Quotient tests and random walks in computational group theory

Alexandre V Borovik Alexei G Myasnikov

11 February 2005

Abstract

For many decision problems on a finitely presented group G, we can quickly weed out negative solutions by using much quicker algorithms on an appropriately chosen quotient group G/K of G. However, the behavior of such "quotient tests" can be sometimes paradoxical. In this paper, we analyze a few simple case studies of quotient tests for the classical identity, word, conjugacy problems in groups. We attempt to combine a rigorous analytic study with the assessment of algorithms from the practical point of view. It appears that, in case of finite quotient groups G/K, the efficiency of the quotient test very much depends on the mixing times for random walks on the Cayley graph of G/K.

1 Decision problems and quotient tests

Let $F = F(X)$ be a free group of rank m with basis X, R be a normal subgroup of F, and $G = F/R$. In this case we say that G is given by a presentation $\langle X \mid R\rangle$. By $\delta : F \to G$ we denote the canonical epimorphism.

Sometimes we will need to represent elements of G by (not necessary reduced) group words in X. To this end, denote by $M = M(X)$ the set of all (not necessary reduced) words in the alphabet $X^{\pm 1} = X \cup X^{-1}$, i.e., M is a free monoid with basis $X^{\pm 1}$. Denote by $\pi : M \to G$ the canonical epimorphism $\pi : M \to F$ and by $\psi : M \to G$ the canonical epimorphism $\psi = \pi \circ \delta$. Thus every word $w \in M$ (as well as $w \in F$) represents an element $\psi(w) \in G$.

A *decision problem* for a finitely generated group G given by a presentation $G = \langle X \mid R\rangle$ is a subset $\mathcal{D} \subset M(X)^k$. A decision problem $\mathcal{D} \subset M(X)^k$ is called *decidable* if there exists an algorithm to decide whether an arbitrary element of $M(X)^k$ belongs to $\mathcal{D}$ or not.

Instead of the problems over $M(X)$ one can consider decision problems only over freely reduced words, that is, decision problems $\mathcal{D} \subset F(X)^k$. Since one can easily (in linear time) reduce a word in $M(X)$ to its reduced form in $F(X)$ these two decision problems are equivalent with respect to time complexity classes. In average-case (see [3]) or generic-case complexity (see [12]), where the measure

on the set of inputs matters, the equivalence of these two approaches needs to be verified.

In this paper we discuss three classical decision problems for a given finitely generated group G given by a presentation $G = \langle X \mid R \rangle$: the *identity problem* (IP), the *word problem* (WP), and the *conjugacy problem* (CP).

(IP) Given a word $w \in M(X)$ determine whether or not w represents the identity element in G (symbolically, $w =_G 1$). Thus

$$IP(G, X) := \{w \in M \mid \psi(w) = 1\}.$$

(WP) Given two words $u, v \in M(X)$ determine whether they represent equal elements of G or not. Thus

$$WP(G, X) := \{(u, v) \in M(X) \times M(X) \mid \psi(u) = \psi(v)\}.$$

(CP) Given two words $u, v \in M(X)$ determine whether they represent conjugate elements of G or not. Thus

$$CP(G, X) := \{(u, v) \in M(X) \times M(X) \mid \psi(u), \psi(v) \text{ are conjugate in } G\}.$$

We refer to the surveys [1, 18, 19] on decision problems for groups.

Sometimes, for a given decision problem $\mathcal{D} \subset M(X)^k$ it is convenient to consider *partial* algorithms. A partial algorithm is a correct algorithm for $\mathcal{D}$ which does not terminate or does not produce an answer on a possibly non-empty set of inputs from $M(X)^k$.

Quotient tests provide one of the key methods for designing decision algorithms for groups. These are partial decision algorithms aiming for the "no" part of the problem. They are based on the following simple idea:

> Let $\phi : G \to H$ be a homomorphism from G onto a group H. Then the following conditions hold for any elements $u, v \in M(X)$:
>
> 1) if $u =_G 1$ then $\phi(u) =_H 1$;
> 2) if $u =_G v$ then $u =_H v$;
> 3) if $\psi(u)$ and $\psi(v)$ are conjugate in G then $\phi(\psi(u))$ and $\phi(\psi(v))$ are conjugate in H.

Now suppose (IP) (or (WP), or (CP)) is decidable for the group H with a decision algorithm B. Suppose the homomorphism ϕ is given by some representatives of the images $\phi(x)$, $x \in X$, i.e., for every $x \in X$ a word $u_x \in M(X)$ is specified such that $\phi(\psi(x)) =_H u_x$. In this event, one can design a partial algorithm A for (IP) (or (WP), or (CP)) which for a given $w \in M(X)$ (given a pair of words $u, v \in M(X)$) makes the substitution $\sigma : x \to u_x$ in w (in u, v) and applies B to the resulting word $\sigma(w)$ (pair $\sigma(u), \sigma(v)$). If B stops on $\sigma(w)$ ($\sigma(u), \sigma(v)$) and says "no", then A stops and says "no". The algorithm A is

called the *quotient test* with respect to the homomorphism $\phi : G \to H$. In fact, in the case of the identity and the word problem we can always assume (replacing H by $\phi(G)$) that ϕ is an epimorphism. In this case we can view H as $H = F/Q$ for some normal subgroup Q of F with $R \leq Q$.

The choice of the testing homomorphism $\phi : G \to H$ may well depend on the group G. But the most typical ones are homomorphisms onto the abelianization $G \to G/[G, G]$ or onto a finite group $G \to K$.

In [12, 13] it has been shown that the classical decision problems ((WP), (CP), and the membership problem) for a wide class of groups are generically very fast, i.e., there are quotient tests which solve the problems "very quickly" (in linear, or quadratic, time) for almost all inputs. In particular, if $\phi : G \to H$ is a homomorphism onto a non-elementary hyperbolic group H then the quotient test with respect to ϕ (generically) solves (WP) in linear time, and (CP) - in quadratic time. As usual, the term "almost all" means that the result holds on inputs from some subset $S \subseteq \mathcal{D}$ such that $\mu(\mathcal{D} - S) = 0$ for some measure (or pseudo-measure) μ defined on $M(X)^k$ or on $F(X)^k$. In [12, 13] the choice of μ is the *asymptotic density* ρ, which allows one to apply some deep results from random walks on infinite groups, in particular, the classical characterization of non-amenable groups in terms of co-growth rates due to Grigorchuk [8].

We have to appreciate, however, the limitations of quotient tests. They are fast because "random" instances of problems almost always have the answer "no" in a "good" quotient test. In case of the answer "yes" for the quotient test we still have to solve the original problem. Still, for problems with random inputs, this means a dramatic improvement in the overall performance of algorithms for solving (IP), (WP) or (CP).

In this paper we focus mostly on the quotient tests based on amenable quotient groups, in particular, infinite cyclic and finite quotient groups. Our choice of the measure μ is the spherical uniform distribution and also the multiplicative measure from [3].

Our main goal is to develop practical criteria to compare effectiveness of various quotient test. Notice, in particular, that (WP) is not reduced to (IP); as one can easily see, they are different in the probabilistic context. For if elements $x, y \in F(x)$ are independently chosen from a probabilistic distribution μ on $F(X)$, the probabilistic distribution of xy^{-1} is the *convolution* $\mu * \mu$ (under a mild and quite natural assumption that the measure μ is preserved under inversion of elements). The behaviour of (CP) is even trickier. Therefore we treat them separately, starting with the easiest one, the identity problem (Section 2) with two probability distributions: spherical, when elements of fixed length k are chosen at random with equal probabilities, and the multiplicative measure [3]. Then we move to the word problem (Section 5) and end with a very brief discussion of the conjugacy problem (Section 6).

2 Quotient test for the identity problem: non-amenable quotient group

In this section we assume that the input elements x of our test are reduced words from $F(X)$ uniformly distributed in the sphere

$$S_k = \{x \in F(X) \mid |x| = k\}.$$

We shall call this probability distribution *spherical distribution* and denote it σ_k. For a subset $S \subseteq F$ we obviously have

$$\sigma_k(S) = \frac{|S \cap S_k|}{|S_k|},$$

the relative frequency of elements of length k from S.

Given a finitely presented group $G = F/R$, we run, for the identity problem in F/R, a quotient test in the smaller and computationally more feasible quotient group F/Q, $R < Q$. The *defect* of this quotient test is the conditional probability

$$P(x \in Q \mid x \in F \setminus R),$$

which, when R is sufficiently 'small' subgroup, that is, a subgroup with 'small' probability $P(x \in R) = \sigma_k(R)$, is, for all practical purposes, the same as the easier to evaluate probability

$$\sigma_k(Q) = P(x \in Q).$$

We denote $n_k = |Q \cap S_k|$ and $q_k = \sigma_k(Q) = n_k/|S_k|$. The robustness of the group F/Q as the quotient test for the identity problem is determined by how small the values of q_k are for k in the practical range of real life computations, say, for $10 \leqslant k \leqslant 1000$. Some asymptotic estimates, however, can be of value for the general guidance.

Denote by

$$q(t) = \sum_{k=0}^{\infty} q_k t^k$$

the *frequency generating function* for Q. Set

$$\alpha = \limsup \sqrt[k]{q_k},$$

then the convergence radius of $q(t)$ is $A = 1/\alpha$. Notice the following classical Cauchy bound for the coefficients q_k:

$$q_k \leqslant \frac{M(A_0)}{A_0^k},$$

where A_0 is any real number, $0 < A_0 < A$ and $M(A_0)$ is the maximum of $|q(z)|$ on the circle $|z| = A_0$. We shall denote $\alpha_0 = 1/A_0$, then the Cauchy bound takes the form

$$q_k \leqslant M(A_0)\alpha_0^k.$$

We shall call α the *cogrowth* of Q. By a classical result of Grigorchuk [8], the group F/Q is not amenable if and only if $\alpha < 1$. This suggests that if F/Q is not amenable (for example, if F/Q contains a non-abelian free group) then it might make a robust fast check. This thesis is supported by anecdotal evidence accumulated in computational group theory.

3 Quotient test for the identity problem: amenable quotient groups

However, it is interesting to look at the other end of the spectrum and see what happens if the quotient group F/Q is relatively small.

An important class of tests for equality consists of mapping the group $G = F/R$ onto abelian groups.

If $G = B_n$ is the braid group given by the band generators, the preimage in F of the commutator subgroup in G is the *co-diagonal subgroup* D of F, that is, the kernel of the homomorphism $F \longrightarrow \mathbb{Z}$ defined by mapping all generators $x_i \in X$ to the generator 1 of $\mathbb{Z}$. Notice that F has rank $m = n(n-1)/2$ since the band generators of B_n correspond to transpositions in the symmetric group Sym_n. Let S be the kernel of the homomorphism of F onto the symmetric group Sym_n which factorizes through the canonical homomorphism of B_n onto Sym_n. We wish to compare the quotient tests F/D and F/S for the problem of identity for elements in B_n presented as words in the free monoid $M(X)$. This naturally leads us to random walks: a random element of length k from $M(X)$ represents a simple random walk on the Cayley graph of groups F/D and F/S (in the latter case the symmetric group S_n being generated by the conjugacy class of transpositions). We immediately encounter a difficulty: the no-return random walk on F/S and simple random walk on F/S do not mix well: a word of length k, reduced or not, is an odd or even permutation depending on parity of k. Therefore we have to modify a random walk on F to make the quotient tests F/D and F/S comparable. Therefore we use a random walk from the classical paper by Diaconis and Shahshahani [5]: at any vertex v of the Cayley graph of F, we stay at the vertex with probability $1/n^2$, or move, with equal probabilities, to one of $n(n-1)$ adjacent vertices $vx_i^{\pm 1}$, $i = 1, \ldots, n(n-1)/2$. In the symmetric group $S_n = F/S$, this corresponds to mixing a pack on n cards: we select cards i and i independently (so, with probability $1/n$, we pick the same card twice and do nothing). This can also be interpreted in terms of the convolution of measures. Let $P = P^1$ be the probability distribution on F concentrated at the points 1 and $x_i^{\pm 1}$, $i = 1, \ldots, n(n-1)/2$, and given by

$$P(x) = \begin{cases} \frac{1}{n} & \text{if} \quad x = 1 \\ (1 - \frac{1}{n})(n(n-1))^{-1} & \text{if} \quad x = x_i^{\pm 1} \end{cases} .$$

Then $P(x)$ is the probability to be at vertex x after the first step of the walk, and the probabilistic distribution after k steps of the random walk is the k-th convolution $P^{*k} = P^{*(k-1)} * P^1$.

We wish to estimate the *return probability* r_k of the Diaconis-Shahshahani walk, that is, the probability to end up precisely at the starting point 1.

For the group $F/D \simeq \mathbb{Z}$ we have a random walk on the one-dimensional lattice, and a well-known result (see [22, Proposition P9 on p. 75]) yields that

$$r_k(F/D) \sim \frac{1}{\sqrt{2\pi k}} \cdot \left(1 - \frac{1}{n^2}\right)^{-\frac{1}{2}}$$

(that is, $r_k(F/D)\sqrt{2\pi k}\left(1 - \frac{1}{n^2}\right) \longrightarrow 1$).

Regarding the group $F/S = \mathrm{Sym}_n$, the probabilistic distribution generated by the Diaconis-Shahshahani walk quickly converges to the uniform distribution:

Fact 3.1 (Diaconis and Shahshahani [5]) *If* $k = \frac{1}{2} \cdot n \log n + cn$, $c > 0$, *then*

$$\|P^{*k} - U\| \leqslant ae^{-2c}$$

for an universal constant a. *Here* U *is the uniform distribution on* Sym_n *and the norm* $\|\ \|$ *is defined as*

$$\|P^{*k} - U\| = \frac{1}{2} \sum_{g \in \mathrm{Sym}_n} \left|P^{*k}(g) - U(g)\right|.$$

In particular, this means that

$$r_k(F/S) \leqslant \frac{1}{n!} + ane^{-\frac{2k}{n}}$$

and $r_k(F/S)$ starts to converge to $1/n!$ exponentially fast. However, within the range of practically useful values,

$$\frac{1}{n!} \ll ane^{-\frac{2k}{n}}$$

and the term $ane^{-\frac{2k}{n}}$ dominates the estimate for $r_k(F/S)$. The unknown constant a is, of course, very worrying. But the good news is that the experimental evidence shows that the convergence is indeed very good, which suggests that the constant a is not that big and is quickly annihilated by the factor $e^{-\frac{2k}{n}}$.

In particular, this means that, for $n \ll k \ll n!$,

$$r_k(F/S) \approx ane^{-\frac{2k}{n}} \ll r_k(F/D) \approx \frac{1}{\sqrt{2\pi k}}.$$

Our conclusion is that, for the sizes of input actually appearing in practical computations, (say, k = 1000 and $n = 100$), the quotient test with the symmetric group quotient $\mathrm{Sym}_n = F/S$ is incomparably better than the one with the cyclic quotient group $\mathbb{Z} = F/Q$.

4 The role of the mixing time

Now we wish to analyze the same quotient tests for the identity problem, F/D and F/S, using a different approach. We shall show the crucial importance of the mixing time of the random walk on the finite quotient group F/K for the suitability of the F/K for a quotient test for identity.

4.1 Multiplicative measures

Setup. Let $F = F(X)$ be a free group with a basis $X = \{x_1, \dots, x_m\}$. In our work [3] we suggested to use, as a random word generator, the following no-return random walk W_s ($s \in (0, 1]$) on the Cayley graph $C(F, X)$ of F with respect to the generating set X. We start at the identity element 1 and either do nothing with probability s (and return value 1 as the output of our random word generator), or move to one of the $2m$ adjacent vertices with equal probabilities $(1 - s)/2m$. If we are at a vertex $v \neq 1$, we either stop at v with probability s (and return v), or move, with probability $\frac{1-s}{2m-1}$, to one of the $2m - 1$ adjacent vertices lying away from 1, thus producing a new freely reduced word $vx_i^{\pm 1}$. In other words, we make random freely reduced words w of random lengths $|w|$ distributed according to the geometric law

$$P(|w| = k) = s(1 - s)^k,$$

in such way that words of the same length k are produced with equal probabilities. Observe that the set of all words of length k in F forms the sphere S_k of radius k in $C(F, X)$ of cardinality $|S_k| = 2m(2m - 1)^{k-1}$. It is easy to see that the resulting probabilistic atomic measure[1] μ_s on F is given by the formula

$$\mu_s(w) = \frac{s(1 - s)^{|w|}}{2m \cdot (2m - 1)^{|w|-1}} \quad \text{for } w \neq 1 \tag{1}$$

and

$$\mu_s(1) = s. \tag{2}$$

Thus, $\mu_s(w)$ is the probability that the random walk W_s stops at w. The mean length L_s of words in F distributed according to μ_s is equal to

$$L_s = \sum_{w \in F} |w| \mu_s(w) = s \sum_{k=1}^{\infty} k(1 - s)^{k-1} = \frac{1}{s} - 1.$$

Hence we have a family of probabilistic distributions $\mu = \{\mu_s\}$ with the stopping probability $s \in (0, 1)$ as a parameter, which is related to the average length L_s as

$$s = \frac{1}{L_s + 1}.$$

[1] Recall that a measure μ on a countable set X is *atomic* if every subset $Y \subseteq X$ is measurable. This is equivalent to saying that every singleton subset $\{x\}$ is measurable. Obviously, $\mu(Y) = \sum_{x \in Y} \mu(x)$.

For any subset $R \subseteq F$, by $\mu(R)$ we denote the function

$$\begin{aligned} \mu(R) : (0,1) &\to \mathbb{R} \\ s &\mapsto \mu_s(R); \end{aligned}$$

we call it *measure* of R with respect to the family of distributions μ. This family of measures is called *multiplicative measures* on F.

Power series expansions of multiplicative measures. In [3] it was shown that, for the co-diagonal subgroup D,

$$\mu(D) = \frac{m-1}{m\sqrt{2-\frac{2}{m}}} \cdot \sqrt{s} + o(\sqrt{s}) = \frac{m-1}{m\sqrt{2-\frac{2}{m}}} \cdot \frac{1}{\sqrt{L}} + o\left(\frac{1}{\sqrt{L}}\right) \qquad (3)$$

where L is the mean length of words produced by our random word generator, while, for any normal subgroup K of finite index, $\mu(K)$ is a rational function of s and

$$\mu(K) = \frac{1}{|F:K|} + O(s) = \frac{1}{|F:K|} + O\left(\frac{1}{L}\right). \qquad (4)$$

For large $|F:K|$ and moderate mean lengths L (remember, that we are looking at words of lengths practically used in computations, which in most cases mean $L \leqslant 1000$), the error term $O(1/L)$ cannot be ignored. Let us look at the next term in the power series expansion for $\mu(K)$.

Theorem 4.1 *Let $K \lhd F$ and assume that the quotient group F/K is finite, $|F:K| = N$. Then, in the power series expansion*

$$\mu(K) = \mu_0 + \mu_1 s + \mu_2 s^2 + \cdots,$$

for $\mu(K)$, the first two coefficients are

$$\mu_0 = \frac{1}{|F:K|}$$

and

$$\mu_1 = \frac{d}{d+1}\frac{1}{|F:K|}\sum_{i=2}^{N}\frac{1}{1-\lambda_i} + \frac{1}{d+1}, \qquad (5)$$

where $d = 2m-1$ and

$$1 = \lambda_1 > \lambda_2 > \cdots > \lambda_N$$

are eigenvalues of the simple random walk on F/K. In particular,

$$\mu_1 < \frac{1}{1-\lambda_2}.$$

Proof. Let n_k^* be the number of non-reduced words from $M(X)$ of length k which are mapped into K. By result of Quenell [25], the generating function $N^*(t) = \sum n_k^* t^k$ is expressed as

$$N^*(t) = \frac{1}{|F:K|} \sum_{i=1}^{N} \frac{1}{1-\lambda_i' t},$$

where λ_i' are the eigenvalues of the the adjacency matrix of the Cayley graph Γ of F/K. We can use Bartholdi's equation [2]

$$\frac{N(t)}{1-t^2} = \frac{N^*\left(\frac{t}{1+dt^2}\right)}{1+dt^2}. \tag{6}$$

and find the number of (reduced) words of given length k in K:

$$N(t) = \left[\frac{1}{|F:K|} \sum_{i=1}^{N} \frac{1}{1-\lambda_i' \frac{t}{1+dt^2}}\right] \frac{1-t^2}{1+dt^2}.$$

Since the biggest eigenvalue λ_1 is the valency of the graph Γ and thus $\lambda_1' = d+1$, after an easy algebraic rearrangement we have

$$N(t) = \frac{1}{|F:K|} \left[\frac{1+t}{1-dt} + \frac{1-t^2}{1+dt^2} \sum_{i=2}^{N} \frac{1}{1-\lambda_i' \frac{t}{1+dt^2}}\right].$$

Taking into account that

$$\mu(K) = \frac{d}{d+1} s N\left(\frac{1-s}{d}\right) + \frac{1}{d+1} s,$$

we have, after expanding $\mu(K)$ into a power series in s,

$$\mu(K) = \frac{1}{|F:K|} + \left[\frac{d}{d+1}\frac{1}{|F:K|} \sum_{i=2}^{N} \frac{1}{1-\frac{\lambda_i'}{d+1}} + \frac{1}{d+1}\right] s + O(s^2).$$

Notice that $\lambda_i = \lambda_i'/(d+1)$ are the eigenvalues of the random walk on Γ, which immediately yields the desired formulae for μ_0 and μ_1. The inequality

$$\mu_1 < \frac{d}{d+1} \cdot \frac{1}{1-\lambda_2} + \frac{1}{d+1} = \frac{d+1-\lambda_2}{(d+1)(1-\lambda_2)} < \frac{1}{1-\lambda_2}$$

is now obvious. □

4.2 Finite cyclic quotient group

Recall that the canonical homomorphism $F \to F/K$ sends every free generator x_i of F to the fixed generator c of the cyclic group $C = F/K$ of order $N = n!$.

The eigenvalues of the simple random walk on the cyclic group of even order N are

$$\lambda_1 = 1, \ldots, \lambda_k = \cos\frac{2\pi(k-1)}{N}, \ldots, \lambda_N = \cos\frac{2\pi(N-1)}{N}.$$

Therefore the crucial part of Equation (5) evaluates as

$$\begin{aligned}
\sum_{k=2}^{N} \frac{1}{1-\lambda_k} &= \sum_{k=1}^{N-1} \frac{1}{1-\cos\frac{2\pi k}{N}} \\
&\geqslant \frac{N}{2\pi} \int_{\frac{2\pi}{N}}^{\frac{2\pi(N-1)}{N}} \frac{dx}{1-\cos x} \\
&= \frac{N}{2\pi} \left[\frac{\sin x}{\cos x - 1}\right]_{\frac{2\pi}{N}}^{\frac{2\pi(N-1)}{N}} \\
&\approx \frac{N}{2\pi} \left[\frac{-2\pi/N}{-(2\pi/N)^2/2} - \frac{2\pi/N}{-(2\pi/N)^2/2}\right] \\
&= \frac{N^2}{\pi^2}.
\end{aligned}$$

Hence

$$\begin{aligned}
\mu_1 &= \frac{d}{d+1}\frac{1}{|F:K|}\sum_{i=2}^{N}\frac{1}{1-\lambda_i} + \frac{1}{d+1}, \\
&\geqslant \frac{d}{d+1}\frac{N}{\pi^2} + \frac{1}{d+1}
\end{aligned}$$

Our crude estimates could be made much more precise. Essentially, they mean that, for the cyclic quotient group of big finite order N, μ_1 is at least of the same order of magnitude as

$$\frac{N}{\pi^2}.$$

Recall that when we compare homomorphisms onto $\mathbb{Z}/n!\mathbb{Z}$ and Sym_n as quotient tests for identity, $N = n!$ is quite big number.

For the symmetric quotient group $F/S \simeq \mathrm{Sym}_n$ one can run (much longer) eigenvalue estimates along the lines of calculations by Diaconis and Shhshahani in [5] (modifications required are minimal) and get a dramatically different result:

$$\mu_1(S) \leqslant 10.$$

4.3 Comparing the two tests using the multiplicative measure

Theorem 4.1 provides some food for thought. The spectral gap $1 - \lambda_2$ is the crucial parameter of the simple random walk on the finite quotient group F/K

which determines its mixing time. This suggests that the quotient test F/K for the identity problem behaves better for those finite quotient groups whose Cayley graphs have better mixing properties of random walks.

Can the terms $o(\,)$ and $O(\,)$ in Equations (3) and (4) be rendered negligible? The answer is, of course, no. Indeed, take a subgroup $K > D$ with a very big index $|F : K|$, so that K is the preimage in F of kernel of the homomorphism of $R/D = \mathbb{Z}$ onto a very big cyclic group $\mathbb{Z}_N$, for example, take $N = 100!$. Then, obviously, for all $k \ll N$, $r_k(F/K) = r_k(F/D)$: a short word equals 1 in R/K if and only if it equals 1 in F/D. A very noticeable property of multiplicative measures is that they are strongly skewed towards shorter elements. In particular, this means that for all mean lengths $L \ll N$ we have $\mu(K) \approx \mu(D)$.

The difference in the behaviour of the finite quotient groups $\mathbb{Z}_N = F/K$ and $\mathrm{Sym}_n = F/S$ as quotient tests for the identity problem shows that the mixing time for the random walk on the Cayley graph of the finite quotient group F/K is the crucial factor for the efficiency of the quotient test. Big cyclic groups, of course, have the worst possible mixing properties.

In particular, for non-recurrent (and, in particular, for all non-amenable) quotient groups F/Q,

$$\mu(Q) = \frac{c}{L} + o\left(\frac{1}{L}\right)$$

and the probability of an element $x \in F$ to belong to Q behaves as $1/|x|$, not as $e^{-c|x|}$ as in the case when x is chosen at random from the sphere S_k and F/Q is non-amenable.

Interestingly, in our particular case of comparing the diagonal and symmetric quotient groups of F, the results of the analysis are the same in the multiplicative and spherical measures. This observation supports the conjecture that 'good' fast checks come from non-recurrent quotient groups.

5 Quotient tests for the word problem

5.1 Asymptotic estimates

Let x and y be any two random elements of length k in F. We want to estimate the probability

$$\pi_k(Q) = P(xy^{-1} \in Q \mid x, y \in S_k).$$

Notice that the probability

$$\pi_k(z) = P(xy^{-1} = z \mid x, y \in S_k)$$

depends only on $|z|$. Of course, $\pi_k(z) = 0$ if $|z| > 2k$ or if $|z|$ is odd.

Notice (although it will not be used in this paper) that π_k is the convolution $\sigma_k * \sigma_k$,

$$\pi_k(z) = \sum_{y \in F} \sigma_k(zy)\sigma_k(y^{-1}).$$

To make notation shorter, set $d = 2m - 1$. It is easy to see that

- The probability that the element xy^{-1} has length exactly $2k$ is
$$\frac{d}{d+1}$$
(because this means the the first element of the word y^{-1} can take d of possible $d+1$ values, to avoid cancellation with the last element of x).

- Similarly, the probability that the element xy^{-1} has length exactly $2k-2$ is
$$\frac{1}{d+1}\cdot\frac{d-1}{d}$$
(one cancellation happened, which gives the factor $1/(d+1)$, but did not spread further, which gives the factor $(d-1)/d$).

- The probability that the element xy^{-1} has length exactly $2k-2i$, $i<k$, is
$$\frac{d-1}{d+1}\cdot\frac{1}{d^i}.$$

- The probability that the element xy^{-1} has length 0 is
$$\frac{1}{(d+1)d^{k-1}}.$$

Now the probability of xy^{-1} to get in Q is

$$\begin{aligned}
\pi_k(Q) &= \frac{d}{d+1}q_{2k}+\frac{d-1}{d+1}\cdot\frac{1}{d}\cdot q_{2k-2}+\cdots+\frac{d-1}{d+1}\cdot\frac{1}{d^{k-1}}\cdot q_2 \\
&\qquad +\frac{1}{(d+1)d^{k-1}}\cdot q_0 \\
&\qquad \Big(\text{after denoting } \delta=1/\sqrt{d} \text{ and using the Cauchy bound}\Big) \\
&\leqslant M_Q\left[\frac{d}{d+1}\alpha_0^{2k}+\frac{d-1}{d+1}\left(\alpha_0^{2k-2}\delta^2+\cdots+\alpha_0^2\delta^{2k-2}\right)\right]+\frac{M_Q}{(d+1)d^{k-1}} \\
&= M_Q\left[\frac{d}{d+1}\alpha_0^{2k}+\frac{d-1}{d+1}\frac{\alpha_0^{2k}-\delta^{2k}}{\alpha_0^2-\delta^2}+\frac{d}{d+1}\delta^{2k}\right] \\
&= M_Q\left[\frac{d}{d+1}\left(\alpha_0^{2k}+\delta^{2k}\right)+\frac{d-1}{d+1}\frac{1}{\alpha_0^2-\delta^2}\left(\alpha_0^{2k}-\delta^{2k}\right)\right].
\end{aligned}$$

If the group $H=F/Q$ is not amenable, then $\alpha<1$ by a theorem of Grigorchuk. In that case α_0 can be chosen so that $\alpha<\alpha_0<1$. Since we also have $\delta<1$, the probability $\pi_k(Q)$ decreases exponentially fast with the growth of k.

We formulate this observation as our first theorem.

Theorem 5.1 *Assume that $Q\lhd F$ and the quotient group $H=F/Q$ is not amenable. Then*
$$\pi_k(Q)\to 0$$
exponentially fast as $k\to\infty$.

5.2 A more detailed computation

We start as before and set $D = 1/\delta = \sqrt{d}$:

$$\begin{aligned}
\pi_k(Q) &= \frac{d}{d+1}q_{2k} + \frac{d-1}{d+1}\cdot\frac{1}{d}q_{2k-2} + \cdots + \frac{d-1}{d+1}\cdot\frac{1}{d^{k-1}}q_2 \\
&\qquad + \frac{1}{(d+1)d^{k-1}}q_0 \\
&= \frac{d}{d+1}\left(q_{2k} + \delta^{2k}\right) + \frac{d-1}{d+1}\left(q_{2k-2}\delta^2 + q_{2k-4}\delta^4 + \cdots + q_2\delta^{2k-2}\right) \\
&= \frac{1}{d+1}\left(q_{2k} + \delta^{2k}\right) + \frac{d-1}{d+1}\left(q_{2k} + q_{2k-2}\delta^2 + \cdots + q_2\delta^{2k-2} + q_0\delta^{2k}\right) \\
&= \frac{1}{d+1}\left(q_{2k} + \delta^{2k}\right) + \frac{d-1}{d+1}\cdot\delta^{2k}\cdot\sum_{i=0}^{k} q_{2i}D^{2i}
\end{aligned}$$

In particular, if the group F/Q is not recurrent, that is, if the series $\sum q_k$ converges, then the series $\sum q_{2k}$ and $\sum \pi_k(Q)$ are also convergent.

Notice that a new character appears on the scene: the relative frequency generating function

$$\tilde{q}(t) = \sum_{i=0}^{\infty} q_{2i}t^{2i}$$

for the subgroup $\tilde{Q} \leqslant Q$ which consists of all elements in Q of even length. Obviously, $|Q : \tilde{Q}| \leqslant 2$ and the cogrowth $\tilde{\alpha}$ of $\tilde{Q}$ does not exceed that of $q(t)$: $\tilde{\alpha} \leqslant \alpha$.

It is well known that $\delta < \tilde{\alpha} \leqslant \alpha$ for a non-trivial normal subgroup $Q \neq 1$.

In this notation,

$$\pi_k(Q) = \frac{1}{d+1}\left(q_{2k} + \delta^{2k}\right) + \frac{d-1}{d+1}q^*_{2k} \tag{7}$$

where

$$q^*_{2k} = \left(q_{2k} + q_{2k-2}\delta^2 + q_{2k-4}\delta^4 + \cdots + q_2\delta^{2k-2} + q_0\delta^{2k}\right) \tag{8}$$

is the coefficient for t^{2k} in the expansion at 0 of the analytic function

$$q^*(t) = \sum_{k=0}^{\infty} q_{2k}t^{2k}\cdot\sum_{k=0}^{\infty}\delta^{2k}t^{2k} = \frac{\tilde{q}(t)}{1-\delta^2t^2} \tag{9}$$

The radii of convergence of the functions $q(t)$ and $\tilde{q}(t)$ are equal $1/\alpha$ and $1/\tilde{\alpha}$, correspondingly. Since

$$\frac{1}{\alpha} \leqslant \frac{1}{\tilde{\alpha}} < \frac{1}{\delta} = \sqrt{d},$$

the radius of convergence of the function $q^*(t)$ is the same as that of $\tilde{q}(t)$, that is, $1/\tilde{\alpha}$.

Now it follows from Cauchy's bound that there exists $C > 0$ such that, for arbitrary $\epsilon > 0$,

$$\pi_k(Q) < C(\alpha + \epsilon)^k,$$

which gives us another proof of Theorem 5.1.

From now on we work without the assumption that F/Q is not amenable. In particular, it is possible that $\alpha = 1$.

Lemma 5.2 *Let $\{a_k\}$ be a sequence of nonegative real numbers. If the limit*

$$\lim_{k\to\infty} a_k$$

exists, then, for any $0 < \delta < 1$, the limit

$$\lim_{k\to\infty} \left(a_k + a_{k-1}\delta + \cdots + a_1\delta^{k-1} + a_0\delta^k\right)$$

exists and equals

$$\frac{1}{1-\delta} \cdot \lim_{k\to\infty} a_k.$$

Proof. An elementary exercise. □

Hence we have the following theorem.

Theorem 5.3 *Let $\tilde{Q}$ be the subgroup of Q consisting of all elements of even length. Then*

$$\lim_{k\to\infty} \pi_k(Q) = \begin{cases} \frac{2}{|G:\tilde{Q}|} & \text{if} \quad |G:Q| < \infty \\ 0 & \text{otherwise .} \end{cases}$$

Proof. We analyze the expression for π_k given in formula (7). Notice that the relative frequencies of elements of length $2k$ from $\tilde{Q}$ are q_{2k}. By Woess [27] the limit $\lim_{k\to\infty} q_{2k}$ exists and equals either 0, if $\tilde{Q}$ (and hence Q) has infinite index, or $2/|G:\tilde{Q}|$. Hence we can apply Lemma 5.2 and conclude that the limit

$$\lim_{k\to\infty} q^*_{2k} = \lim_{k\to\infty} \left(q_{2k} + q_{2k-2}\delta^2 + q_{2k-4}\delta^4 + \cdots + q_2\delta^{2k-2} + q_0\delta^{2k}\right)$$

exist and equal

$$\lim_{k\to\infty} q^*_{2k} = \frac{1}{1-\delta^2} \cdot \lim_{k\to\infty} q_{2k}.$$

Now $\lim_{k\to\infty} \pi_k(Q)$ exists and after a simple rearrangement we have

$$\begin{aligned} \lim_{k\to\infty} \pi_k(Q) &= \left(\frac{1}{d+1} + \frac{d-1}{d+1} \cdot \frac{1}{1-\delta^2}\right) \cdot \left(\lim_{k\to\infty} q_{2k}\right) \\ &= \lim_{k\to\infty} q_{2k} \\ &= \begin{cases} \frac{2}{|G:\tilde{Q}|} & \text{if} \quad |G:Q| < \infty \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

as required. □

6 Quotient tests for the conjugacy problem

The conjugacy problem suffers the most when is transplanted from a group to its quotient group. For example, if we use the quotient test $F/S \simeq \mathrm{Sym}_n$ for the conjugacy of elements in the braid group B_n, we immediately discover that random permutations happen to be conjugate so frequently that there is no need to run sophisticated analysis based on random walks or the multiplicative measure. Indeed, in the uniform probability distribution on Sym_n, a random element is a n-cycle with probability $1/n$; two independent random elements happen to simultaneously be n-cycles (and therefore conjugate) with probability $1/n^2$. Hence the probability for two random permutations to be conjugate is at least $1/n^2$.

The quotient test with the cyclic quotient group G/D of the same order $n!$ as Sym_n distinguishes non-conjugate elements much better: in that case, the conjugacy is just equality, and results of the previous section apply.

Notice in passing that a surprisingly crude estimate for the probability of two permutations to be conjugate cannot be much improved. Indeed, Laci Babai and Laci Pyber communicated to us the following result.

Fact 6.1 (L. Babai and L. Pyber) *The probability p_n for two random elements of the symmetric group Sym_n to be conjugate asymptotically behaves as*

$$p_n \sim \frac{c}{n^2}$$

where $c > 3$ is a computable constant. Moreover,

$$\frac{1}{n^2} \leqslant p_n \leqslant \frac{1}{n-1}$$

for all n.

The proof is elementary and beautiful; we do not give it here since the sharper estimate does not add much to the crude lower bound $\geqslant 1/n^2$.

References

[1] S. I. Adian and V. G. Durnev, *Algorithmic problems for groups and semigroups,* Uspekhi Mat. Nauk **55** (2000), no. 2, 3–94; translation in Russian Math. Surveys **55** (2000), no. 2, 207–296

[2] L. Bartholdi, *Counting paths in graphs*, Ensignment Math. **45** (1999), 83–131.

[3] A. V. Borovik, A. G. Myasnikov and V. N. Remeslennikov, *Multiplicative measures on free groups*, Int. J. Algebra Comp. **13** no. 6 (2003), 705–731.

[4] J. M. Cohen, *Cogrowth and amenability of discrete groups*, J. Fuct. Anal. **48** (1982), 301–309.

[5] P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrscheinlichkeitstheorie Verw. Gebiete **57** (1981), 159–179.

[6] D. B. A. Epstein, with J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson and W. P. Thurston, "Word processing in groups". Jones and Bartlett, Boston-London (1992).

[7] P. Flajolet and R. Sedgwick, "Analytic Combinatorics: Functional Equations, Rational and Algebraic Functions", Res. Rep. INRIA RR4103, January 2001, 98 pp.

[8] R. I. Grigorchuk, *Symmetrical random walks on discrete groups*, in "Multicomponent random systems" (R. L. Dobrushin and Ya. G. Sinai, eds.), Dekker, New York, 1980, pp. 285–325.

[9] M. Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. **53** (1981), 53–78.

[10] S. P. Humphries, *Cogrowth of groups and the Dedekind-Frobenius group determinant*, Math. Proc. Cambridge Phil. Soc. **121** (1997), 193-217.

[11] I. Kapovich, *The non-amenability of Schreier graphs for infinite index quasiconvex subgroups of hyperbolic groups*, preprint.

[12] I. Kapovich, A. G. Myasnikov, P. E. Schupp and V. Shpilrain, *Generic-case complexity, decision problems in group theory and random walks*, J. Algebra **264** no. 2 (2003), 665–694.

[13] I. Kapovich, A. G. Myasnikov, P. E. Schupp and V. Shpilrain, *Average-case complexity and decision problems in group theory*, Adv. Math. **190** no. 2 (2005), 343–359.

[14] I. Kapovich and P. E. Schupp, *Genericity, the Arzhantseva-Ol'shanskii method and the isomorphism problem for one-relator groups*, Math. Ann. **331** no. 1 (2005), 1–19.

[15] I. Kapovich and P. E. Schupp, *Delzant's T-invariant, Kolmogorov complexity and one-relator groups*, Commentari Math. Helv., to appear.

[16] D. Kouksov, *On rationality of the cogrowth series*, Proc. Amer. Math. Soc. **126** (1998), 2845–2847.

[17] D. Kuksov, *Cogrowth series of free products of finite and free groups*, Glasgow Math. J. **41** (1999), 19–31.

[18] C. F. Miller III, *On Group-theoretic Decision Problems and their Classification,* Ann. of Math. Studies, **68** (1971) Princeton University Press, Princeton.

[19] C. F. Miller III, *Decision problems for groups- Survey and reflections.* in *Algorithms and Classification in Combinatorial Group Theory*, (1992), Springer, 1–60.

[20] D. E. Muller and P. E. Schupp, *Groups, the theory of ends, and context-free languages*, J. Comp. Syst. Sci. **26** (1983), 295–310.

[21] S. Northshield, *Cogrowth of regular graphs*, Proc. Amer. Math. Soc. **116** (1992), 203–205.

[22] F. Spitzer, "Principles of Random Walk", Springer-Verlag, New York, 2001.

[23] R. Szwarc, *The ratio and generating function of cogrowth coefficients of finitely generated groups*, Studia Mathematica **131** (1998), 89–94.

[24] R. P. Stanley, "Enumerative Combinatorics", vol. 2, Cambridge University Press, 1999.

[25] G. Quenell, *Combinatorics of free product graphs*, Contemp. Math. **173** (1994), 257–281.

[26] N. Varopoulos, L. Saloff-Coste and T. Coulhon, "Analysis and Geometry on Groups", Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992.

[27] W. Woess, *Cogrowth of groups and simple random walks*, Arch. Math. **41** (1983), 363–370.

[28] W. Woess, *Context-free languages and random walks on groups*, Discrete Math. **67** (1987), 81–87.

[29] W. Woess, "Random Walks on Infinite Graphs and Groups", Cambridge University Press, Cambridge, 2000.

Alexandre V. Borovik
School of Mathematics, PO Box 88, The University of Manchester, Sackville Street, Manchester M60 1QD, United Kingdom
borovik@manchester.ac.uk
http://www.ma.umist.ac.uk/avb/

Alexei G. Myasnikov
Department of Mathematics and Statistics, McGill University, Montreal, QC, Canada, H3A2K6
alexeim@att.net
http://home.att.net/~alexeim/index.htm