EuroComb 2005 DMTCS proc. **AE**, 2005, 111–116

On the $\frac{3}{4}$-Conjecture for Fix-Free Codes

Christian Deppe and Holger Schnettler

Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D-33501 Bielefeld
Germany

In this paper we concern ourself with the question, whether there exists a fix-free code for a given sequence of codeword lengths. We focus mostly on results which shows the $\frac{3}{4}$-conjecture for special kinds of lengths sequences.

Keywords: Fix-free Codes, Kraft inequality, $\frac{3}{4}$-Conjecture

Contents

1 Introduction

A *fix-free code* is a code, which is prefix-free and suffix-free, i.e. any codeword of a fix-free code is neither a prefix, nor a suffix of another codeword. Fix-free codes were first introduced by Schützenberg (4) and Gilbert and Moore (5), where they were called *never-self-synchronizing* codes. Ahlswede, Balkenhol and Khachatrian propose in (6) the conjecture that a Kraftsum of a lengths sequence smaller than or equal to $\frac{3}{4}$, imply the existence of a fix-free code with codeword lengths of the sequence. This is known as the $\frac{3}{4}$-conjecture for fix-free codes. Harada and Kobayashi generalized in (7) all results of (6) for the case of q-ary alphabets and infinite codes.

Over the last years many attempts were done to prove the $\frac{3}{4}$-conjecture either for the general case of a q-ary alphabet or at least for the special case of a binary alphabet. In this paper we focus mostly on results which shows the $\frac{3}{4}$-conjecture for special kinds of lengths sequences.

The $\frac{3}{4}$-conjecture holds for finite sequences, if the numbers of codewords on each level is bounded by a term which depends on q and the smallest codeword length which occurs in the lengths sequence. This

theorem was first shown by Kukorelly and Zeger in (10) for the binary case. We generalize this theorem to q-ary alphabets.

If the Kraftsum of the first level which occurs in a lengths sequence together with the Kraftsum of the following level is bigger than $\frac{1}{2}$, then from Yekanins theorem (8) follows, that the $\frac{3}{4}$-conjecture holds. Yekanins theorem is only for the binary case. We give a generalization of the theorem. For the proof of the theorem and its generalization, we introduce π-systems, which are special kinds of fix-free codes with Kraftsum $\lceil \frac{q}{2} \rceil q^{-1}$. We show, that π-systems with only two neighbouring levels and $L \cdot \lceil \frac{q}{2} \rceil$ codewords on the first level exist, if and only if there exists a $\lceil \frac{q}{2} \rceil$-regular subgraph of the directed de Bruijn graph $\mathcal{B}_q(n)$ with n edges over a q-ary alphabet with L vertices. Furthermore we show that arbitrary one level π-systems exist. Since there exist cycles of arbitrary length in $\mathcal{B}_2(n)$, we obtain Yekhanin's original theorem with the π-system extension theorem. However, in the generalization of Yekhanin's theorem to the q-ary case, an extra condition for the existence of $\lceil \frac{q}{2} \rceil$-regular subgraph in $\mathcal{B}_q(n)$ occurs.

The last part is about the binary version of the $\frac{3}{4}$-conjecture. We obtain some new results for the binary case of the $\frac{3}{4}$-conjecture with the help of quaternary fix-free codes.

2 The $\frac{3}{4}$-conjecture for q-ary fix-free codes

This section is about the cases, where the $\frac{3}{4}$-conjecture can be shown for an arbitrary finite alphabet $\mathcal{A}$. We give a generalization of a theorem from Kukorelly and Zeger (10), which was shown for the binary case originally. This theorem shows, that the $\frac{3}{4}$-conjecture holds for finite codes, if the number of codewords on each level, expect the maximal level, is bounded by a term which depends on the minimal level.

We write a sequence $(\alpha_l)_{l\in\mathbb{N}}$ of nonnegative integers fits to a code $\mathcal{C} \subseteq \mathcal{A}^*$ if $|\mathcal{C} \cap \mathcal{A}^l| = \alpha_l$ for all $l \in \mathbb{N}$.

Theorem 1 *Let $|\mathcal{A}| = q \geq 2$, $(\alpha_l)_{l\in\mathbb{N}}$ be a sequence of nonnegative integers with $\sum_{l=l_{min}}^{l_{max}} \alpha_l q^{-l} \leq \frac{3}{4}$ and*
$l_{min} := \min\{l\,|\alpha_l \geq 0\}$,
$l_{max} := sup\,\{l\,|\alpha_l \geq\} \leq \infty$. If $l_{min} \geq 2$, $l_{max} < \infty$ and $\alpha_l \leq q^{l_{min}-2} \lfloor \frac{q}{2} \rfloor^2 \lceil \frac{q}{2} \rceil^{l-l_{min}}$ for all $l \neq l_{max}$, then there exists a fix-free Code $\mathcal{C} \subseteq \mathcal{A}^*$ which fits to $(\alpha_l)_{l\in\mathbb{N}}$.

3 Fix-free codes obtained from π-systems

We give a generalization of a theorem of Yekhanin (8), which shows that the $\frac{3}{4}$-conjecture holds for binary codes if the Kraftsum of the first level which occurs in the code together with it neighboring level is bigger than $\frac{1}{2}$.

For an arbitrary set $\mathcal{C} \subseteq \mathcal{A}^*$ the *prefix-, suffix- and bifix-shadow* of $\mathcal{C}$ on the n-th level are defined as:

$$\begin{array}{rcll} \Delta_P^n(\mathcal{C}) & := & \bigcup_{l=0}^{n} (\mathcal{C} \cap \mathcal{A}^l)\mathcal{A}^{n-l} & \subseteq \mathcal{A}^n\,, \\ \Delta_S^n(\mathcal{C}) & := & \bigcup_{l=0}^{n} \mathcal{A}^{n-l}(\mathcal{C} \cap \mathcal{A}^l) & \subseteq \mathcal{A}^n\,, \\ \Delta_B^n(\mathcal{C}) & := & \Delta_P^n(\mathcal{C}) \cup \Delta_S^n(\mathcal{C}) & \subseteq \mathcal{A}^n\,. \end{array}$$

For proving the theorem, Yekhanin introduced in (8) a special kind of fix-free codes, which he called π-systems:

Definition 1 *Let $|\mathcal{A}| = 2$, we say $\mathcal{D} \subseteq \bigcup_{l=1}^{n} \mathcal{A}^l$ is a π_2-system if $\mathcal{D}$ is fix-free with Kraftsum $\frac{1}{2}$ and*

$$|\Delta_S^n(\mathcal{D})| = |\Delta_P^n(\mathcal{D})| = |\mathcal{A}^{-1}\Delta_P^n(\mathcal{D})| = |\Delta_S^n(\mathcal{D})\mathcal{A}^{-1}| \tag{1}$$

To prove a generalization for arbitrary finite alphabets, we give a more general definition of π-systems.

Definition 2

Let $|\mathcal{A}| = q \geq 2$, $1 \leq k \leq q$ and $n \in \mathbb{N}$. We call a set $\mathcal{D} \subseteq \bigcup_{l=1}^{n} \mathcal{A}^l$ a $\pi_q(n;k)$-system if $\mathcal{D}$ is fix-free, and there exists a partition of $\mathcal{D}$ into k sets $\mathcal{D}_1, \ldots, \mathcal{D}_k$ for which the following three equivalent properties holds.

(1): For all $1 \leq i \leq k$ holds:

$$\begin{aligned} q^{n-1} &= |\Delta_P^n(\mathcal{D}_i)| = |\mathcal{A}^{-1}\Delta_P^n(\mathcal{D}_i)| \\ &= |\Delta_S^n(\mathcal{D}_i)| = |\Delta_S^n(\mathcal{D}_i)\mathcal{A}^{-1}| \end{aligned}$$

(2): $S(\mathcal{D}) = \frac{k}{q}$ and for all i with $1 \leq i \leq k$ holds:

$$|\Delta_P^n(\mathcal{D}_i)| = |\mathcal{A}^{-1}\Delta_P^n(\mathcal{D}_i)| \text{ and } |\Delta_S^n(\mathcal{D}_i)| = |\Delta_S^n(\mathcal{D}_i)\mathcal{A}^{-1}|$$

(3): For all $1 \leq i \leq k$ the set $\mathcal{A}^{-1}\mathcal{D}_i$ is maximal prefix-free, $\mathcal{D}_i\mathcal{A}^{-1}$ is maximal suffix-free and $|\mathcal{A}^{-1}\mathcal{D}_i| = |\mathcal{D}_i\mathcal{A}^{-1}| = |\mathcal{D}_i|$.

The sets $\mathcal{D}_1, \ldots, \mathcal{D}_k$ are called a π-partition of $\mathcal{D}$

For $\alpha_1, \ldots, \alpha_n \in \mathbb{N}$ we call a $\pi_q(n;k)$-system $\mathcal{D}$ a $\pi_q(\alpha_1, \ldots, \alpha_n; k)$-system if $\left|\mathcal{D} \cap \mathcal{A}^l\right| = \alpha_l$ for all $1 \leq l \leq n$.

(1)-(3) in the definition are all equivalent.
For $1 \leq k < q$ let

$$\gamma_k := \begin{cases} \frac{1}{2} + \frac{k}{2q} & \text{for } 1 \leq k \leq \left\lfloor \frac{q}{2} \right\rfloor \\ \left(\frac{q-k}{q}\right)^2 + \frac{k}{q} & \text{for } \left\lfloor \frac{q}{2} \right\rfloor < k < q \,. \end{cases}$$

Especially we have $\gamma_{\lceil \frac{q}{2} \rceil} \geq \frac{3}{4}$. We obtain the following theorem for fix-free extensions of π-systems:

Theorem 2 (π-system extension Theorem) *Let $|\mathcal{A}| = q \geq 2$, $1 \leq k < q$, $(\alpha_l)_{l \in \mathbb{N}}$ be a sequence of nonnegative integers with $\sum_{l=1}^{\infty} \alpha_l q^{-l} \leq \gamma_k$ and $n \in \mathbb{N}$,*
$1 \leq \beta \leq \alpha_n$ such that $\beta q^{-n} + \sum_{l=1}^{n-1} \alpha_l q^{-l} = \frac{k}{q}$. Then for every $\pi_q(\alpha_1, \ldots, \alpha_{n-1}, \beta; k)$-system there exists a fix-free-extension which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

Let $\mathcal{A} = \{0, \ldots, q-1\}$. The directed de Bruijn graph $\mathcal{B}_q(n)$ has $\mathcal{A}^n$ as its vertex set and for every $a, b \in \mathcal{A}$, $w \in \mathcal{A}^{n-1}$ there is an edge $aw \to wb$ in $\mathcal{B}_q(n)$ which can be labelled by the word $awb \in \mathcal{A}^{n+1}$. By examining the existence of $\pi_q(n+1;k)$-systems with codewords on the n-th and $n+1$-th level but no other codeword lengths, we obtain that such a system exists if and only if there exists a k-regular subgraph in $\mathcal{B}_q(n-1)$ with the number of edges equal to the number of codewords of length n. Especially for such a $\pi_q(n+1;k)$ system the codewords of the n-th level are the edges of a k-regular subgraph of $\mathcal{B}_q(n-1)$ and the codewords of the $n+1$-level are given by $\bigcup_{i=1}^{k} \bigcup_{a \in \mathcal{A}} a\mathcal{V}^c\varphi_i(a)$, where $\mathcal{V}^c$ is the complement of the vertex set of the k-regular subgraph of $\mathcal{B}_q(n-1)$ and $\varphi_1, \ldots, \varphi_k$ are permutations of $\mathcal{A}$ with the property $\varphi_i(a) \neq \varphi_j(a)$ for $i \neq j$, $a \in \mathcal{A}$. Furthermore the codewords of a one-level $\pi_q(n)$-system are the edges of a k-factor of $\mathcal{B}_q(n-1)$ and vice versa. Thus we obtain with Theorem 2 the following generalization of Yekhanin's Theorem for arbitrary finite alphabets:

Theorem 3 *Let $|\mathcal{A}| = q \geq 2, 1 \leq k < q$ and $(\alpha_l)_{l \in \mathbb{N}}$ be a sequence of nonnegative integers with $\sum_{l=1}^{\infty} \alpha_l q^{-l} \leq \gamma_k$.*

(i) If $\frac{\alpha_n}{q^n} + \frac{\alpha_{n+1}}{q^{n+1}} \geq \frac{k}{q}$, $\alpha_n = kL$ for some $1 \leq L < q^{n-1}$ and there exists a k-regular subgraph in $\mathcal{B}_q(n-1)$ with L vertices, then there exists a fix-free code which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

(ii) If $\frac{\alpha_n}{q^n} \geq \frac{k}{q}$ then there exists a fix-free code which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

Since Lempel has shown in (11), that there are cycles of arbitrary length in $\mathcal{B}_q(n)$, we obtain for the binary case Yekhanin's original theorem.

By examining π_q-systems with more than two levels, we obtain with Theorem 2.

Theorem 4 *Let $|\mathcal{A}| = q \geq 2, 1 \leq d < q, k \leq \min\{d, q-d\}$ and $(\alpha_l)_{l \in \mathbb{N}}$ be a sequence of nonnegativ integers with $\sum_{l=1}^{\infty} \alpha_l q^{-l} \leq \gamma_k$.*

(i) Let $n \geq 2$. If $\alpha_1 = 0, \alpha_l = kd(q-d)^{l-2}$ for $2 \leq l < n$ and $\alpha_n \geq kq(q-d)^{n-2}$ then there exists a fix-free code which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

(ii) Let $n \geq 3$. If $\alpha_1 = \alpha_2 = 0, \alpha_l = kd(q-d)^{l-2} + k(q-d)d^{l-2}$ for $3 \leq l < n$ and $\alpha_n \geq kq(q-d)^{n-2} + kqd^{n-2}$ then there exists a fix-free code which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

4 The $\frac{3}{4}$-conjecture for binary fix-free codes

In this section we examine the $\frac{3}{4}$-conjecture for the special case $|\mathcal{A}| = 2$. If we identify quaternary fix-free codes with binary fix-free codes in the natural way we obtain from the theorems above that the following statements hold for the binary case:

Theorem 5 *Let $\mathcal{A} := \{0, 1\}$ and $(\alpha_l)_{l \in \mathbb{N}}$ be a sequence of nonnegative integers with $\sum_{l=1}^{\infty} \alpha_l \left(\frac{1}{2}\right)^l \leq \frac{3}{4}$.*

(i) If there exists an $n \geq 2$ such that $\alpha_2 = \alpha_{2l+1} = 0$ for all $l \in \mathbb{N}_0$, $\alpha_{2l} = 2^l$ for all $2 \leq l < n$, $\alpha_{2n} \geq 2^{n+1}$ and $\alpha_{2l} \in \mathbb{N}_0$ for all $l > n$, then there exists a fix-free code $\mathcal{C} \subseteq \mathcal{A}^+$ which fits to $(\alpha_l)_{l \in \mathbb{N}}$.

(ii) *If there exists an* $n \geq 3$ *such that* $\alpha_2 = \alpha_4 = \alpha_{2l+1} = 0$ *for all* $l \in \mathbb{N}_0$, $\alpha_{2l} = 2^{l+1}$ *for all* $2 \leq l < n$, $\alpha_{2n} \geq 2^{n+2}$ *and* $\alpha_{2l} \in \mathbb{N}_0$ *for all* $l > n$, *then there exists a fix-free code* $\mathcal{C} \subseteq \mathcal{A}^+$ *which fits to* $(\alpha_l)_{l \in \mathbb{N}}$.

(iii) *If there exists an* $n \in \mathbb{N}$ *such that* $\alpha_2 = \alpha_4 = \ldots = \alpha_{2n-2} = \alpha_{2l+1} = 0$ *for all* $l \in \mathbb{N}_0$, α_{2n} *is even,* $\frac{\alpha_{2n}}{2^{2n}} + \frac{\alpha_{2n+2}}{2^{2n+2}} \geq \frac{1}{2}$ *and there exists a 2-regular subgraph of* $\mathcal{B}_4(n-1)$ *with* $\frac{\alpha_{2n}}{2}$ *vertices, then there exists a fix-free code* $\mathcal{C} \subseteq \mathcal{A}^+$ *which fits to* $(\alpha_l)_{l \in \mathbb{N}}$.

(iv) *Let* $l_{min} := \min\{l \,|\, \alpha_l \neq 0\}$ *and* $l_{max} := \sup\{l \,|\, \alpha_l \neq 0\}$. *If* $l_{max} < \infty$, $4 \leq l_{min}$ *is even,* $\alpha_{2l+1} = 0$ *for all* $l \in \mathbb{N}_0$ *and* $\alpha_{2l} \leq 2^{\frac{l_{min}}{2}-2+l}$ *for all* $2l \neq l_{max}$, *then there exists a fix-free code* $\mathcal{C} \subseteq \mathcal{A}^+$ *which fits to* $(\alpha_l)_{l \in \mathbb{N}}$.

References

[1] L.G. Kraft, A device for quantizing, grouping and coding amplitude modulated pulses, Master's thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., 1949.

[2] B. McMillan, Two inequalities implied by unique decipherability, IRE Trans. Inform. Theory, vol. IT-2, pp. 115-116, (1956).

[3] D. Huffman, A method for construction of minimum redundancy codes, Proc. of the IRE, vol. 40, pp. 1098-1101, (1952).

[4] M. P. Schützenberg, On an application of semigroup methods to some problems in coding, IRE. Trans. Inform. Theory, vol. IT-2, pp 47-60, (1956).

[5] E. N. Gilbert and E. F. Moore, Variable-length binary encodings, Bell Syst. Tech. J., vol. 38, pp. 933-968, July (1959).

[6] R. Ahlswede, B.Balkenhol and L.Khachatrian, Some Properties of Fix-Free Codes, Proc. 1^{st} Int. Sem. on Coding Theory and Combinatorics, Thahkadzor, Armenia, pp. 20-33, (1996).

[7] K. Harada and K. Kobayashi, A Note on the Fix-Free Property, IEICE Trans. Fundamentals, vol. E82-A, no 10, pp.2121-2128, October (1999).

[8] S. Yekhanin, Sufficient Conditions of Existence of Fix-Free Codes, Proc. Int. Symp. Information Theory, Washington, D.C., p.284, June (2001).

[9] S. Yekhanin, Improved upper bound for the redundancy of fix-free codes, IEEE Tran. Inform. Theory., vol. 50, Issue 11, pp. 2815-2818, Nov. (2004)

[10] Z. Kukorelly and K. Zeger, Sufficient Condition for Existence of Binary Fix-Free Codes, submitted to IEEE Trans. Inform. Theory. October 15, (2003).

[11] A. Lempel, m-Ary closed sequences, J. Cobin. Theorey 10, pp. 253-258, (1971).

[12] S.W. Golomb, Shift Register Sequences, Aegean Park Press, Laguna Hills, CA, (1982).