Random operator approach for word enumeration in braid groups

1998 J. Phys. A: Math. Gen. 31 5609

(http://iopscience.iop.org/0305-4470/31/26/003)

J. Phys. A: Math. Gen. **31** (1998) 5609–5630. Printed in the UK PII: S0305-4470(98)88958-0

Random operator approach for word enumeration in braid groups

Alain Comtet† and Sergei Nechaev†‡

† Institut de Physique Nucléaire, Division de Physique Théorique§, 91406 Orsay Cedex, France
‡ L D Landau Institute for Theoretical Physics, 117940, Moscow, Russia

Received 3 November 1997

Abstract. We investigate analytically the problem of enumeration of nonequivalent primitive words in the locally free, $\mathcal{LF}_n$ and braid groups B_n for $n \gg 1$ by analysing the random word statistics and target spaces on these groups. We develop a 'symbolic dynamics' method for exact word enumeration in locally free groups and give arguments in support of the conjecture that the number of very long primitive words in the braid group is not sensitive to the precise local commutation relations. We touch briefly the connection of these problems with conventional random operator theory, localization phenomena and statistics of systems with quenched disorder. We also discuss the relation of particular problems of random operator theory to the theory of modular functions.

1. Introduction: problems and motivations

Recent years have been marked by the emergence of more and more problems related to the consideration of physical processes on noncommutative groups. In trying to classify such problems, we distinguish between the following categories in which the noncommutative origin of phenomena appear with perfect clarity.

(1) Problems connected with the spectral properties of the Harper–Hofstadter equation [1] dealing with the electron dynamics on the lattice in a constant magnetic field, groups of magnetic translations [2, 3] and properties of quantum planes [4].

(2) Problems of classical and quantum chaos on hyperbolic manifolds: spectral properties of dynamical systems and derivation of trace formulae [5–7] as well as construction of probability measures for random walks on modular groups [8].

(3) Problems giving rise to application of quantum group theory in physics: deformations of classical Abelian objects such as harmonic oscillators [9] and standard random walks [10].

(4) Problems of knot theory and statistical topology: construction of non-Abelian topological invariants [11, 12], consideration of probabilistic behaviour of the words on the simplest noncommutative groups related to topology (such as braid groups) [13], statistical properties of 'anyonic' systems [14].

(5) Classical problems of random matrix and random operator‖ theory and localization phenomena: determination of Lyapunov exponents for products of random noncommutative

§ Unité de Recherche des Universités Paris XI et Paris VI associée au CNRS.

‖ Following Pastur, we will distinguish between random matrices and matrix representations of random operators. To the random operators we associate an $n \times n$ table having of order n random entries; if the number of random entries grows faster than n when $n \to \infty$, we call such a table a random matrix.

0305-4470/98/265609+22$19.50

matrices [15–17], study of the spectral properties and calculation of the density of states of large random matrices [18, 19].

Certainly, such a division of problems into these categories is very speculative and reflects to a marked degree the authors' personal point of view. However, we believe that the enumerated items reflect, at least partially, the currently growing interest in theoretical physics of the ideas of noncommutative analysis.

Let us stress that we do not touch upon the pure mathematical aspects of noncommutative analysis in this paper and the problems discussed in this work mainly concern the points (4) and (5) of the list above.

A preliminary analytical and numerical study of the statistics of random walks (Markov chains) on braid and so-called 'locally free' groups† (see definition below) was recently undertaken in works [20, 21]. In the case of the braid group, the rather complicated group structure prevents us from applying the simple geometrical pictures of the free group Γ_2 (see [22]). Nevertheless the problem of the limit distribution for random walks on B_n can be reduced to the problem of a random walk on some graph [20, 21]. In the case of the group B_3 we were able to construct this graph explicitly, whereas for the group B_n ($n \geqslant 4$) we gave only an upper estimate for the limit distribution of random walks analysing statistics of Markov chains on 'locally free groups'.

The study of problems dealing with the limit distributions of Markov chains on braid group B_n requires examination of a 'target space' of this group, i.e. a space where the random walk takes place. The structure of the target space is uniquely determined by the group relations.

In this work we study the target space of the braid group B_n when $n \gg 1$, trying to develop a new 'statistical approach' for word enumeration in this group.

We should stress that our presentation offers a mathematical analysis which is far from being rigorous, and the ideas expressed here are mainly supported by numerical simulations. Moreover, we skip here some important but difficult questions, such as the problem of 'word identity' in the braid group (deep advances concerning this subject can be found in recent work [23]). Our aim is to describe a constructive algorithm which, of course, has to be be justified and verified later.

The structure of the paper is as follows. In the next section we give some necessary definitions concerning braid and 'locally free' groups and describe the model under consideration; section 3 is devoted to developing a 'symbolic dynamics' method for word enumeration in the locally free group $\mathcal{LF}_n$ (for $n \gg 1$). The target space of the braid group is studied in section 4 by means of a statistical approach based on the concept of a 'locally free group with errors'. In the discussion we express a geometrical point of view on the problem of word enumeration; while in the appendix we pay attention to additional links between the above mentioned problems and conventional random matrix theory, localization phenomena and statistics of systems with 'quenched' disorder. We also briefly discuss the possible relations between particular problems of random operator theory and the theory of modular functions.

2. Basic definitions and statistical model

First we recall some useful notation concerning the braid and 'locally free' groups.

† This notation has been introduced by Vershik in [20].

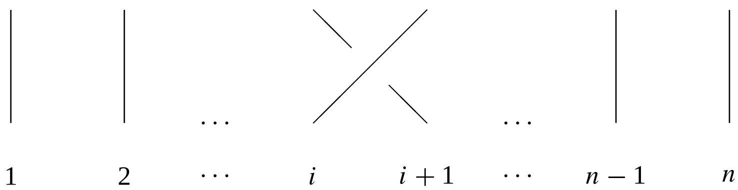
2.1. Braid group

The braid group B_n of 'n-strings' has $n-1$ generators $\{\sigma_1, \sigma_2, \dots, \sigma_{n-1}\}$ with the following commutation relations:

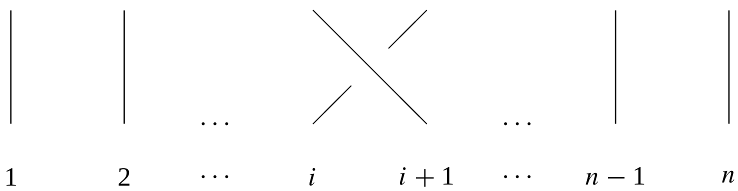
$$
\begin{aligned}
&\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \qquad (1 \leqslant i < n-1) \\
&\sigma_i\sigma_j = \sigma_j\sigma_i \qquad (|i-j| \geqslant 2) \\
&\sigma_i\sigma_i^{-1} = \sigma_i^{-1}\sigma_i = e.
\end{aligned}
\tag{2.1}
$$

Let us mention that:

—A word written in terms of 'letters'—generators from the set $\{\sigma_1, \dots, \sigma_{n-1}, \sigma_1^{-1}, \dots, \sigma_{n-1}^{-1}\}$ gives a particular *braid*. Schematically the generators σ_i and σ_i^{-1} may be represented as follows



$= \sigma_i$



$= \sigma_i^{-1}$

—The *length* of the braid is the total number of letters used, while the *minimal irreducible length* hereafter referred to as 'primitive word' is the shortest noncontractible length of a particular braid remaining after all possible group relations (2.1) are applied. Diagrammatically, the braid can be represented as a set of crossed strings going from the top to the bottom after gluing the braid generators.

—The closed braid appears after gluing the 'upper' and the 'lower' free ends of the braid on the cylinder.

—Any braid corresponds to some knot or link; hence the possibility to use the braid group representation for the construction of topological invariants of knots and links. Note, however, that the correspondence between braids and knots is not mutually single valued and each knot or link can be represented by an infinite series of different braids.

2.2. Locally free group

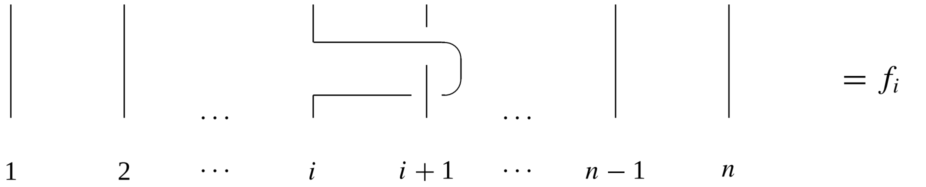
The group $\mathcal{LF}_n(d)$ is called *locally free* if the generators, $\{f_1, \dots, f_{n-1}\}$ obey the following commutation relations:

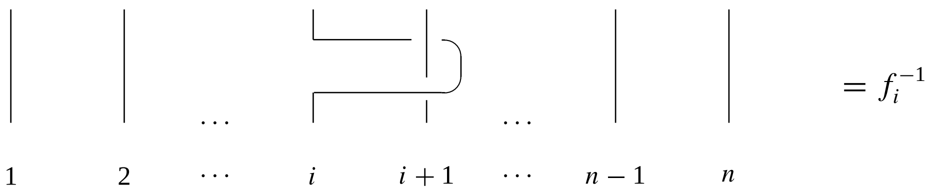
(a) each pair (f_j, f_k) generates the free subgroup of the group $\mathcal{LF}_n(d)$ if $|j-k| < d$;

(b) $f_j f_k = f_k f_j$ for $|j-k| \geqslant d$.

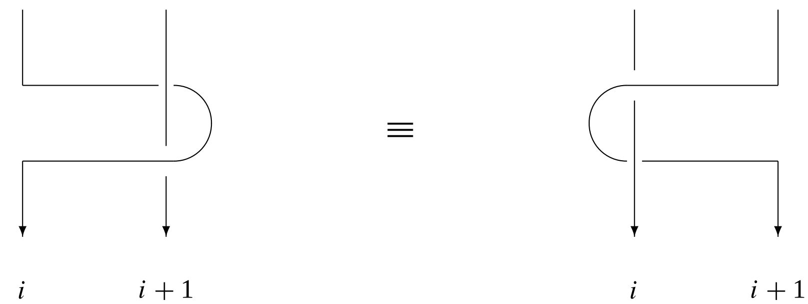
We will be concerned mostly with the case $d = 2$ for which we define $\mathcal{LF}_n(2) \equiv \mathcal{LF}_n$.

—The *length* of the word written in terms of letters $\{f_1, \dots, f_{n-1}, f_1^{-1}, \dots, f_{n-1}^{-1}\}$ is the total number of generators used, and the 'primitive word' is the shortest noncontractible length of a particular word after applying all relations of the group $\mathcal{LF}_n(2)$ (cf the case of the braid group). The graphical representation of generators g_i and g_i^{-1} is also rather similar to that of the braid group:





It is easy to see that the following geometrical identity is valid:



hence, it is unnecessary to distinguish between 'left' and 'right' operators f_i.

It can be seen that the only difference between the braid and locally free groups consists of the elimination of the Yang–Baxter relations (first line in equation (2.1)).

2.3. Statistical model

Our aim is to calculate a specific 'partition function', $V_n(\mu, d = 2)$, giving the number of all nonequivalent primitive words of length μ in the groups $\mathcal{LF}_{n+1}(d = 2)$ and B_{n+1} for $n \gg 1$.

Remark. To have a geometrical picture of the group $\mathcal{LF}_{n+1}$ let us describe the recursion procedure of raising the graph (the 'target space') associated with this group.

Take the *free* group Γ_n with generators $\{\tilde{f}_1, \dots, \tilde{f}_n\}$ where all $\tilde{f}_i$ $(1 \leqslant i \leqslant n)$ do not commute. It is well known that the group Γ_n has the structure of a $2n$-branching Cayley tree, $C(\Gamma_n)$ (see figure 1(a)) where the number of distinct primitive words of length μ is equal to the number $\tilde{V}_n(\mu)$ of vertices of the tree $C(\Gamma_n)$ lying at a distance of μ steps from the origin:

$$\tilde{V}_n(\mu) = 2n(2n-1)^{\mu-1}. \tag{2.2}$$

The graph $C(\mathcal{LF}_{n+1})$ corresponding to the group $\mathcal{LF}_{n+1}$ can be constructed from the graph $C(\Gamma_n)$ in accordance with the following recursion procedure.

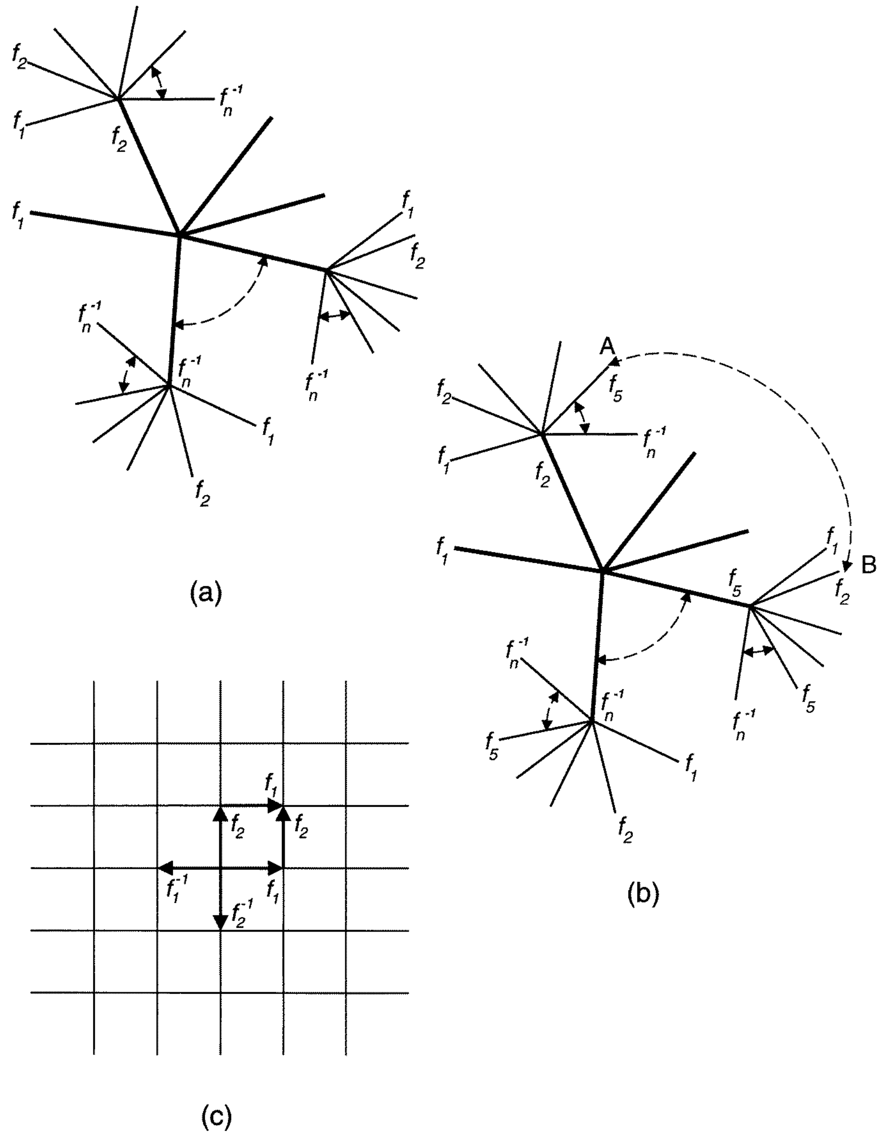


Figure 1. Graphs, corresponding to: (*a*) free group Γ_n; (*b*) locally free group $\mathcal{LF}_{n+1}$; (*c*) complete commutative group. In the case of locally free group the vertices *A* and *B* should be glued because they represent the same word in the group $\mathcal{LF}_{n+1}$.

(a) Take the root vertex of the graph $C(\Gamma_n)$ and consider all vertices at a distance $\mu = 2$ from it. Identify those vertices which correspond to the equivalent words in the group $\mathcal{LF}_{n+1}$ (see example in figure 1(*b*)).

(b) Repeat this procedure taking all vertices at the distance $\mu = (1, 2, \ldots)$ and 'gluing' them at the distance $\mu + 2$ according to the definition of the locally free group.

By means of the procedure described, we raise a graph ('target space') corresponding to the locally free group $\mathcal{LF}_{n+1}$. Now our main problem can be reformulated as follows: how many *distinct* vertices has the graph $C(\mathcal{LF}_{n+1})$ at a distance of μ steps from the origin (for $n \gg 1$)?

In the next section we give an exact answer to that question, which we use, in turn, as the basis for consideration of the much trickier case of the braid group.

It is worthwhile mentioning that the graph $C(\mathcal{E}_{n+1})$ of the complete commutative group $\mathcal{E}_{n+1}$ (all generators of $\mathcal{E}_{n+1}$ commute with each other) is the lattice embedded in $\mathbb{R}^n$ (see figure 1(*c*)). Hence, the number of nonequivalent words $V_n^{\text{comm}}(\mu)$ of length μ can be roughly estimated as the number of lattice points lying on the surface of the n-dimensional sphere, i.e.

$$V_n^{\text{comm}}(\mu) \simeq \text{constant } \mu^{n-1}. \tag{2.3}$$

Comparing (2.2) and (2.3) we obtain

$$\begin{aligned} &\lim_{\mu\to\infty} \frac{1}{\mu} \ln \tilde{V}_n(\mu) = \ln(2n-1) > 0 \\ &\lim_{\mu\to\infty} \frac{1}{\mu} \ln V_n^{\text{comm}}(\mu) = 0. \end{aligned} \tag{2.4}$$

Naively we could expect that the behaviour $\lim_{\mu\to\infty} \frac{1}{\mu} \ln V_n(\mu) = 0$ for 'locally free' and braid groups remains unchanged (i.e. the same as in the completely commutative case) because for $n \gg 1$ we have of order $\sim n^2$ commutative relations and only of order $\sim n$ noncommutative ones. However, it is proven for $\mathcal{LF}_n$ and conjectured for B_n that

$$\lim_{\substack{\mu\to\infty \\ n=\text{constant}\gg 1}} \frac{1}{\mu} \ln V_n(\mu) = \text{constant} > 0$$

which clearly reflects the hyperbolic character of these groups.

3. Exact words enumeration in locally free groups

We derive here an explicit expression for the number $V_n(\mu, d)$ of all nonequivalent primitive words of length μ in the group $\mathcal{LF}_{n+1}(d)$ (when $d = 2$ and $n \gg 1$) on the basis of the so-called 'normal order' representation of words proposed by Vershik in [24] and developed in [20, 21] which is reminiscent of the enumeration of 'partially commutative monoids' arising in combinatorics [25].

3.1. 'Normal order' representation of words

Let us represent each primitive word W_p of length μ in the group $\mathcal{LF}_{n+1}(d)$ in the *normal order* similar to the so-called 'symbolic dynamics' appearing in the context of chaotic systems (see, for instance, [7])

$$W_p = (f_{\alpha_1})^{m_1} (f_{\alpha_2})^{m_2} \dots (f_{\alpha_s})^{m_s} \tag{3.1}$$

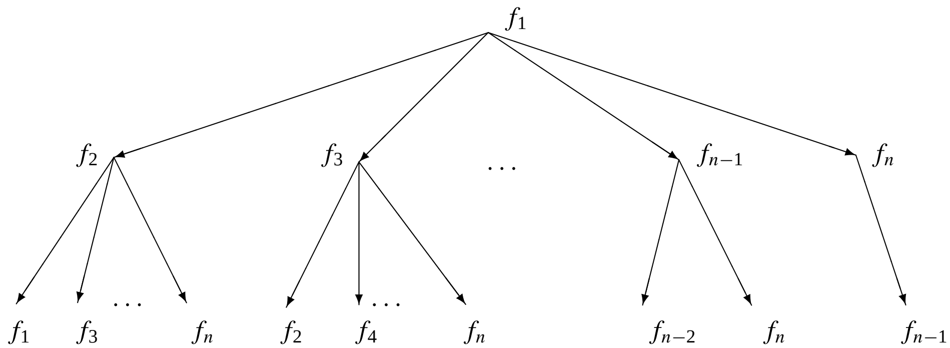
where $\sum_{i=1}^{s} |m_i| = \mu$ ($m_i \neq 0 \forall i$; $1 \leqslant s \leqslant \mu$) and the sequence of generators f_{α_i} in equation (3.1) *for all distinct* f_{α_i} satisfies the following local rules [20]:

(i) if $f_{\alpha_i} = f_1$, then $f_{\alpha_{i+1}} \in \{f_2, f_3, \dots f_n\}$;

(ii) if $f_{\alpha_i} = f_k$ $(1 < k \leqslant n-1)$, then $f_{\alpha_{i+1}} \in \{f_{k-d+1}, \dots, f_{k-1}, f_{k+1}, \dots f_n\}$;

(iii) if $f_{\alpha_i} = f_n$, then $f_{\alpha_{i+1}} \in \{f_{k-d+1}, \dots, f_{n-1}\}$.

These local rules may be represented diagrammatically as follows.



The rules (i)–(iii) give the prescription to encode and enumerate all distinct primitive words in the group $\mathcal{LF}_{n+1}(d)$. If the sequence of generators in the primitive word W_p does not satisfy the rules (i)–(iii), we commute the generators in the word W_p until the normal order is restored. Hence, the normal order representation enables one to give the unique coding of all nonequivalent primitive words in the group $\mathcal{LF}_{n+1}(d)$.

Example 1. Take an arbitrary primitive word of length $\mu = 10$ in the group $\mathcal{LF}_{8+1}(2)$:

$$\begin{aligned} W_p &= f_5^{-1} f_3 f_8 f_1^{-1} f_2 f_4 f_8 f_8 f_4 f_7 \\ &\equiv (f_5)^{-1}(f_3)(f_8)(f_1)^{-1}(f_2)(f_4)(f_8)^2(f_4)(f_7). \end{aligned} \tag{3.2}$$

To represent the word W_p in the 'normal order' we have to push all generators with smaller indices to the left when it is allowed by the commutation relations of the locally free group $\mathcal{LF}_9(2)$. We obtain

$$W_p = (f_1)^{-1}(f_3)^1(f_2)^1(f_5)^{-1}(f_4)^2(f_8)^3(f_7)^1 \tag{3.3}$$

(the 'normal order' for this word is the sequence of used generators: $\{1, 3, 2, 5, 4, 8, 7\}$).

To compute the number of different primitive words of length $\mu = 10$ with the same normal order as in equation (3.3), we have to sum up all the words like

$$W_p = (f_1)^{m_1}(f_3)^{m_2}(f_2)^{m_3}(f_5)^{m_4}(f_4)^{m_5}(f_8)^{m_6}(f_7)^{m_7} \tag{3.4}$$

under the condition $\sum_{i=1}^{7} |m_i| = 10$; $m_i \neq 0 \forall m_i \in [1, 7]$.

The calculation of the number of distinct primitive words, $V_n(\mu)$, of a given length μ is now rather straightforward:

$$V_n(\mu, d) = \sum_{s=1}^{\mu} R_n(s, d) \sum_{\{m_1,\ldots,m_s\}}{}' \Delta\left[\sum_{i=1}^{s} |m_i| - \mu\right] \tag{3.5}$$

where:

- $R_n(s, d)$ is the number of all distinct sequences of s generators taken from the set $\{f_1, \ldots, f_n\}$ and satisfying the local rules (i)–(iii);
- the second sum gives the number of all possible representations of the primitive path of length μ *for the fixed sequence of generators* (see the example above); 'prime' means that the sum runs over all $m_i \neq 0$ for $1 \leqslant i \leqslant s$; $\Delta(x)$ is the Kronecker function: $\Delta(x) = 1$ for $x = 0$ and $\Delta(x) = 0$ for $x \neq 0$;
- special attention should be paid to the sequences built on the basis of one generator only, i.e. for primitive words of type $W_p = (f_k)^\mu \forall k \in [1, n]$ (see definition of $R_n(s, d)$ below).

To obtain the partition function $R_n(s, d)$ let us mention that the local rules (i)–(iii) define a generalized Markov chain with the states given by the $n \times n$ 'incidence' matrix $\hat{M}_n(d)$, the rows and columns of which correspond to the generators $f_1, \ldots, f_n$ (the case $d = 2$ is shown below)

$$\hat{M}_n(2) = \begin{array}{|c||c|c|c|c|c|c|c|} \hline & f_1 & f_2 & f_3 & f_4 & \ldots & f_{n-1} & f_n \\ \hline\hline f_1 & 0 & 1 & 1 & 1 & \ldots & 1 & 1 \\ \hline f_2 & 1 & 0 & 1 & 1 & \ldots & 1 & 1 \\ \hline f_3 & 0 & 1 & 0 & 1 & \ldots & 1 & 1 \\ \hline f_4 & 0 & 0 & 1 & 0 & \ldots & 1 & 1 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \hline f_{n-1} & 0 & 0 & 0 & 0 & \ddots & 0 & 1 \\ \hline f_n & 0 & 0 & 0 & 0 & & 1 & 0 \\ \hline \end{array}. \tag{3.6}$$

The matrix $\hat{M}_n(d)$ has a rather simple structure: above the diagonal we put everywhere '1' and below diagonal we have $d - 1$ subdiagonals completely filled by '1'; in all other places we have '0'.

The number of all distinct normally ordered *sequences of words* of length s with allowed commutation relations is given by the following partition function

$$R_n(s, d) = \tilde{\boldsymbol{v}}_{\text{in}}[\hat{M}_n(d)]^{s-1}\boldsymbol{v}_{\text{out}} \tag{3.7}$$

where

$$\tilde{\boldsymbol{v}}_{\text{in}} = (\overbrace{1\ 1\ \ldots\ 1}^{n}) \qquad \text{and} \qquad \boldsymbol{v}_{\text{out}} = \left.\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\right\} n. \tag{3.8}$$

For $s = 1$ we have $R_n(1, d) = \tilde{\boldsymbol{v}}_{\text{in}}\boldsymbol{v}_{\text{out}} = n$ as it should be.

The remaining sum in equation (3.5) is independent of $R_n(s, d)$, so its calculation is very simple:

$$\sum_{\{m_1,\ldots,m_s\}}{}' \Delta\left[\sum_{i=1}^{s} |m_i| - \mu\right] = 2^s C_{\mu-1}^{s-1} \tag{3.9}$$

where $C_{\mu-1}^{s-1}$ is the binomial coefficient: $C_{\mu-1}^{s-1} = \frac{(\mu-1)!}{(s-1)!(\mu-s)!}$.

Substituting equations (3.9) and (3.7) into equation (3.5) we find

$$\begin{aligned} V_n(\mu) \equiv V_n(\mu, d) &= \sum_{s=1}^{\mu} 2^s C_{\mu-1}^{s-1} R_n(s, d) \\ &= 2\tilde{\boldsymbol{v}}_{\text{in}}(2\hat{M}_n(d) + \hat{I})^{\mu-1}\boldsymbol{v}_{\text{out}} \end{aligned} \tag{3.10}$$

where $\hat{I}$ is the identity matrix.

Such a quantity is rather difficult to evaluate exactly. A reasonable approximation is to replace (3.10) by

$$V_n^*(\mu) = 2\sum_{i=1}^{n} (2\lambda_i + 1)^{\mu-1} \tag{3.11}$$

where λ_i are the eigenvalues of the matrix $\hat{M}_n$ which can be shown *all to be real* (see the later discussion). In order to check the validity of approximation (3.11) we have considered the case $\{n = 3, d = 2\}$ where the exact value reads

$$V_3(\mu) = \left(\frac{15 + 7\sqrt{5}}{5}\right)(2 - \sqrt{5})^{\mu-1} + \left(\frac{15 - 7\sqrt{5}}{5}\right)(2 + \sqrt{5})^{\mu-1}$$

whereas the approximation (3.11) gives

$$V_3^*(\mu) = 2(2 - \sqrt{5})^{\mu-1} + 2(2 + \sqrt{5})^{\mu-1} + 2(-1)^{\mu-1}.$$

It can be seen that this approximation works reasonably well even for small values of μ.

The value $V_n(\mu, d)$ grows exponentially fast with μ and the 'rate' of this growth is clearly represented by the fraction

$$q(d) = \left.\frac{V_n(\mu + 1, d)}{V_n(\mu, d)}\right|_{\mu\gg 1} \tag{3.12}$$

which is the effective coordination number of the graph $C(\mathcal{LF}_n)$.

In the next section we present calculations of the asymptotic expression of (3.11) when $n \gg 1$.

3.2. Calculation of eigenvalues of matrix $\hat{M}_n(2)$

Consider the determinant

$$a_n(\lambda) = \det(\hat{M}_n - \lambda\hat{I}) = \begin{pmatrix} -\lambda & 1 & 1 & \dots \\ 1 & -\lambda & 1 & \dots \\ 0 & 1 & -\lambda & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.13}$$

It satisfies the recursion relation

$$a_n(\lambda) = -(\lambda + 1)a_{n-1}(\lambda) - (\lambda + 1)a_{n-2}(\lambda) \tag{3.14}$$

with the boundary conditions

$$\begin{aligned} a_0(\lambda) &= 1 \\ a_1(\lambda) &= -\lambda. \end{aligned} \tag{3.15}$$

For $\lambda > -1$ one may set

$$a_n(\lambda) = (\lambda + 1)^{\frac{n-1}{2}}(-1)^n\varphi_n(\lambda) \tag{3.16}$$

which gives

$$\varphi_n(\lambda) = \sqrt{\lambda + 1}\varphi_{n-1}(\lambda) - \varphi_{n-2}(\lambda). \tag{3.17}$$

The general solution of (3.17) satisfying the previously defined boundary conditions (3.15) is given in terms of Chebyshev's polynomials of the second kind

$$\varphi_n(\lambda) = \mathcal{U}_{n+1}(\cos\vartheta) \tag{3.18}$$

where

$$\cos\vartheta = \frac{\sqrt{\lambda + 1}}{2} \qquad \left(0 < \vartheta < \frac{\pi}{2}\right). \tag{3.19}$$

Therefore

$$a_n(\lambda) = (-1)^n(\lambda+1)^{\frac{n-1}{2}}\mathcal{U}_{n+1}(\cos\vartheta)$$
$$= (-1)^n(\lambda+1)^{\frac{n-1}{2}}\frac{\sin(n+2)\vartheta}{\sin\vartheta}. \tag{3.20}$$

The last expression enables us to obtain all the eigenvalues of the matrix $\hat{M}_n$. In fact, it is convenient to distinguish them according to the parity of n:

(1) $n = 2m+1$

$$\lambda_0 = (-1)\ (m \text{ such values}) \qquad \lambda_k = 4\cos^2\frac{k\pi}{2m+3} - 1\ (k = [1, m+1]) \tag{3.21}$$

(2) $n = 2m$

$$\lambda_0 = (-1)\ (m \text{ such values}) \qquad \lambda_k = 4\cos^2\frac{k\pi}{2m+2} - 1\ (k = [1, m]). \tag{3.22}$$

Since in each case we have exactly n states, this exhausts the complete set of eigenvalues, showing they *are all real* in the interval $[-1, 3]$. One also recovers the result obtained earlier in [20, 21] for the asymptotes of the highest eigenvalue of matrix $\hat{M}_n$ (in the limit $n \gg 1$):

$$\lambda_{\max} = 4\cos^2\frac{\pi}{n+2} - 1\Big|_{n\gg 1} \approx 3 - \frac{4\pi^2}{n^2} \qquad (k=1). \tag{3.23}$$

Now we are in position to compute the number of nonequivalent words $V_n^*(\mu)$ of primitive length μ in the locally free group $\mathcal{LF}_n(2)$ for $n \gg 1$, n = constant (see equation (3.11)). Using the definition (3.11) and equations (3.21), (3.22) we find for $n = 2m+1$

$$V_n^*(\mu) = (n-1)(-1)^{\mu-1} + 2\sum_{k=1}^{\frac{n+1}{2}}\left(8\cos^2\frac{k\pi}{n+2} - 1\right)^{\mu-1}. \tag{3.24}$$

We define $\varphi_n(\mu)$ as follows

$$V_n^*(\mu) = (n-1)(-1)^{\mu-1} + 2\varphi_n(\mu).$$

One can prove the following identity for $\mu < n$:

$$\varphi_n(\mu) = \frac{1}{2}\sum_{k=1}^{n+1}\left(8\cos^2\frac{k\pi}{n+2} - 1\right)^{\mu-1}$$
$$= \frac{(n+2)}{2\pi}\int_0^{\pi}(8\cos^2 x - 1)^{\mu-1}\,\mathrm{d}x - \frac{1}{2}7^{\mu-1}. \tag{3.25}$$

For $\mu \gg 1$, n = constant $\gg 1$ the last integral is evaluated by a saddle-point approximation which yields

$$\varphi_n(\mu) = \frac{(n+2)}{4\pi}\sqrt{\frac{\pi}{2\mu}}7^{\mu-1} - 7^{\mu-1}. \tag{3.26}$$

Thus, for the number of nonequivalent words in the locally free group $\mathcal{LF}_n(2)$ we have the following limiting behaviour:

$$\lim_{\substack{\mu\to\infty \\ n=\text{constant}\gg 1}} \frac{1}{\mu}\ln V_n^*(\mu) = \ln 7 \tag{3.27}$$

and equation (3.12) gives

$$q(d=2) = 7.$$

Hence, the effective graph corresponding to the locally free group can be viewed as a tree with the branching number $q = 7$.

4. Approximate statistical approach to word enumeration in the braid group B_n

The construction of an effective algorithm for the enumeration of the words in the braid group B_n for $n > 2$ is one of the most intriguing problems in group theory.

In this section we propose an *approximate statistical approach* for the enumeration of all distinct primitive words in the group B_n for $n \gg 1$ which exploits some properties of locally free groups $\mathcal{LF}_n$ considered above.

The main idea is as follows. Let us deal with the sequences of words in the braid group B_n from the point of view of the locally free group $\mathcal{LF}_n^{\text{err}}$ 'with errors'. To be more specific let us start with the following example.

Example 2. Write a random word W in the group B_7. Let this word be, for instance,

$$W = (\sigma_1)^{-1}\sigma_4(\sigma_5)^{-1}(\sigma_6)^{-1}\sigma_5\sigma_1\sigma_6(\sigma_2)^{-1}.$$

We reduce this word to the primitive one in two steps.

(1) On the first step we act in the same way as in the case of the locally free group $\mathcal{LF}_7$ and push all generators with smaller indices to the left *assuming that nearest neighbours do not commute at all*. We find,

$$W_{\text{reduced}} = (\sigma_2)^{-1}\sigma_4(\sigma_5)^{-1}\underbrace{(\sigma_6)^{-1}\sigma_5\sigma_6}_{\sigma_5\sigma_6(\sigma_5)^{-1}}.$$

(2) Now we can apply the Yang–Baxter relations to the triple $(\sigma_6)^{-1}\sigma_5\sigma_6$ and obtain after the cancellation of $(\sigma_5)^{-1}$ and σ_5 the primitive word

$$W_p = (\sigma_2)^{-1}\sigma_4\sigma_6(\sigma_5)^{-1}.$$

The first step of the braid contracting procedure exactly coincides with what we did for the locally free group, while the second step we could regard (approximately, of course) as follows.

Consider some pair, for instance, $(\sigma_6)^{-1}\sigma_5$*. We commute it with a probability* p*. Such commutation we denote as an* error*. The probability to meet the letter* σ_i *in the Markov chain with uniform distribution over the generators in the braid group* B_n *is of order of* $p = \frac{1}{2n}$*. Later on we consider the more general case taking* p *as a variational parameter*.

Let $V_n^{\text{braid}}(\mu)$ be the number of all primitive words of length μ in the braid group B_n. Our main idea is as follows: we would like to relate the quantity $V_n^{\text{braid}}(\mu)$ to the number of primitive words in the 'group' $\mathcal{LF}_n^{\text{err}}(2)$ averaged over the uniform distribution of 'errors' in commutation relations.

4.1. *Statistics of words with 'errors' in locally free groups*

The techniques for the analysis of disordered systems are rather well developed, especially those used in the *spin-glass models* [26].

Central to these methods is the concept of *self-averaging* which can be explained as follows. Take some additive function F (the free energy, for instance) of some disordered system. The function F is a self-averaging quantity if the observed value, F_{obs}, of any macroscopic sample of the system coincides with the value F_{av} averaged over the ensemble of disorder realizations:

$$F_{\text{obs}} = \langle F\rangle_{\text{av}}.$$

The phenomenon of self-averaging takes place in systems with sufficiently weak long-range correlations: only in this case can F be considered as a sum of contributions from different volume domains, containing statistically independent realizations of disorder (for more details see [18]).

The main technical problem of systems with *quenched* disorder is the calculation of the free energy $F(\mu)$ averaged over the randomly distributed quenched pattern. In our case we could associate the number of topologically different words with the partition function, hence the free energy would be $F(\mu) = -\langle \ln V_n^{\rm err}(\mu)\rangle$ and the 'quenched pattern' is just the set of 'errors' in commutation relations.

A problem closely related to that mentioned above arises when averaging the correlation functions of some statistical system over disorder. In this case we should find the averaged density of states of some random matrix with a prescribed distribution of random entries. Below we show that the calculation of the mean value $\langle V_n^{\rm err}(\mu)\rangle$ belongs precisely to this class of problems.

Before formulating our main conjecture (which is verified hereafter in a self-consistent way) let us regard the limiting cases $p = 1$ and $p = 0$.

(i) For $p = 1$ we have a complete commutative group and the obvious inequality is fulfilled

$$V_n^{\rm braid}(\mu) \geqslant V_n(\mu, d = 2, p = 1). \tag{4.1}$$

(ii) For $p = 0$ we return to the locally free group $\mathcal{LF}_2$, for which we should have

$$V_n^{\rm braid}(\mu) \leqslant V_n(\mu, d = 2, p = 0) \tag{4.2}$$

where $V_n(\mu, d = 2, p)$ is the number of all distinct primitive words of length μ with the '*errors*' in the commutation relations in the locally free group $\mathcal{LF}_n(2)$ and we allow to commute the neighbouring generators to commute with the probability p.

Conjecture. The number of nonequivalent primitive words, $V_n^{\rm braid}(\mu)$, of length μ in the braid group B_n can be estimated in the limits $n = \text{constant} \gg 1$, $\mu \gg 1$ as follows

$$V_n^{\rm braid}(\mu) \approx \langle V_n(\mu, d = 2, p)\rangle \tag{4.3}$$

while the averaging is performed over the uniform probability distribution of 'errors'.

The question concerning the choice of p is considered below.

In support of our conjecture we invoke the numerical computations performed in the work [21], where we have constructed the (right-hand) random walk (the random word) on the group $\mathcal{G}_n = \{\mathcal{LF}_n^{\rm err}, B_n\}$ with a uniform distribution over generators $\{g_1, \ldots, g_{n-1}, g_1^{-1}, \ldots, g_{n-1}^{-1}\} \in \mathcal{G}_n$. It means that with a probability $\frac{1}{2n-2}$ we have added the element g_{α_N} or $g_{\alpha_N}^{-1}$ to the given word of $N-1$ generators (letters) from the right-hand side. In [21] the following question has been raised: what is the averaged length $\langle\mu\rangle$ of the primitive path for the N-step random walk on the group $\mathcal{G}_n$?

In table 1 we show the results of numerical simulations carried on in [21] for the expectation value $\langle\mu\rangle/N$ of the N-step random walk on the 'locally free structure with errors', $\mathcal{LF}_n^{\rm err}(2)$, and compare them with the same value for the N-step random walk on the braid group B_n.

We find asymptotically a very good correspondence between the mean values $\langle\mu\rangle/N$ for the braid group and the 'locally free structure with errors' for $p = \frac{1}{5}$.

Our conjecture has a 'mean-field' nature because we allow two consecutive generators $(\sigma_{\alpha_k}, \sigma_{\alpha_{k+1}})$ with nearest neighbour indices $\alpha_{k+1} = \alpha_k \pm 1$ to commute with probability p regardless of the value of the generator $\sigma_{\alpha_{k+2}}$ in the sequence of letters in the word. Hence the problem of choosing the correct value of p appears.

Table 1.

	$\mathcal{LF}_n^{\text{err}}$ with $p = \frac{1}{5}$	B_n
n	$\langle\mu\rangle/N$	$\langle\mu\rangle/N$
5	0.55	0.49
10	0.58	0.56
20	0.59	0.59
50	0.60	0.61
100	0.60	0.61
200	0.61	0.61

In the frameworks of the mean-field approximation and taking into account equation (4.3) we claim:

$$\langle V_n(\mu, d = 2, p)\rangle|_{p\to 1^-} \leqslant V_n^{\text{braid}}(\mu) \leqslant V_n(\mu, d = 2, p = 0). \tag{4.4}$$

We prove below that in the limit $n \to \infty$, $\mu \to \infty$ the following equality is satisfied

$$\langle V_n(\mu, d = 2, p)\rangle|_{p\to 1^-} = V_n(\mu, d = 2, p = 0). \tag{4.5}$$

Thus, for the value $V_n^{\text{braid}}(\mu)$ we have the following asymptotic behaviour

$$\lim_{\substack{\mu\to\infty \\ n\to\infty}} \frac{1}{\mu} \ln V_n^{\text{braid}}(\mu) = \lim_{\substack{\mu\to\infty \\ n\to\infty}} \left[\frac{1}{\mu} \ln\langle V_n(\mu, d = 2, p)\rangle\right]_{\substack{\text{independent} \\ \text{on } p}} = \ln 7. \tag{4.6}$$

It is easy to understand that the number of nonequivalent primitive words $V_n(\mu, d = 2, p)$ in the 'locally free group with errors' can be calculated by means of averaging equation (3.10) if we slightly change the matrix $\hat{M}_n$ replacing it by the random incidence matrix $\hat{M}_n$:

$$\hat{M}_n^{\text{err}}(d = 2) = \begin{array}{|c||c|c|c|c|c|c|c|}
\hline
 & f_1 & f_2 & f_3 & f_4 & \ldots & f_{n-1} & f_n \\
\hline\hline
f_1 & 0 & 1 & 1 & 1 & \ldots & 1 & 1 \\
\hline
f_2 & x_{n-1} & 0 & 1 & 1 & \ldots & 1 & 1 \\
\hline
f_3 & 0 & x_{n-2} & 0 & 1 & \ldots & 1 & 1 \\
\hline
f_4 & 0 & 0 & x_{n-3} & 0 & \ldots & 1 & 1 \\
\hline
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
\hline
f_{n-1} & 0 & 0 & 0 & 0 & \ddots & 0 & 1 \\
\hline
f_n & 0 & 0 & 0 & 0 & & x_1 & 0 \\
\hline
\end{array}. \tag{4.7}$$

It is now parametrized by the random sequence of '0' or '1', i.e.

$$\{x^{(n)}\} = \{x_{n-1}, x_{n-2}, \ldots, x_2, x_1\} \tag{4.8}$$

where

$$\begin{aligned} &\text{Prob}(x_j = 1) = 1 - p \\ &\text{Prob}(x_j = 0) = p. \end{aligned} \tag{4.9}$$

4.2. Density of states of random operator and averaged number of nonequivalent words

The determinant of the matrix $\hat{M}_k^{\text{err}}(d = 2) - \lambda \hat{I}$ (see equation (4.7)) satisfies the random recursion relation

$$a_{k+1} + a_k(\lambda + x_k) + a_{k-1}(1 + \lambda)x_k = 0. \tag{4.10}$$

Introducing the Ricatti-like variable

$$\rho_k = \frac{a_{k+1}}{a_k}$$

we arrive at the recursion relation

$$\rho_k = -(\lambda + x_k) - \frac{(1 + \lambda)x_k}{\rho_{k-1}} \qquad k \in [0, n] \tag{4.11}$$

with the boundary condition $\rho_0 = -\lambda$.

For the continuous sequence of $\{1\}$, i.e. $\{x^{(n)}\} = \{1\ 1\ 1\ 1\ \ldots\ 1\}$ we have the following (nonrandom) transformation

$$\rho_{k+1} = -(\lambda + 1)\left(1 + \frac{1}{\rho_k}\right). \tag{4.12}$$

As soon as a zero appears in the random sequence, ρ_k in (4.11) is set to $-\lambda$ which coincides precisely with the initial value ρ_0. Since one returns to the initial value, the process can be easily iterated for arbitrary random sequences $\{x^{(n)}\}$ when $n \gg 1$. Such a property of the map (4.11) is equivalent to the factorization of the determinant $a_n(\lambda)$ of the random matrix $\hat{M}_n^{\text{err}}(d = 2) - \lambda \hat{I}$. Consider the sequence $\{x^{(n)}\} = (\underbrace{1\ 1\ \ldots\ 1}_{l_1}\underbrace{0\ 0\ \ldots\ 0}_{m_1}\ \underbrace{1\ \ldots\ 1}_{l_2})$, the corresponding determinant $a_n(\lambda)$ factorizes: $a_n(\lambda) = \prod_{\{l_j\}} a_{l_j}(\lambda) \prod_{\{m_j\}} (-\lambda)^{m_j - 1}$.

It is worth pointing out that a recursion relation similar to (4.12) also appears in the study of the binary product of random 2×2 matrices where one of the matrices is singular [30]. Such a structure also occurs in the case of the Ising model in a random magnetic field, see [31]. Following Derrida and Hilhorst [31], one can write down the invariant measure associated with (4.11):

$$\mathcal{P}(\rho) = \sum_{k=1}^{\infty} p(1 - p)^k \delta(\rho - \rho_k) + (1 - p) \sum_{k=1}^{\infty} p^k \delta(\rho - \rho_0). \tag{4.13}$$

The first (second) term comes from the complete sequences of $\{1\}$ ($\{0\}$) of arbitrary length $k = 1, 2, \ldots$. From the invariant measure one can compute $\langle \ln a(\lambda) \rangle$ which can be interpreted either as a Lyapunov exponent or as the free energy (depending on the physical context). Hence we find:

$$\langle \ln a(\lambda) \rangle = \int \mathcal{P}(\rho) \ln \rho \, \mathrm{d}\rho = \sum_{k=1}^{\infty} p^2 (1 - p)^{k-1} \ln a_k(\lambda). \tag{4.14}$$

Returning to our problem, we may use this expression to write down the averaged density of states

$$\langle \rho(\lambda) \rangle = \frac{1}{\pi} \frac{\partial}{\partial \lambda} \mathrm{Im}\, \langle \ln a(\lambda) \rangle \equiv \sum_{k=1}^{\infty} p^2 (1 - p)^{k-1} \rho_k(\lambda) \tag{4.15}$$

where $\rho_n(\lambda) \equiv \frac{1}{\pi} \mathrm{Im} \ln a_n(\lambda)$ is the density of states of a pure system (i.e. without randomness) of length n. From the density of states we can find the average number of words in the limit $n \to \infty$.

Define $\langle V^*(\mu, p)\rangle$ as follows

$$\langle V^*(\mu, p)\rangle = \lim_{n\to\infty} \frac{\langle V_n^*(\mu)\rangle}{n} = 2\int (2\lambda + 1)^{\mu-1}\langle\rho(\lambda)\rangle\, \mathrm{d}\lambda \tag{4.16}$$

where the integration over λ runs over the whole spectrum. Let us repeat once more that the function $\langle\rho(\lambda)\rangle$ is the density of states of the random matrix $\hat{M}_n(d = 2)$ $(n \to \infty)$ averaged over the disordered pattern $\{x^{(n)}\}$.

Some additional aspects concerning the relation of these problems with disordered systems are discussed in the appendix.

The physical content of equation (4.15) is as follows. The density of states $\langle\rho(\lambda)\rangle$ can be obtained by averaging the spectrum of the regular case 'weighted' with associated sequences of $\{1\}$:

$$\mathrm{Prob}\left\{x = (0\,\underbrace{1\;1\;1\;1\;\ldots\;1\;1\;1}_{\text{complete set of }\{1\}}\;0)\right\} = p^2(1-p)^n. \tag{4.17}$$

One should also add the contribution coming from the zero-energy state corresponding to entire sequences of $\{0\}$. The resulting expression reads

$$\begin{aligned}\langle\rho(\lambda)\rangle = \sum_{m=1}^{\infty} p^2(1-p)^{2m-1}\left\{m\delta(\lambda+1) + \sum_{k=1}^{m}\delta\left(\lambda + 1 - 4\cos^2\frac{k\pi}{2m+2}\right)\right\}\\ + \sum_{m=0}^{\infty} p^2(1-p)^{2m}\left\{m\delta(\lambda+1) + \sum_{k=1}^{m+1}\delta\left(\lambda + 1 - 4\cos^2\frac{k\pi}{2m+3}\right)\right\}\end{aligned} \tag{4.18}$$

which may be rewritten as

$$\langle\rho(\lambda)\rangle = \frac{1-p}{2-p}\delta(\lambda+1) + \sum_{n=0}^{\infty} p^2(1-p)^n \sum_{k=1}^{[\frac{n+2}{2}]}\delta\left(\lambda + 1 - 4\cos^2\frac{k\pi}{n+3}\right) \tag{4.19}$$

where $[x]$ denotes the integer part of x.

Using equation (4.19) we may check that the function $\langle\rho(\lambda)\rangle$ is properly normalized:

$$\int_{-\infty}^{+\infty}\langle\rho(\lambda)\rangle\,\mathrm{d}\lambda = \int_{-1}^{3}\langle\rho(\lambda)\rangle\,\mathrm{d}\lambda = 1.$$

Returning to (4.16) we find

$$\langle V^*(\mu, p)\rangle = 2\left(\frac{1-p}{2-p}\right)(-1)^{\mu-1} + 2\sum_{n=0}^{\infty} p^2(1-p)^n S_{n+2}(\mu) \tag{4.20}$$

where

$$S_n(\mu) = \sum_{k=1}^{[\frac{n}{2}]}\left(8\cos^2\frac{k\pi}{n+1} - 1\right)^{\mu-1} \tag{4.21}$$

(cf (3.24)).

In order to check the algebra we have computed $\langle V^*(\mu, p)\rangle$ for small values of μ. One finds

$$\langle V^*(1)\rangle = \langle V^*(2)\rangle = 2$$

which can be readily obtained through a direct calculation of $\lim_{n\to\infty} 2\tilde{\boldsymbol{v}}_{\mathrm{in}}(2\hat{M}_n + \hat{I})^{\mu-1}\boldsymbol{v}_{\mathrm{out}}$. We are, however, mainly interested in the limit $\mu \to \infty$.

Using (3.25) we may resum the series (4.20) by isolating the contribution $n < \mu$ and $n > \mu$. After some algebra one obtains for $\mu < n+1$

$$S_n(\mu) = (n+1)\sum_{p=1}^{\mu-1} C_{\mu-1}^{p} 2^p C_{2p-1}^{p-1} - \tfrac{1}{2}\{7^{\mu-1} + (-1)^\mu\} - \left[\frac{n}{2}\right](-1)^\mu \tag{4.22}$$

where $C_{\mu-1}^{p}$ and $C_{\mu-1}^{p-1}$ are the binomial coefficients.

Equation (4.22) gives

$$S_n(\mu)|_{\mu\gg 1} \simeq \begin{cases} 7^\mu & n_0 < n < \mu \\ (n+1)7^\mu & \mu < n \end{cases} \tag{4.23}$$

where n_0 is some constant ($n_0 \approx 5$). The curve $\ln S_n(\mu)$ is plotted in figures 2(*a*), (*b*).

Thus, we can rewrite equations (4.20), (4.21) as follows

$$\langle V^*(\mu, p)\rangle = 2\left(\frac{1-p}{2-p}\right)(-1)^{\mu-1} + \sum_{0}^{\mu} p^2(1-p)^n S_{n+2}(\mu) + \sum_{\mu}^{\infty} p^2(1-p)^n S_{n+2}(\mu). \tag{4.24}$$

The corresponding behaviour of the function $Q(p|\mu)$ where

$$Q(p|\mu) = \frac{\ln\langle V^*(\mu, p)\rangle}{\mu} \qquad (0 < p < 1) \tag{4.25}$$

is shown in figure 3 for few fixed values $\mu = \{10,\ 30,\ 150\}$.

The plot in figure 3 enables us to come to the following conclusions.

(i) If the number of 'errors' is small ($p \to 0^+$), the volume of the group grows exponentially with the Lyapunov exponent $\ln 7$ (for $\mu \to \infty$).

(ii) For an arbitrary number of 'errors', p, the corresponding Lyapunov exponent approaches the same value $\ln 7$ for all $p < 1$ in the limit $\mu \to \infty$ and exhibits *a singular behaviour* just at the point $p = 1$ (which corresponds to a completely commutative group).

The asymptotic expression (4.24) allows us to conclude that the limit behaviour of the function $V^*(\mu)$ is independent of p, $\forall p \in]0, 1[$ and is the same as for the locally free group $\mathcal{LF}_n$ without any errors. This fact proves relation (4.6) and hence supports our conjecture (4.3).

It should be emphasized that these results are expected to hold only in the thermodynamic limit $n \to \infty$. It would be more desirable to consider the limit in which the number of generators n is kept fixed and the length of the word, μ, is much larger than n.

5. Discussion

5.1. The geometrical view of the word enumeration problem

The number of primitive words in the locally free or braid groups allows a rather straightforward geometrical description, namely, the matrix $\hat{M}_n(d = 2)$ can be regarded as the transfer matrix for the model of a 'biased Levy-flight'-like ('BLF'-like) one-dimensional random walk on the finite support. Actually, let us compute the statistical sum of the process described below. Take n integers on the line: $1, 2, \ldots, n$ and consider the random walk when the walker can jump with equal probabilities from the vertex with the coordinate m_1 ($1 \leqslant m_1 < n$) to:

(i) each vertex with the coordinate m_2 ($m_2 \in [m_1 + 1, n]$);

(ii) the vertex with the coordinate $m_2 = m_1 - 1$ (i.e. one step back).

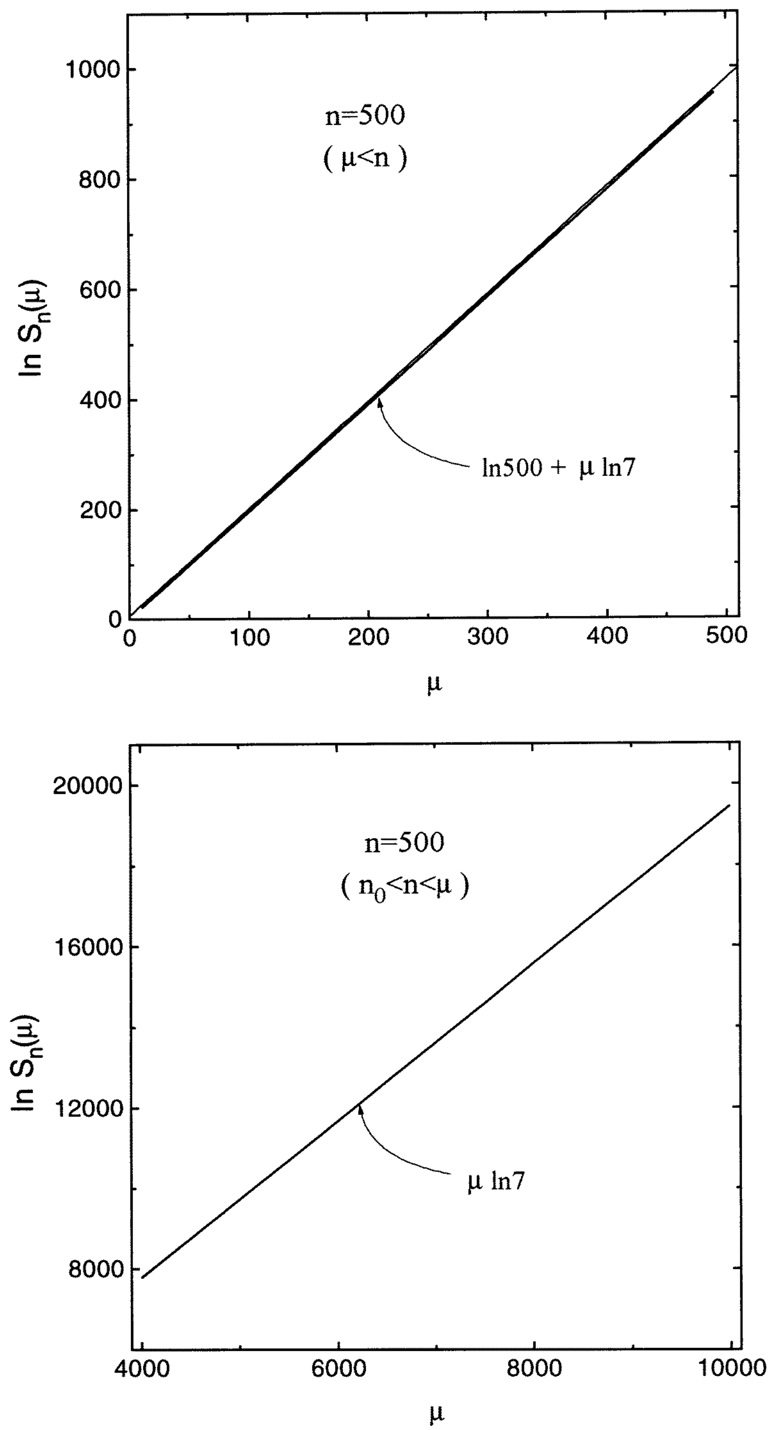


Figure 2. Plot of the function $\ln S_n(\mu)$ in two regimes (equation (4.22)).

The corresponding process is represented schematically in figure 4.

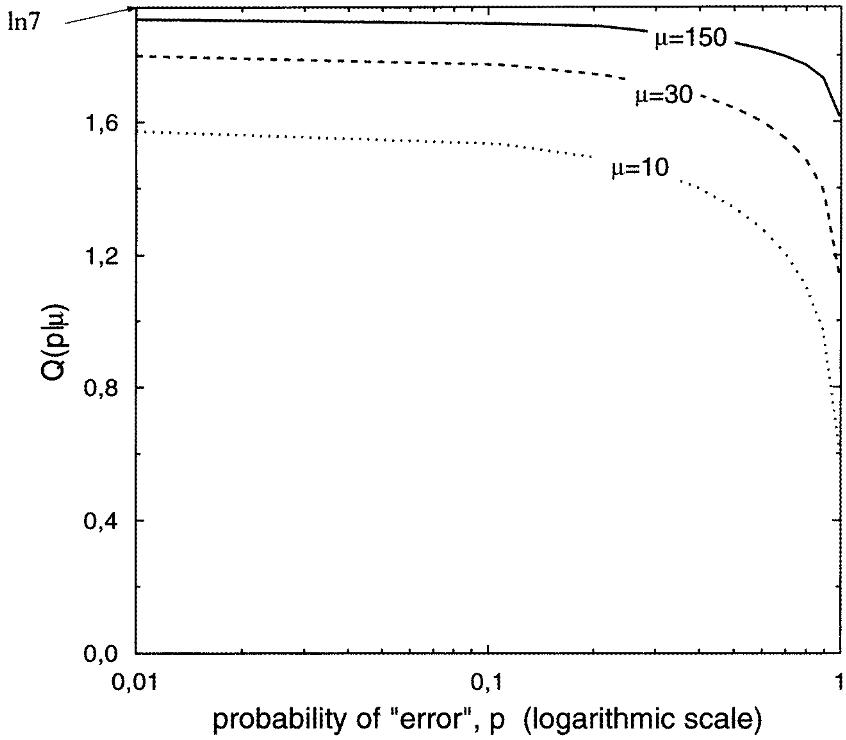


Figure 3. Plot of the function $Q(p|\mu)$ for three fixed values of $\mu = \{10, 30, 150\}$—see equation (4.25).

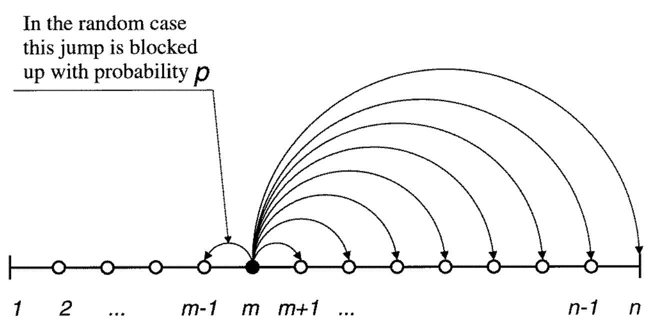


Figure 4. Schematic representation of the process associated with the 'biased Levy-flight' (BLF)-like random walk.

Analogously, we can associate the random operator $\hat{M}_n^{\mathrm{err}}$ with the transfer matrix of the generalized BLF-like random process which is described via the same rules (i) and (ii) but with the additional requirement that the jump (ii) is blocked with probability p and allowed with probability $1 - p$, independently of the position of the vertex.

5.2. Final remarks

We have proposed a statistical method for enumerating the primitive words in the braid group B_n based on the consideration of *locally free groups with errors in the commutation relations*. We invoked self-consistent arguments in support of the conjecture that the number of long primitive words in the braid group is not sensitive to the precise local commutation relations.

We have found an intriguing connection between the above mentioned problems: conventional random operator theory, localization phenomena and the theory of automorphic functions—see the appendix for details.

We believe that the problem of discovering the integrable models associated with the proposed locally free groups and developing the corresponding conformal field theory could help establish a bridge between the statistics of random walks on noncommutative groups, spectral theory on multiconnected Riemann surfaces, and topological field theory.

Acknowledgments

We are very grateful to J Desbois for elucidating to us many questions concerning the limiting behaviour of random walks on locally free and braid groups; we would also like to thank M Tsypin for useful remarks and for help in the numerical confirmation of some of our conjectures. SN acknowledges fruitful discussions with S Fomin, L Pastur, Ya Sinai and A Vershik on many aspects of the work. We highly appreciate the assistance of D Dean and O Martin in the final preparation of the paper and critical reading of the manuscript.

Appendix. Functional equation, continued fractions and invariant measure

The behaviour of the spectral density of our model is very similar to the one encountered in the study of harmonic chains with binary random distribution of masses. This problem, which goes back to Dyson has been investigated by Domb *et al* [27] and then thoroughly discussed by Nieuwenhuizen and Luck [28]. One considers a chain of oscillators where the masses can take two values:

$$\begin{array}{ll} m & \text{with the probability } 1-p \\ M > m & \text{with the probability } p. \end{array}$$

In the limit $M \to \infty$ the system breaks into islands, each of which consist of n light masses surrounded by two infinite heavy masses. The probability of occurrence of such an island is $p^2(1-p)^n$. There is clearly a mapping to our model if one replaces the sequences of heavy and light masses by the sequences of '0' and '1'. Many results may therefore be borrowed from the works [27, 28]. In particular, by adapting the calculations of Nieuwenhuizen and Luck to our case one may rewrite the integrated density of states

$$\langle \mathcal{N}(\lambda) \rangle = \int_{-\infty}^{\lambda} \langle \rho(\lambda') \rangle \, \mathrm{d}\lambda'$$

in the form

$$\langle \mathcal{N}(\lambda) \rangle = 1 - \frac{p}{(1-p)^2} \sum_{n=1}^{\infty} (1-p)^{\mathrm{Int}(\frac{n\pi}{\vartheta})} \tag{A.1}$$

where the relation between λ and ϑ is given in equation (3.19).

For $\lambda \to -1$ one obtains $\langle \mathcal{N}(\lambda) \rangle \to \frac{1-p}{2-p}$ which corresponds to the contribution of the states $\lambda = -1$ at the bottom of the spectrum.

At the upper edge of the spectrum (namely for $\lambda \to 3^-$) one obtains $\langle \mathcal{N}(\lambda) \rangle \to 1$ which means that all the states are counted. Equation (A.1) shows that the behaviour around $\lambda = 3$ (corresponding to $\vartheta = 0$) is in fact dominated by the first term ($n = 1$) of the series. One has

$$\begin{aligned} \langle \mathcal{N}(\lambda) \rangle &\simeq 1 - \frac{p}{(1-p)^2}(1-p)^{\frac{2\pi}{\sqrt{3-\lambda}}} \\ &\equiv 1 - \frac{p}{(1-p)^2} \exp\left[\frac{2\pi}{\sqrt{3-\lambda}} \ln(1-p)\right]. \end{aligned} \tag{A.2}$$

The behaviour (A.2) signals the appearance of Lifshits' singularity in the density of states. A more precise analysis shows that this result is in fact modulated by a periodic function [28].

Equation (A.1) displays many interesting features. In particular, the function $\langle \mathcal{N}(\lambda) \rangle$ occurs in the mathematical literature as a generating function of the continued fraction expansion of $\frac{\pi}{\vartheta}$.

Let us briefly sketch this connection. Consider the continued fraction expansion

$$\frac{\pi}{\vartheta} = \cfrac{1}{c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cdots}}} \tag{A.3}$$

where all c_n are natural integers. Truncating this expansion at level n we obtain a rational number $\frac{p_n}{q_n}$ which converges to $\frac{\pi}{\vartheta}$ when $n \to \infty$.

A theorem of Böhmer [29] states that the generating function of the integer part of $\frac{\pi}{\vartheta}$

$$G(z) = \sum_{n=1}^{\infty} z^{\mathrm{Int}(\frac{n\pi}{\vartheta})}$$

is given by the continued fraction expansion

$$G(z) = \frac{z}{1-z} \cfrac{1}{A_0 + \cfrac{1}{A_1 + \cfrac{1}{A_2 + \cdots}}}$$

where

$$A_n(z) = \frac{(\frac{1}{z})^{q_n} - (\frac{1}{z})^{q_{n-2}}}{(\frac{1}{z})^{q_{n-1}} - 1}$$

and q_n is the denominator of the fraction $\frac{p_n}{q_n}$ approximating the value $\frac{\pi}{\vartheta}$.

In order to make a connection with our problem it is sufficient to set $z = 1 - p$ and express $\langle \mathcal{N}(\lambda) \rangle$ in terms of $G(z)$.

Equation (A.1) shows that the invariant measure is a very singular object. However, it satisfies a simple functional equation reminiscent of that which arises in the theory of automorphic forms. The equation can be derived either by a Dyson–Schmidt approach (see, for instance, [18]) or just by looking at the explicit expression of $\mathcal{P}(\rho)$.

It is in fact simpler to work with the rescaled variable

$$z_n = -\frac{1}{\sqrt{\lambda + 1}} \rho_n$$

which satisfies the recursion relation equivalent to equation (4.12)

$$z_n = h - \frac{1}{z_{n-1}} \tag{A.4}$$

where $h = \sqrt{\lambda + 1}$. The transformation (A.4) may be obtained from the two matrices belonging to the group $SL(2, \mathbb{R})$

$$\hat{T} = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \qquad \hat{S} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The $SL(2, \mathbb{R})$-transformation $\hat{T}\hat{S} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ acts on z by the fractional linear transformation

$$z_n = \frac{az_{n-1} + b}{cz_{n-1} + d} = h - \frac{1}{z_{n-1}}. \tag{A.5}$$

The invariant measure of (A.4), which may be rewritten as

$$\mathcal{P}(z) = \sum_{n=0}^{\infty} p(1-p)^n \delta(z - [\hat{S}\hat{T}]^n z_0) \tag{A.6}$$

can easily be shown to satisfy the fundamental equation

$$\mathcal{P}(z) = \frac{1}{1-p}(cz + d)^2 \mathcal{P}(\hat{S}\hat{T}z) \tag{A.7}$$

up to some singular terms.

By suitable rescaling it is in fact possible to absorb the prefactor $1/(1-p)$ and rewrite equation (A.7) as

$$\mathcal{P}(\hat{g}z) = (cz + d)^{-2}\mathcal{P}(z) \tag{A.8}$$

where $\hat{g} \in SL(2, \mathbb{R})$.

An analytical continuation of this expression into the complex z-plane would eventually permit one to interpret $\mathcal{P}$ as an automorphic form. From the theory of automorphic functions it is well known that equation (A.8) is satisfied by the Dedekind modular function $\eta(z)$:

$$\eta(z) = \frac{1}{(cz + d)^2}\eta\left(\frac{az + b}{cz + d}\right) \qquad (ad - bc = 1).$$

Such objects although perfectly smooth in the upper half-plane $\mathrm{Im}\, z > 0$, display highly nontrivial fractal behaviour on the boundary $\mathrm{Im}\, z = 0$ (see, for instance, [7, 32]).

Another interesting connection which would be worth investigating is the fact that $\hat{S}$ and $\hat{T}$ generate the so-called Hecke group $\Gamma(h)$ for $h = 2\cos\frac{\pi}{q}$ ($q \geqslant 3$ is integer). Surprisingly, these values of h coincide with a subset of the spectrum of the matrix $\hat{M}_n$ (see equation (3.21)).

References

[1] Hofstadter D 1976 *Phys. Rev.* B **14** 2239
Guillement J, Helffer B and Treton P 1989 *J. Physique* **50** 2019
[2] Bellissard J 1986 *Lect. Note. Phys.* **257** 99
[3] Wiegmann P and Zabrodin A 1994 *Phys. Rev. Lett.* **72** 1890
Wiegmann P and Zabrodin A 1994 *Nucl. Phys.* B **422** 495
[4] Bellissard J and Simon B 1982 *J. Funct. Anal.* **48** 408
[5] Gutzwiller M 1990 *Chaos in Classical and Quantum Mechanics* (Heidelberg: Springer)
[6] Terras A 1979 *Harmonic Analysis on Symmetric Spaces and Applications* (Berlin: Springer)

[7] Hejhal D 1976 *The Selberg Trace Formula for $PSL(2, R)$ (Lecture Notes in Mathematics 548)* (Berlin: Springer)
Hejhal D 1983 *The Selberg Trace Formula for $PSL(2, R)$ (Lecture Notes in Mathematics 1001)* (Berlin: Springer)
[8] Chassaing P, Letac G and Mora M 1983 *Probability Measures on Groups (Lecture Notes in Mathematics 1064)*
[9] Gómez C and Sierra G 1993 *J. Math. Phys.* **34** 2119
[10] Majid S 1993 *Int. J. Mod. Phys.* A **8** 4521
[11] Jones V F R 1989 *Pac. J. Math.* **137** 311
[12] Kauffman L H 1991 *Knots and Physics* (Singapore: World Scientific)
[13] Nechaev S K 1996 *Statistics of Knots and Entangled Random Walks* (Singapore: World Scientific)
[14] Lerda A 1992 *Anyons (Lecture Notes in Physics 14)* (Berlin: Springer)
Khare A 1998 *Fractional Statistics and Quantum Theory* (Singapore: World Scientific)
[15] Fürstenberg H 1963 *Trans. Am. Math. Soc.* **198** 377
Tutubalin V 1965 *Prob. Theor. Appl.* **10** 15 (in Russian)
[16] Nechaev S and Sinai Ya 1991 *Bol. Sci. Bras. Mat.* **21** 121
[17] Koralov L, Nechaev S and Sinai Ya 1993 *Prob. Theor. Appl.* **38** 331
[18] Lifshitz I M, Gredeskul S A and Pastur L A 1982 *Introduction to the Theory of Disordered Systems* (Moscow: Nauka)
[19] Mirlin A and Fyodorov Y 1991 *J. Phys. A: Math. Gen.* **24** 227
Fyodorov Y 1995 Basic features of Efetov's supersymmetry approach *Mesoscopic Quantum Physics (Proc. Les Houches, Session LXI, 1994)* (Amsterdam: Elsevier)
[20] Nechaev S K, Grosberg A Yu and Vershik A M 1996 *J. Phys. A: Math. Gen.* **29** 2411
[21] Desbois J and Nechaev S 1997 *J. Stat. Phys.* **88** 201
Desbois J and Nechaev S 1998 *J. Phys. A: Math. Gen.* **31** 2767
[22] Kesten H 1959 *Trans. Am. Math. Soc.* **92** 336
[23] Birman J and Ki Hyoung Ko 1997 A polynomial time algorithm for the word problem in the braid groups *Preprint*
[24] Vershik A M Private communication
[25] Viennot G X 1989 *Ann. NY Acad. Sci.* **576** 542
[26] Mezard M, Parisi G and Virosoro 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[27] Domb C, Maradudin A A, Montroll E W and Weiss G H 1959 *Phys. Rev.* **115** 18
[28] Nieuwenhuizen Th M and Luck J M 1985 *J. Stat. Phys.* **41** 745
[29] Böhmer P E 1925 *Math. Ann.* **96** 368
Borwein J M and Borwein P B 1993 *J. Number Theor.* **45** 293
[30] Lima R and Rahibe M 1994 *J. Phys. A: Math. Gen.* **27** 3427
[31] Derrida B and Hilhorst H J 1983 *J. Phys. A: Math. Gen.* **16** 2641
[32] Berry M 1996 *J. Phys. A: Math. Gen.* **29** 6617