Сбалансированные слова и динамические системы

А. Л. ЧЕРНЯТЬЕВ
Центр дополнительного образования
«Дистантное обучение», Москва

УДК 512+517.987

Ключевые слова: слова Штурма, сбалансированные слова, символическая динамика, комбинаторика слов.

Аннотация

Работа посвящена описанию непериодических сбалансированных слов над произвольным алфавитом. Бесконечное слово W называется сбалансированным, если для любых двух его подслов u_1, u_2 одинаковой длины количество символов одного сорта отличается в них не больше чем на 1. Сбалансированные непериодические слова над произвольным алфавитом являются естественным обобщением слов Штурма. В работе получено описание сбалансированных непериодических слов в терминах одномерной динамической системы.

Abstract

A. L. Chernyatiev, Balanced words and dynamical systems, Fundamentalnaya i prikladnaya matematika, vol. 13 (2007), no. 5, pp. 213—224.

This article is devoted to the description of all nonperiodic balanced words with n different letters. A superword W is called balanced if the numbers of equal letters in any two of its factors (subwords) u_1 and u_2 of equal length differ by at most 1. Balanced words are one of the possible generalizations of Sturmian words. We give a geometric interpretation of nonperiodic balanced sequences over an n-letter alphabet.

1. Введение

Понятие *сбалансированности* играет существенную роль при изучении комбинаторных свойств слов, а также в теории бильярдов и теории динамических систем. Для бесконечного слова $W = \{w_n\}_{n\in\mathbb{Z}}$ над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$ свойство сбалансированности (или «равномерной перемешанности») означает, что в любых двух подсловах слова W одинаковой длины количество символов одного сорта (например, символов a_1) отличается не более чем на 1. Бесконечные сбалансированные непериодические слова над алфавитом из двух символов — это в точности слова Штурма.

Основным результатом о бинарных сбалансированных словах является известная теорема эквивалентности [7,9], из которой следует, что любое бесконечное сбалансированное непериодическое слово $W = (w_n)$ порождается сдвигом

Фундаментальная и прикладная математика, 2007, том 13, № 5, с. 213—224.

на единичной окружности, т. е. может быть представлено следующим образом:

$$w_n = \begin{cases} a, & \text{если} \quad T_\alpha^n(x_0) \in U, \\ b, & \text{если} \quad T_\alpha^n(x_0) \notin U, \end{cases}$$

где $T_\alpha \colon x \to x + \alpha \mod 1$ — сдвиг единичной окружности на величину α, U — дуга длины α, x_0 — начальная точка.

Целью работы является обобщение данного результата на сбалансированные слова над алфавитом из произвольного числа символов. Для произвольного сбалансированного непериодического слова над n-буквенным алфавитом мы построим динамическую систему, которая будет порождать данное слово.

Теорема 1.1 (основная теорема). *Пусть W — сбалансированное непериодическое слово над алфавитом A. Тогда для W существует динамическая система (M, f), удовлетворяющая следующим условиям.*

1. *Пространство M гомеоморфно $\mathbb{S} \times \mathbb{Z}_m$ как топологическое пространство.*
2. *Отображение $f \colon M \to M$ есть композиция поворота на α в $\mathbb{S}$ и сдвига на 1 в $\mathbb{Z}_m$. Длина $\mathbb{S}^1$ равна m.*
3. *Каждая компонента $\mathbb{S}^1 \times \{k\}$, $k = 1, \ldots, n$, разбита на $2m$ дуг: m красных и m синих. Все красные имеют длину α, все синие — $1 - \alpha$, красные и синие дуги чередуются.*
4. *Синий цвет имеет l оттенков, красный — k оттенков, $k + l = |A|$ — число букв в алфавите. Все середины дуг данного оттенка образуют вершины правильного многоугольника («правильный 1-угольник» — это точка на окружности, «правильный 2-угольник» — пара диаметрально противоположных точек).*
5. *При переходе от компоненты $\mathbb{S}^1 \times \{k\}$ к компоненте $\mathbb{S}^1 \times \{k + 1\}$ $(\mathbb{S}^1 \times \{m+1\} = \mathbb{S}^1 \times \{1\})$ порядок расположения оттенков внутри красных и синих компонент сохраняется, а сами красные и синие дуги («рулетки») проворачиваются относительно друг друга на 1, так что преобразование f приводит к смещению на α относительно красных компонент и на $1 - \alpha$ (в обратную сторону) относительно синих.*

Отметим, что описание сбалансированных слов в арифметической форме было получено Грэхемом [4], позднée описание в форме бинарных слов Штурма было получено Хубертом [5].

2. Основные конструкции и определения

2.1. Пространство слов

Пусть A — конечный алфавит, т. е. непустое множество элементов (символов). Через A^+ обозначим множество всех конечных последовательностей символов, или *слов*. Конечное слово всегда может быть единственным образом представлено в виде $w = w_1 \ldots w_n$, где $w_i \in A$, $1 \leqslant i \leqslant n$. Число n называется *длиной* слова w и обозначается $|w|$

Множество A^+ всех конечных слов над A образует простую полугруппу, где полугрупповая операция определяется как конкатенация (приписывание). Если к множеству слов добавить элемент Λ (пустое слово), то получим свободный моноид A^* над A. Длина Λ по определению равна 0.

Слово u есть *подслово* слова w, если существуют слова $p, q \in A^+$, такие что $w = puq$.

Пусть $w = w_1 \ldots w_n$, $w_i \in A$, $i = l, \ldots, n$, — некоторое слово. Натуральное число q называется *периодом* w, если $w_i = w_i + q$ для всех $i \in [1, \ldots, n-q]$. Мы будем обозначать через p_w (или просто через p) минимальный период слова w.

Слово w называется *периодическим*, если $p \leqslant [|w|/2]$ (здесь $[x]$ обозначает целую часть x). Через $\mathbb{N}$ ($\mathbb{N}_+$) мы будем обозначать неотрицательные (положительные) целые числа. *Одностороннее* (*двустороннее*) *бесконечное слово* над алфавитом A — это отображение $w \colon \mathbb{N}_+ \to A$ (соответственно $w \colon \mathbb{Z} \to A$). Для каждого n мы полагаем $w_n = w(n)$ и обозначаем $W = w_1 w_2 \ldots$. Слово $u \in A^+$ — *конечное подслово* W, если существуют такие $i, j \in \mathbb{N}$ ($1 \leqslant i \leqslant j$), что $u = w_i \ldots w_j$. Последовательность $w[i, j] = w_i \ldots w_j$ назовём *вхождением* слова u в W.

Через $F(W)$ обозначим множество всех конечных подслов слова W, через $F_n(W)$ — множество всех конечных подслов длины n.

Назовём слова W_1 и W_2 *эквивалентными* (обозначение $W_1 \sim W_2$), если множества их конечных подслов равны, т. е. $F(W_1) = F(W_2)$.

Плотностью символа $a_i \in A$ в слове W называется предел

$$\rho(a_i) = \lim_{n \to \infty} \frac{|w_1 \ldots w_n|_{a_i}}{n},$$

где $|w_1 \ldots w_n|_{a_i}$ обозначает количество вхождений символа a_i в подслово $w_1 \ldots w_n$ (если этот предел существует).

2.2. Слова, порождаемые динамическими системами

Пусть M — компактное метрическое пространство, $U \subset M$ — его открытое подмножество, $f \colon M \to M$ — гомеоморфизм компакта в себя и $x \in M$ — начальная точка.

По последовательности итераций можно построить над бинарным алфавитом бесконечное слово $W = (w_n)$,

$$w_n = \begin{cases} a, & \text{если } \ f^{(n)}(x_0) \in U, \\ b, & \text{если } \ f^{(n)}(x_0) \notin U, \end{cases}$$

которое называется *эволюцией* точки x_0 (мы также будем говорить, что динамика (M, f) *порождает* слово W). Символическая динамика исследует взаимосвязь свойств динамической системы (M, f) и комбинаторных свойств слова W.

Для слов над алфавитом, состоящим из большего числа символов, нужно рассмотреть несколько характеристических множеств $U_1, \ldots, U_n$.

Заметим, что эволюция точки корректно определена только в случае, когда траектория точки не попадает на границу характеристических множеств $\partial U_1, \partial U_2, \ldots$. Для того чтобы рассматривать траекторию произвольной точки, мы введём понятие *существенной эволюции*.

Конечное слово v^f называется *существенной конечной эволюцией* точки x^*, если в любой окрестности точки x^* существует открытое множество V, любая точка $x \in V$ из которого обладает эволюцией v^f. Бесконечное слово W называется *существенной эволюцией* точки x^*, если любое его начальное подслово — существенная конечная эволюция точки x^*.

Под *эволюцией точки*, когда это не вызывает недоразумений, будем понимать существенную эволюцию. Отметим, что точка может иметь несколько существенных эволюций.

Предложение 2.1. *Пусть V — конечная эволюция точки x динамики (M, f, U). Тогда множество точек с конечной существенной эволюцией V замкнуто. Аналогичное утверждение верно для бесконечного слова W.*

Доказательство можно найти в [1].

Из построения следует, что если начальная точка принадлежит множеству U_i, то её эволюция начинается с символа a_i. Рассмотрим образы множеств U_i при отображениях $f^{(-1)}, f^{(-2)}, \ldots$, $n \in \mathbb{N}$. Ясно, что если точка принадлежит множеству

$$f^{(-k)}(U_{i_k}) \cap f^{-(k-1)}(U_{i_{k-1}}) \cap \ldots \cap f^{(-1)}(U_{i_1}) \cap U_{i_0},$$

то эволюция начинается со слова $a_{i_0} a_{i_1} \ldots a_{i_k}$. Количество различных существенных эволюций длины $k+1$ равно количеству разбиений множества M на непустые подмножества границами подмножеств ∂U_i и их образами при отображениях $f^{(-1)}, f^{(-2)}, \ldots, f^{(-k)}$.

2.3. Слова Штурма и теорема эквивалентности

Пусть W — бесконечное (в обе стороны) слово над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$. *Функция сложности* $T_W(n)$ определяется как количество различных подслов длины n слова W. *Слово Штурма* — это бесконечное слово с функцией сложности $T_W(n) = n + 1$, $n \geqslant 1$. Для конечного подслова $u \subset W$ обозначим через $|u|$ длину слова u и через $|u|_{a_i}$ количество вхождений в подслово u символов a_i. Ясно, что $|u| = \sum\limits_{a_i \in A} |u|_{a_i}$.

Бесконечное слово W над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$ называется m-сбалансированным, если для любых двух подслов $u, v \subset W$ и любого символа $a_i \in A$ выполняется неравенство

$$\big| |u|_{a_i} - |v|_{a_i} \big| \leqslant m.$$

В случае, когда слово 1-сбалансированно, будем просто говорить, что оно сбалансированно.

Предложение 2.2 [6]. *Пусть для бесконечного слова W существует такое натуральное k, что для любых двух подслов u, v длины k и любого символа a_i выполняется $|u|_{a_i} = |v|_{a_i}$. Тогда слово W периодично.*

Фактически, сбалансированность означает минимальное отклонение от периодичности, так же как и минимальный рост.

Хорошо известна следующая теорема [7, 9].

Теорема 2.3 (теорема эквивалентности). *Следующие условия на слово W эквивалентны*:

1) *слово W имеет функцию сложности $T_W(n) = n + 1$;*
2) *слово W сбалансированно и непериодично*;
3) *слово W порождается системой $(\mathbb{S}^1, U, T_\alpha)$ с иррациональным α.*

Наша цель — обобщение этого результата на случай алфавита из n символов.

3. Слова, порождаемые поворотом окружности

В этом разделе будут сформулированы общие результаты о словах, порождаемых произвольным сдвигом окружности. Пусть $M = \mathbb{S}^1$ — единичная окружность, $U \subset \mathbb{S}^1$ — интервал окружности, $T_\alpha \colon x \to x + \alpha$ — поворот окружности на величину α. Если α иррационально, то траектория произвольной точки всюду плотна (лемма Вейля—Кронекера).

Отметим также, что так как отображение T равномерно (см. [2]), каждый символ в слове, порождённом системой $(\mathbb{S}^1, T_\alpha, U_1, U_2, \ldots, U_n)$ обладает плотностью, причём $\rho(a_i) = \mu(U_i)$, где μ — мера Лебега.

Предложение 3.1. *Для любых двух точек слова, порождаемые этой динамикой (т. е. существенные эволюции) эквивалентны, т. е. множества их конечных подслов совпадают.*

Доказательство можно найти в [1].

Предложение 3.2. *Пусть $W = (w_n)$ — слово, которое порождается динамикой $(\mathbb{S}^1, T_\alpha, U, x_0)$, и пусть $\beta = \alpha/q + r/q$, $U_q = \{x \mid qx \in U\}$, $y_0 = x_0/q$, где q, r — целые. Тогда эволюция точки x_0, порождённая динамикой $(\mathbb{S}^1, T_\beta, U_q, y_0)$, будет совпадать с W.*

Доказательство. Непосредственно проверяется, что $T_\alpha^n(x_0) \in U$ тогда и только тогда, когда $T_\beta^n(y_0) \in U_q$. $\square$

Такую операцию над динамикой мы будем называть *q-размножением*. Ясно также, что U — симметричное подмножество в том смысле, что оно является инвариантным относительно сдвига на $1/q$.

Лемма 3.3. *Пусть два слова $W_1 = (w^1{}_n)$ и $W_2 = (w^2{}_n)$ порождены динамиками $(\mathbb{S}^1, T_\alpha, U, x_0)$ и $(\mathbb{S}^1, T_\beta, V, y_0)$ соответственно и для некоторого $k \in \mathbb{Z}$ слова совпадают в каждой k-й позиции, т. е. $w^1{}_{kn} = w^2{}_{kn}$ для всех n. Тогда существуют целые p и q, такие что множества U и V при соответствующих размножениях совпадают с точностью до поворота, т. е. $U_p = T_\delta(V_q)$ для некоторого δ.*

Доказательство. Покажем сначала, что 1, α и β линейно зависимы над $\mathbb{Q}$, т. е. существуют целые m, n, k, такие что $m\alpha+n\beta = k$. Предположим противное. Пусть α, β и 1 линейно независимы над $\mathbb{Q}$. Рассмотрим сдвиг на двумерном торе $\mathbb{T}^2$

$$h\colon (x, y) \to (x + k\alpha, y + k\beta) \mod 1.$$

Согласно лемме Вейля [14] траектория любой точки при таком сдвиге всюду плотна. Тогда существует такое l, что $h^l(x_0, y_0) \in U \times (\mathbb{S}^1/V)$, но это означает, что в lk-м символе слова различаются. Противоречие. Значит, $\alpha = p/q\,\beta + r/s$, где p, q, r, s — целые. Покажем, что данные p и q искомые. Положим $\alpha_p = \alpha/p = \beta/q + r/sp$, $\beta_q = k\beta/q$, тогда

$$kps\alpha_p = kps\beta_q = \gamma \mod 1.$$

Совместим начальные точки сдвигом, пусть величина сдвига равна δ. Рассмотрим сдвиг на γ. Поскольку γ иррационально, то траектория любой точки при сдвиге на γ всюду плотна. Поэтому если $U_p \neq V_q$, то существует момент, когда траектория точки попадет в $U_p \triangle V_q$, т. е. слова в данной позиции различаются. $\square$

Следствие 3.4. *Пусть слово W порождается системой $(\mathbb{S}^1, T_\alpha, U, x_0)$ и является словом Штурма. Тогда $U = \Delta_1 \cup \ldots \cup \Delta_k$, где Δ_i — интервал длины λ, и характеристическое множество U периодично, т. е. инвариантно относительно сдвига на $1/k$. Более того, α и λ связаны одним из двух возможных соотношений: $\alpha = \lambda + l/k$ или $\alpha = -\lambda + l/k$, где l — некоторое целое.*

Пусть теперь окружность разбита на k дуг:

$$\{I_1 = [\beta_1, \beta_2),\ I_2 = [\beta_2, \beta_3), \ldots,\ I_n = [\beta_n, \beta_{n+1})\}, \quad \beta_{n+1} = \beta_1.$$

Для произвольного символа $a \in A$ мы будем обозначать характеристическое множество I_a. Рассмотрим сдвиг окружности на иррациональную величину α и построим соответствующее слово W над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$.

Предложение 3.5. *Конечное слово $u = x_0x_1x_2\ldots x_k$ является подсловом W тогда и только тогда, когда $\bigcap\limits_{i=0}^{k} T_\alpha^{-i}(I_{x_i}) \neq \varnothing$.*

Доказательство. Это утверждение является следствием леммы Кронекера. Действительно, если пересечение непусто, то из того, что траектория плотна всюду, следует, что траектория попадёт в пересечение (конечное пересечение полуинтервалов — интервал, или отрезок, или полуинтервал). Фрагмент эволюции с момента попадания траектории в пересечение с длиной, равной $k + 1$, будет совпадать со словом u. $\square$

4. Сбалансированные слова над алфавитом из n символов

4.1. Построение динамической системы

Пусть $W = (w_n)_{n\in\mathbb{Z}}$ — бесконечное в обе стороны непериодическое сбалансированное слово над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$. Построим слова $W_1, W_2, \ldots, W_n$ над бинарными алфавитами $A_1 = \{a_1, \bar{a}_1\}$, $A_2 = \{a_2, \bar{a}_2\}$,..., $A_n = \{a_n, \bar{a}_n\}$ следующим образом:

$$W_i = (w_n^i)_{n\in\mathbb{Z}}, \quad w_n^i = \begin{cases} a_i, & \text{если } \ w_n = a_i, \\ \bar{a}_i, & \text{если } \ w_n \neq a_i. \end{cases}$$

Предложение 4.1. *Слово W_i являются сбалансированным над алфавитом A_i для любого i.*

Доказательство. Из сбалансированности W следует сбалансированность W_i по символу a_i. Действительно, если $u = w_j w_{j+1} \ldots w_k \subset W$ и $u' = w_j^i w_{j+1}^i \ldots w_k^i$ — соответствующее подслово W^i, то $|u|_{a_i} = |u'|_{a_i}$. Сбалансированность по одному символу для слов над бинарным алфавитом влечёт сбалансированность в целом. □

Согласно теореме эквивалентности 2.3 для каждого слова W_i существует динамика $(S^1, T_{\alpha_i}, \Delta_i, x_i)$, $i = 1, 2, \ldots, n$, которая порождает данное слово. Заметим, что для любого k среди символов $w_k^1, w_k^2, \ldots, w_k^n$ должен быть ровно один символ из алфавита $A = \{a_1, a_2, a_3, \ldots\}$ и ровно $n-1$ символ из алфавита $\bar{A} = \{\bar{a}_1, \bar{a}_2, \bar{a}_3, \ldots\}$.

Предложение 4.2. *Элементы 1, α_i, α_j линейно зависимы над $\mathbb{Q}$ для любых i и j.*

Доказательство. Предположим противное, т. е. что 1, α_i, α_j линейно независимы над $\mathbb{Q}$. Рассмотрим сдвиг на вектор (α_i, α_j) на двумерном торе $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$. Согласно лемме Кронекера—Вейля траектория любой точки при таком сдвиге плотна. Тогда плотна и траектория точки (x_i, x_j), следовательно, на каком-то шаге она попадёт в множество $U = \{(x, y \mid x \in \Delta_i,\ y \in \Delta_j)\}$. Это означает, что на n-м месте в слове W одновременно стоят символы a_i и a_j, что невозможно. Противоречие. □

Мы можем представить динамическую систему, которая порождает слово W, следующим образом:

$$\big(\mathbb{T}^n, T_\gamma, U_1, U_2, \ldots, U_n, (x_1, x_2, \ldots, x_n)\big),$$

где $T_\gamma\colon \mathbb{T}^n \to \mathbb{T}^n$ — сдвиг на n-мерном торе на вектор $\gamma = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, а $U_i = \{(y_1, y_2, \ldots, y_n) \in \mathbb{T}^n \mid y_i \in \Delta_i\}$. Однако ясно, что в силу попарной рациональной зависимости над $\mathbb{Q}$ динамика реализуется на меньшем множестве. Обозначим через M замыкание траектории точки $x = (x_1, x_2, \ldots, x_n)$ при действии T_γ.

Предложение 4.3. *M гомеоморфно множеству $\mathbb{S}^1 \times \{1, 2, \ldots N\}$.*

Доказательство. Так как для любых $1 \leqslant i, j \leqslant n$ элементы 1, α_i, α_j линейно зависимы над $\mathbb{Q}$, то существует такое α (например, α_1), что

$$\alpha_1 = \frac{r_1}{s_1}\alpha + \frac{p_1}{q_1}, \quad \alpha_2 = \frac{r_2}{s_2}\alpha + \frac{p_2}{q_2}, \ldots, \quad \alpha_n = \frac{r_n}{s_n}\alpha + \frac{p_n}{q_n},$$

где $r_i, s_i, p_i, q_i \in \mathbb{Z}$. Тогда

$$\gamma = \alpha\left(\frac{r_1}{s_1}, \ldots, \frac{r_n}{s_n}\right) + \left(\frac{p_1}{q_1}, \ldots, \frac{p_n}{q_n}\right) = \alpha\vec{\rho}_1 + \vec{\rho}_2,$$

где $\vec{\rho}_1$, $\vec{\rho}_2$ — рациональные векторы. Так как сдвиги между собой коммутируют, то замыкание траектории есть прямая, направленная вдоль вектора $\vec{\rho}_1$, и все её образы при сдвиге на вектора вида $m\vec{\rho}_2$ ($m \in \mathbb{Z}$). Поскольку $\vec{\rho}_1$ — рациональный вектор, прямая на торе будет замкнута, т. е. гомеоморфна окружности, а поскольку $\vec{\rho}_2$ рационален, различных образов этой окружности при сдвигах на $m\vec{\rho}_2$ будет конечное число. □

Теперь видно, что динамика реализуется на множестве $M = \mathbb{S} \times \{1, 2, \ldots, N\}$, а отображение имеет вид $f\colon (x, k) \to (x + \alpha,\ k + 1 \mod N)$. Будем теперь понимать под U_i характеристическое множество для символа a_i, которое лежит на замыкании траектории, и пусть U_i^k обозначает часть характеристического множества, которое лежит на k-й компоненте связности (окружности) M. Из того, что траектория всюду плотна на M, следует, что каждая точка принадлежит какому-то множеству U_i^k. Для составления общей картины нам необходимо описание характеристических множеств.

Предложение 4.4. *Каждое множество U_i^l является объединением k_i дуг, периодически разбивающих окружность* (*т. е. инвариантное относительно сдвига на* $1/k_i$), *причём k_i не зависит от номера компоненты l.*

Доказательство. Рассмотрим слово, порождаемое на l-й компоненте и характеристическом множестве U_i^l, которое мы построим следующим образом. Возьмём произвольную точку траектории x_0, которая попадает на l-ю компоненту. Через N шагов преобразования f она снова попадёт на эту же окружность, сместившись на вектор $N\alpha$, следующая итерация — такое же смещение на $N\alpha$ и т. д. При попадании в U_i^l записываем a_i, при непопадании — $\bar{a}_i$. Легко видеть, что это слово совпадает с некоторым словом вида $\{v_k = w_{kN}^i\}$, составленным из N-х символов слова W_i, т. е. совпадает со словом, порождённым динамикой $(\mathbb{S}^1, T_N\alpha_i, \Delta)$. Из предложения 4.1 и следствия 3.4 следует, что U_i^k есть объединение k_i дуг, периодически разбивающих окружность. □

Будем обозначать длины дуг, которые образуют U_i, через λ_i.

Предложение 4.5. *Для каждого i выполняется $\alpha = \lambda_i + l_i/k_i$ или $\alpha = -\lambda_i + l_i/k_i$.*

Доказательство. Утверждение является следствием леммы 3.3. □

Лемма 3.3 утверждает, что если поворот окружности порождает слово Штурма, то этот поворот может быть получен путём q-размножения стандартной динамической системы.

Теперь разобьём характеристические множества на два класса: те U_i, для которых $\alpha = \lambda_i + l_i/k_i$, для удобства окрасим в красный цвет (различать различные множества будем по оттенкам), остальные аналогично окрасим в синий цвет. Заметим, что из соотношения следует, что если из левого конца красной дуги сдвинуться на α, то попадём в правый конец дуги того же цвета (и оттенка), а если из правого конца дуги синего цвета сместиться на α, то попадём в левый конец дуги того же цвета (левый и правый концы определяются согласно ориентации окружности).

Предложение 4.6. *На одной компоненте дуги, имеющие общую границу, имеют разный цвет.*

Доказательство. Предположим противное. Пусть дуги (x_1, x_2), (x_2, x_3) красные. Предположим также, что другая соседняя с дугой (x_1, x_2) дуга (x_0, x_1) синяя (случай двух синих дуг рассматривается аналогично). Рассмотрим образы точек x_1, x_2 при сдвиге на α. Так как $\alpha = \lambda + l_i/k_i$ для дуг красного цвета, то $x_1 + \alpha$ — конец дуги того же цвета, что и (x_1, x_2), а $x_2 + \alpha$ — конец дуги того же цвета, что и (x_2, x_3). Так как длины красных дуг равны λ, то точка $x_1 + \alpha$ является левым концом дуги того же цвета, что и (x_2, x_3). Но x_1 также является правым концом дуги синего цвета, следовательно, $x_1 + \alpha$ — левый конец дуги этого же цвета, значит, x_1 одновременно является левым концом двух разных дуг. Противоречие. □

Предложение 4.7. *Все дуги одного цвета имеют одинаковую длину.*

Доказательство. Рассмотрим три идущие подряд интервала разбиения (x_0, x_1), (x_1, x_2), (x_2, x_3), такие что первый и третий имеют красный цвет, а второй синий. Тогда $x_0 + \alpha$ является концом интервала того же цвета, что и (x_0, x_1), $x_2 + \alpha$ является концом интервала того же цвета, что и (x_0, x_1), а также концом интервала того же цвета, что и (x_1, x_2). Между этими интервалами красного цвета располагается интервал синего цвета, имеющий такую же длину, что и (x_1, x_2). Получаем два интервала синего цвета, между которыми находится один интервал красного цвета, имеющие одинаковую длину. Из произвольности выбора следует, что все интервалы синего цвета имеют одинаковую длину. Аналогично доказывается утверждение о красных интервалах. □

Итак, мы получили, что на одной компоненте разбиение на характеристические множества, или раскраска, выглядит следующим образом:

1) каждое характеристическое множество U_i представляет собой объединение дуг одной длины, на каждой компоненте середины дуг образуют правильный k_i-угольник (т. е. инвариантны при сдвиге на $1/k_i$);
2) дуги разбиты на два цвета, красный и синий. Дуги одного цвета имеют одинаковую длину;
3) на одной компоненте (окружности) красные и синие дуги чередуются.

Осталось понять, как меняется раскраска при переходе от i-й к $(i+1)$-й компоненте. Пусть на каждой компоненте m красных, а значит, и m синих дуг.

Предложение 4.8. *При переходе от i-й к $(i+1)$-й компоненте дуги красного цвета остаются на месте, дуги синего цвета поворачиваются на $1/m$. Количество компонент N кратно m.*

Доказательство. Пусть a_i соответствует красному характеристическому множеству. При «склейке» всех цветов, кроме i-го, мы получим систему, которая порождает слово Штурма W_i. С другой стороны, то же самое слово порождается на любой из компонент M_k системой $(M_k, T'{}_\alpha, U_i^k)$, где T'_α — поворот на α на k-й компоненте. Поскольку траектория любой точки для каждой из этих систем всюду плотна, то $(x,k) \in U_i^k$ тогда и только тогда, когда $(x, k+1) \in U_i^{k+1}$. Переход к синему варианту происходит рассмотрением сдвига в отрицательном направлении: сдвиг на α в отрицательном направлении соответствует сдвигу на $1/m - \alpha$ в положительном направлении. □

4.2. Основная теорема о непериодических сбалансированных словах над произвольным алфавитом

Для удобства формулировки основной теоремы будем считать, что окружность, на которой реализуется динамика, имеет длину m, длины всех интервалов характеристических множеств также автоматически умножаются на m.

Теорема 4.9 (основная теорема). *Пусть W — сбалансированное непериодическое слово над алфавитом A. Тогда для W существует динамическая система (M, f), удовлетворяющая следующим условиям.*

1. *Пространство M гомеоморфно $\mathbb{S} \times \mathbb{Z}_m$ как топологическое пространство.*
2. *Отображение $f\colon M \to M$ есть композиция поворота на α в $\mathbb{S}$ и сдвига на 1 в $\mathbb{Z}_m$. Длина $\mathbb{S}^1$ равна m.*
3. *Каждая компонента $\mathbb{S}^1 \times \{k\}$, $k = 1, \ldots, n$, разбита на $2m$ дуг: m красных и m синих. Все красные имеют длину α, все синие — $1-\alpha$, красные и синие дуги чередуются.*
4. *Синий цвет имеет l оттенков, красный — k оттенков, $k + l = |A|$ — число букв в алфавите. Все середины дуг данного оттенка образуют вершины правильного многоугольника («правильный 1-угольник» — это точка на окружности, «правильный 2-угольник» — пара диаметрально противоположных точек).*
5. *При переходе от компоненты $\mathbb{S}^1 \times \{k\}$ к компоненте $\mathbb{S}^1 \times \{k+1\}$ ($\mathbb{S}^1 \times \{m+1\} = \mathbb{S}^1 \times \{1\}$) порядок расположения оттенков внутри красных и синих компонент сохраняется, а сами красные и синие дуги («рулетки») проворачиваются относительно друг друга на 1, так что преобразование f приводит к смещению на α относительно красных компонент и на $1-\alpha$ (в обратную сторону) относительно синих.*

Замечание. Из вида динамической системы следует, что при «склейке» всех оттенков красного в один цвет, а всех оттенков синего в другой получится бинарное слово Штурма. При «склейке» всех дуг оттенков одного цвета (например, красного) получится также сбалансированное слово над алфавитом из $l+1$ символа.

Следствие 4.10. *Пусть W — непериодическое сбалансированное слово над алфавитом из трёх или более букв. Тогда найдутся два символа одинаковой плотности.*

Доказательство. Поскольку все середины дуг одного цвета находятся в вершинах правильного m-угольника и оттенки внутри цвета также являются вершинами правильных k_i-угольников, то по теореме Ньюмана—Знама о покрытии множества целых арифметическими прогрессиями (см. [10]) найдутся два одинаковых многоугольника. Это означает, что если $l > 1$ ($k > 1$), то среди оттенков синего (красного) найдутся два, характеристические множества которых есть объединения одинакового числа интервалов, а значит, они имеют одинаковую меру Лебега. Следовательно, соответствующие символы обладают одинаковой плотностью. □

4.3. Замечания о периодических сбалансированных словах над произвольным алфавитом

В рассуждениях о непериодических сбалансированных словах мы пользовались тем, что траектория для каждого слова W_i всюду плотна. В периодическом случае эти рассуждения не проходят. Однако из имеющейся динамической системы можно сконструировать периодическое сбалансированное слово.

Предложение 4.11. *Пусть динамическая система удовлетворяет условиям основной теоремы с той лишь разницей, что α — рациональное число. Тогда соответствующее слово является сбалансированным периодическим.*

Доказательство. Покажем сбалансированность по произвольному символу a_i. В нашей динамической системе «склеим» все цвета, отличные от цвета, соответствующего a_i. Нетрудно убедиться, что в этом случае система будет эквивалентна системе, подобной системе, порождающей слова Штурма, но с рациональной величиной сдвига. По теореме 4.9 это слово является периодическим сбалансированным. □

Возникает естественный вопрос: любое ли периодическое слово может быть получено таким путём? Ответ на этот вопрос отрицательный. Действительно, по следствию 4.10 в случае алфавитов, содержащих более чем два символа, для сбалансированных слов, порождающихся такой динамической системой, обязательно существуют символы, имеющие одинаковую плотность. Но уже для алфавита из трёх символов существует слово с попарно различными плотностями символов:

$$W = (abacaba)^{\infty}.$$

Несложно проверить, что W является сбалансированным. Эту конструкцию можно рекуррентно продолжить на алфавиты из большего числа символов:

$$W_k = (U_{k-1} a_k U_{k-1}), \quad U_3 = a_1 a_2 a_1 a_3 a_1 a_2 a_1. \tag{1}$$

В связи с этим сформулируем гипотезу Френкеля.

Гипотеза. Единственное (с точностью до перестановок) сбалансированное слово над алфавитом из $k \geqslant 3$ символов с попарно различными плотностями есть слово (1).

В настоящий момент гипотеза доказана для $k = 3, 4, 5, 6, 7$.

Литература

[1] Белов А., Кондаков Г. Обратные задачи символической динамики // Фундамент. и прикл. мат. — 1995. — Т. 1, вып. 1. — С. 71—79.

[2] Синай Я. Г. Введение в эргодическую теорию. — М.: ФАЗИС, 1996.

[3] Berstel J. Recent results on Sturmian words // Developments in Language Theory. II. — World Scientific, 1996. — P. 13—24.

[4] Graham R. L. Covering the positive integers by disjoints sets of the form $\{ [n\alpha + \beta]: n = 1, 2, \ldots \}$ // J. Combin. Theory. Ser A. — 1973. — Vol. 15. — P. 354—358.

[5] Hubert P. Well balanced sequences // Theoret. Comput. Sci. — 2000. — Vol. 242. — P. 91—108.

[6] Lothaire M. Combinatorics on words // Encyclopedia of Mathematics and Its Applications. Vol. 17. — Reading: Addison-Wesley, 1983.

[7] De Luca A. Sturmian words: Structure, combinatorics and their arithmetics // Theoret. Comput. Sci. — 1997. — Vol. 183. — P. 45—82.

[8] De Luca A., Varricchio S. Combinatorial properties of uniformly recurrent words and an application to semigroups // Internat. J. Algebra Comput. — 1991. — Vol. 1, no. 2. — P. 227—246.

[9] Morse M., Hedlund G. A. Symbolic dynamics. II. Sturmian trajectories // Amer. J. Math. — 1940. — Vol. 62. — P. 1—42.

[10] Newman M. Roots of unity and covering sets // Math. Ann. — 1971. — Vol. 191. — P. 279—282.

[11] Tijdeman R. Decomposition of the integers as a direct sum of two subsets // Number Theory. Number Theory Seminar Paris 1992—1993 / S. David, ed. — Cambridge: Cambridge Univ. Press, 1995. — P. 261—276.

[12] Tijdeman R. Fraenkel's conjecture for six sequences // Discrete Math. — 2000. — Vol. 222, no. 1—3. — P. 223—234.

[13] Vuillon L. Balanced words // Bull. Belg. Math. Soc. Simon Stevin. — 2003. — Vol. 10. — P. 787—805.

[14] Weyl H. Über der Gleichverteilung von Zahlen mod 1 // Math. Ann. — 1916. — Bd. 77. — S. 313—352.