}) &= \prod_{l=1}^{t} \sum_{\mathbf{j} \in \mathbf{P}^{(h)}} F_{\mathbf{j}}(\mathbf{x}_{i_l}) B(l)_{\mathbf{j}} \\ &= \sum_{\mathbf{j}_1, \ldots, \mathbf{j}_t} F(w)_{\mathbf{j}}(\mathbf{x}_1, \ldots, \mathbf{x}_k) B(w)_{\mathbf{j}} \end{aligned}$$

where for $\mathbf{j} = (\mathbf{j}_1, \ldots, \mathbf{j}_t)$

$$F(w)_{\mathbf{j}}(\mathbf{X}_1, \ldots, \mathbf{X}_k) = F_{\mathbf{j}_1}(\mathbf{X}_{i_1}) \ldots F_{\mathbf{j}_t}(\mathbf{X}_{i_t}),$$

$$B(w)_{\mathbf{j}} = B(1)_{\mathbf{j}_1} \ldots B(t)_{\mathbf{j}_t}.$$

Proposition 1 *Let $c \in G$ and suppose that $c = w(\mathbf{h})$ for some $\mathbf{h} \in G^{(k)}$. Then*

$$\left| f_w^{-1}(c) \right| \geq p \, |G|^k \, \varepsilon(G).$$

Proof. Let's take the elements of G as basis for the regular representation. Then for $g \in G$ we have

$$g = c \Longleftrightarrow g_{1c} = 1,$$

where g_{1c} denotes the $(1, c)$-entry of the matrix g.

Define a map $\psi : \mathbf{P}^{hk} \to \mathbb{F}_p$ by

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_k) = 1 - w(\mathbf{b}^{\mathbf{x}_1}, \ldots, \mathbf{b}^{\mathbf{x}_k})_{1c}.$$

Lemma 3 shows that ψ is equal to a polynomial of total degree at most $m-1$, since for $|\mathbf{j}| \geq m$ we have

$$B(w)_{\mathbf{j}} \in V_{|\mathbf{j}|} = 0.$$

Also $\psi(\mathbf{z}_1, \ldots, \mathbf{z}_k) = 0$ where $(\mathbf{b}^{\mathbf{z}_1}, \ldots, \mathbf{b}^{\mathbf{z}_k}) = \mathbf{h}$. Identifying $\mathbf{P}$ with $\mathbb{F}_p$, we can now apply the **Chevalley-Warning theorem** to infer that ψ has at least p^{hk-m+1} zeros in $\mathbf{P}^{hk}$. Each one corresponds to a solution of $w(\mathbf{b}^{\mathbf{x}_1}, \ldots, \mathbf{b}^{\mathbf{x}_k}) = c$, giving the result since $\varepsilon(G) = p^{-m}$.

Lemma 4 *If $G = AB$ is finite where A and B are proper subgroups of G then*

$$\varepsilon(G) \leq \varepsilon(A)\varepsilon(B). \tag{5}$$

Proof. Say p is the largest prime factor of both $|G|$ and $|A|$, and q is the largest prime factor of $|B|$. Then $q \leq p$ so

$$p^{|G|} \geq p^{|A|+|B|} \geq p^{|A|}q^{|B|}.$$

Case 2. Suppose $G = P_1 \times \cdots \times P_r$ is nilpotent, where P_i is a p_i-group, $p_1, \ldots, p_r$ distinct primes. Let w be a positive generalized word over G.

If $c_i \in P_i$ and $c = c_1 \ldots c_r \in Gf_w$ then $c_i \in P_i f_w$ for each i. So Proposition 1 gives

$$\left|f_w^{-1}(c)\right| = \prod \left|f_w^{-1}(c_i)\right| \geq \prod p_i \left|P_i\right|^k \varepsilon(P_i)$$
$$\geq \prod p_i \cdot |G|^k \varepsilon(G). \qquad (6)$$

In particular, taking w to be an ordinary word (which we may assume to be positive), we see that

$$P(G, w = c) = \frac{\left|f_w^{-1}(c)\right|}{|G|^k} > \varepsilon(G),$$

which completes the proof of Theorem 1(ii).

Case 3. Fix a positive word w. Suppose that G is *soluble, but not nilpotent*. Put

$$\begin{aligned} N &= \mathsf{Fit}(G) \\ K/N &\lhd_{\mathsf{min}} G/N \\ P &\in Syl_p(K) \\ H &= \mathsf{N}_G(P) \end{aligned}$$

where K/N is a p-group. Then $H < G$ because K is not nilpotent, and by the Frattini argument

$$\begin{aligned} K &= NP \\ G &= KH = NH. \end{aligned}$$

Arguing by induction on the group order, we may suppose that

$$\left|f_w^{-1}(1)\right| > |H|^k \varepsilon(H).$$

Now fix $\mathbf{h} \in H^{(k)}$ such that $w(\mathbf{h}) = 1$. There is a generalized word $w'_{\mathbf{h}}$ over N such that

$$w(\mathbf{a} \cdot \mathbf{h}) = w'_{\mathbf{h}}(\mathbf{a})w(\mathbf{h})$$

for all $\mathbf{a} \in N^{(k)}$, where

$$\mathbf{a} \cdot \mathbf{h} = (a_1 h_1, \ldots, a_k h_k).$$

Apply (6) to the group N:

$$\left| f_{w'_{\mathbf{h}}}^{-1}(1) \right| > |N|^k \, \varepsilon(N).$$

So: the number of pairs $(\mathbf{a}, \mathbf{h}) \in N^{(k)} \times H^{(k)}$ for which $w(\mathbf{a} \cdot \mathbf{h}) = 1$ exceeds

$$\begin{aligned} |H|^k \, \varepsilon(H) \cdot |N|^k \, \varepsilon(N) &= |H \cap N|^k \, |G|^k \, \varepsilon(H)\varepsilon(N) \\ &\geq |H \cap N|^k \, |G|^k \, \varepsilon(G). \end{aligned}$$

The fibres of the map $(\mathbf{a}, \mathbf{h}) \mapsto \mathbf{a} \cdot \mathbf{h}$ each have size $|H \cap N|^k$. Therefore: $w(\mathbf{g}) = 1$ for more than $|G|^k \, \varepsilon(G)$ elements $\mathbf{g} \in G^{(k)}$. So

$$P(G, w) > \varepsilon(G).$$

Corollary 1 *Let G be a finite group. Then G is soluble if and only if for every sufficiently large n, every n-generator one-relator group maps onto G.*

Proof. 1. If G is *not soluble*.

Let Q be a just-non-soluble quotient of G, and let $n \in \mathbb{N}$. By Lemma 1, there exists a word w in n variables such that

$$w(\mathbf{g}) = 1 \Longrightarrow \langle g_1, \ldots, g_n \rangle \neq Q.$$

The one-relator group $\langle x_1, \ldots, x_n; w \rangle$ then does *not* map onto Q, and a fortiori it doesn't map onto G.

2. If G is *soluble*.

w a word in n variables. Then the probability that $\mathbf{g} \in G^{(n)}$ satisfies *both* $w(\mathbf{g}) = 1$ and $\langle g_1, \ldots, g_n \rangle = G$ is

$$\pi_n(w) := P(G, w) + P(G, n) - 1.$$

Now:

$$P(G,n) \geq 1 - m2^{-n}$$

where m denotes the number of maximal subgroups of G and

$$P(G,w) > \varepsilon(G).$$

So as long as $m2^{-n} \leq \varepsilon(G)$ we have

$$\pi_n(w) > 1 - m2^{-n} + \varepsilon(G) - 1 \geq 0.$$

Thus $w(\mathbf{g}) = 1$ for at least one generating set $\{g_1, \ldots, g_n\}$ for G, and $\langle x_1, \ldots, x_n; w\rangle$ maps onto G by $x_i \mapsto g_i$ $(i = 1, \ldots, n)$.

Conjecture. (A. Amit) If G is a finite nilpotent group and w is any word then

$$P(G,w) \geq \left|G^{-1}\right|.$$