THE DENSITY OF INJECTIVE ENDOMORPHISMS OF A FREE GROUP

A. MARTINO, E.C.TURNER, AND E. VENTURA

Abstract. We show that among the endomorphisms of the free (non-abelian) group F_r of rank r, the set of monomorphisms (the injective endomorphisms) has density one. This contrasts with the known fact that the set of automorphisms has density zero. We show more generally that in the set of homomorphisms from one free group to another, the set of monomorphisms has density one whereas the set of epimorphisms has density zero.

0. Introduction

If $F_r = F(x_1, \cdots, x_r)$ is the free group of rank r, then by $Aut(F_r)$, $Mono(F_r)$ and $End(F_r)$ we mean the sets of automorphisms, monomorphisms and endomorphisms of F_r respectively. Clearly

$$Aut(F_r) \subset Mono(F_r) \subset End(F_r).$$

One can measure the density of each set in the next—see §1 for details. Denoting by $\delta_{Aut(F_r)}$ and $\delta_{Mono(F_r)}$ the densities of $Aut(F_r)$ and $Mono(F_r)$ in $End(F_r)$, then our main theorem is the following.

Theorem 1. *For all r,*

$$\delta_{Aut(F_r)} = 0 \quad \textit{and} \quad \delta_{Mono(F_r)} = 1.$$

This is certainly the result that one would expect, and the fact that $\delta_{Aut(F_r)} = 0$ is essentially the result of [1].

One can express Theorem 1 by saying that in the sequence

$$Aut(F_r) \subset Mono(F_r) \subset End(F_r),$$

The authors gratefully acknowledge the hospitality of the Centre de Recerca Matemàtica in Bellaterra, Spain. The second author was partially supported by Spanish government grant SAP2003-0253. The third author was partially supported by Spanish government grant DGI2003-06613.

$Aut(F_r)$ is 'sparse' in $End(F_r)$ and $Mono(F_r)$ is 'dense' in $End(F_r)$. (It follows easily that $Aut(F_r)$ is sparse in $Mono(F_r)$.) We can denote this schematically as

$$Aut(F_r) \ll Mono(F_r) \doteq End(F_r), \qquad (*)$$

where "$\ll$" means density zero and "$\doteq$" means density one.

But we point out a curious reversal in the relationship between monomorphisms and automorphisms when considering their fixed subgroups. If one considers the sets

$$\begin{aligned} 1\text{-}Autofixed(F_r) &= \{H \subset F_r \mid H = \mathrm{Fix}(\alpha) \text{ for some automorphism } \alpha\} \\ 1\text{-}Monofixed(F_r) &= \{H \subset F_r \mid H = \mathrm{Fix}(\mu) \text{ for some monomorphism } \mu\} \\ 1\text{-}Endofixed(F_r) &= \{H \subset F_r \mid H = \mathrm{Fix}(\varphi) \text{ for some endomorphism } \varphi\}, \end{aligned}$$

then of course

$$1\text{-}Autofixed(F_r) \quad \subset \quad 1\text{-}Monofixed(F_r) \quad \subset \quad 1\text{-}Endofixed(F_r).$$

But in fact 1-$Autofixed(F_r)$ =1-$Monofixed(F_r)$ (this follows from [5]) but 1-$Monofixed(F_r)$ is strictly smaller than 1-$Endofixed(F_r)$ [7] (and in fact seems likely to have density zero). Schematically, in contrast to $(*)$,

$$1\text{-}Autofixed(F_r) \quad = \quad 1\text{-}Monofixed(F_r) \quad \ll \quad 1\text{-}Endofixed(F_r).$$

In §2 we extend the techniques of §1 to study homomorphisms from one free group to another. We denote by $Homo(F_s, F_r)$, $Mono(F_s, F_r)$ and $Epi(F_s, F_r)$ the sets of homomorphisms, monomorphisms and epimorphisms from F_s to F_r. Again there is a natural notion of density (see §2): denote by $\delta_{Mono(F_s,F_r)}$ and $\delta_{Epi(F_s,F_r)}$ the densities of monomorphisms and epimorphisms respectively. We prove the following somewhat counter-intuitive theorem[1].

Theorem 2. *For all r and s,*

$$\delta_{Mono(F_s,F_r)} = 1 \qquad \textit{and} \qquad \delta_{Epi(F_s,F_r)} = 0.$$

In a forthcoming paper, we will establish a much more precise estimate of the density of monomorphisms from one free group to another.

[1]Thus most maps from F_{1000} to F_2 are injective but not surjective.

1. Proof of Theorem 1

We begin with definitions.

Definition 1. *Let $F_r = F(x_1, \cdots, x_r)$ be the free group generated by variables x_i and*

$$Aut(F_r) \subset Mono(F_r) \subset End(F_r)$$

where

$$\begin{aligned} End(F_r) &= \{\varphi : F_r \to F_r \mid \varphi \text{ is an endomorphism}\}, \\ Mono(F_r) &= \{\mu : F_r \to F_r \mid \mu \text{ is a monomorphism}\}, \\ Aut(F_r) &= \{\alpha : F_r \to F_r \mid \alpha \text{ is an automorphism}\}. \end{aligned}$$

Relative to the usual length function

$$|w| = \#\{\text{letters in the reduced word } w\}$$

on F_r, the ball of radius n in F_r is

$$B_n(F_r) = \{w \mid |w| \le n\}.$$

Similarly, we have

$$\begin{aligned} End_n(F_r) &= \{\varphi \in End(F_r) \quad \mid \varphi(x_i) \in B_n(F_r) \quad \text{for all } i\}, \\ Mono_n(F_r) &= \{\mu \in Mono(F_r) \mid \mu(x_i) \in B_n(F_r) \quad \text{for all } i\}, \\ Aut_n(F_r) &= \{\alpha \in Aut(F_r) \quad \mid \alpha(x_i) \in B_n(F_r) \quad \text{for all } i\}. \end{aligned}$$

An endomorphism is really an ordered r-tuple of words in F_r. By the *density* of automorphisms or monomorphisms, we mean the probability that a random endomorphism, considered as an r-tuple all of whose entries have length less than some very large bound, will be an automorphism or a monomorphism.

Definition 2. *The density*[2] $\delta_{Aut(F_r)}$ *of automorphisms and* $\delta_{Mono(F_r)}$ *of monomorphisms of F_r are*

$$\delta_{Aut(F_r)} = \lim_{n\to\infty} \left\{ \frac{\#\left(Aut_n(F_r)\right)}{\#(End_n(F_r))} \right\}$$

and

$$\delta_{Mono(F_r)} = \lim_{n\to\infty} \left\{ \frac{\#\left(Mono_n(F_r)\right)}{\#(End_n(F_r))} \right\}.$$

[2]Technically speaking we should prove the existence of these limits before making the definition. Theorem 1 will establish existence.

Counting the number of elements of $End_n(F_r)$ is easy since there is no restriction on the words

chosen except that they have length bounded by n.

An element φ of $End_n(F_r)$ is the ordered r-tuple

$$\{w_1, \cdots, w_r\} = \{\varphi(x_1), \cdots, \varphi(x_r)\}.$$

(Sets of words will always be considered as ordered.) For $n \geq 1$, the number of words of length exactly n is $2r(2r-1)^{n-1}$ (since the first letter can be any of the $\{x_1^{\pm 1}, \cdots, x_r^{\pm 1}\}$ and each succeeding choice must avoid the inverse of the previous letter). Thus the number of words of length less than or equal to n is

$$\begin{array}{rcl} 1 + 2r + 2r(2r-1) + \cdots + 2r(2r-1)^{n-1} & = & \\ 1 + 2r\left\{1 + (2r-1) + \cdots + (2r-1)^{n-1}\right\} & = & \\ 1 + 2r\left\{\frac{(2r-1)^n - 1}{2r-1}\right\} & \geq & (2r-1)^n. \end{array}$$

Making this estimate for each i,

$$\#(End_r(n)) \geq (2r-1)^{rn}.$$

That $\delta_{Aut(F_r)} = 0$ is a corollary of the theorem of Burillo and Ventura [1]. The endomorphism determined by the set $\{w_1, \cdots, w_r\}$ is an automorphism if and only if the set is a basis; in particular, w_1 must be a primitive. Burillo and Ventura prove that the proportion of primitive words in the n-ball goes to zero as n goes to infinity from which it follows easily that $\delta_{Aut(F_r)} = 0$.

Our main result is that $\delta_{Mono(F_r)} = 1$. The main idea is to show that the number of r-tuples which correspond to non-monomorphisms grows exponentially with a base strictly smaller than $2r-1$.

Imagine selecting the $w_i's$ in the r-tuple

$$\{w_1, \cdots, w_r\} = \{\varphi(x_1), \cdots, \varphi(x_r)\}$$

in sequence and at each point asking whether the subgroup generated by $\{w_1, \cdots, w_k\}$ has rank k. If not, then no matter how we continue selecting $\{w_{k+1}, \cdots, w_r\}$, the subgroup can't have rank r. The following terminology will be useful.

Definition 3. *In F_r, a subset $\mathcal{W} = \{w_1, \cdots, w_k\}$ is* viable *if the rank of $\langle \mathcal{W} \rangle$ is k. Given a viable set $\mathcal{W}$, a word w is a $\mathcal{W}$-viable choice (or a viable choice if context makes clear what $\mathcal{W}$ is) if the rank of $\langle \mathcal{W}, w \rangle$ is $k+1$. Thus,*

$$\begin{array}{rcrcl} \mathcal{W} \text{ is viable} & \Longleftrightarrow & rank\langle \mathcal{W} \rangle & = & k \\ w \text{ is } \mathcal{W}\text{-viable} & \Longleftrightarrow & rank\langle \mathcal{W}, w \rangle & = & k+1. \end{array}$$

To estimate the size of $Mono_n(F_r)$, we build viable sets in $B_n(F_r)$ one viable choice at time and show that as n gets large, the proportion of non-viable choices approaches zero. The first two steps are particularly easy to understand. The only non-viable set of cardinality one is $\{1\}$, so at the first step, the number of non-viable choices is 1 and the proportion of non-viable choices is less than $\frac{1}{(2r-1)^n}$. If $\mathcal{W} = \{w_1\}$ is a viable set of cardinality one (i.e., $w_1 \neq 1$), then let v_1 be the *root* of w_1; i.e. w_1 is a power of v_1 but not of any shorter word. Then the non-viable choices in $B_n(F_r)$ are precisely those powers of v_1 that have length at most n, of which there can be no more than $2n$. Therefore for any viable set of cardinality one, the number of non-viable choices is at most $2n$ and the proportion of non-viable choices is less than $\frac{2n}{(2r-1)^n}$. The key observation involved in continuing this process is the following.

Proposition 1. *Suppose $\mathcal{W} = \{w_1, \cdots, w_k\}$ is a viable set of words in F_r, $r \geq 2$ and $k < r$. Then for large n, the number of non-viable choices w for $\mathcal{W}$ of length bounded by n is at most $K(2n2+n+1)(2k)^n$ for some universal constant K (independent of $\mathcal{W}$ and k).*

The theorem follows rather quickly from the proposition. Determining a monomorphism φ involves making r choices, all of which must be viable for the endomorphism to be a monomorphism. Denote by $NV_{n,k}$ the number of non-viable sets

$\{w_1, \cdots, w_r\}$ of words of length at most n so $\{w_1, \cdots, w_{k-1}\}$ is viable but $\{w_1, \cdots, w_k\}$ is not: i.e., the sequence of selections first becomes non-viable at step k. Then

$$\#\{\text{non-viable } r\text{-tuples}\} = NV_{n,1} + \cdots + NV_{n,r}$$

so it suffices to show

$$k \leq r \qquad \Longrightarrow \qquad \lim_{n\to\infty} \frac{NV_{n,k}}{(2r-1)^{nr}} = 0.$$

In determining the ratio of $NV_{n,k}$ to the size of the n-ball in $Endo_r$, we can clearly ignore the factors corresponding to $\{w_{k+1}, \cdots, w_r\}$ since they are the same in both numerator and denominator. In other words, if

$$NV'_{n,k} = \# \left\{\text{non-viable } \{w_1,\cdots,w_k\} \subset B_n(F_r)^k \mid \{w_1, \cdots, w_{k-1}\} \text{ is viable}\right\}$$

then

$$\frac{NV_{n,k}}{(2r-1)^{nr}} = \frac{NV'_{n,k}}{(2r-1)^{nk}}.$$

The cases $k = 1$ and $k = 2$ follow from the comments preceding Proposition 1 so suppose $k > 2$. By Proposition 1, for any viable set $\mathcal{W} = \{w_1, \cdots, w_{k-1}\}$ there are at most $K(2n2+n+1)(2k-2)^n$ non-viable choices for w_k. Therefore NV'_k is less than $K(2n2+n+1)(2k-2)^n$ times the number of sets $\{w_1, \cdots, w_{k-1}\}$ (regardless of viability). That is

$$NV'_{n,k} \leq (2r(2r-1)^{n-1})^{k-1} K(2n2+n+1)(2k-2)^n.$$

Then since

$$\begin{aligned} \tfrac{(2r(2r-1)^{n-1})^{k-1}}{(2r-1)^{nk}} &= \left(\tfrac{2r}{2r-1}\right)^{k-1} \left(\tfrac{1}{2r-1}\right)^n \\ &\leq \left(\tfrac{2r}{2r-1}\right)^{r-1} \left(\tfrac{1}{2r-1}\right)^n \end{aligned}$$

we have

$$\begin{aligned} \frac{NV'_{n,k}}{(2r-1)^{nk}} &\leq \frac{(2r(2r-1)^{n-1})^{k-1}}{(2r-1)^{nk}} K(2n2+n+1)(2k-2)^n \\ &\leq K'(2n2+n+1)\left(\frac{2r-2}{2r-1}\right)^n \end{aligned}$$

where the constant $\left(\tfrac{2r}{2r-1}\right)^{r-1}$ has been absorbed into K' and we have used the fact that $k \leq r$.

Finally, if λ_r is any number so that

$$\left(\frac{2r-2}{2r-1}\right) < \lambda_r < 1,$$

then for large n,

$$(2n2+n+1)\left(\frac{2r-2}{2r-1}\right) < \lambda_r^n.$$

This follows from the fact that if $1 < \alpha < \beta$ and p is a polynomial, then

$$\lim_{n\to\infty}\left\{\frac{p(n)\alpha^n}{\beta^n}\right\} = 0.$$

□

The remainder of the section is devoted to a proof of Proposition 1. We begin by recalling some relatively standard terminology.

Definition 4. *Suppose H is a subgroup of F_r and X_H is the corresponding covering space of the wedge of r circles.*

a) *The core C_H of H is the smallest subgraph of X_H containing the base point $*$ which has fundamental group H.*
b) *The reduced core $\widetilde{C}_H$ of H is C_H with a maximal tail τ to the base point $*$ removed.*
c) *The point at which τ enters $\widetilde{C}$ is the* reduced base point *denoted $*'$.*

Let C be the core of the subgroup $H = \langle w_1, \cdots, w_k \rangle$ with associated reduced core $\widetilde{C}$, tail τ and reduced base point $*'$. Since $\mathcal{W}$ is viable, the rank of (the fundamental group) of C is k. To decide whether a word w is $\mathcal{W}$-viable, we adjoin a circle labeled with w at the base point of C and then *fold* [8] the ends of the added circle into C as far as possible in both directions, followed by any necessary folds within C to produce the core C' for the subgroup $\langle \mathcal{W}, w \rangle$. If any part of the added loop survives the folding, in that it does not get identified into C, then $rank(C') = k + 1$ and w is viable. In other words, the only way that w can be non-viable is that the entire added loop is folded into C. In this case, suppose $w = uv^{-1}$ where u and v are folded into C starting from the base point. If this can be done in several ways, always choose u as long as possible so that w determines the pair (u, v). This means that C contains paths starting at the base point $*$, one labeled by u and the other by v. Suppose these paths end at points p and q respectively—then C' is obtained from C by identifying p to q and then folding to get a core graph. In this way, every non-viable w gives rise to a pair of points p and q in C and a path from p to q—labeled by $u^{-1}v$—that visits the base point and has length at most n. Note that the paths may be non-reduced, since there may be cancellation between u^{-1} and v. Note also that $p = q$ if and only if $w \in \langle w_1, \cdots, w_k \rangle$, clearly a non-viable choice. Conversely, a path in C from p to q of length at most n can visit the base point at most n times and so can come from at most n non-viable words w.

We have therefore proven the following claim[3].

Claim 1. *The number of non-viable words of length at most n for $\mathcal{W}$ is at most n times the number of (not necessarily reduced) paths in C of length at most n.*

For technical reasons discussed below, we need to argue further when $C \neq \widetilde{C}$: i.e., the degree of $*$ is one. In this case, the paths u and v must agree with τ for a while—so $u = \tau_1 u'$ and $v = \tau_2 v'$ where τ_1 and τ_2 are non-empty initial segments of τ. Then $u^{-1}v = {u'}^{-1}v'$ and is a path of one of several types[4];

a) a path in $\widetilde{C}$ that visits $*'$,
b) a path in $\widetilde{C}$ starting at $*'$ preceded by a non-empty subpath of τ,
c) a path in $\widetilde{C}$ ending at $*'$ followed by a non-empty subpath of τ,

[3]We suspect that this is a very weak estimate, since we are counting many reduced paths (see §2 in this regard), counting paths that don't visit the base point and many of the selected paths will in fact correspond to viable choices.

[4]Recall that $*'$ is the vertex at which the tail τ enters $\mathfrak{C}$.

d) a closed loop at $*'$ preceded and/or followed by non-empty subpaths of τ starting or ending at $*'$,
e) a subpath of τ.

If N is the number of paths in $\widetilde{C}$ of length at most n, the numbers of paths of types a), b), c), d) and e) of length at most n are bounded by N, nN, nN, n^2N and $n(n-1)$ (which is less than $n(n-1)N$)) respectively for a total bounded by

$$(1 + n + n + n2 + n(n-1))\, N = (2n2 + n + 1)N.$$

Thus we have proven the following claim.

Claim 2. *The number of non-viable words of length at most n for $\mathcal{W}$ is at most $2n2 + n + 1$ times the number of (not necessarily reduced) paths in $\widetilde{C}$ of length at most n.*

The proof of Proposition 1 is completed by bounding the number of such paths by a multiple of $2k^n$. The key tools are the *adjacency matrix* of C and the notion of *spectral radius* (see, e.g., [4]), the basic properties of which we now review.

The spectral radius of a graph: review

In this section, all graphs are undirected.

Definition 5. *The* adjacency matrix of a graph C *with v vertices is the matrix A of size $(v \times v)$ whose entry A_{ij} in the (i,j) position is the number of edges between the i^{th} vertex and the j^{th} vertex.*

In this regard, a loop will be considered as two edges which are inverse (in a directed sense) but a non-loop as one. Thus A has even diagonal entries, is symmetric (so all its eigenvalues are real) and each edge in C contributes two to the sum of the entries of A. Each row and column sum of A is the degree of the corresponding vertex of C. By an *n-path* in C we mean an edge path in C of length n. A path is *reduced* if no two successive edges are inverses.

Fact: *Suppose A is the adjacency matrix of C, and u is the column vector whose transpose is $u^t = (1, \cdots, 1)$. The $(i,j)^{th}$ entry A^n_{ij} of A^n counts the number of n-paths in C from vertex i to vertex j. The number of all n-paths from any one vertex to any other is*

$$\#(n\text{-paths in } C) = u^t A^n u.$$

This is because the $(i,j)^{th}$ entry of A^n is the sum of terms of the form

$$a_{k_1k_2}a_{k_2k_3}\cdots a_{k_nk_{n+1}} \qquad i=k_1, k_{n+1}=j$$

which counts the number of paths from the i^{th} vertex to the j^{th} that visit the vertices indexed by k_2,k_3, etc in sequence, proving a).

Definition 6. *The spectrum of a matrix M is its set of eigenvalues and the spectral radius $\rho(M)$ is the absolute value of the largest eigenvalue.*

A square matrix M of size $(n \times n)$ is *reducible* if $n \geq 2$ and there is a permutation matrix P so that

$$PMP^{-1} = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

and is *irreducible* otherwise. For any matrix M there is permutation matrix P so that

$$PMP^{-1} = \begin{pmatrix} M_1 & * & * & * \\ 0 & M_2 & * & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & M_3 \end{pmatrix}$$

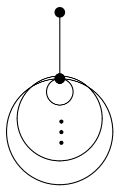
where the blocks are irreducible. Note that all (1×1) matrices are irreducible, so some of the diagonal blocks might be zeroes.

The Perron-Frobenius Theorem ([4], p508 or [3], p28) says that if M is a non-zero, irreducible, non-negative matrix M, then $\rho(M)$ is positive and has a unique positive eigenvector (up to multiplication by a constant) which is larger in absolute value than all the other eigenvalues. If C a connected graph, then its adjacency matrix A is irreducible since $A + A2 + \cdots + A^{v-1}$ has strictly positive entries.

Definition 7. *The spectral radius $\rho(C)$ of a graph C is the spectral radius of its adjacency matrix A.*

We will see shortly that if the degree of the base point of a connected graph C is not 1 then its spectral radius is bounded by twice its rank. That this is not always true if the base point has degree 1 is shown by the following example. This is why we work in $\widetilde{C}$ rather than C.

Example 1. *The 'earring' graph E_k has two vertices, base point $*$ of degree 1, one other vertex $*'$ of degree $2k+1$ and adjacency matrix A_k (below). The rank of E_k is k and its spectral radius is $\rho = k + \sqrt{k2+1}$.*

E_k  $A_k = \begin{pmatrix} 2k & 1 \\ 1 & 0 \end{pmatrix}$ $\rho(E_k) = k + \sqrt{k2+1}$

The following propositions collect the facts that we need about spectral radii, Proposition 2 for matrices and Proposition 3 for graphs. Good sources for the properties of spectral radii of matrices and graphs are [4], [3] and [2]. If A and B are matrices, by $A \leq B$ we mean that A and B are the same size and the inequality is true in every entry. By $A < B$, we mean $A \leq B$ and the inequality is strict in at least one coordinate. If C and D are graphs, then by $C \leq D$ we mean that C is a subgraph of D in the sense that all the vertices and edges of C are vertices and edges respectively of D, and by $C < D$ we mean $C \leq D$ but $C \neq D$.

Proposition 2. *Let A and B be non-negative matrices of the same size.*

a) *If A is irreducible and m and M are the minimum and maximum row sums of A, then*
$$m \leq \rho(A) \leq M,$$
with strict inequalities if $m < M$.

b) *For irreducible matrices,*
$$A < B \implies \rho(A) < \rho(B).$$

c) *In general,*
$$A \leq B \implies \rho(A) \leq \rho(B).$$

d) *If B is obtained from A by deleting corresponding rows and columns, then $\rho(B) \leq \rho(A)$. If A is irreducible and any non-zero entry is removed, the inequality is strict.*

Proposition 3. *Let C and D be connected graphs.*

a) *If m and M are the minimum and maximum degrees of the vertices of C, then*

$$m \leq \rho(C) \leq M,$$

with strict inequalities if $m < M$.

b) *In general,*

$$C < D \implies \rho(C) < \rho(D).$$

c) *If C has no degree 1 vertices and the graph D is obtained from C by subdividing an edge with a new degree two vertex, then $\rho(D) \leq \rho(C)$ with equality only when C is a circle*[5].

d) *In general,*

$$1 \leq rank(C) \implies 2 \leq \rho(C)$$

with equality only for the circle.

e) *If the degree of the base point of C is at least two, then*

$$rank(C) = k \implies \rho(C) \leq 2k$$

with $\rho(C) = 2k$ only when C is a k-regular graph (all vertices of degree k).

Proof of Proposition 2: Most of these properties are well known and are collected here for convenience.

Suppose m and M are the smallest and largest rows sums of A respectively. If the positive eigenvector for ρ is x, then

$$Ax = \begin{pmatrix} a_{11} & \cdots & a_{1v} \\ \vdots & \vdots & \vdots \\ a_{v1} & \cdots & a_{vv} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_v \end{pmatrix} = \rho \begin{pmatrix} x_1 \\ \vdots \\ x_v \end{pmatrix}$$

and for each i,

$$\sum_{j=1}^{v} A_{ij} x_j = \rho x_i.$$

Thus

$$\rho = \sum_{j=1}^{v} A_{1j} \left(\frac{x_j}{x_1} \right) = \sum_{j=1}^{v} A_{2j} \left(\frac{x_j}{x_2} \right) = \cdots = \sum_{j=1}^{v} A_{vj} \left(\frac{x_j}{x_v} \right) .$$

[5]This statement is false if the subdivided edge is part of a path to a degree one vertex, in which case the spectral radius increases. This is one reason that we work in $\mathfrak{C}$ rather then C.

Reorder the coordinates for convenience so that x_1 is the smallest and x_v the largest. Then

$$m \leq \sum_{j=1}^{v} A_{1j} \leq \sum_{j=1}^{v} A_{1j} \left(\frac{x_j}{x_1}\right) = \rho = \sum_{j=1}^{v} A_{vj} \left(\frac{x_j}{x_v}\right) \leq \sum_{j=1}^{v} A_{vj} \leq M.$$

If $m < M$ then both inequalities are strict. This establishes a).

The most useful tools in understanding spectral radii of matrices are items b) and c): see [3], p30 and [4] p491. One can derive c) from b) by arguing in the various diagonal blocks.

If B is obtained from A by deleting corresponding rows and columns, then we can think of first reducing entries to be removed to 0, rearranging rows and columns so the deleted ones are the last few and then throwing away the rows and columns. The first steps reduces the value of ρ and the last leaves it unchanged—since ρ is positive, removing eigenvectors with value zero does not change it. This proves d) and completes the proof of Proposition 2. □

Proof of Proposition 3: Part a) follows immediately from Proposition 2a) since row sums of an adjacency matrix are the corresponding vertex degrees. Part b) follows immediately from Proposition 2b).

Item d) follows from a) since if the $rank(C) \geq 1$, then C contains a circle and a circle has spectral radius 2.

The row sums are the degrees of the vertices of C, so it suffices by a) to show that if the base point has degree at least two, then all vertices of C have degree at most $2k$. Denote the vertices of C by $\{z_1, \cdots, z_v\}$ and the number of edges of C by e. Then

$$\sum_{i=1}^{v} (deg(z_i) - 2) = \sum_{i=1}^{v} deg(z_i) - 2v = 2e - 2v = 2k - 2.$$

This uses the well-known facts that the sum of the degrees of the vertices of any graph is twice the number of edges and that for a connected graph of rank k,

$$v - e = 1 - k = \chi(C), \text{ the Euler characteristic of } C.$$

If all vertices have degree at least two, then each of the summands $deg(v_i) - 2$ is non-negative and each must be at most their sum; i.e.,

$$deg(z_i) - 2 \leq 2k - 2 \quad \Longrightarrow \quad deg(z_i) \leq 2k.$$

□

We now return to the proof of Proposition 1: see Claim 2 and the comment that follows it.

The number of paths in $\widetilde{C}$ (not necessarily reduced) from the i^{th} vertex to the j^{th} vertex is $A^n_{i,j}$, the $(i,j)^{th}$ entry of A^n. Thus the number of paths in $\widetilde{C}$ of length n from any vertex to any other is

$$u^t A^n u = \sum_{i,j=1}^{v} A^n_{ij}$$

where u is the column vector whose transpose is $u^t = (1, \cdots, 1)$. Now let $\{\alpha_1, \cdots, \alpha_v\}$ be a basis of real eigenvectors for A (which exists since A is symmetric) with corresponding eigenvalues $\rho_1 > \rho_2 \geq \rho_3 \geq \cdots \geq \rho_v$, and write

$$u = b_1\alpha_1 + \cdots + b_v\alpha_v.$$

Then

$$\begin{aligned}
\textstyle\sum_{i,j=1}^{v} A^n_{ij} &= u^t A^n u \\
&= u^t A^n(b_1\alpha_1 + \cdots + b_v\alpha_v) \\
&= u^t(b_1 A^n\alpha_1 + \cdots + b_v A^n\alpha_v) \\
&= u^t(b_1\rho_1^n\alpha_1 + \cdots + b_v\rho_v^n\alpha_v) \\
&= (b_1 u^t\alpha_1)\rho_1^n + \cdots + (b_v u^t\alpha_v)\rho_v^n \\
&= c_1\rho_1^n + \cdots + c_n\rho_v^n.
\end{aligned}$$

where $c_i = b_i u^t \alpha_i$. But since $\rho = \rho_1$ is strictly larger than the other $\rho_i' s$, this is dominated by the first term if $c_1 \neq 0$. In any case, it is dominated by the first non-zero term and for $K = \sum |c_i|$

$$\sum_{i,j=1}^{v} A^n_{ij} \leq K\rho^n.$$

Since $rank(C) = k$, by Proposition 3f), $\rho \leq 2k$. This completes the proof of Proposition 1. $\square$

2. Proof of Theorem 2

In this section we show how to extend the methods of §1) to compute the density of monomorphisms from one free group to another. The density of epimorphisms follows immediately, as follows. If $s < r$, then $Epi(F_s, F_r)$ is empty. If $s = r$, $Epi(F_s, F_r) = Aut(F_r)$ which was treated in §1. If $s > r$, then $Epi(F_s, F_r)$ is contained in the complement of $Mono(F_s, F_r)$ (since F_s and F_r are not isomorphic). Therefore if $Mono(F_s, F_r)$ has density one, then $Epi(F_s, F_r)$ has density zero. We devote the rest of the section to monomorphisms.

We begin the section with definitions.

Definition 8. *If F_s and F_r are two free groups, then*

$$Mono(F_s, F_r) \subset Homo(F_s, F_r), \qquad Epi(F_s, F_r) \subset Homo(F_s, F_r)$$

where

$$\begin{aligned} Homo(F_s, F_r) &= \{\varphi : F_s \to F_r \mid \varphi \text{ is a homomorphism}\}, \\ Mono(F_s, F_r) &= \{\mu : F_s \to F_r \mid \mu \text{ is a monomorphism}\}, \\ Epi(F_s, F_r) &= \{\varepsilon : F_s \to F_r \mid \varepsilon \text{ is an epimorphism}\}. \end{aligned}$$

Relative to the length function on F_r with balls $B_n(F_r)$ of radius n, we have

$$\begin{aligned} Homo_n(F_s, F_r) &= \{\varphi \in Homo(F_s, F_r) \mid \varphi(x_i) \in B_n(F_r), \quad \forall i\} \\ Mono_n(F_s, F_r) &= \{\mu \in Mono(F_s, F_r) \mid \mu(x_i) \in B_n(F_r), \quad \forall i\}, \\ Epi_n(F_s, F_r) &= \{\varepsilon \in Epi(F_s, F_r) \mid \varepsilon(x_i) \in B_n(F_r), \quad \forall i\}, \end{aligned}$$

The respective densities are

$$\delta_{Mono(F_s,F_r)} = \lim_{n\to\infty} \left\{ \frac{\#\left(Mono_n(F_s, F_r)\right)}{\#(Homo_n(F_s, F_r))} \right\}$$

and

$$\delta_{Epi(F_s,F_r)} = \lim_{n\to\infty} \left\{ \frac{\#\left(Epi_n(F_s, F_r)\right)}{\#(Homo_n(F_s, F_r))} \right\}.$$

The homomorphism $\varphi : F_s \to F_r$ is the ordered s-tuple

$$\mathcal{W} = \{w_1, \cdots, w_s\}$$

of words in F_r and φ is a monomorphism if and only if $\{\mathcal{W}\}$ is viable. If $s \leq r$, then the proof of §1 can be used verbatim. If $s > r$, however, the ranks of the graphs C encountered are no longer necessarily less than r and the estimates on the vertex degree and the spectral radius no longer hold. On the other hand, the following claim gives an upper bound on the vertex degree that we encounter for given r.

Claim 3. *If C is the core graph of a subgroup of F_r, then all vertices are of degree at most $2r$.*

This is clear since each vertex can have at most one out-edge labeled by each of $\{x_1^{\pm 1}, \cdots, x_r^{\pm 1}\}$.

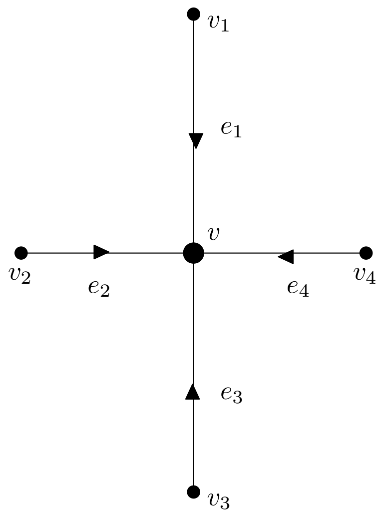
Counting all words labeling paths in C is clearly not going to give the required upper bound because the number of unreduced words grows like $(2r)^n$ if all vertices have maximum degree, which is possible. To count reduced words, we need the following refinement of the undirected graph C and of its adjacency matrix A.

Definition 9. *Let C be an undirected graph whose edges are inverse pairs $\{e, \overline{e}\}$. Define the directed graph C_d as follows.*

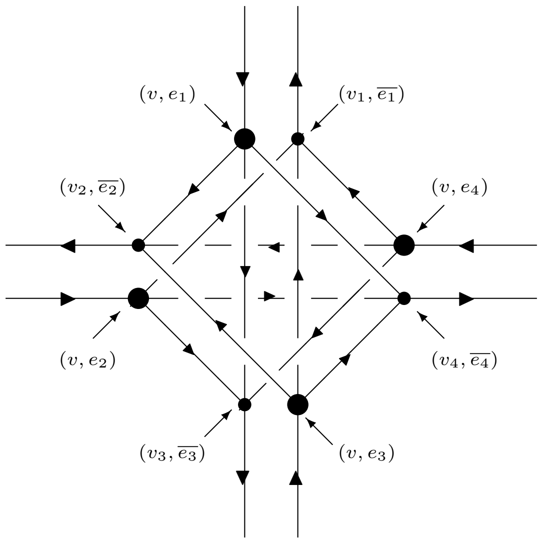
a) *The vertices of C_d are ordered pairs (v, e) where e is an edge whose tip is v,*
b) *There is a directed edge in C_d from (v_1, e_1) to (v_2, e_2) if e_2 is an edge of C from v_1 to v_2 and $e_2 \neq \overline{e_1}$.*

The vertices of C_d can be considered as the vertices of C with a 'memory' of which edge one arrived at v on. The out edges at vertices associated with v are then chosen to avoid the arrival edge. Then directed paths in C_d correspond to reduced paths in C. (If C has a vertex v_0 of degree 1, a reduced path beginning at v_0 does not have a directed analogue in C_d. This will cause us no problems.) Notice that the out degree of (v, e) in C_d is one less than the degree of v in C. If one thinks of C as a system of two-way roads, then C_d is the same system with each two-way road replaced by a divided highway with lanes in each direction and with interchanges that permit all transitions except for 180 dergee turns.

Example:



C



C_d

Definition 10. *The* adjacency matrix of a directed graph C_d *with* v *vertices is the matrix* A *of size* $(v \times v)$ *whose entry* A_{ij} *in the* (i, j) *position is the number of directed edges from the* i^{th} *vertex to the* j^{th} *vertex. The spectral radius of a directed graph is the spectral radius of its adjacency matrix.*

Proposition 4. *Suppose* C *is a graph whose minimum and maximum vertex degrees are* m *and* M *respectively.*

a) *For some constant* K*,*

$$\#\{\text{reduced } n\text{-paths in } C\} \leq K(M-1)^n.$$

b) *If* $m < M$*, then for some* $\widetilde{M} < M - 1$ *and constant* K*,*

$$\#\{\text{reduced } n\text{-paths in } C\} \leq K\widetilde{M}^n.$$

Proof: The graph C_d will have maximum out degree $M - 1$ and the associated matrix will have row sums between $m - 1$ and $M - 1$. The reasoning in the proof of Proposition 1 shows that the spectral radius of $\rho(C_d)$ is less than or equal to $M - 1$ and a) follows. If $m < M$, then $\rho(C_d) < M - 1$ and b) follows. □

Proof of Theorem 2: We proceed exactly as in §1, building viable sets one element at a time and bounding the number of non-viable choices at each stage in terms of paths in the cores. Recall that every non-viable w determines a pair of reduced paths (u, v) in C out of the base point. In §1 it sufficed to consider the possibly unreduced path $u^{-1} \cdot v$ and to bound the number of unreduced paths, clearly a very rough bound. In this section it is necessary to consider the pairs (u, v) and to count pairs of combined length at most n. We also need to deal separately with cores of finite index subgroups.

Let $\mathcal{W} = \{w_1, \cdots, w_k\}$ be a viable set in F_r. There are three possibilities for the core C of $\mathcal{W}$. Recall that since C is the core of a subgroup of F_r, the maximum vertex degree is $2r$.

Possible core graphs

a) The base point of C has degree at least 2 and there are vertices of degree less than $2r$.
b) The base point of C has degree one.
c) All vertices of C have degree $2r$. These cores correspond to subgroups of finite index.

Claim 4. *There is a number $\gamma_k < 2r - 1$ and a constant K so that for cores of type a) and b),*

$$\#\{\textit{reduced } n\textit{-paths in } C\} \le K\gamma_k^n.$$

Cores of type c) (where the statement is false) occur with asymptotic probability 0.

Theorem 2 follows from Claim 3 as follows. The number of non-viable choices w is bounded by the number of pairs (u, v) of reduced paths in C which is bounded by

$$\sum_{n_1+n_2 \le n} (K\gamma_k^{n_1})(K\gamma_k^{n_2}) \le K2\binom{n}{2}\gamma_k^n.$$

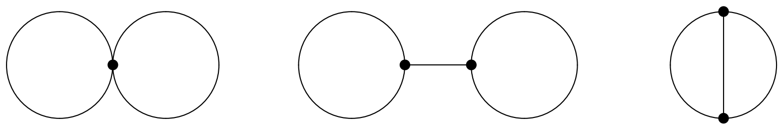
Then since

$$\lim_{n\to\infty} \binom{n}{2}\left(\frac{\gamma_k}{2r-1}\right)^n = 0,$$

the result follows as in §1. □

Proof of Claim 3.

Consider a core graph of type a). Since the rank k of C is less than s, there are only finitely many topological possibilities for C. For example, if $k = 2$ there are the three shown below.



Topological types of rank 2 cores

Let $\widehat{C}$ be the graph topologically equivalent to C with no degree two vertices. Since C has vertices of degree less than $2r$, the growth of reduced n-paths in $\widehat{C}$ is exponential in powers of some number $\gamma(\widehat{C}) < 2r - 1$. Any n-path in C is an initial piece which is a part of an edge of $\widehat{C}$, then a reduced n-path in $\widehat{C}$ followed by a terminal piece which is again part of an edge of $\widehat{C}$. (Note that whether a path is reduced does not depend on whether it is considered in $\widehat{C}$ or in C.) The number of edges of $\widehat{C}$ is at most $3s - 6$ (by a standard Euler characteristic argument) each with two ends and the initial and terminal pieces can each have length at most n. Therefore, for appropriate constant K,

$$\begin{aligned} \#\{\text{reduced } n\text{-paths in } C\} &\leq (3s-6)(2n)2\#\{\text{reduced } n\text{-paths in } \widehat{C}\} \\ &\leq K\gamma(\widehat{C})^n. \end{aligned}$$

Let γ_k be the largest $\gamma(\widehat{C})$ for core graphs C of rank k. The claim is then established for core graphs of type a).

If C is a core graph of type b), then let C' be the graph obtained by a removal of the tail to the base point—then C' is of type a). A path of length n is C is a path of length at most n in C' preceded and/or followed by paths on the tail to the base point. These tails can each be at most length n yielding the same estimate as in case a).

Finally consider cores of type c). Each such core has rank less than s and represents a subgroup of finite index v in F_{2r}, where v is the number of vertices of C. From the Schreier formula,

$$k - 1 = v(2r-1) \Longrightarrow v = \frac{k-1}{2r-1} \leq \frac{s-2}{2r-1}.$$

This universal bound on the number of vertices implies that there are finitely many such C, including labeling the edges with generators. It is easy to see that no matter what w is chosen, the path will be completely folded into C: there is nowhere else to go. In at least one case, the rose (the wedge of r circles labelled with the generators) which corresponds to the whole group, *every* such w is obviously fact non-viable. (Topological roses some of whose edges are labelled with words of length 2 or more, were considered in case a).)

In the case of the rose, $k = r$ and $\{w_1, \cdots, w_k\}$ is a basis for F_r. But the result of Burillo and Ventura [1] shows that this happens with asymptotic probability zero. Basically the same observation works for any other core of type c). The set $\{w_1, \cdots, w_k\}$ must be a basis for $\langle w_1, \cdots w_k\rangle$. Using the measure of length of words in $\langle w_1, \cdots w_k\rangle$ relative to any particular basis, the density of such bases goes to zero as n goes to infinity. But the metric of this length in $\langle w_1, \cdots w_k\rangle$ and the metric of length in F_r relative to the standard basis are related by a scale factor (are Lipschitz equivalent). The result follows.

□

References

[1] Burillo, J. and Ventura, E. *Counting primitive elements in free groups*, Geom. Dedicata, **93** (2002) 231–255.

[2] Cvetkovič, Dragoš, Doob, M. and Sachs, H. *Spectra of Graphs*, Series in Pure and Applied Math. **87**, Academic Press (1979).

[3] Dicks, W. and Ventura, E. *The group fixed by a family of injective endomorphisms*, Contemporary Mathematics **195**, American Math. Soc. (1996).

[4] Horn, R. A. and Johnson, C. R. *Matrix Analysis*, Cambridge University Press, Cambridge (1985).

[5] Imrich, W. and Turner, E. *Endomorphisms of free groups and their fixed points*, Math. Proc. Cam. Phil. Soc. **105**, (1989), 421–422.

[6] Kapovich, I. and Myasnikov, A. *Stallings foldings and subgroups of free groups*, Journal of Algebra **248** (2002), 608–668.

[7] Martino, A. and Ventura, E. *Example of retracts in free groups that are not the fixed subgroup of any automorphism*, J. Algebra **87**, (2003), 735–747.

[8] Stallings, J. *Topology of finite graphs*, Invent. Math. **71** (1983), 551–565.