

## MEASURING CONES AND OTHER THICK SUBSETS IN FREE GROUPS

ELIZAVETA FRENKEL\* AND VLADIMIR N. REMESLENNIKOV

Communicated by Evgeny Vdovin

**ABSTRACT.** In this paper we investigate the special automata over finite rank free groups and estimate asymptotic characteristics of sets they accept. We show how one can decompose an arbitrary regular subset of a finite rank free group into disjoint union of sets accepted by special automata or special monoids. These automata allow us to compute explicitly generating functions,  $\lambda$ -measures and Cesaro measure of thick monoids. Also we improve the asymptotic classification of regular subsets in free groups.

### 1. Introduction

This paper continues the series of papers written by different authors [2, 1, 6, 7, 8]. More specifically, we expand the results of [9] and give their proofs. We return to the question of asymptotic classification of regular subsets in finite rank free groups, thus being motivated by needs of universal algebraic geometry. Namely, having in mind the notion of an  $A$ -dimension function over arbitrary algebraic structure introduced by the second author and its applications in different algebraic systems (see [3]), we have started to prepare the algebraic and algorithmic foundations for a suitable dimension function in group theory. In particular, in a sequel paper we are going to present such an algorithm for regular subsets of finite rank free groups over certain group  $A$ . However, the existing asymptotic classification of sets, appeared first in [2] and then refined in [1] does not allow us to fulfill this task. To reveal the problem, we formulate these results (see Section 2 for definitions):

**Theorem 1.1.** [2, 1] *Let  $F$  be a finite rank free group. Then*

- 1) *every regular subset of  $F$  is either thick or exponentially negligible;*

---

MSC(2010): Primary: 20F05; Secondary: 05C05.

Keywords: Free group,  $\lambda$ -measure, regular subset, special automaton, thick monoid.

Received: 30 March 2016, Accepted: 06 May 2017.

\*Corresponding author.

2) a regular subset of  $F$  is thick if and only if its prefix closure contains a cone.

As we shall see below, all the necessary computations can be easily done in the case of regular exponentially negligible sets. The missing bit consists in more specific characterisation of thick sets and in finding the way to distinguish between them in a finer way. The present paper covers these problems.

The following theorem adds to our knowledge on how does the thick sets look like; these new details turn out to be crucial as we shall show in section 4:

**Theorem 3.8.** *A regular subset  $R$  of  $F$  is thick if and only if it contains a subset  $w \circ T$ , with  $T$  being a thick monoid and  $w \in F$ .*

Another important results of the current paper concerns new algorithms for the computation of the generating functions and Cesaro measure of sets recognised by so-called special automata and thick monoids. The algorithms we suggest appears to be easier with respect to the older ones.

Now, a few words on the structure of the paper. In Section 2 we give some basics on regular sets and recall techniques for measuring subsets in a free group  $F$  and the asymptotic classification of regular sets. In subsection 2.2 we also provide Algorithm I for computation of the generating function of a regular set by means of linear algebra.

Section 3 starts with the definition of a special automaton over monoid and group. Further on we prove that every regular subset  $L$  of a finite rank free group  $F$  can be represented as a finite disjoint union of languages accepted by certain type of automata (see Proposition 3.1), which is going to be crucial property of the special automata in a context of both current paper and the construction of a dimension function in free groups. In Lemma 3.3 we also show how one can split the sets accepted by special automata, which leads to the notion of a thick monoid.

Further in this section we analyse the structure and compute the most important asymptotic characteristics of thick monoids: the generating function and the Cesaro measure among the most important of them (see Proposition 3.6 and Lemma 3.7). In Theorem 3.8 we improve already mentioned result on the asymptotic classification of regular sets. We conclude Section 3 with Algorithm II computing  $\lambda$ -measure of a negligible set accepted by a special automaton.

The next Section 4 is dedicated to computations and also it reassumes the results from above. Preliminary calculations made in Lemma 4.1 allows us to compute the generating function of an arbitrary double-based cone (see Theorem 4.2). We want to emphasize that this is a crucial theorem for all the paper, interesting *per se*, applied in Lemma 3.7, and having a lot of structural and computational consequences. In particular, Theorem 4.2 reduces Algorithm I to much more straightforward combinatorial calculations and formulae that does not use linear algebra methods.

## 2. Regular sets in free groups

In this section we recall the main definitions and tools of particular interest for our purposes.

**2.1. Regular sets: some properties.** We assume that the reader is familiar with basic facts on regular sets in monoids and groups (described in details, for example, in [4, 12]). Let  $X = \{x_1, \dots, x_m\}$

be an alphabet and define  $\Sigma$  to be the letters of  $X$  with their formal inverses:  $\Sigma = X \cup X^{-1}$ . Let  $F = F(X)$  be the free group generated by  $X$ . A *finite state automaton*  $\mathcal{A}$  is a quintuple  $(S, \Sigma, \delta, I, Z)$ , where  $S$  is a finite set of states,  $\Sigma$  is an alphabet,  $I \subset S$  is the (non-empty) set of initial states,  $Z \subseteq S$  is the set of final states, and  $\delta$  is a set of arrows with labels in the enlarged alphabet  $\Sigma \cup \varepsilon$  (here  $\varepsilon$  is assumed not to lie in  $\Sigma$ ). Further, a *deterministic automaton* can be considered a special case of a finite state automaton, with no arrows labelled  $\varepsilon$ , the only one initial state and each state being the source of exactly one arrow with any given label from  $\Sigma$ . By the Kleene-Rabin-Scott theorem, all regular subsets over  $\Sigma$  (i.e. the closure of finite subsets of free monoid over  $\Sigma$  under the rational operations) are exactly the sets accepted by a finite state automaton over  $\Sigma \cup \varepsilon$ , or, equivalently, accepted by a deterministic automaton over  $\Sigma$ . The language accepted by an automaton  $\mathcal{A}$  we shall denote by  $L = L(\mathcal{A})$ .

**2.2. Multiplicative measures: basics and first algorithms.** We denote by  $|f|$  the length of an element  $f \in F$ , and let  $S_k = \{w \in F \mid |w| = k\}$  denote the sphere of radius  $k$  in  $F$ . We consider a subset  $R$  of  $F$ , and denote by  $f_k(R) = \frac{|R \cap S_k|}{|S_k|}$  the *frequency* of elements from  $R$  among the words of length  $k$  in  $F$ .

$\lambda$ - **measure.** An important measuring tool in  $F$  is the so-called frequency measure, introduced in [2] and studied in [6] and [7]. By definition,

$$\lambda(R) = \sum_{k=0}^{\infty} f_k(R).$$

A subset  $R \subseteq F$  is called  $\lambda$ -*measurable*, if  $\lambda(R) < \infty$ , and *exponentially  $\lambda$ -measurable* if there exists a positive constant  $\delta < 1$  such that  $f_k(R) < \delta^k$  for big enough  $k$ . We adjust this measure to obtain  $\lambda^*(R) = \frac{2m}{2m-1} \lambda(R)$ .

**Generating function.** One can consider the (*frequency*) *generating function* for  $R$  as a formal series in  $R[[t]]$ :  $g_R(t) = \sum_{k=0}^{\infty} f_k(R)t^k$ . We shall also use the *adjusted* version of this function:  $g_R^*(t) = \frac{2m}{2m-1} \cdot g_R(t)$ . In case of regular subsets of  $F$  the generating function can be described in a very concise form:

**Theorem 2.1.** *For a regular set  $R \subseteq F$  the function  $g_R(t)$  is a rational function of  $t$  with rational coefficients and either*

- *has no singularity at  $t = 1$  (in this case  $R$  is exponentially  $\lambda$ -measurable) or*
- *has a simple pole at  $t = 1$  (in this case  $R$  is thick<sup>1</sup>).*

*In particular,*

$$(2.1) \quad Res_1 g_R(t) = -\mu_0(R).$$

---

<sup>1</sup>The rationality of  $g_R(t)$  for regular sets is well known (see for instance [5]; it follows also from Algorithm I for  $g_R(t)$  below), while statements about asymptotic properties of  $R$  follow from asymptotic classification of regular sets shown in [2, 1].

Recall that a regular set is called *thick* if the parameter  $\mu_0(R)$  defined by formula (2.1) is strictly positive. This parameter  $\mu_0(R)$  is called *Cesaro density* of  $R$ . We use often the following simple properties of the generating function: suppose  $R_1$  and  $R_2$  are regular subsets of  $F$ . Then

- (1.) If  $R = R_1 \cup R_2$ , then the corresponding generating function can be computed as  $g_R(t) = g_{R_1}(t) + g_{R_2}(t) - g_{R_1 \cap R_2}(t)$ .
- (2.) If  $R = R_1 \circ R_2$ , then  $g_R(t) = g_{R_1}(t)g_{R_2}^*(t)$ .

Now we describe the first algorithm for calculation of the generating function for an arbitrary regular subset of a finite rank free group  $F$ . This algorithm is previously known (see, for example, [2]), although it was not directly formulated there.

**Algorithm I computing the frequency generating function  $g_R(t)$  for a regular set  $R$ .** Indeed, let  $\mathcal{A} = (S, \Sigma, \delta, I, Z)$  be an automaton such that  $|S| = n$  and let  $A$  be it's adjacency matrix, i.e.  $n \times n$  matrix with entries  $a_{ij}$  such that each  $a_{ij}$  corresponds to the number of arrows from the state  $i$  to the state  $j$ . Clearly, the number of different paths of length  $k$  from  $i$  to  $j$  is equal to  $(A^k)_{i,j}$ . Denote by  $R$  the subset of  $F$  accepted by  $\mathcal{A}$ .

**Algorithm I:**

1. Given an automaton  $\mathcal{A}$ , compute the entries  $a_{ij}$ ,  $i, j = 1, \dots, n$  of the adjacency matrix  $A$ .
2. Compute the entries  $b_{ij}$  of the fundamental matrix  $B = tA(E - tA)^{-1}$  of  $\mathcal{A}$ , with the entries  $b_{ij}$  from the ring of formal power series  $R[[t]]$ .
3. The generating function  $g_R(t)$  is equal to  $\sum_{i \in I, j \in Z} b_{ij}$ .

One of the disadvantages of the Algorithm I is that step [2.] involves the matrix inversion, and it makes the algorithm hardly implementable with the size of automaton  $n$  big enough. However, computation of the generating function can be significantly simplified for a wide class of regular sets. In what follows we introduce this type of sets and describe their structure along with the improved algorithm for computation of  $g(t)$ . Now, using Algorithm I and the properties above, we compute generating function for certain regular subsets of  $F$ . We also calculate the corresponding values of Cesaro density  $\mu_0(R)$  (defined by formula (2.1)).

**Example 2.2.** (1) The entire free group  $F$  has  $g_F(t) = -\frac{1}{t-1}$  and  $\mu_0(F) = 1$ .

(2) For a set  $R = F^\sharp = F \setminus \{1\}$  we have  $g_R(t) = \frac{-t}{t-1}$  while  $\mu_0(F^\sharp) = 1$ .

(3) Let  $R$  be a cone<sup>2</sup>  $C(w)$  or  $R = C[w]$ , and let  $|w| = r$ . Then  $g_R(t) = \frac{1}{2m(2m-1)^{r-1}} \cdot \frac{-t^r}{t-1}$ , and

$$\mu_0(R) = \frac{1}{2m(2m-1)^{r-1}}.$$

(4) If  $R = F \setminus B_{r-1}$ , then  $g_R(t) = \frac{-t^r}{t-1}$ , and  $\mu_0(R) = 1$ .

(5) For a subgroup  $H < F(X)$  of all words of even length direct calculations of frequency generating functions gives  $g_H(t) = \frac{1}{1-t^2}$ , and therefore  $\mu_0(H) = \frac{1}{2}$ .

---

<sup>2</sup>We recall the notions of cones in 3.4

### 3. Special automata over free groups and monoids

In this section we investigate one of the central concepts of this paper, i.e. special automata over monoids and groups. We show in Proposition 3.1 that every regular set in a free group can be decomposed into finite union of subsets accepted by special automata.

**3.1. Definitions.** Let  $\mathcal{A} = (S, \Sigma, \delta, i_0, Z)$  be a deterministic automaton.  $\mathcal{A}$  is called *special* over the monoid  $\Sigma^*$  if

- a. The initial vertex has no inedges;
- b. There is only one final state  $z_0 \in Z$ ;
- c.  $\mathcal{A}$  does not contain inaccessible states;
- d. For every state  $s \in S$  there is a direct path from  $s$  to the final state  $z_0$ ;
- e. For any state  $s \in S$ , all arrows which enter  $s$  have the same label  $x \in \Sigma$  (we shall say,  $s$  has type  $x$ ).

In order to adjust the notion of speciality to groups, we impose an additional constraint on automata. Namely, let  $F$  be the free group, and  $\mathcal{A}$  be a special automaton. Suppose also that

- f. For any state  $s$  of type  $x$  in  $\mathcal{A}$ , all arrows exiting from  $s$  cannot have label  $x^{-1}$ .

$\mathcal{A}$  is a *special* automaton over the group  $F$ , if it satisfies the conditions (a)–(f).

In what follows we also shall use a notion of a special monoid. Namely, a monoid  $M$  is called *special* if it is accepted by a finite automata  $\mathcal{A}$  with  $i_0 = z_0$ , satisfying conditions (b) – (f).

### 3.2. Decomposition into special automata.

**Proposition 3.1.** *Let  $L$  be a regular language in  $F$ . Then there exist a finite number of automata  $\mathcal{A}_0, \dots, \mathcal{A}_k$  such that*

- $L$  is a disjoint union of languages  $L_0 = L(\mathcal{A}_0), \dots, L_k = L(\mathcal{A}_k)$  in  $F$ :  $L = L_0 \sqcup L_1 \sqcup \dots \sqcup L_k$ ;
- every  $L_i$  is either accepted by a special automaton or a special monoid.

*Proof.* Since  $L$  is regular in  $F$ , it is accepted by a finite automaton  $\mathcal{A}$ . Although we can assume that  $\mathcal{A}$  is deterministic automaton, it will be more convenient for us to start with a non-deterministic one, which accepts  $L$  as a language of reduced words, satisfies (c) (it is always possible, see [10]), but, probably, has  $\varepsilon$ -transitions and more than one initial state. Therefore,  $\mathcal{A}$  has a form  $\mathcal{A} = (S, \Sigma \cup \varepsilon, \delta, I, Z)$ . We begin with an application of Rabin-Scott powerset construction (see [13] for details). As an output of this procedure, we obtain an automaton  $\mathcal{A}' = (S', \Sigma, \delta', i_0, Z')$ , which does not have  $\varepsilon$ -transitions and has only one initial state  $i_0$  without inedges, as required. As a by-product of the construction, we have conditions (c) and (d) satisfied. Further, because we have started from the automaton  $\mathcal{A}$  which does not have consecutive  $x, x^{-1}$  ( $x \in \Sigma$ ) transitions,  $\mathcal{A}'$  does not have these transitions as well. Nevertheless, it might happen that  $\mathcal{A}'$  has more than one final state and some states of  $S'$  have incoming edges with different labels. In the latter case we split the states of  $\mathcal{A}'$  as it is shown on figure 1:

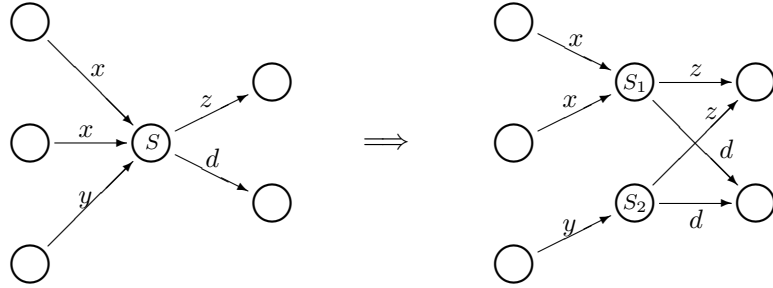


FIGURE 1. Splitting the states of the automaton  $\mathcal{A}'$ .

The output of the splitting procedure we shall call  $\mathcal{A}'' = (S'', \Sigma, \delta'', i_0, Z'')$ . If  $\mathcal{A}''$  has only one final state  $z_0 \neq i_0$ , then it is special over  $F$ . If  $Z'' = \{z_0\}$  and  $i_0 = z_0$ , then  $L(\mathcal{A}'')$  is a special monoid by definition and due to (f). Suppose now  $Z'' = \{z_0, \dots, z_k\}$ , with  $k \geq 1$ . For every  $z_i \in Z''$  consider the maximal connected subgraph  $\mathcal{A}_i = (S_i, \Sigma, \delta_i, i_0, z_i)$  of  $\mathcal{A}''$  such that  $S_i \subset S''$  and  $\delta_i \subset \delta''$  induced by the paths of arrows from  $i_0$  to  $z_i$ ; obviously, there are two options for  $\mathcal{A}_i$ : either  $\mathcal{A}_i$  has distinct initial and final state and therefore  $\mathcal{A}_i$  is special, or initial and final states coincide and so  $L(\mathcal{A}_i)$  is a monoid. Since  $L = L(\mathcal{A}'')$  and  $\mathcal{A}''$  satisfies (c), (d), clearly,  $L = L(\mathcal{A}_0) \cup L(\mathcal{A}_2) \cup \dots \cup L(\mathcal{A}_k)$ . Moreover, this union is disjoint since  $L_i \cap L_j \neq \emptyset$  implies existence of paths of arrows  $p_1, p_2$  such that  $p_1 = i_0 v_1 \dots v_s z_i \in \mathcal{A}''$  and  $p_2 = i_0 u_1 \dots u_r z_j \in \mathcal{A}''$  for  $z_i \neq z_j$ , with the label  $\delta''(p_1) = \delta''(p_2)$ , a contradiction with  $\mathcal{A}''$  being deterministic.  $\square$

**Remark 3.2.** Notice that the number  $k$  and subsets  $L_i$  for different decompositions of  $L$  can vary. On the other hand, suppose  $L = L_0 \sqcup L_2 \sqcup \dots \sqcup L_k$  and  $L = M_0 \sqcup M_2 \sqcup \dots \sqcup M_s$  are different decomposition of a regular set  $L$  as in Proposition 3.1. Then by property (1.) of the generating functions we have  $g_{L_0}(t) + \dots + g_{L_k}(t) = g_L(t) = g_{M_0}(t) + \dots + g_{M_s}(t)$ .

**3.3. Further splitting of subsets in free groups.** A special automaton satisfying (a)–(f) in turn admits further splitting:

**Lemma 3.3.** Let  $R = R(\mathcal{A})$  and  $\mathcal{A} = (S, \Sigma, \delta, i_0, z_0)$  be a special automaton over  $F$ . Then there exist regular languages  $R_1, R_2, R_3 \subset F$  such that  $R_j$  are accepted by  $\mathcal{A}_j = (S, \Sigma_j, \delta_j, i_j, z_j)$ ,  $\mathcal{A}_1$  is special over  $F$  and

1. if  $\mathcal{A}$  has at least one arrow exiting  $z_0$ , then  $R_2$  is non-empty and  $i_2 = z_2$ , while  $i_3 \neq z_3$  and

$$(3.1) \quad R = R_1 \circ R_2 \text{ is unambiguos;}$$

$$(3.2) \quad R_2 = 1 \sqcup R_3 \sqcup (R_3 \circ R_3) \sqcup (R_3 \circ R_3 \circ R_3) \sqcup \dots;$$

$$(3.3) \quad g_R(t) = g_{R_1}(t)g_{R_2}^*(t); \quad \lambda(R) = \lambda(R_1)\lambda^*(R_2).$$

2. if there is no arrows exiting  $z_0$ , then  $R_2 = R_3 = \emptyset$ ,  $R = R_1$ ,  $\lambda(R) = \lambda(R_1)$ , and  $g_R(t) = g_{R_1}(t)$ .

*Proof.* Although the construction of sets  $R_1, R_2, R_3$  and their  $\lambda$ -measures appears in [2] and [7], we shall widely use these sets and automata in what follows, and therefore we repeat briefly the necessary computations (see also Example 3.5 and its illustrations in figures 2, 3, 4(a), 4(b)).

Suppose that the final state of  $\mathcal{A}$  does not have exiting arrows. Then we leave  $\mathcal{A}$  as it is, and, clearly, [2.] holds.

Let now  $z_0$  has at least one exiting arrow. In this case the special automaton  $\mathcal{A}_1$  accepting  $R_1$  can be obtained from  $\mathcal{A}$  by removing all arrows exiting from  $z_0$ ; we take  $i_1 = i_0$  and  $z_1 = z_0$ . Let us consider the automaton  $\mathcal{A}_2$  accepting  $R_2 \neq \emptyset$  formed by all states accessible from the state  $z_0$ , with the same arrows between them as in  $\mathcal{A}$ ; we take  $z_0$  for the both  $i_2$  and  $z_2$ . If now  $u \in R_1$  and  $v \in R_2$ , then the word  $uv$  is reduced and  $\lambda(uv) = \lambda(u)\lambda^*(v)$ . Therefore, the presentation of  $R$  in the form  $R = R_1 \circ R_2$  is unambiguous. Indeed, let  $w \in R$  can be written in two different forms as  $u_1 \circ v_1$  and  $u_2 \circ v_2$ , where  $u_1, u_2 \in R_1$  and  $v_1, v_2 \in R_2$ . Assume that  $|u_1| > |u_2|$  (otherwise consider the pair  $v_1$  and  $v_2$ ), and let  $h = u_2^{-1}u_1 \in F$  be readable in  $\mathcal{A}$ . Notice that  $h$  starts at  $z_0$  since  $u_2$  is accepted by  $\mathcal{A}_1$  and ends at  $z_0$  because  $\mathcal{A}_1$  accepts  $u_1$ . Therefore,  $h$  is accepted by  $\mathcal{A}_2$ , a contradiction with the construction of  $\mathcal{A}_1$ . The estimates on  $\lambda(R)$  and  $g_R(t)$  now follow immediately from the construction (frequencies assigned to arrows in  $\mathcal{A}_1, \mathcal{A}_2$  the same as they were in  $\mathcal{A}$ ) and formula (2.).

Further, we transform the automaton  $\mathcal{A}_2$  by splitting the final state  $z_2 = z_0$  into separate initial state  $i_3$  (with no arrows entering it, and those arrows which were exiting  $z_2$  now exiting  $i_3$ ), and the final state  $z_3$  (with no arrows exiting  $z_3$ , and those arrows which were entering  $z_0$  now entering  $z_3$ ). Then, clearly, (3.2) holds. □

**Corollary 3.4.** *Let  $R = R(\mathcal{A})$  and  $\mathcal{A} = (S, \Sigma, \delta, i_0, z_0)$  be a special automaton over  $F$ , and let  $R_1, R_2, R_3 \subset F$  be regular languages such that  $R_1$  is accepted by a special automaton over  $F$ ,  $R_2$  is non-empty set such that its initial and the final state coincide,  $R_3$  is accepted by a special automaton over  $F$ ;  $R = R_1 \circ R_2$ , and  $R_2$  satisfies (3.2) as in lemma above. Then the subset  $R_2$  of  $F$  is the free special monoid generated by  $\{w_i | i \in I\}$ , where  $w_i \in F$  are words in  $R_3$  and  $w_i$  can be computed effectively by  $\mathcal{A}$ .*

*Proof.* The automaton  $\mathcal{A}_2$  constructed in the proof of Lemma 3.3 has  $i_2 = z_2$ , and its final vertex  $z_2$  is of  $x$ -type, for some  $x \in \Sigma$ . The condition (f) provided by the speciality of  $\mathcal{A}$  guarantees that the arrow labelled  $x^{-1}$  cannot exit from  $i_2$ . Therefore, if  $u_1, u_2$  are accepted by  $\mathcal{A}_2$ , then  $u_1u_2 = u_1 \circ u_2$ . In particular, using the further splitting of  $R_2$ , one can express every  $u_i$  as a reduced product of  $w_i$ 's accepted by (non-empty)  $R_3$ . Since the identity belongs to  $R_2$ ,  $\mathcal{A}_2$  accepts the free special monoid with generators  $w_i, i \in I$ . □

The subsets and automata described in Lemma 3.3, claim 1. are of particular interest for us. Regular sets  $R \subseteq F$  of such form we shall call *saturated*. Sets  $R_1, R_2$ , and  $R_3$  in the splitting defined in this lemma we shall call a *set of first, second, and third type*, respectively. In what follows, we use the notations  $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$  for the splitting of arbitrary automaton  $\mathcal{A}$  and  $R_1, R_2, R_3$  for the corresponding

regular sets exclusively in a sense of Lemma 3.3. We provide an example of such automata and sets below.

**Example 3.5.** Let  $\Sigma$  be an alphabet  $x, y, t, z$  and the inversion is given by the rule  $t \rightarrow x^{-1}, z \rightarrow y^{-1}$  (so  $X = \{x, y\}$ ). Consider the special automaton  $\mathcal{A}$  (the arrow with a tale corresponds to the initial state, and the finale state is drawn as a double circle).

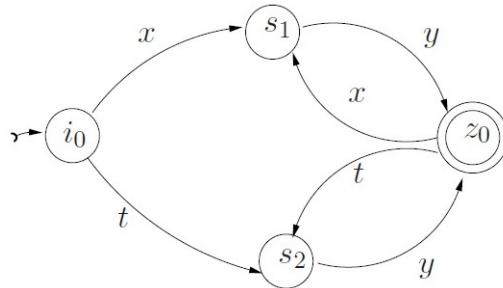


FIGURE 2. The special automaton  $\mathcal{A}$ .

Clearly,  $R = R(\mathcal{A})$  is generated by the following regular expression:

$$R = xy ((x^{-1}y)^*(xy)^*)^* \cup xy ((xy)^*(x^{-1}y)^*)^* \sqcup \sqcup x^{-1}y ((xy)^*(x^{-1}y)^*)^* \cup x^{-1}y ((x^{-1}y)^*(xy)^*)^* .$$

The set of first type  $R_1$  can be read off by the automaton  $\mathcal{A}_1$  shown in figure 3; therefore,  $R_1 = xy \cup x^{-1}y$ .

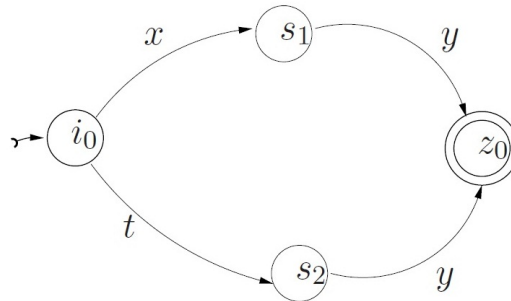


FIGURE 3. The special automaton  $\mathcal{A}_1$ .

The sets of second type  $R_2 = ((xy)^*(x^{-1}y)^*)^*$  and third type  $R_3 = xy \cup x^{-1}y$  with their automata  $\mathcal{A}_2$  and  $\mathcal{A}_3$ . Clearly, the elements  $w_1 = xy, w_2 = x^{-1}y$  provides a set of generators for the monoid  $R_2$  (see Corollary 3.4).

**3.4. Thick semigroups and cones.** We want to classify the subsets accepted by special automata over  $F$  by modulo of their measure. This classification requires recalling of the notion of  $X$ -complete automaton, which was already used in [2] and [6] for analogous purposes. Let  $\mathcal{A} = (S, \Sigma, \delta, i_0, z_0)$  be a special automaton satisfying the conditions (a)–(f).  $\mathcal{A}$  is called  $\Sigma$ -complete if for every state



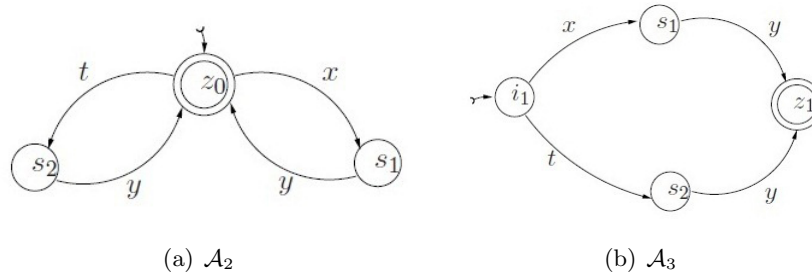


FIGURE 4. The automata  $\mathcal{A}_2, \mathcal{A}_3$ .

$s \in S \setminus \{i_0\}$  of type  $x \in \Sigma$  every label from  $\Sigma \setminus \{x^{-1}\}$  is present on one of the arrows exiting from  $s$  and exactly  $|\Sigma|$  arrows exits from  $i_0$ . Further, let  $R_2$  be a regular set of the second type and  $\mathcal{A}_2 = (S, \Sigma, \delta, z_0, z_0)$  be the corresponding automaton.  $\mathcal{A}_2$  is called  $\Sigma$ -complete if for every state  $s \in S$  of type  $x$  all arrows labeled by  $\Sigma \setminus \{x^{-1}\}$  exit from  $s$ . Otherwise  $\mathcal{A}$  (or  $\mathcal{A}_2$ ) is not  $\Sigma$ -complete (for instance, the automaton  $\mathcal{A}$  in Example 3.5, as well as  $\mathcal{A}_1, \mathcal{A}_2$ , and  $\mathcal{A}_3$  are not  $\Sigma$ -complete). The following proposition shows that the  $\lambda$ -measure can be easily estimated in the latter case.

**Proposition 3.6.** *Let  $\mathcal{A}$  be a special automaton satisfying the conditions (a)–(f), and  $R = L(\mathcal{A})$  be a saturated set such that  $R = R_1 \circ R_2$  is the splitting of the form (3.1), with  $R_1 = L(\mathcal{A}_1)$  and  $R_2 = L(\mathcal{A}_2)$ . If  $\mathcal{A}_2$  is not  $\Sigma$ -complete, then  $R$  is exponentially  $\lambda$ -measurable.*

*Proof.* If  $\mathcal{A}$  is  $\Sigma$ -complete, then only one of  $\mathcal{A}_1, \mathcal{A}_2$  can be  $\Sigma$ -complete since  $R_1 \circ R_2$  is unambiguous by Lemma 3.3. Moreover, if  $\mathcal{A}_2$  is  $\Sigma$ -complete for the corresponding  $\Sigma$ -complete automaton  $\mathcal{A}$ , then  $\mathcal{A}_1$  is not  $\Sigma$ -complete at the state  $z_0$ . Thus, we can assume that precisely  $\mathcal{A}_2$  is not  $\Sigma$ -complete. Then  $R = L(\mathcal{A})$  is  $\lambda$ -measurable by Theorem 3.4 in [2]. Since  $R$  is regular, it is exponentially  $\lambda$ -measurable by asymptotic classification of regular sets ([2, 1]). □

If, on the other hand,  $\mathcal{A}$  is  $\Sigma$ -complete, then we can improve some previously known results on classification of regular subsets in free groups. Namely, let  $R_2$  be a regular subset of  $F$  of second type accepted by the automaton  $\mathcal{A}_2$ . According to Corollary 3.4,  $R_2 = L(\mathcal{A}_2)$  forms a (special) monoid, and if  $\mathcal{A}_2$  is  $\Sigma$ -complete, we shall call  $R_2$  thick. An interesting fact about thick monoids is that we can describe them in terms of double-based cones. We recall that the cone  $C(w)$  is the set of all elements in  $F$  containing  $w$  as initial subword. In what follows we also shall be interested in a symmetric notion of a cone with a right-hand side handle, i.e. the set of all words in  $F$  that terminates with  $w$  (we denote this sort of cones by  $C[w]$ ). Another member of this family is the double-based cone with (nontrivial) handles  $w_1, w_2$ , consisting of all words in  $F$  of the form  $w_1 \circ f \circ w_2, f \in F$ . Notice that all three types of cones are regular in  $F$  (see, for example, Corollary 3.15 [8]). Let us consider the generalized  $x$ -cone  $C(Y, x), x \in \Sigma$ , i.e. the union of double-based cones of the form  $\bigsqcup_{y \in \Sigma: y \neq x^{-1}} C(y, x)$ .

The following technical observation regarding generalized cones give us first examples of thick monoids:

**Lemma 3.7.** *Let  $C(Y, x)$  be the generalized  $x$ -cone,  $x \in \Sigma$  and  $M = C(Y, x) \cup \{1\}$ . Then*

1.  $C(Y, x) = C[x] \setminus C(x^{-1}, x)$ , and  $M$  is a thick monoid;
2.  $g_M(t) = \frac{(2m-1)t^2}{4m^2(1-t)} + 1 + \frac{t}{2m} + \frac{t^2}{4m^2} + \frac{t^3}{2m(2m-1-t)}$ , and
3.  $\mu_0(M) = \frac{2m-1}{4m^2}$ .

*Proof.* The proof of 1. follows from the definitions of cones and thick monoids, while Example 2.2 (3) and Lemma 4.1 [3.], [4.] below provide estimates on the generating function and the Cesaro density given in [2.] and [3.].  $\square$

Now we are ready to refine the asymptotic classification of regular sets in  $F$  (see Introduction and Theorem 3.4 [2] for comparison).

**Theorem 3.8.** *A regular subset  $R$  of  $F$  is thick if and only if it contains a subset  $w \circ T$ , with  $T$  being a thick monoid and  $w \in F$ .*

*Proof.* Clearly, every set of the form  $w \circ T$  is regular and thick (where  $1 \circ T$ , by definition, stands for  $T$ ). Suppose now  $R$  is regular and thick. We decompose  $R$  into a finite number of subsets as in Lemma 3.1. Since a finite union of exponentially  $\lambda$ -measurable subsets is exponentially  $\lambda$ -measurable (see, for example, Proposition 4.1 [6]), without loss of generality one can suppose that  $R$  is accepted by a special automaton or forms a special monoid. In the latter case,  $R$  is a thick monoid itself, so suppose  $R$  is a set accepted by a special automaton. In this case we apply Lemma 3.3 to procure a pair of sets  $R_1$  and  $R_2$  of corresponding types, with  $R_2$  being  $\Sigma$ -complete by Proposition 3.6. Since  $R = R_1 \circ R_2$ , the set  $R$  contains a subset  $w \circ R_2$ , with  $w \in R_1$ . This completes the proof.  $\square$

**3.5. Computing  $\lambda$ -measure of regular sets.** Another immediate consequence of Lemma 3.3 and Proposition 3.6 is an algorithm for computation of  $\lambda$ -measure of exponentially negligible regular set  $R$  accepted by a special automaton  $\mathcal{A}$ . We assume that our reader is familiar with the concept of discrete-time Markov chain and refer to [11] as one of the fundamental manuals on this subject.

Let  $\mathcal{A} = (S, \Sigma, \delta, i_0, z_0)$  be a special automaton over  $F$  and let  $R = R(\mathcal{A})$  be a  $\lambda$ -measurable regular set. We split  $\mathcal{A}$  into  $\mathcal{A}_1$ ,  $\mathcal{A}_2$  and  $\mathcal{A}_3$ , obtaining regular sets  $R_1$ ,  $R_2$  and  $R_3$  (without loss of generality, one can consider the case when all these sets are non-empty). Further, due to formula (3.3) and Proposition 3.6, it is enough to calculate the value of  $\lambda$ -measure for  $R_3$ , accepted by the special automaton  $\mathcal{A}_3 = (S_3, \Sigma, \delta_3, i_1, z_1)$ .

Consider a finite Markov chain  $\mathcal{M}$  with the same states as in  $\mathcal{A}_3$  together with an additional dead state  $D$ . We set transition probabilities from  $z_1$  to  $z_1$  and from  $D$  to  $D$  being equal 1. Every arrow from a state  $s$  in  $\mathcal{A}_3$  gives the corresponding transition from the state  $s$  in  $\mathcal{M}$  which we assign the transition probability  $\frac{1}{2m-1}$ . If at some state  $s$  of type  $x$  in  $\mathcal{A}_3$  there is no exiting arrow labeled  $y \in \Sigma \setminus \{x^{-1}\}$ , we make a transition from  $s$  to  $D$  in  $\mathcal{M}$  assigning it the probability  $\frac{1}{2m-1}$ . We take the stochastic vector being zero everywhere except the state  $i_1$  (so it have the only nontrivial entry 1 at the state  $i_1$ ). This complete the description of the Markov chain  $\mathcal{M}$ . Clearly, the states  $z_1$  and  $D$  of Markov chain  $\mathcal{M}$  are absorbing, and all other states are transient. Obviously,  $P(z_1) = \lambda(R_3)$ ,

and it was shown in [2, 6] that  $\lambda(R_3) < 1$  for any  $\lambda$ -measurable set  $R$ . Therefore, one can calculate  $\lambda(R_2)$  using formula (3.2). A similar argument allows to compute  $\lambda(R_1)$ , and so we are done. Thus, the Markov chain  $\mathcal{M}$  provides us with the following algorithm for computation of  $\lambda(R)$ .

**Algorithm II:** Let  $\mathcal{A}$  be a special automaton and  $R = R(\mathcal{A})$  be  $\lambda$ -measurable.

1. Split  $\mathcal{A}$  into  $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$  as in Lemma 3.3.
2. Construct Markov chains for  $\mathcal{A}_1 = (S_1, \Sigma, \delta_1, i_0, z_0)$  and  $\mathcal{A}_3 = (S_3, \Sigma, \delta_3, i_1, z_1)$ .
3. Calculate the probabilities  $P(z_0)$  and  $P(z_1)$ ; so  $\lambda(R_1) = P(z_0)$  and  $\lambda(R_3) = P(z_1)$ .
4. Compute  $\lambda(R_2) = \sum_{i=0}^{\infty} P^i(z_1) < \infty$ .
5. Finally, compute  $\lambda(R) = \lambda(R_1) \cdot \lambda^*(R_2)$ .

### 4. Computations

In this section we carry out all necessary measurements of double-based cones and thick monoids.

**4.1. Generating functions and Cesaro density of double-based cones.** This technical but crucial lemma will supply us with data about generating function and values of Cesaro density for double-based cones.

**Lemma 4.1.** *Let  $C(a, b)$  be a double-based cone with both handles  $a, b$  in  $\Sigma$ . Then following holds:*

1.  $f_k(C(a, b)) = f_k(C(c, d))$  and therefore  $g_{C(a,b)}(t) = g_{C(c,d)}(t)$  for all  $a, b, c, d$  in  $\Sigma$  such that  $ab \neq 1, cd \neq 1$ . Further,  $f_k(C(a, a^{-1})) = f_k(C(b, b^{-1}))$  for arbitrary  $a, b \in \Sigma$ .
2.  $f_k(C(a, a^{-1})) = (2m - 1)f_k(C(a, a)) - \frac{1}{2m(2m - 1)^{k-1}}$ , for  $k \geq 3$ ,
3.  $g_{C(a,a)}(t) = \frac{t^2}{4m^2(1-t)} + \frac{t^2}{4m^2(2m-1)} + \frac{t^3}{2m(2m-1)(2m-1-t)}$ , and  
 $g_{C(a,a^{-1})}(t) = \frac{t^2}{4m^2(1-t)} - \frac{t^2}{4m^2} - \frac{t^3}{2m(2m-1-t)}$ ,
4.  $\mu_0(C(a, b)) = \mu_0(C(c, d)) = \frac{1}{4m^2}$  for all  $a, b, c, d \in \Sigma$ .

*Proof.* Notice first, that  $n_k(C(a, a^{-1})) = n_k(C(b, b^{-1}))$  (recall that  $n_k(L) = |L \cap S_k|$ , i.e. the number of elements of length  $k$  in  $L$ ), for all  $a, b \in \Sigma$ . The same equalities holds between the other double-based cones:  $n_k(C(a, b)) = n_k(C(c, d))$  for all  $a, b, c, d \in \Sigma$  such that  $ab \neq 1$  and  $cd \neq 1$ . This proves the first claim.

To prove 2., 3., and 4. we are going to construct a bijective map  $\psi : C(a, a) \rightarrow C(a, a^{-1}) \sqcup \{a^n\}_{n>0}$ . For every element  $u \in C(a, a)$  of the form  $u = a^l \circ f_0 \circ a^m$ , where  $l$  and  $m$  maximal, i.e.  $f_0 \neq 1$  does not starts with  $a$  and does not end with  $a$ , define  $\psi(u) = a^l \circ f_0 \circ a^{-m}$ . If, on the hand,  $u \in C(a, a)$  has a form  $u = a^l$ , then take  $\psi(a^l) = a^l$ . Clearly,  $\psi$  is bijective and therefore  $f_k(C(a, a)) = f_k(C(a, a^{-1})) + f_k(\{a^n\}_{n>0})$  for  $k > 2$ . Since  $C(a) = \bigsqcup_{b \in \Sigma} C(a, b)$ , and due to the equality  $f_k(C(a)) = \frac{1}{2m}$ , we have

$$\frac{1}{2m} = (2m - 1)f_k(C(a, a)) + f_k(C(a, a^{-1})) \text{ for } k > 2.$$

But  $f_k(C(a, a^{-1})) = f_k(C(a, a)) - \frac{1}{2m(2m - 1)^{k-1}}$ , and therefore

$$\frac{1}{2m} = 2mf_k(C(a, a)) - \frac{1}{2m(2m-1)^{k-1}} \text{ for } k > 2.$$

To compute generating functions of corresponding sets, we multiply  $f_k$  with  $t^k$  and take an infinite sum of these products. As a result we obtain:

$$\sum_{k=2}^{\infty} f_k(C(a))t^k = (2m-1) \sum_{k=2}^{\infty} f_k(C(a, a))t^k + \sum_{k=3}^{\infty} f_k(C(a, a^{-1}))t^k$$

and therefore

$$\frac{1}{2m} \sum_{k=2}^{\infty} t^k = (2m-1)f_2(C(a, a))t^2 + 2m \sum_{k=3}^{\infty} f_k(C(a, a))t^k,$$

from which follows

$$\frac{t^2}{2m(1-t)} = 2mg_{C(a, a)}(t) - \frac{t^2}{2m(2m-1)} - \sum_{k=3}^{\infty} \frac{t^k}{2m(2m-1)^{k-1}}.$$

Hence,

$$g_{C(a, a)}(t) = \frac{t^2}{4m^2(1-t)} + \frac{t^2}{4m^2(2m-1)} + \frac{t^3}{2m(2m-1)(2m-1-t)},$$

and therefore

$$g_{C(a, a^{-1})}(t) = \frac{t^2}{4m^2(1-t)} - \frac{t^2}{4m^2} - \frac{t^3}{2m(2m-1-t)}.$$

Applying Corollary 2.1, from the last two equalities we deduce

$$\mu_0(C(a, a)) = -\lim_{t \rightarrow 1} (t-1)g_{C(a, a)}(t) = \frac{1}{4m^2}$$

as well as  $\mu_0(C(a, a^{-1})) = \frac{1}{4m^2}$ . □

Lemma 4.1 can be easily generalized to the case of an arbitrary double-based cone in  $F$ .

**Theorem 4.2.** *Let  $R = C(u, v)$  be a double-based cone with handles  $u, v$  in  $F$  such that  $u = u_0 \circ a$ ,  $v = b \circ v_0$ , where  $u_0, v_0 \in F$  and  $a, b \in \Sigma$ . Then*

1.  $g_R(t) = g_{C(a, b)}(t) \cdot \lambda^*(u_0) \cdot \lambda^*(v_0)$ ;
2.  $\mu_0(R) = \frac{\lambda^*(u_0) \cdot \lambda^*(v_0)}{4m^2}$ .

*Proof.* The proof follows immediately from Lemma 4.1 and definitions of generating function and  $\lambda$ -measure. □

It remains to show how one can compute both generating function and Cesaro measure of a thick monoid.

**Theorem 4.3.** *Let  $\mathcal{A}$  be a special automaton over  $F$  such that  $\mathcal{A}_2$  and  $\mathcal{A}_3$  are automata from the decomposition (3.1), (3.2) of Lemma 3.3. Suppose  $T = L(\mathcal{A}_2)$  is a thick monoid and  $z_3$  is of type  $x$ . Then*

- 1)  $\bigsqcup_{i=1}^k T \cdot w_i = \bigsqcup_{j=1}^l C(\mathcal{A}, x) \cdot v_j$  with  $C(\mathcal{A}, x)$  being generalized  $x$ -cone,  $w_i, v_j \in F$  and  $k, l < \infty$ ;
- 2)  $g_T(t)$  and  $\mu_0(T)$  can be computed effectively by  $\mathcal{A}_3$ .

*Proof.* Let  $\bar{T}$  be a prefix closure of  $T$ . Then

$$(4.1) \quad \bar{T} = \sqcup_{w_i \in W} T \cdot w_i$$

for a set  $W$  such that  $W = \{w_i \in F : \text{there is a simple path } p_i \text{ in } \mathcal{A}_3 \text{ such that } p_i \text{ starts at } s \in S, \text{ ends at } z_3 \text{ and } w_i^{-1} \text{ is a label of } p_i\}$ . Clearly,  $W$  is finite. On the other hand,

$$(4.2) \quad \bar{T} = \sqcup_{v_j \in V} C(A, x) \cdot v_j$$

for some finite set  $V$  of words in  $F$ , defined by  $\mathcal{A}_3$ . Then claim 1) follows from (4.1) and (4.2), while claim 2) follows from 1) and Lemma 3.7.  $\square$

### Acknowledgments

The second author was supported by following grants: results of section 4 and Theorem 4.2 in particular was supported by RNF grant 14-11-00085 and the rest of the paper was written with a support of RFFI grant 14-01-00068.

### REFERENCES

- [1] Ya. S. Averina and E. V. Frenkel, On strictly sparse subsets of a free group, *Sib. elektron. Mat. Izv.*, **2** (2005) 1–13, <http://semr.math.nsc.ru>.
- [2] A. V. Borovik, A. G. Myasnikov and V. N. Remeslennikov, Multiplicative measures on free groups, *Internat. J. Algebra Comput.*, **13** no. 6 (2003) 705 – 731.
- [3] E. Yu. Daniyarova, A. G. Myasnikov and V. N. Remeslennikov, Dimension in universal algebraic geometry, (Russian), translated from *Dokl. Akad. Nauk*, **457** no. 3 (2014) 265–267.
- [4] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson and W. Thurston, *Word Processing in Groups*, Jones and Bartlett, Boston, 1992.
- [5] P. Flajolet and R. Sedgwick, *Analytic Combinatorics: Functional Equations*, Rational and Algebraic Functions. In: INRIA, **RR 4103** 2001.
- [6] E. Frenkel, A. G. Myasnikov and V. N. Remeslennikov, *Regular sets and counting in free groups*, Combinatorial and Geometric Group Theory, Trends in Mathematics, Birkhauser Verlag, Basel/Switzerland, 2010 93–118.
- [7] E. Frenkel, A. G. Myasnikov and V. N. Remeslennikov, Amalgamated products of groups: measures of random normal forms, *Fundam. Prikl. Mat.* **16** no. 8 (2010) 189–221.
- [8] E. Frenkel and V. N. Remeslennikov, Double cosets in free groups, *Internat. J. Algebra Comput.*, **23** no. 5 (2013) 1225–1241.
- [9] E. Frenkel and V. N. Remeslennikov, Cones and thick monoids in free groups, Materials of International Workshop 'Almaz-2', OmGTU press, Omsk, 2015 64–68.
- [10] R. Gilman, Formal languages and their application to combinatorial group theory, *Groups, Languages Algorithms*, 1–36, Contemp. Math., **378**, Amer. Math. Soc., Providence, RI, 2005.
- [11] J. G. Kemeny and J. L. Snell, *Finite Markov chains*, The University Series in Undergraduate Mathematics D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960.
- [12] M. V. Lawson, *Finite automata*, Chapman & Hall, CRC Press, 2004.
- [13] M. O. Rabin and D. Scott, Finite automata and their decision problems, *IBM J. Res. Develop.* **3** no.2 (1959) 114–125.

**Elizaveta Frenkel**

Faculty of Further Education, Moscow State University Moscow, Russia

Email: [lizzy.frenkel@gmail.com](mailto:lizzy.frenkel@gmail.com)

**Vladimir N. Remeslennikov**

Sobolev Institute of Mathematics, IM SB RAS Omsk, Russia

Email: [remesl@ofim.oscsbras.ru](mailto:remesl@ofim.oscsbras.ru)