Московский государственный университет
им. М.В. Ломоносова

На правах рукописи
УДК 512+519.17

Чернятьев Александр Леонидович

Нормальные базисы и символическая динамика

Специальность 01.01.06 — математическая логика,
алгебра и теория чисел

Диссертация на соискание ученой степени кандидата
физико-математических наук

Научные руководители:
доктор физико-математических наук,
профессор А.В. Михалев,
доктор физико-математических наук,
А.Я. Белов

Москва
2007

Содержание

1 Введение. 4

2 Пространство слов и символическая динамика. 19
2.1 Пространство слов. 19
2.2 Рекуррентность и равномерная рекуррентность 20
2.3 Специальные подслова 22
2.4 Морфизмы 22
2.5 Вопросы периодичности 23
2.6 Графы Рози подслов. 24
2.7 Слова, порождаемые динамическими системами. 26
2.8 Функция рассогласования. 28
2.9 Соответствие между словами и разбиениями множества. . 28
2.10 Перекладывания отрезков. 29
2.11 Слова Штурма. 30
2.12 Структура слов Штурма. 31
2.13 Слова Арно-Рози. 38

3 Сбалансированные слова и динамические системы 40
3.1 Введение и постановка проблемы. 40
3.2 Основные конструкции и определения. 40
3.3 Конечные сбалансированные слова 41
3.4 Расширение сбалансированных слов. 43
3.5 Слова, порождаемые повороотом окружности. 44
3.6 Сбалансированные слова над алфавитом из n символов. . . 46
3.7 Основная теорема о непериодических сбалансированных словах над произвольным алфавитом 50
3.8 Замечания о периодических сбалансированных словах над произвольным алфавитом 52

4 Слова с минимальной функцией роста. 54
4.1 Постановка задачи. 54
4.2 Свойства слов с минимальной функцией роста 56
4.3 Эволюции k-графов слов медленного роста 56
4.4 Частоты подслов в словах медленного роста 58
4.5 Конструкция динамической системы 59

5 Перекладывания отрезков. 61
5.1 Введение и постановка задачи. 61
5.2 Необходимые условия для порождаемости слова перекладыванием отрезков 61
5.3 Построение динамической системы 65

6 Эквивалентность множества р.р слов, порождаемых кусочно-непрерывным преобразованием, множеству слов, порождаемых перекладыванием отрезков. 68

1 Введение.

Актуальность темы. Комбинаторика слов находит свое применение в самых разных разделах математики. Например, в алгебре при изучении базисов и нормальных форм, в алгебраической топологии, в символической динамике. Ряд проблем, относящихся к комбинаторике слов находится на стыке алгебры и теории динамических систем. Многие проблемы комбинаторики слов представляют самостоятельный интерес.

Комбинаторика слов широко используется в задачах комбинаторной теории групп ([19],[1]), в теории алгебр Ли, вопросах бернсайдовского типа и задачах, связанных с мономиальными алгебрами. Комбинаторная техника, относящаяся к теории групп, развивалась в работах М. Дэна, Е. С. Голода и И. Р. Шафаревича, П. С. Новикова, С. И. Адяна, И. Рипса, М. Громова, А. Ю. Ольшанского, М. В. Сапира.

Е. С. Голод и И. Р. Шафаревич [9], [10] построили конечно порожденную бесконечную периодическую группу (с неограниченной экспонентой) на основе рассмотрения нормальных форм алгебр и оценки функий роста. П. С. Новиков и С. И. Адян [1] провели детальное исследование свойств периодичности, находящее свое применение в самых разных разделах математики. Ими было впервые построены примеры бесконечных конечно порожденных периодических групп ограниченой экспоненты (т.е. решена проблема Бернсайда) см. обзор [2]. Свойства периодичности такаже находятся в центре внимания настоящей работы.

В основе замечательных работ М. Громова и А. Ю. Ольшанского лежит техника диаграмм Ван-Кампена, возникшая в комбинаторной топологии. Подробнее (а также литературу по этой теме) см. в монографии [23].

Комбинаторные соображения, возникшие в символической динамике (автоматные группы), нашли свое применение в работах С. В. Алешина [3, 4, 5] и и Р. И. Григорчука [11], [12], [13] при решении проблемы Милнора – посторении групп промежуточного роста (при этом группы Григорчука периодичны). Впервые автоматные полугруппы были построены в работах С. В. Алешина. (Изложение примера С. В. Алешина – см. в книге [14]) Автоматные конструкции активно используются в самых разных ситуациях (см. [7], [26], [24], [18]). Возникают они и у нас (графы Рози).

Комбинаторика слов активно используется в алгебрах Ли, особенно в проблемах бернсайдовского типа [17]. В теории алгебр Ли описание базиса дается в терминах так называемых “правильных слов” (базис Холла-

Ширшова). Слово называется *правильным* если оно лексикографически больше любого его циклически сопряженного (слова v_1, v_2 *циклически сопряжены*, если $v_1 = u_1u_2$, $v_2 = u_2u_1$ для некоторых u_1 и u_2). (А запись слова по циклу используется в теории групп, тесно связанной с теорией алгебр Ли.) Только в правильном слове (причем однозначно) можно расставить лиевы скобки так, чтобы при их раскрытии исходное слово оказалось старшим членом получившегося полинома. Тем самым правильные слова задают базис свободной алгебры Ли (базис Холла–Ширшова). Применив методы символической динамики (равномерно рекуррентные слова и соображения компактности) Д. Бэкелин установил, что любое сверхслово содержит подслово вида uvu, где u и v – правильные слова, получив, тем самым, короткое доказательство локальной нильпотентности подалгебры алгебры Ли, порожденной сэндвичами, упростив соответствующие работы А. И. Кострикина [26].

Применив гомологическое соображение, связанное с невозможностью зацепления правильного слова за самого себя, А. И. Ширшов показал алгоритмическую разрешимость проблемы равенства в алгебрах Ли с одним определяющим соотношением.

С помощью комбинаторики слов А. И. Ширшов [30] доказал теорему о свободе подалгебры свободной алгебры Ли. Для супералгебр это обобщил А. А. Михалев [20], [21]. Он показал алгоритмическую разрешимость проблемы равенства для алгебр ли с одним определяющим соотношением.

Комбинаторика слов активно используется в проблемах бернсайдовского типа, в теории PI-алгебр, достаточно упомянуть знаменитую теорему Ширшова о высоте [27], [28] утверждающую возможность приведения слов к кусочно периодическому виду.

Теорема А.И.Ширшова о высоте. *Пусть A – конечно порожденная PI-алгебра степени m. Тогда существует конечный набор элементов Y и число $h \in \mathbb{N}$ такие, что A линейно представима (то есть порождается линейными комбинациями) множеством элементов вида:*

$$w = {u_1}^{k_1}{u_2}^{k_2}\cdots{u_r}^{k_r},\ \textit{где}\ u_i \in Y\ \textit{и}\ r \leq h.$$

При этом в основе оригинальных доказательств А. И. Ширшова (как теоремы о свободе так и теоремы о высоте) лежала техника, связанная с преобразованием алфавита путем подстановок. Эта же техника используется при работе с равномерно–рекуррентными словами и в символической динамике.

Последующие доказательства теоремы о высоте [37] и ее обобщение для алгебр Ли [22] использовали анализ свойств периодичности.

Понятие *роста* в алгебре является важным комбинаторным инвариантом, ему посвящена монография [51] (см. также [26], [7]). Если размерность пространства, порожденного словами степени не выше n от образующих A растет как n^λ, то величина λ называется *размерностью Гельфанда–Кириллова алгебры* A. Размерность Гельфанда–Кириллова может быть равной 0, 1, а также любому числу ≥ 2, ∞ или не существовать. То обстоятельство, что она не может принимать промежуточные значения на интервале $(1, 2)$ составляет содержание знаменитой Bergman gap theorem. Ассоциативная алгебра размерности Гельфанда–Кириллова 0 конечномерна. Л. Смолл показал, что ассоциативная алгебра размерности Гельфанда–Кириллова 1 является PI-алгеброй. Базисы ассоциативных алгебр размерности Гельфанда–Кириллова больше 1 с минимальной функцией роста исследовались в работах А. Т. Колотова [15], [16]. Их описание дается в терминах так называемых *последовательностей Штурма* находящейся в центре внимания данной работы.

Обобщение понятия роста на бесконечномерный случай является понятие *ряда коразмерностей*, введенное А. Регевым. Первоначальное доказательство А. Регева об экспоненциальной оценке ряда коразмерности относительно свободных алгебр было улучшено В. Н. Латышевым с помощью числа полилинейных n-разбиваемых слов на основе теоремы Дилуорса [**?**]. Само же понятие *n-разбиваемого слова* возникло у А. И. Ширшова в его теореме о высоте. Ряды коразмерности исследовались также в работах В. Н. Латышева, С. П. Мищенко, М. В. Зайцева, A. Giambruno. Понятию роста посвящена монография [51].

Комбинаторика слов работает в теории полугрупп. Слеует указать работы Екатеринбургской школы Л. Н. Шеврина, в часности, работы М. В. Сапира, О. Г. Харлампович. Они активно применяли методы символической динамике к теории полугрупп.

В теории мономиальных алгебр комбинаторика слов имеет основополагающее значение и находится в центре внимания работы [7].

Вопросы, связанные с мономиальными алгебрами, приводят изучению бесконечных (в одну или обе стороны) слов или *сверхслов*. Фундаментальным понятием в теории сверхслов является понятие *равномерно-рекуррентного* слова, введенное Х. Фюрстенбергом [44]. Слово W называется *равномерно-реккурентным*, если для каждого подслова $v \subset W$ существует натуральное $N(v)$, такое, что для любого подслова $u \subset W$

длины не менее, чем $N(v)$, v является подсловом u.

Имеет место следующая

Теорема. *Пусть W – бесконечное сверхслово. Тогда существует такое равномерно-рекуррентное слово $\hat{W}$, все подслова которого являются подсловами W.*

Эта теорема исключительно важна в комбинаторике слов, поскольку очень часто позволяет свести изучение произвольных слов к изучению равномерно-реккурентных слов.

В терминах равномерно-рекуррентных слов строится теория радикалов мономиальных алгебр ([7]). Мономиальная алгебра называется *мономиально почти простой*, если фактор по идеалу, порожденному по любому моному, нильпотентен.

Множество ненулевых слов в почти простой мономиальной алгебре есть множество всех подслов некоторого равномероно-рекуррентного слова.

Пересечение же идеалов с почти простым фактором совпадает с нильрадикалом мономиальной алгебры, а также ее радикалом Джекобсона (см. [36]).

В терминах равномерно-рекуррентных слов также получается описание слабо нетеровых мономиальных алгебр. Каждое ненулевое слово слабо нетеровой мономиальной алгебры есть подслово из набора (сверх)слов, удовлетворяющего следующему условию: каждое слово из этого набора либо конечное, либо является бесконечным (односторонними или двухсторонними) словом, которое при выбрасывании некоторого конечного куска распадается на равномерно-рекуррентные части.

Существует разные подходы к изучению сверхслов:

1. Непосредственно комбинаторные свойства слов
2. Графы подслов, или *графы Рози*
3. Топологическая динамика

Классическими работами по теории комбинаторики слов являются монографии [54, 55], [60].

Другим инструментом изучения сверхслов является понятие графов подслов или *графов Рози*. Если W – бесконечное сверхслово над алфавитом A, то *k-графом Рози* называется граф, вершины которого соответствуют различным подсловам W длины k. Из вершины w_1 в вершину w_2

ведет стрелка, если максимальный суффикс w_1 совпадает с максимальным префиксом w_2, то есть $w_1 = a_1 u$, $w_2 = u a_2$, где $a_1, a_2 \in A$.

Общий подход, связанный с описанием слов с помощью динамических систем, следующий. Пусть $W = \{w_n\}$ – бесконечное слово. $\tau(\{w_n\}) = \{w_{n+1}\}$ – оператор сдвига. Рассмотрим замыкание траектории слова относительно метрики Хэмминга $X \in A^*$. Прямые задачи символической динамики связаны с получением информации о динамической системе (X, τ) по информации о слове W.

Известно, что если слово W *равномерно-рекуррентно*, то полученная динамическая система минимальна, то есть не содержит нетривиальных замкнутых инвариантных траекторий.

Также стоит отметить работу [8], изучающую слова, получаемые из взятия дробных частей многочленов со старшим иррациональным коэффициентом в целых точках.

Обратно, пусть имеется дискретная топологическая динамическая система, то есть задано компактное топологическое пространство M, гомеоморфизм $f : M \to M$ и несколько открытых подмножеств

$$U_1, U_2, \ldots, U_{n-1}.$$

Положим также

$$U_n = M \setminus U_1 \cup U_2 \cup \ldots \cup U_{n-1}.$$

Рассмотрим начальную точку $x \in M$ и последовательность итераций $f^{(-1)}(x), x, f(x), \ldots$. По этой последовательности можно построить слово $W = \{w_n\}$ над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$ следующим образом:

$w_i = a_k$, если $f^{(i)}(x) \in U_k$. По свойствам динамической системы (размерность множества M, эргодичность) можно получить информацию о слове W.

Важным примером в изучении сверхслов динамическим подходом являются слова с предельной функцией роста.

Хорошо известно, что если функция роста слова $V(n)$ (то есть размерность пространства, порожденного словами степени не выше n) при некотором n удовлетворяет неравенству $V(n) < n(n+3)/2$, то алгебра имеет линейный рост.

В работе Колотова [15] построены алгебры с "предельной" функцией роста (то есть когда $V(n) = n(n+3)/2$), которые описаны в терминах поворота окружности. А именно, все такие алгебры, кроме счетного множества, строятся как алгебры A_W, где $W = \{w_i\}$ – слово над алфавитом $\{0, 1\}$, задаваемая иррациональными $\alpha, \beta \in (0, 1)$): $w_i = g(i+1) - g(i)$,

где $g(i) = [\alpha i + \beta]$. В комбинаторике слов чаще под функцией роста понимается функция $T(n)$ – количество различных подслов длины n. И, таким образом, $V(n) = \sum_k T_k$.

Слова с предельной функцией роста $T(n) = n+1$ образуют класс так называемых *слов Штурма (Sturmian words)*, другое название – *слова Бетти (Beaty words)*, которые впервые были упомянуты в работе [56] Hedlund, Morse "Symbolic dynamics II. Sturmian trajectories".

Классическая теория слов Штурма описана в обзорах [40], [52].

К наиболее важным результатам в теории слов Штурма относится так называемая *теорема эквивалентности*, в которой утверждается эквивалентность трех классов сверхслов над двубуквенным алфавитом:

Теорема эквивалентности.([56]) *Следующие условия на слово W эквивалентны:*

1. *Слово W имеет фунцию 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 углом вращения α.*

Последние продвижения в теории слов Штурма описаны в обзоре [41] J. Berstel "Recent results in Sturmian words".

Естественными обобщениями слов Штурма являются слова с минимальным ростом, то есть слова с функцией роста $T(n) = n + K$, начиная с некоторого n. Для двубуквенных алфавитов они носят название *квазиштурмовых* слов. Слова с функцией роста, удовлетворяющей соотношению $\lim_{n\to\infty} T(n)/n = 1$ изучены в работе [34] A.Aberkane "Words whose complexity satisfies $\lim p(n)/n = 1$".

Другим обобщением слов Штурма является обобщение, связанное с понятием *сбалансированности*, а также *m-сбалансированности*. Сбалансированные непериодические слова над n-буквенным алфавитом изучены в работе [45] Graham "Covering the Positive Integers by disjoints sets of the form $\{n\alpha + \beta\} : n = 1; 2; \ldots$". Построение порождающей динамической системы для сбалансированных непериодических слов является одним из результатов данной работы.

Описание периодических сбалансированных слов связано с *гипотезой Френкеля (Fraenkel's conjecture)*, утверждающей, что все сбалансированные периодические слова над алфавитом $A = \{a_1, \ldots, a_n\}$ из n символов имеют вид

$$W = (U_n)^\infty,$$

где U_n задается рекуррентно:

$$U_n = (U_{n-1}a_nU_{n-1}),\ U_3 = a_1a_2a_1a_3a_1a_2a_1.$$

Для 3-х буквенного алфавита гипотеза была доказана Р. Тайдеманом ([63, 64]). В настоящий момент гипотеза доказана для алфавитов состоящих не более чем из 7 символов.

Продвижение в задачах символической динамики для слов с линейной функцией роста получено в работе [35] P. Arnoux, G.Rauzy "Representation geometrique des suites the complexite $2n+1$". В этой работе построена динамическая система для слов с функцией роста $T(n) = 2n+1$, обладающих дополнительным комбинаторным свойством. В работе [61] G.Rote, "Sequences with subword complexity 2n" в терминах эволюции графов Рози описаны слова с функцией роста $2n$.

Одним из основных результатов данной работы является обобщение этих результатов на слова с линейной функцией роста, то есть с функцией роста $T(n) = kn + l$, для $n > N$.

Перекладывания отрезков естественным образом служат обобщением вращения круга. Эти преобразования были введены Осиледецом [33], следовавшим идее Арнольда [31], (см. также [32]). Рози [59] впервые показал, что связь между вращениями круга и последовательностями Штурма обобщается если рассматривать перекладывания отрезков. В связи с этим он задал вопрос описания последовательностей, связанных с перекладываниями отрезков.

Такие последовательности являются еще одним естественным обобщением слов Штурма. В частном случае, для $k = 3$ отрезков, описание таких последовательностей было получено в работе [47], а работе [49] были изучены частные случаи последовательностей, порождаемых перекладыванием 4-х отрезков.

В случае произвольного числа отрезков также получен ряд интересных результатов. В работе [48] получен комбинаторный критерий на порождаемость слов, получаемых симметричным перекладыванием отрезков, то есть перекладыванием, связанным с перестановкой $1 \to k, 2 \to k-1, \ldots, k \to 1$.

В работе [50] получен критерий порождаемости слов преобразованием перекладывания отрезков, удовлетворяющих следующему условию: траектория каждой концевой точки отрезка перекладывания не попадает на концевую отрезка перекладывания, в том числе сама на себя. В этом случае, как не сложно видеть, слова будут иметь функцию сложности $T(n) = (k-1)n + 1$.

Отметим, что во всех этих работах изучаются перекладывания, не меняющие ориентацию отрезков, а характеристические множества совпадают с отрезками перекладывания. В нашей работе с помощью языка графов Рози мы сначала изучаем слова, порождаемые кусочно-непрерывным преобразованием отрезка, а затем показываем эквивалентность множества таких слов и слов, порождаемых перекладываниями. Этот метод дает возможность описать р.р. слова, связанные с произвольным перекладыванием отрезка, более того, мы не требуем, чтобы характеристические множества, соответсвующие символам алфавита, совпадали с перекладываемыми отрезками.

Цель работы. Данная диссертация посвящена исследованию нормальных базисов алгебр медленного роста, а также исследованию взаимосвязи между комбинаторными свойствами слов, цепочками порождающих их автоматов (графов Рози) и порождающими их динамическими системами.

Научная новизна. Все основные результаты являются новыми. Среди них можно отметить следующие:

1. Асимптотическое описание нормальных базисов алгебр с минимальной функцией роста $T(n) = n + K$.
2. Построение общего критерия порождаемости слова преобразованием перекладывания отрезков в автоматных терминах.
3. Построение динамической системы для сбалансированных непериодических слов над произвольным алфавитом.

Основные методы исследования. Основными инструментами исследования в работе являются исследование цепочки автоматов (графов Рози), порождающих сверхслово, а также техника подстановок, восходящих к А. И. Ширшову. Мы пользуемся различными результатами теории равномерно-рекуррентных слов и слов Штурма. Для построения динамической системы, порождающей сбалансированные слова используются результаты эргодической теории, а также лемма Кронекера-Вейля об иррациональных сдвигах тора.

Практическая и теоретическая ценность. Работа носит теоретический характер. Результаты диссертации могут быть полезны в комбинаторной теории колец и полугрупп, в частности, при изучении мономиальных алгебр, а также в символической динамике.

Апробация результатов. Основные результаты диссертации докладывались на следующих семинарах:

1. "Кольца и модули" кафедры высшей алгебры МГУ в 2000-2006 гг.
2. "Арифметика и геометрия" кафедры теории чисел МГУ в 2004-2005г.
3. Dynamics seminar, Einstein Institute of Mathematics под руководством профессора Х.Фюрстенберга (H.Furstenberg) в 2004г.
4. Семинаре под руководством профессора А.Френкеля (A. Freankel) в 2004г., Weizman Institute of Science.

Структура и объем работы. Диссертация состоит из оглавления, введения, четырех глав и списка литературы, который включает 58 наименований.

Благодарности. Автор глубоко благодарен своем научным руководителям — доктору физико-математических наук профессору Алексею Яковлевичу Белову и доктору физико-математических наук Александру Васильевичу Михалеву за постановку задач, обсуждение результатов и постоянное внимание к работе.

Также автор хотел бы поблагодарить за внимание и обсуждения работы доктора физико-математических наук, профессора Виктора Николаевича Латышева, доктора физико-математических наук Николая Германовича Мощевитина и доктора физико-математических наук Андрея Михайловича Райгородского.

Автор выражает свою отдельную благодарность профессору университета Вайсмана А.Френкелю (A. Freankel) за детальное обсуждение работы.

Краткое содержание работы. **Первая глава** посвящена обзору и доказательству базовых результатов комбинаторики слов.

Отдельно рассматриваются результаты теории слов Штурма, представлено доказательство теоремы эквивалентности, которое используется в следующих главах. Также доказываются основные результаты из теории графов Рози.

Вторая глава посвящена изучению сбалансированных слов над произвольным алфавитом. Изучаются свойства слов, порождаемых иррациональными сдвигами тора.

Пусть W – слово над бинарным алфавитом, порождаемое сдвигом одномерного тора, то есть динамикой $(\mathbb{S}^1, U, T_\alpha, x_0)$.

Обозначим $\beta = \alpha/q + r/q$, $U_q = \{x | qx \in U\}$, $y_0 = x_0/q$, где q, r — целые. Несложно показать, что динамика $(\mathbb{S}^1, T_\beta, U_q, y_0)$ порождает то же слово W. Переход от первой динамики ко второй мы назовем q–*размножением*

Доказана следующая теорема об эквивалентности динамик:

Теорема 1.1 *Пусть два слова W_1 и W_2 порожденные динамиками $(\mathbb{S}^1, T_\alpha, U, x_0)$ и $(\mathbb{S}^1, T_\beta, V, y_0)$ совпадают. Тогда существуют 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 $\delta > 0$.*

Далее проводим редукцию от сбалансированного слова W над n-буквенным алфавитом к n бинарным словам Штурма:

Построим слова $W_1, W_2, \ldots, W_n$ над бинарными алфавитами

$$A_1 = \{a_1, \bar{a_1}\}, A_2 = \{a_2, \bar{a_2}\}, \ldots, A_n = \{a_n, \bar{a_n}\}$$

следующим образом:

$W_i = (w_n^i)_{n \in \mathbb{Z}}$

$$w_n^i = w_n = \begin{cases} a_i, & w_n = a_i \\ \bar{a_i}, & w_n \neq a_i \end{cases}$$

Каждому слову W_i соответствует одномерная динамика $(\mathbb{S}^1, T_{\alpha_i}, \Delta_i, x_i)$, ее порождающая, следовательно слово W порождается свигом на n-мерном торе с вектором сдвига $\gamma = (\alpha_1, \ldots, \alpha_n)$.

Обозначим через M замыкание траектории начальной точки $x = (x_1, x_2, \ldots, x_n)$ при действии T_γ.

Доказано следующее предложение:

Предложение 1.2 *M гомеоморфно множеству $\mathbb{S}^1 \times \{1, 2, \ldots, N\}$.*

Теперь видно, что динамика реализуется на множестве
$M = \mathbb{S}^1 \times \{1, 2, \ldots, N\}$, а отображение имеет вид:
$f : (x, k) \to (x + \alpha, k + 1 \bmod N)$.

Будем теперь понимать под U_i характеристическое множество для символа a_i, которое лежит на замыкании траектории, и пусть также U_i^k обозначает часть характеристического множества, которое лежит на k-ой компоненте связности (окружности) M. Далее изучаются соответствующие характеристические множества.

Основным результатом данной главы является

Теорема 1.3 *Пусть W – сбалансированное непериодическое слово над алфавитом A. Тогда для W существует динамическая система (M, f), удовлетворяющая следующим условиям:*

1. *$M = \mathbb{S}^1 \times \mathbb{Z}_m$ как топологическое пространство.*
2. *$f : M \to M$ есть композиция поворота на α в $\mathbb{S}^1$ и сдвига на 1 в $\mathbb{Z}_m$. Длина $\mathbb{S}^1$ равна m.*
3. *Каждая компонента $\mathbb{S} \times \{k\}$ $k = 1, \ldots, n$ 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 $\mathbb{S} \times \{k\}$ к компоненте $\mathbb{S} \times \{k + 1\}$ ($\mathbb{S} \times \{m + 1\} = \mathbb{S} \times \{1\}$) 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.*

Третья глава посвящена другому обобщению слов Штурма – словам *медленного роста*. Слово W называется *словом медленного роста*, если $F_W(n) = n + K$, для всех $n \geq N$. В словах Штурма $F_W(n) = n + 1$ и $F_W(n + 1) - F(n) = 1$. Естественным обобщением будут слова, для

которых $F(n) = n + K$, то есть выполняется $F(n+1) - F(n) = 1$ для достаточно больших n.

Показывается, что для каждого такого слова существует граф Рози, соответствующий некоторому слову Штурма. Действительно, если начиная с некоторого $n > N$ выполняется $T(n) = n + K$, то это означает, что для каждого $n > N$ в слове W существует только одно правое и одно левое специальное слово длины n. То есть в графах Рози таких слов есть одна входящая и одна выходящая развилка, а значит, существует слово Штурма с такой же эволюцией k-графов, что и данное. Дальше просходит редукция к теореме эквивалентности. Доказана следующая

Теорема 1.4 *Пусть W – рекуррентное слово над произвольным конечным алфавитом A. Тогда следующие условия на слово W эквивалентны:*

1. *Существует такое натуральное N, что функция роста слова W равна $T_W(n) = n + K$, для $n \geq N$ и некоторого постоянного натурального K.*

2. *Существуют такое иррациональное α и целые $n_1, n_2, \ldots n_m$, что слово W порождается динамической системой $(\mathbb{S}^1, T_\alpha, I_{a_1}, \ldots, I_{a_n}, x)$, где $T\alpha$ – сдвиг окружности на иррациональную величину α, I_{a_i} – объединение дуг вида $(n_j\alpha, n_{j+1}\alpha)$.*

Четвертая глава посвящена описанию слов, связанных с перекладыванием отрезков. Рассматриваются случаи, когда ориентация отрезков сохраняется и когда нет. В терминах графов Рози дается комбинаторное описание сверхслов порождающихся перекладыванием отрезков. В центре внимания слова, имеющие функцию роста $T_W(n) = kn + l$. Если слово обладает такой функцией роста, то $T_W(n+1) - T_W(n) = k$.

Известно, что если перекладывание k отрезков *регулярно*, то есть траектория любого из концов отрезков перекладывания не попадает на другой конец любого отрезка, то эволюция любой точки является словом с ростом $T_W(n) = kn + l$.

Мы ищем условия на слово W, при которых оно порождалось бы преобразованием перекладываний отрезка, не обязательно являющегося регулярным. Отметим, что сдвиг окружности, по сути, является перекладыванием двух отрезков с сохранением ориентации.

Рассмотрим соответствие между подсловами и подмножествами M. Легко видеть, что если начальная точка принадлежит множеству U_i, то ее эволюция начинается с символа a_i. Рассмотрим образы множеств

U_i при отображениях $f^{(-1)}, f^{(-2)}, \ldots$. Ясно, что если точка принадлежит множеству

$$T^{(-n)}(U_{i_n}) \cap T^{-(n-1)}(U_{i_{n-1}}) \cap \ldots \cap T^{(-1)}(U_{i_1}) \cap U_{i_0},$$

то эволюция начинается со слова $a_{i_0}a_{i_1}\cdots a_{i_n}$. Соответственно, количество различных существенных эволюций длины $n+1$ равно количеству разбиений множества M на непустые подмножества границами подмножств ∂U_i и их образами при отображениях $f^{-1}, f^{-2}, \ldots, f^{-n+1}$. Обозначим через I_u множество разбиения, которое соответствует слову u. Ясно, что специальным подсловам соответствуют те интервалы, которые делятся образами концов перекладываемых интервалов. Для данного слова u назовем слово v *левым (соотв. правым) потомком*, если u – суффикс (соотв. префикс) слова v, в соответствии с этим будем называть вершину в G_n левым (соотв. правым) потомком вершины в G_k, $n > k$. Прообраз конца интервала может являться граничной точкой только для двух интервалов, соответственно, специальные подслова могут иметь валентность только равную 2. Сформулируем

Правило 1. Для того, чтобы бесконечное слово W порождалось системой $(I, T, U_1, \ldots, U_k)$ необходимо, чтобы любое специальное слово имело валентность 2.

Таким образом, мы можем наложить условие на эволюцию графов Рози: начиная с некоторого k все k-графы Рози имеют входящие и исходящие развилки степени 2. Предположим, что некоторому подслову w соответствует характеристический интервал, полностью лежащий внутри интервала перекладывания. Пусть точка $A \in [0,1]$ делит I_w на два интервала, образы которых лежат в I_{a_k} и I_{a_l} соответственно, а точка $B \in [0,1]$ – делит на интервалы, прообразы которых лежат в I_{a_i} и I_{a_j} соответственно.

Выбор минимального невстречающегося слова, а, значит, удаляемого ребра, определяется взаиморасположением точек A и B, а также сохранением или сменой ориентации отображения на этих множествах. Итого, имеется 8 вариантов, которые разбиваются на четыре пары, соответствующие одинаковым наборам слов. Например, слову $a_i w a_k$ соответствует ситуация

$$B < A, T^{-1}([x_w, B]) \subset I_{a_i}, T([x_w, A] \subset I_{a_k}).$$

Введем понятие *размеченного графа Рози*. Граф Рози называется *размеченным*, если

1. Ребра каждой развилки помечены символами l (“left”) и r (“right”)

2. Некоторые вершины помечены символом “–”.

Последователем размеченного графа Рози назовем ориентированный граф, являющийся его последователем как графа Рози, разметка ребер которого определяется по правилу:

1. Ребра, входящие в развилку должны быть помечены теми же символами, как и ребра, входящие в любого левого потомка этой вершины;

2. Ребра, выходящие из развилки должны быть помечены теми же символами, как и ребра, выходящие из любого правого потомка этой вершины;

3. Если вершина помечена знаком “–”, то все ее правые потомки также должны быть помечены знаком “–”.

Замечание. Поясним смысл разметки графа. Пусть ребра входящей развилки соответствуют a_i и a_j, символы l и r соответствуют левому и правому множеству в паре $(T(I_{a_i}), T(I_{a_j}))$. Если символы a_k и a_l соответствуют ребрам исходящей развилки, то символы l и r ставятся в соответствии с порядком “лево-право” в паре (I_{a_k}, I_{a_l}). Знак “–” ставится в вершине, если характеристическое множество, ей соответствующее, принадлежит интервалу перекладывания, на котором меняется ориентация.

Условие для перехода от графа G_n к G_{n+1}:

Правило 2.

1. Если в графе нет двойных развилок, соответствующих биспециальным подсловам, то при переходе от G_n к G_{n+1} имеем $G_{n+1} = D(G_n)$;

2. Если вершина, соответствующая биспециальному слову не помечена знаком “–”, то ребра, соответствующие запрещенным словам выбираются из пар lr и rl

3. Если вершина помечена знаком “–”, то удаляемые ребра должны выбираться из пары ll или rr.

Назовем эволюцию размеченных графов Рози *правильной*, если **правила 1** и **2** выполняются для всей цепочки эволюции графов, начиная с G_1, назовем эволюцию *асимптотически правильной*, если **правила 1** и **2** выполняются, начиная с некоторого G_n. Будем говорить, что эволюция

размеченных графов Рози *ориентированна*, если в k-графах нет вершин, помеченных знаком "–".

Теорема 1.5 *Равномерно-рекуррентное слово W*

1. *Порождается перекладыванием отрезков, тогда и только тогда, когда слово обеспечивается асимптотически правильной эволюцией размеченных графов Рози.*

2. *Порождается перекладыванием отрезков с сохранением ориентации тогда и только тогда, когда слово обеспечивается асимптотически правильной ориентированной эволюцией размеченных графов Рози.*

2 Пространство слов и символическая динамика.

2.1 Пространство слов.

В этой части мы определим основные понятия комбинаторики слов. В дальнейшем будет обозначать конечный алфавит, то есть непустое множество элементов (символов). Через A^+ обозначим множество всех конечных последовательностей, символов, или *слов*.

Конечное слово всегда может быть единственным образом представлено в виде

$w = w_1 \cdots w_n$, где $w_i \in A, 1 \leq i \leq n$. 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* (или *фактор*) слова w, если существуют слова $p, q \in A^+$ такие, что $w = puq$.

Если слово p (или q) равно Λ, то u называется *префиксом* (или *суффиксом*) слова w.

Мы будем обозначать $\mathrm{Pref}(w)$ ($\mathrm{Suf}(w)$) множество всех префиксов (соответственно суффиксов) слова w.

Для любой пары целых чисел таких, что $1 \leq i \leq j \leq n$ мы обозначим через $w[i, j]$ подслово $w[i, j] = w_i \cdots w_j$.

Если $w = w_1 \cdots w_n$ $w_i \in A$, $i = l, \cdots n$ – слово, то *обратное слово* $\tilde{w}$ для w определяется как $\tilde{w} = w_n \cdots w_l$. Более того, мы полагаем, что $\tilde{\Lambda} = \Lambda$.

Слово называется *палиндромом*, если $\tilde{w} = w$. Множество всех палиндромов обозначается PAL.

Пусть $w = w_1 \cdots w_n$ $w_i \in A$, $i = l, \ldots, n$ – 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 p_w, или просто p, минимальный период w.

Слово w называется *периодическим*, если $p \leq [|w|/2]$ (здесь $[x]$ обозначает целую часть x) Через $\mathbb{N}$ (соответственно $\mathbb{N}_+$) мы будем обозначать неотрицательные (положительные) целые числа.

Одностороннее (соответственно *двустороннее*) *бесконечное слово* над алфавитом A – это отображение

$w : \mathbb{N}_+ \to A$, (соответственно $w : \mathbb{Z} \to A$).

Для каждого n, мы полагаем $w_n = w(n)$ и обозначим $W = w_1 w_2 \cdots$.

Слово $u \in A^+$ – конечное подслово W, если существуют такие $i, j \in N$, $(1 \le i \le j)$, что $u = w_i \cdots w_j$. Последовательность $w[i,j] = w_i \cdots w_j$ назовем вхождением u в W.

Слово W назовем в *существенно периодическим*, если его можно представить как $W = uv^\infty = uvvv \cdots$, где $u \in A^*$, $v \in A^+$.

Бесконечное в обе стороны слово назовем просто *периодичным*, если оно имеет вид

$$W = v^\infty = \cdots vvv \cdots$$

Через $F(W)$ обозначим множество всех подслов (конечных и бесконечных) слова W.

Определение 2.1 *Два бесконечных слова W и V над алфавитом A назовем* **эквивалентными**, *если $F(W) = F(V)$.*

Морфизмами на пространстве конечных слов мы будем называть обычные полугрупповые морфизмы, то есть отображение $\varphi : A^* \to A^*$ – морфизм слов, если оно сохраняет операцию конкатенации (приписывания): $\varphi(uv) = \varphi(u)\varphi(v)$ для любых $u, v \in A^*$. Это определение естественным образом продолжается для морфизмов бесконечных слов.

Понятия *мономорфизма, эпиморфизма и изоморфизма* вводятся обычным путем.

2.2 Рекуррентность и равномерная рекуррентность

Определение 2.2 *Слово W называется* **рекуррентным**, *если каждое его подслово встречается в нем бесконечно много раз (в случае двустороннего бесконечного слова, каждое подслово встречается бесконечно много раз в обоих направлениях). Слово W называется равномерно-рекуррентным или (р.р словом), если оно рекуррентно и для каждого подслова v существует натуральное $N(v)$, такое, что для любого подслова W и длины не менее, чем $N(v)$, v является подсловом u.*

Пусть W – бесконечное слово. Для любого подслова v можно определить множество *возвращаемых* слов $\mathrm{Ret}_W(v)$, а именно, слово u – *возвращаемое* для v, если vuv – подслово W и v – не является подсловом u. Ясно, что для рекуррентных слов множество возвращаемых слов $\mathrm{Ret}(v)$ каждого подслова v будет непустым, а в случае равномерной рекуррентности множество длин слов из $\mathrm{Ret}(v)$ будет ограниченным.

Определение 2.3 *Для произвольного слова W можно определить функцию роста:*

$$T_W(n) = \operatorname{Card} F_n(W)$$

Ясно, что если $T_W(n) = 0$ для какого-то n, то W – конечное слово. В противном случае бесконечное.

Теорема 2.4 *Следующие два свойства бесконечного слова W равносильны:*

а) для любого k можно найти $N(k)$ такое, что любой участок W длины k можно найти в любом участке W длины $N(k)$;

б) если все конечные куски слова V являются конечными кусками слова W, то и все конечные куски слова W являются конечными кусками слова V. □

Лемма 2.5 (о компактности) *Пусть $\mathcal{M}$ — множество слов неограниченной длины над конечным алфавитом $\mathcal{A}$. Тогда существует бесконечное слово V, все подслова которого являются подсловами слов из $\mathcal{M}$.* □

Следствие 2.6 *Если u — подслово слова W, и в слове W есть сколь угодно длинные подслова, не содержащие u, то существует слово W', которое беднее кусками, чем W. При этом можно выбрать W так, чтобы u не входило в W'.* □

Следствие 2.7 *Для любой убывающей в смысле отношения $\sqsupseteq$ цепочки сверхслов найдется инфимум.* □

Следствие 2.8 *Для слова W найдется беднейшее по подсловам слово W' такое, что $W \sqsupseteq W'$.* □

Мы имеем следующую теорему, означающую, что из кусков любого слова можно составить равномерно-рекуррентное слово.

Теорема 2.9 *Пусть W — бесконечное слово. Тогда существует равномерно рекуррентное слово $\widehat{W}$, все подслова которого являются подсловами W.* □

Эта теорема исключительно важна в комбинаторике слов, т.к. очень часто позволяет сводить изучение произвольных слов к изучению р.р. слов.

2.3 Специальные подслова

Определение 2.10 *Рассмотрим бесконечное (одностороннее или двустороннее) слово над алфавитом A. Пусть v – его подслово и $x \in A$. Тогда*

1. *Символ x – левое (правое) расширение v, если xv (соотв. vx) принадлежит $F(W)$.*
2. *Подслово v называется левым (правым) специальным подсловом, если для него существуют два или более левых (правых) расширения.*
3. *Подслово v называется биспециальным, если оно является и левым, и правым специальнм подсловом одновременно.*
4. *Количество различных левых (правых) расширений подслова назовем левой (правой) валентностью этого подслова.*

2.4 Морфизмы

Классическим методом получения бесконечных (в одну сторону) слов является итерация *подстановок* (или *морфизмов*). Рассмотрим отображение множества конечных подслов над алфавитом A в себя: $\varphi : A^* \to A^*$, которое сначала определим на буквах алфавита:

$$\begin{aligned} &a_1 \to \varphi(a_1) \\ &a_2 \to \varphi(a_2) \\ &\vdots \\ &a_n \to \varphi(a_n). \end{aligned}$$

Продолжим это отображение на конечные слова:

$$\varphi(x_1 x_2 \cdots x_k) = \varphi(x_1)\varphi(x_2)\cdots\varphi(x_k)$$

.

Для бесконечных (вправо) слов отображение продолжается естественным образом:

$$\varphi(x_1 x_2 x_3 \cdots) = \varphi(x_1)\varphi(x_2)\varphi(x_3)\cdots$$

В случае, если для какого-то символа $a \in A$ $\varphi(a)$ начинается с a, то

имеем последовательность префиксов:

$$\begin{aligned}
&\varphi(a) = au,\ u \in A^* \\
&\varphi^2(a) = \varphi(au) = \varphi(a)\varphi(u) \\
&\varphi^3(a) = \varphi(\varphi^2(a)) = \varphi^2(a)\varphi^2(u) \\
&\vdots
\end{aligned}$$

И таким образом, последовательность $a, \varphi(a), \varphi^2(a), \ldots$ задает бесконечное слово W такое, что $\varphi(W) = W$. Такие слова называются *инвариантными относительно подстановки* φ .

Пример 1. Слово Фиббоначи. Рассмотрим последовательность слов над алвафитом $A = \{a, b\}$:

$f_0 = b,\ f_1 = a, f_2 = ab,\ f_3 = aba, \ldots,\ f_n = f_{n-1}f_{n-2}$. Предел этой последовательности слов, т.е. слово $W = abaababaabaab\cdots$, называется *словом Фиббоначи*. Оно инвариантно относительно подстановки $\varphi(a) = ab$, $\varphi(b) = a$

Пример 2. Слово Туе-Морса. Рассмотрим подстановку $\varphi(a) = ab$, $\varphi(b) = ba$. Тогда:

$$\begin{aligned}
&\varphi^2(a) = \varphi(ab) = abba \\
&\varphi^3(a) = \varphi(abba) = abbabaab \\
&\vdots
\end{aligned}$$

Слово, инвариантное относительно φ, называется *словом Туе-Морса (Thue-Morse word)*. Это слово обладает следующим свойством: оно свободно от кубов, то есть не содержит подслов вида u^3. Легко проверить, что не существует бесконечных слов над двубуквенным, свободных от квадратов, то есть подслов вида u^2.

Пример 3. Слово Триббоначи. Рассмотрим слово над трехбуквенным алфавитом $A = \{a, b, c\}$, инвариантное относительно подстановки $\varphi(a) = ab$, $\varphi(b) = ac$, $\varphi(c) = a$, называется *словом Триббоначи*.

2.5 Вопросы периодичности

Самым простым примером бесконечного слова являются периодические, или существенно периодические (с предпериодом) слова. Имеет место следующая

Теорема 2.11 *Пусть W – бесконечное вправо слово над конечным алфавитом A ($\mathrm{Card}(A) \geq 2$). Тогда следующие условия эквивалентны:*

1. *Слово W существенно периодично.*

2. *В слове W конечное число правых специальных подслов.*

3. *Существует n такое, что для каждого подслова длины большей n существует ровно одно возвращаемое слово*

4. *Существует n такое, что $T_W(n) = T_W(n+1)$*

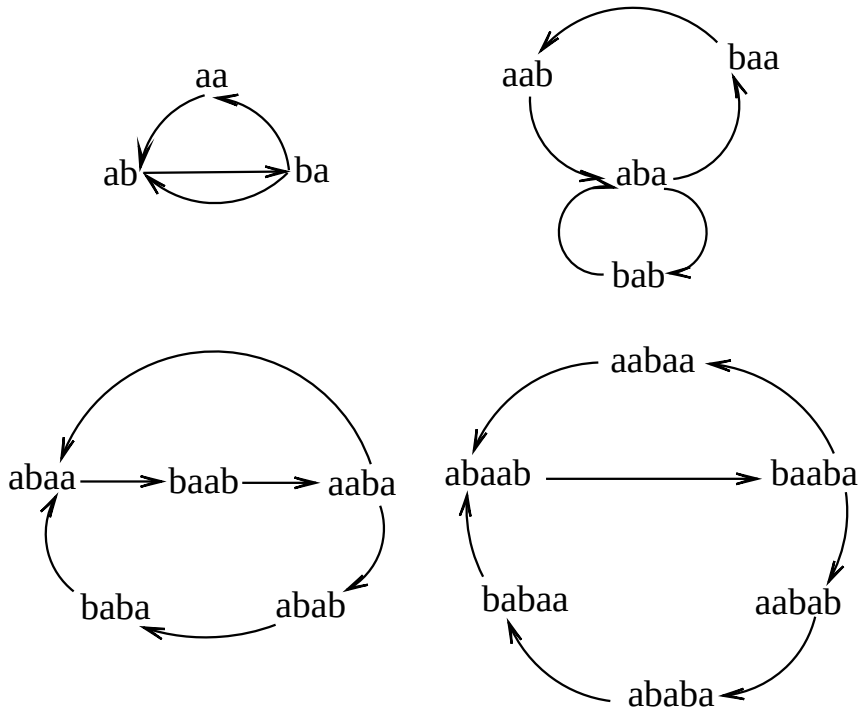
Предложение 2.12 *Множество L конечных слов является множеством подслов бесконечного слова $W = w_1w_2w_3\cdots$ тогда и только тогда, когда*

1. *Каждое слово длины n в L содержится в некотором слове длины $n+1$ в L и*

2. *Подмножества $L_1, L_2, L_3, \ldots$, где $L_n = w \in L : |w| = n$,*

2.6 Графы Рози подслов.

Удобным инструментом для описания слова W является *графы подслов*, или графы Рози (Rauzy's graphs), введенные Рози [57], которые строятся следующим образом: *k-граф* слова W – ориентированный граф, вершины которого взаимнооднозначно соответствуют подсловам длины k слова W, из вершины A в вершину B ведет стрелка, если в W есть подслово длины $k+1$, у которого первые k символов – подслово соответствующее A, последние k символов – подслово, соответствующее B. таким образом, ребра k-графа биективно соответствуют $(k+1)$-подсловам слова W.



2-,3-,4-,5-графы Рози для слова Фиббоначи.

Ясно, в k-графе G слова W правым специальным словам соответствуют вершины, из которых выходит (соотв. в которые входит) больше одной стрелки. Такие вершины мы будем называть развилками. Граф G будем называть *сильно связным*, если из любой вершины в любую вершину можно перейти по стрелкам.

Последователем ориентированного графа G будем называть ориентированный граф $\mathrm{Fol}(G)$ построенный следующим образом: вершины графа G биективно соответствуют ребрам графа G, из вершины A в вершину B ведет стрелка, если в графе G конечная вершина ребра A является начальной вершиной ребра B.

Связность графов Рози и рекуррентность соответсвующего слова связаны естественным образом. Имеет место следующее

Предложение 2.13 *Пусть W – бесконечное (в одну сторону) слово. Следующие условия эквивалентны:*

1. *Слово W – рекуррентно.*
2. *Для всех k соответствующий k-граф слова W является сильно связным.*
3. *Каждое подслово W встречается не меньше двух раз.*

4. Любое подслово является продолжаемым слева.

Доказательство. Достаточно доказать следующие 4 импликации:

$(i) \to (ii)$. Пусть w и v – два подслова длины k слова W. Так как W – рекуррентно, то в W существует подслово вида $wAvBw$, где A, B – некоторые подслова. Соответственно, в G_k существует путь из вершины, соответствующей w в вершину v и путь из вершины, соответствующей v в вершину, соответствующую w. Значит, G_k сильно связанный.

$(ii) \to (iii)$. Предположим противное. Пусть v – подслово, встречающееся ровно раз в W, тогда существует префикс v', который содержит v и который тоже встречается ровно раз в W. Но тогда G_k (где k – длина v') не является строго связанным.

$(iii) \to (iv)$. Пусть v – подслово W и слово v встречается в W минимум дважды. Ясно, что v является продолжаемым (от второго вхождения к первому).

$(iv) \to (i)$. Предположим противное, пусть W не рекуррентно. Пусть v подслово, которое встречается конечное число раз в W. Тогда пусть v' – префикс W достаточно большой длины, то есть содержащий все вхождения v. Слово v' больше не встречается в W следовательно, не является продолжаемым слева.

□

В терминах графов Рози можно выразить такие важные понятия, как рост подслов, множество запрещенных подслов, минимальные запрещенные слова и т.д.

2.7 Слова, порождаемые динамическими системами.

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

По последовательности итераций можно построить бесконечное слово над бинарным алфавитом:

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

которое называется *эволюцией* точки x_0. Символическая динамика исследует взаимосвязь свойств динамической системы (M, f) и комбинаторных свойств слова W_n.

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 введем понятие *существенной эволюции.*

Определение 2.14 *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 вызывет недоразумений, будем понимать, существенную эволюцию. Отметим, что точка может иметь несколько существенных эволюций.

Предложение 2.15 *Пусть V – конечное слово. Тогда множество точек с конечной существенной эволюцией V замкнуто. Аналогичное утверждение верно для бесконечного слова W.*

Доказательство. См. [8]

Определение 2.16 *Морфизмом двух динамик $G : (M_1, f_1) \to (M_2, f_2)$ называется непрерывное отображение $g : M_1 \to M_2$ такое, что диаграмма*

$$\begin{array}{ccc} M_1 & \xrightarrow{f_1} & M_2 \\ \downarrow g & & \downarrow g \\ M_1 & \xrightarrow{f_2} & M_2 \end{array}$$

коммутативна.

Понятие мономорфизма, эпиморфизма и изоморфизма вводится обычным образом.

Фактор динамика естественным образом определяется на фактор-топологии тогда и только тогда, когда f переставляет классы эквивалентности отображения f. Отметим, что прообразы точек при морфизмах замкнуты.

2.8 Функция рассогласования.

Вместе с метрикой Хэмминга, разделяющей любые два различные слова можно рассмотреть другую меру, называемую *плотностью рассогласования.*

Определение 2.17 *Пусть $V = \{v_n\}$ и $W\{w_n\}$ – два бесконечных слова над алфавитом A. Функция $\rho = \rho(V, W)$ называется плотностью рассогласования и определяется по формуле:*

$$\rho(U, W) = \lim_{n\to\infty} \frac{\sum_{i=-n}^{n} \rho(i)}{2n}$$

где

$$\rho(i) = \begin{cases} 0, & v_i = w_i \\ 1, & v_i \neq w_i \end{cases}$$

есть функция рассогласования.

2.9 Соответствие между словами и разбиениями множества.

Теперь посмотрим соответствие между словами и подмножествами M. Из построения следует, что если начальная точка принадлежит множеству U_i, то ее 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

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

то эволюция начинается со слова $a_{i_0}a_{i_1}\cdots a_{i_n}$.

Соответственно, количество различных существенных эволюций длины $n+1$ равно количеству разбиений множества M на непустые подмножества границами подмножств ∂U_i и их образами при отображениях $f^{(-1)}, f^{(-2)}, \ldots, f^{(-n)}$.

Замечание. Количество конечных существенных эволюций напрямую связано с топологической размерностью множества M. Ясно, например, что если гомеоморфно отрезку или окружности, то одна точка может являться граничной только для двух открытых подмножеств M и, соответственно, иметь только две существенные эволюции. Если M гомеоморфно части плоскости $\mathbb{R}^2$ – то существенных эволюций может быть сколь угодно много.

2.10 Перекладывания отрезков.

Перекладывание отрезков является естественным обобщением сдвига окружности – в случае разбиения окружности на дуги длины α и $1-\alpha$ и величины сдвига, равной α, это преобразование совпадает с перекладыванием двух отрезков.

Более того, перекладывание отрезков является очень важным преобразованием в эргодической теории, теории динамических систем и теории чисел.

Рассмотрим общий случай:

Пусть отрезок $[0,1]$ разбит на полуинтервалы длин $\lambda_1, \lambda_2, \ldots, \lambda_n$ и $\sigma \in S_n$ – перестановка на множестве $\{1, 2, \ldots, n\}$.

Интервалы разбиения могут быть представлены через длины отрезков следующим образом:

$$X_i = \left[\sum_{i<j} \lambda_j, \sum_{i \leq j} \lambda_j \right)$$

Перекладывание отрезков "переставляет" отрезки $(X_1, X_2, \ldots, X_n)$ разбиения между собой, в результате получается новое разбиение

$$(X_\sigma(1), X_\sigma(2), \ldots, X_\sigma(n))$$

.

Более точно, преобразование T ставит в соответствие каждой точке $x \in X_i$:

$$T(x) = x + a_i$$

где

$$a_i = \sum_{k<\sigma^{-1}(i)} \lambda_{\sigma(k)} - \sum_{k<i} \lambda_k$$

Если же преобразование переворачивает некоторый отрезок, то все точки дополнительно симметрично отражаются относительно середины этого отрезка. В случае, если перекладывание не переворачивает отрезки, мы будем говорить, что такое перекладывание сохраняет ориентацию.

Определение 2.18 *Перекладывание отрезков T называется регулярным, если для любой точки a_i, где $X_i = [a_i, a_{i+1})$, $T^n(a_i) \neq a_j$*

Сформулируем следующую важную теорему:

Теорема 2.19 *Перекладывание отрезков является регулярным тогда и только тогда, когда траектория любой точки всюду плотна в* $[0,1]$

Теорема 2.20 *Пусть слово* W *является эволюцией некоторой точки при регулярном перекладывании* k *отрезков. Тогда любое подслово* W *имеет ровно* k *возвращаемых слов.*

Изучение свойств слов, порождаемых перекладываниями отрезков осуществляется теми же методами, что и при сдвиге единичной окружности. Основным инструментом здесь является рассмотрение отрицательных орбит концов интервалов перекладывания:

$$0 = a_1 < a_2 < \ldots < a_{k+1} = 1,$$

где
$X_i = [a_i, a_{i+1})$, $(i \in \{1, \ldots, k\})$.

Множество концов интервалов перекладывания $\{a_i | 1 \leq i \leq k+1\}$ обозначим X^1. Слово w длины n является подсловом эволюции точки x, то есть бесконечного слова $U(x)$, тогда только тогда, когда существует такой интервал $I_w \subset [0,1]$ и точка $y \in I_w$ такие, что слово w равно конкатенации следующих символов:

$$\mathcal{I}(x)\mathcal{I}(T(x)) \cdots \mathcal{I}(T^{n-1}(x)) = w,$$

где $\mathcal{I}(x) = a_i \in A$ тогда и только тогда, когда $x \in X_i$

Предложение 2.21 *Пусть* T *– регулярное перекладывание* k *отрезков. Тогда эволюция* $U(x)$ *произвольной точки* x *имеет функцию сложности* $T_{U(x)}(n) = n(k-1)+1$, *для любого* $n \in \mathbb{N}$.

2.11 Слова Штурма.

В этой части мы дадим основные определения и конструкции, связанные со словами Штурма (также называемыми последовательностями Бетти, сбалансированными словами и т.д.), а также сформулируем и докажем теорему эквивалентности.

Определение 2.22 *Бесконечное вправо (соответственно, двустороннее бесконечное) слово* $W = (w)_{n \in \mathbb{N}}$ *(соответственно,* $W = (w)_{n \in \mathbb{Z}}$ *) над конечным алфавитом* A *называется словом Штурма, если его функция сложности равна* $T_W(n) = n+1$ *для всех* $n \geq 0$.

Примечание 1. Так как $T_W(1) = 2$, то слово Штурма по определению является словом над двухсимвольным алфавитом. Здесь и далее в этой главе мы будем рассматривать слова только над бинарными алфавитами.

Предложение 2.23 *Если $W = (w)_{n\in\mathbb{N}}$ – слово Штурма, то W рекуррентно.*

Доказательство. Предположим противное и некоторое подслово u встречается в W конечное число раз. Пусть оно не встречается, начиная с N-ой позиции. Рассмотрим слово $V = (v)_{n\in\mathbb{N}}$: $v_k = w_k + N$. Тогда все подслова V являются подсловами W и в V нет подслова u, значит $T_V(n) \leq n$ для какого-то n, а следовательно, с какого-то момента периодично. □

Примечание 2. В случае бесконечного в обе стороны слова при использовании аналогичного определения слова Штурма возникают исключительные классы слов, например:

$V_1 = \cdots aaabbb\cdots$, $V_2 = \cdots aaabaaa\cdots$. Поэтому в определении двусторонних слов Штурма мы будем также предполагать рекуррентность.

Пример. Слово Фиббоначи. Рассмотрим последовательность слов над алфавитом $A = \{a, b\}$:

$f_0 = b$, $f_1 = a$,$f_2 = ab$, $f_3 = aba, \ldots$, $f_n = f_{n-1}f_{n-2}$. Напомним (см. 2.11), что предел этой последовательности слов, т.е. слово $W = abaababaabaab\cdots$, называется *словом Фиббоначи.*

Для подслов небольшой длины нетрудно заметить:

$$T_W(1) = \mathrm{Card}\{a, b\} = 2$$
$$T_W(2) = \mathrm{Card}\{aa, ab, ba\} = 3,$$
$$T_W(3) = \mathrm{Card}\{aba, baa, bab, aab\} = 4,$$
$$T_W(4) = \mathrm{Card}\{abaa, abab, baab, baba, aaba\} = 5$$

Предложение 2.24 *Слово Фиббоначи является словом Штурма.*

□

2.12 Структура слов Штурма.

Предложение 2.25 *Слово Штурма содержит ровно одно из двух подслов: aa или bb.*

Доказательство. Действительно, $T_W(2) = 3$, однако, если слово содержит и aa и bb, то оно обязано, в силу рекуррентности, содержать и ab и ba, что приводит к противоречию. □

Определение 2.26 *Мы будем говорить, что слово Штурма имеет тип "a", если в нем встречается подслово aa, и типа "b", если встречается подслово bb.*

Теперь перейдем к определению слов Штурма через другое понятие – свойство сбалансированности.

Определение 2.27 *Слово W называется сбалансированным, если для любых его двух подслов u, v одинаковой длины выполняется неравенство:*

$$||u|_a - |v|_a| \leq 1 \tag{1}$$

Ясно, что из неравенства (1) автоматически следует $||u|_b - |v|_b| \leq 1$

Естественным обобщением понятия сбалансированности является *m-сбалансированность*:

Определение 2.28 *Слово W над алфавитом A называется m-сбалансированным по символу $a \in A$, если для любых подслов $u, v \subset W$ таких что $|u| = |v|$ выполняется $||u|_a - |v|_a| \leq m$. Слово называется m-сбалансированным, если оно сбалансировано по каждому символу алфавита.*

Свойство сбалансированности означает, что для непериодического слова W символы a и b перемешаны "наилучшим образом". Если для какого-то k будет выполняться более сильное условие, то есть для любых двух слов длины k количество символов a (а соответственно и b) было бы одинаковым, то слово W было бы периодичным с периодом k.

Докажем теперь следующее важное

Предложение 2.29 *Пусть слово W не является сбалансированным. Тогда существует такое слово u (возможно, пустое), что слова aua и bub являются подсловами W.*

Доказательство. Если слово W не сбалансировано, то существуют два подслова одинаковой длины v_1 и v_2 такие, что $||v_1|_a - |v_2|_a| \geq 2$. Будем также считать, что эти слова имеют наименьшую длину среди всех подслов, обладающих данным свойством.

Обозначим v_1^k, v_2^k – префиксы длины k слов v_1 и v_2 .

Пусть также $d_k = ||v_1^k|_a - |v_2^k|_a|$. Ясно, что $d_n \geq 2$ (где $n = |v_1| = |v_2|$) и $|d_{k+1} - d_k|$ равно 0 или 1.

Положив для удобства $d_0 = 0$ и рассматривая последовательность d_k от $k = 0$ до $k = n$, мы обнаружим первый момент, когда $d_k = 2$. В силу минимальности длины слов v_1 и v_2, имеем $k = n$

Пусть теперь $v_1 = a_1 a_2 \cdots a_n$ и $v_2 = b_1 b_2 \cdots b_k$. Тогда ясно, что $a_n \neq b_n$, иначе при обрезании n-х символов мы опять получили бы более короткие слова с этим же свойством. Аналогично, $a_1 \neq b_1$. Также ясно, что $a_1 = a_n$, $b_1 = b_n$, так как в противном случае слова $a_2 \cdots a_{n-1}$ и $b_2 \cdots b_{n-1}$ также обладали бы этим свойством. Пусть $a_1 = a_n = a$ и $b_1 = b_n = b$.

Но тогда $a_2 = b_2$, в противном случае опять получатся более короткие слова с разницей символов, бóльшей 1.

Рассматривая также далее последовательность префиксов, индуктивно приходим к равенствам $a_k = b_k$, $k = 2, \ldots, n-1$, следовательно, слово $u = a_2 \cdots a_{n-1}$ – искомое. □

Следующее утверждение является первой частью общей теоремы эквивалентности:

Теорема 2.30 *Слово W является словом Штурма тогда и только тогда, когда оно существенно непериодично и сбалансированно.*

Доказательство. Докажем, сначала **импликацию** $\Rightarrow$. Предположим противное, то есть слово W не является словом Штурма. Покажем тогда, что оно не сбалансированно. Рассмотрим такое наименьшее n_0, что $T_W(n_0) \geq n_0 + 2$. Ясно, что $n_0 \geq 2$, так как $T_W(1) = 2$ по определению. Так как n_0 – минимальное, то $T_W(n_0 - 1) = n_0 + 1$, это значит, что существует как минимум два подслова u и v, которые имеют по два разных продолжения справа. Из минимальности n_0 опять же следует, что u и v различаются только в первых символах, иначе более короткие суффиксы также обладали бы этим свойством, следовательно существует подслово w такое, что $u = aw$ и $v = bw$. Так как u, v имеют по два разных продолжения, то в W встретятся подслова awa и bwb, значит W не сбалансированно.

Теперь докажем **импликацию** $\Leftarrow$. Предположим , что W не сбалансированно. Тогда по предложению 2.29, найдется такое подслово $u = w_0 w_1 \cdots w_n$, что aua и bub – также подслова W. Предположим также, что u имеет минимальную длину из всех слов с таким свойством.

Отметим, что, исходя из этого предположения, следует, что если пара слов одинаковой длины u_1, u_2 не сбалансирована, то слова имеют длину, не меньше, чем $n + 3$. Действительно, из доказательства предложения 2.29, следует, что минимальная длина такого u не больше наименьшей

длины не сбалансированной пары уменьшенной на 2. Поэтому любая не сбалансированная пара имеет длину не меньше $|u| + 2 = n + 3$.

Заметим, что по предложению u не пусто, иначе в слове W встретились бы слова aa и bb.

По этой же причине, $w_0 = w_n$, $w_1 = w_{n-1}$ и более обще, $w_k = w_{n-k}$, то есть *u–палиндром.*

Теперь, мы знаем, что различных слов длины $n+1$ ровно $n+2$. Слово u может быть расширено двумя путями как влево, так и вправо. Соответственно, ровно одно из слов au или bu (мы предположим, что au) может быть расширено вправо двумя путями. Следовательно, aua, aub и bub – подслова W, а bua – не является подсловом. Пусть i – первый момент, когда подслово bub встречается в W. Докажем следующее

Предложение 2.31 *Слово au не является подсловом слова*

$$v = w_i w_{i+1} \cdots w_{i+2n+3}.$$

Доказательство. Длина v равна $2n + 4$, длина bub равна $n + 3$ и длина au равна $n + 2$, следовательно, в предложении в точности утверждается, что первый символ вхождения au не может встречаться в bub. Предположим противное, то есть что начало au перекрывается с bub. Ясно, что перекрытие происходит не в первой позиции bub, если перекрытие происходит позиции w_k, то это означает, что $aw_0 \cdots w_{n-k} = w_k \cdots w_n b$ и это означает, что $w_k = a$ и $w_{n-k} = b$, что противоречит тому, что u – палиндром. □

Завершение доказательства Теоремы 2.30

Понятно, что в слове $v = w_i w_{i+1} \cdots w_{i+2n+3}$ встречается ровно $n + 3$ подслова длины $n + 2$. Но из предложения следует, что среди них нет подслова au, следовательно, какое-то подслово встречается как минимум два раза. Все подслова из v могут быть продолжены вправо единственным образом, так как двумя способами продолжается вправо только au. Следовательно, сверхслово W существенно периодично. □

Теперь отметим еще одну важную характеристику слов Штурма – плотность вхождения символа.

Сформулируем следующее предложение – определение.

Предложение 2.32 *Определим плотность вхождения символа a в слове $W = (w)_{\mathbb{N}}$ как предел*

$$\rho(a) = \lim_{n\to\infty} |w_0 \cdots w_n|_a / n$$

Если W – слово Штурма, то предел существует и $\rho(a)$ иррационально.

Доказательство. Сначала докажем, что этот предел существует, если W – слово Штурма. Пусть a_n – минимальное количество символов в подлслове W длины n. Из сбалансированности W следует, что максимальное количество вхождений равно $a_n + 1$. Ясно, что достаточно доказать существование предела последовательности a_n/n и его иррациональность.

Слово длины $kq + r$ может быть разбито на k слов длины q и одно слово длины r, из чего немедленно следует, что

$$ka_q \leq a_{kq+r} \leq k(a_q + 1) + r \tag{2}$$

Если рассмотреть достаточно большие n, то есть $n > q^2$, мы можем записать $n = kq + r$, где $k \geq q$ и $0 \leq r < q$. Так как $r < k$, то нетрудно проверить, что $r/n < k/n < 1/q$, следовательно, из второй части неравенства (2) имеем : $a_n/n \leq a_q/q + 2/q$.

Далее, так как $n \geq a_n \geq ka_q > ra_q$, то из первой части неравенства (1) вытекает $(a_q - 1)n/q = ka_q + (ra_q - n)/q \leq a_n$ значит, после деления на n имеем:

$$\frac{a_q}{q} - \frac{1}{q} \leq \frac{a_n}{n} \leq \frac{a_q}{q} + \frac{2}{q}.$$

Значит, последовательность a_n/n является последовательностью Коши, а, значит, имеет предел.

Предположим теперь, что этот предел является рациональным числом p/q. Из неравенства $ka_n \leq a_k n < a_k n + 1 \leq k(a_n + 1)$, следует, что если n' делит n, то $a_n/n \leq a_{n'}/n' < (a_{n'} + 1)/n' \leq (a_n + 1)/n$.

В частности, последовательность $(a_{2^n q}/2^n q)_{n\in\mathbb{N}}$ является возрастающей, а $(a_{2^n q} + 1/2^n q)_{n\in\mathbb{N}}$ – убывающей.

Тогда имеем неравенства:

$$\frac{a_q}{q} \leq \frac{a_{2^n q}}{2^n q} < \frac{a_{2^n q} + 1}{2^n q} \leq \frac{a_q + 1}{q}.$$

Но эти последовательности стремятся к p/q, что возможно только при $a_q = p$ и $a_{2^n q} = 2^n p$ для всех n

Другой замечательной характеристикой слов Штурма является их описание через динамику на единичной окружности.

Предложение 2.33 *Пусть $\mathbb{S}^1$ – окружность единичной длины, $U \subset \mathbb{S}^1$ – дуга длины α, T_α – сдвиг окружности на иррациональную величину α. Тогда динамика произвольной точки порождает слово с функцией сложности $T_W(n) = n + 1$, т.е. слово Штурма.*

Набросок доказательства (подробнее см. [54, 55]). Рассмотрим образы U при итерациях $T^{(-1)}(U), T^{(-2)}(U), \cdots$. Поскольку длина дуги совпадает с величиной сдвига, то при первой такой итерации появится ровно одна новая граничная точка и, соответственно, одно новое множество разбиения, при каждой следующей итерации также будет появляться ровно одна дополнительная граничная точка, остальные граничные точки будут переходить друг в друга. Так как количество множеств разбиения – это и есть количество подслов данной длины, а изначально было два множества, то функция сложности $T_W(n) = n + 1$

Слова, получаемые такими динамическими системами, иногда называют *механическими словами* (*mechanical words*).

Возникает естественный

Вопрос: *Любое ли слово Штурма может быть получено из сдвига окружности.*

Ответ на него положительный. Таким образом, можно сформулировать следующую **теорему эквивалентности**:

Теорема 2.34 *Следующие условия на слово W эквивалентны:*

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

□

Рассмотрим подстановку над алфавитом $A = \{a, b\}$: $\sigma(a) = ab$ и $\sigma(b) = a$. Тогда последовательность $a, \sigma(a), \sigma^2(a) \ldots$. задает *слово Фиббоначи*(см. 2.11):

$$W = abaababaabaab \cdots$$

Подстановка σ называется *морфизмом Фиббоначи*, так как длина $\sigma^n(a)$ равна n-му числу Фиббоначи. Действительно, легко проверить, что подстановка удовлетворяет рекуррентному соотношению $\sigma^{n+2}(a) = \sigma^{n+1}(a)\sigma^n(a)$, а значит, и рекуррентному соотношению на длины.

Подстановка σ обладает еще одним замечательным свойством: если слово W над бинарным алфавитом является словом Штурма, то $\sigma(W)$ также является словом Штурма.

Дадим теперь более общее

Определение 2.35 *Морфизм $\varphi : A^* \to A^*$ называется* **морфизмом Штурма**, *если для произвольного слова Штурма W слово $\varphi(W)$ также является словом Штурма.*

Отметим, что свойство быть морфизмом Штурма является локальным, то есть его можно проверить на конечных словах. Справедлива следующая

Теорема 2.36 (S.Berstel, P.Séébold) *Морфизм φ является морфизмом Штурма тогда и только тогда, когда образ слова $u = abbabbababbaba$ является сбалансированным.*

Для примера, морфизм Фиббоначи переводит слово u в слово $\sigma(u) = abaaabaaabaabaaabaab$, которое является сбалансированным.

Теперь введем следующее

Определение 2.37 *Слово Штурма W называется характеристическим (или стандартным), если все его префиксы являются левыми специальными подсловами.*

Согласно теореме эквивалентности, для заданного α существует бесконечно много слов Штурма с плотностью символов α и $1-\alpha$ соответственно, а именно слова, порожденные динамиками $(\mathbb{S}^1, T_\alpha, \Delta, x_0)$, которые отличаются выбором начальной точки x_0. Стандартное слово Штурма отвечает начальной точке $x_0 = 0$.

Как показано в [54], стандартные слова Штурма могут быть получены как предел последовательности s_n конечных слов, задаваемых рекуррентным соотношением: $s_n = s_{n-1}^{d_n} s_{n-2}$, где $n \geq 1$, $s_{-1} = b$, $s_0 = a$ и последовательность целых d_n удовлетворяет условию $d_0 \geq 0$, $d_n > 0$, при $n > 0$. Последовательность (d_n) связана с величиной сдвига α следующим образом: непрерывная дробь $[0, 1 + d_0, d_1, d_2, \ldots]$ равна в точности α.

Еще одной характеристикой слов Штурма является их описание через количество подслов-палиндромов. Напомним, что конечное слово $w = w_1 w_2 \cdots w_n$ называется *палиндромом*, если $w = \tilde{w} = w_n w_{n-1} \cdots w_1$.

Палиндромическим замыканием конечного слова u называется наименьшее по длине слово – u^+, которое имеет u в качестве префикса.

Пример. Для слова $u = abaabab$ палиндромическое замыкание равно $u^+ = abaababaaba$.

Следующая характеристика стандартных слов Штурма получена A. de Luca [52]:

Предложение 2.38 *Бесконечное вправо слово W над бинарным алфавитом $A = \{a, b\}$ является характеристическим словом Штурма тогда и только тогда, когда существует бесконечное слово $\Delta(W) = a_0 a_1 \cdots$, в котором оба символа встречаются бесконечно много раз такое, что:*

$$W = \lim_{n\to\infty} u_n,$$

где $u_0 = \varepsilon$ *и* $u_{n+1} = (u_n a_n)^{(+)}$, $n \geq 1$.

Более общая теорема была доказана A.Droubay и J.Pirillo (см.[42]):

Предложение 2.39 *Бесконечное слово* W *является словом Штурма тогда и только тогда, когда для любого натурального* n *слово* W *имеет ровно одно подслово-палиндром длины* n*, если* n *– четное и ровно два подслова-палиндрома длины* n*, если* n *– четное.*

Напомним, что для рекуррентного слова W и его подслова u *возвращаемые слова* – это максимальные по длине слова, находящиеся между двумя последовательными вхождениями u в W.

Для слов Штурма имеет место следующее характеристическое свойство:

Предложение 2.40 *Равномерно рекуррентное слово* W *является словом Штурма, если и только если каждый его префикс имеет ровно два возвращаемых слова.*

2.13 Слова Арно-Рози.

Естественными обобщениями слов Штурма являются сбалансированные слова над произвольным алфавитом, и слова с медленным ростом, о которых пойдет речь в главах 3 и 4. Другой подход к обобщению слов Штурма был исследован в работах П. Арно и Г.Рози (см. [58]), полученный класс слов носит название *слов Арно-Рози (Arnoux-Rauzy words)* или *AR-слов*.

Бесконечное слово W над k-буквенным алфавитом называется *AR-словом*, если для каждого натурального $n \geq 1$ существует ровно одно левое и ровно одно правое специальное подслово длины n, причем валентность этих специальных подслов равна в точности k.

Ясно, что слова Штурма являются AR-словами. Примером AR-слова над алфавитом более чем из двух символов является слово Трибоначи (2.4). Оно строится над алфавитом из 3-х символов.

Графы Рози для AR-слов над k-буквенным алфавитом имеют вид. схожий с графами Рози для слов Штурма. Они имеют ровно одну вершину в которую входят k ребер и ровно одну, из которой выходят k ребер.

В случае, если соответствующее вершине специальное подслово является палиндромом, то оно является биспециальным, то есть и правым и левым одновременно. Остальные вершины графа имеют одно входящее и одно исходящее ребро. Все AR-слова также являются равномерно-рекуррентными словами.

AR-слово W называется *характеристическим*, если все его префиксы являются левыми специальными подсловами. В частности, слово Трибоначчи является характеристическим словом.

Характеристические AR-слова могут быть описаны с помощью так называемого *правила Рози (Rauzy's rules)*. Правило заключается в способе получения бесконечного (право) слова посредством итерации последовательности морфизмов определенного вида. Точнее, для заданного алфавита A рассмотрим семейство морфизмов $\tau_a : A^* \to A^*$, $a \in A$:

$$\tau_a(b) = \begin{cases} a, & b = a \\ ab, & b \neq a \end{cases}$$

Для заданного слова $w = a_1 \cdots a_n$ определим $\tau_w = \tau_{a_1} \cdots \tau_{a_n}$. Например, $\tau_{abc}(a) = \tau_{ab}(ca) = \tau_a(bcba) = abacaba$.

Имеет место следующее

Предложение 2.41 *Произвольное характеристическое AR-слово над алфавитом A может быть получено как предел последовательности $((\tau_{d_n}(a))_{n\geq 1}$ для некоторого $a \in A$, где d_n есть префикс длины n некоторого бесконечного слова Δ, которое содержит бесконечно много вхождений каждого символа из A.*

Следующий результат является обобщением теоремы 2.38 для характеристических AR-слов был получен [43]:

Теорема 2.42 *Произвольное характеристическое AR-слово может быть получено как предел последовательности $(u_n)_{n\geq 0}$ слов, где $u_{n+1} = (u_n a_n)^+$ – палиндромическое замыкание $u_n a_n$, $u_0 = \varepsilon$, и бесконечное слово $\Delta = a_0 a_1 a_2 \cdots$ содержит бесконечно много вхождений каждого символа из A.*

3 Сбалансированные слова и динамические системы

3.1 Введение и постановка проблемы.

Понятие сбалансированности играет существенную роль при изучении комбинаторных свойств слов, а также в теории биллиардов и теории динамических систем.

Для слов над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$ свойство сбалансированности (или "равномерной перемешанности") означает, что в любых двух подсловах одинаковой длины количество символов одного сорта (например, символов a_1) почти одно и то же. Более точно, для любого $1 \leq i \leq n$, и любых двух подслов одинаковой длины, количество символов a_i в них отличается не более, чем на 1.

Бесконечные непериодические слова над алфавитом из двух символов – это в точности слова Штурма.

Из теоремы эквивалентности 2.34 следует, что сбалансированное непериодическое слово порождается сдвигом на единичной окружности, то есть может быть получено следующим образом:

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

где $T_\alpha : x \to x + \alpha (\mathrm{mod}\, 1)$ – сдвиг на величину α, U – дуга длины α и x_0 – начальная точка.

Наша цель – обобщение данного результата на сбалансированные слова над более чем двухбуквенным алфавитом, то есть построение для каждого сбалансированного слова такой динамической системы, которая порождала бы данное слово.

Отметим, что описание сбалансированных слов в арифметической форме были получены Р.Грэхемом [45] , а позднее описание в форме бинарных слов Штурма было получено П.Хубертом [46].

Также отметим, что описание периодических слов Штурма связано с известной проблемой Френкеля (Fraenkel conjecture, [64]) о покрытии множества целых чисел последовательностями Бетти (другое название последовательностей Штурма).

3.2 Основные конструкции и определения.

Пусть $A = \{a_1, a_2, \ldots, a_n\}$ – конечный алфавит, W – бесконечное (в обе стороны) слово, то есть произвольная последовательность символов из A.

Напомним, что *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, $n \geq 1$.

Для конечного подслова $u \subset W$ обозначим величины $|u|$ – длина слова и $|u|_{a_i}$ – количество вхождений в подслово u символов a_i. Ясно, что $|u| = \sum_{a_i \in A} |u|_{a_i}$.

Определение 3.1 *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*

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

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

Отметим, что 0-сбалансированность ведет к периодичности:

Предложение 3.2 *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*

Доказательство. Фактически, сбалансированность означает минимальное отклонение от периодичности, также, как и минимальный рост.

3.3 Конечные сбалансированные слова

Предложение 3.3 *Пусть $k \geq 2$, тогда для любого конечного слова u, u^2 сбалансированно тогда и только тогда, когда сбалансированно u^k.*

Доказательство. 1. Ясно, что если u^2 не сбалансированно, то и u^k не сбалансированно.

2. Пусть u^k не сбалансированно, тогда пусть v и v' – два подслова u^k минимальной длины такие, что $||v|_a - |v'|_a| \geq 2$.

Если $|v| \geq |u|$, то $v = xy$, $v' = x'y'$ и $|x| = |x'| = |u|$. Слова x и x' сопряжены с u и, значит, $|x|_a = |x'|_a = |u|_a$. Из этого следует, что $||y|_a - |y'|_a| \geq 2$, что приводит к противоречию с минимальностью $|v|$. Следовательно, $|v| < |u|$ и слово v является подсловом u^2. $\square$

Для любого конечного сбалансированного слова w над алфавитом $A = \{a, b\}$, существует такое целое n, что между двумя последовательными вхождениями символа b находится ровно n или $n+1$ символ a.

Более того, каждое такое сбалансированное слово w может начинаться и оканчиваться не более, чем $n+1$ символом a и, значит, может быть представлено в виде $w = SA_1A_2 \cdots A_kP$, где $k \geq 0$, $S = a^l$ $(0 \leq l \leq n+1)$, $A_i \in \{ba^n, ba^{n+1}\}$ $(1 \leq i \leq k)$ и $P = ba^l$ $(0 \leq l \leq n+1)$. Если слово w (не обязательно сбалансированное) может быть представлено в таком виде, мы будем говорить, что оно является *n-сводимым* или, просто, *сводимым*.

Для заданного сводимого слова $w = SA_1A_2 \cdots A_kP$ мы можем построить новое слово, которое обозначим $\mathrm{Red}(w) = sa_1a_2 \cdots a_kp$ такое, что:

- $s = \varepsilon$, если $S \neq a^{n+1}$, иначе $s = b$
- $a_i = a$, если $A_i = ba^n$
- $a_i = b$, если$A_i = ba^{n+1}$
- $p = \varepsilon$, если $P \neq ba^{n+1}$, иначе $s = b$

Если $n \geq 1$, то операция сведения нетривиальна, действительно, тогда $|\mathrm{Red}(w)| \leq |w|/2$ (так как $|s| \leq |S|/2$, $p \leq |P|/2$ и $|a_i| \leq |A_i|/2$).

С предложенной операцией сведения связана следующая

Лемма 3.4 *Пусть конечное слово w над алфавитом A содержит не менее одного символа b. Тогда следующие утверждения эквивалентны:*

1. *Слово w сбалансированно.*
2. *Слово w n-сводимо (для некоторого целого $n > 0$) и слово $\mathrm{Red}(w)$ сбалансированно.*

Доказательство. Так как любое сбалансированное слово сводимо, достаточно показать, что n-сводимое слово не сбалансированно, если и только если не сбалансированно слово $\mathrm{Red}(w)$. Сначала предположим, что w не сбалансированно. Из предложения 3 следует, что существует такое подслово v, что ava и bvb – подслова w.

Если $|v|_b = 0$, то $v = a^n$ или $v = a^{n+1}$. В этом случае a^{n+2} является подсловом ava и тогда w не является n-сводимым.

Тогда существует такое целое l, что $a^l b$ является префиксом слова v. Так как bvb – подслово w, то $l = n$ или $l = n+1$. Так как a^{n+2} не является подсловом w и aa^l подслово av, мы имеем $l = n$.

Аналогично, ba^n является суффиксом v, значит $v = a^n h^n(u) b a^n$, где $u \in \{a; b\}^*$. Так как $ba^n h^n(u) b a^n b$ – подслово w, то aua – подслово $\mathrm{Red}(w)$.

Так как $a^{n+1} h^n(u) b a^{n+1}$ – посдлово w, то bub – подслово $\mathrm{Red}(w)$. Следовательно, $\mathrm{Red}(w)$ не сбалансированно.

Теперь, предположим, что $\mathrm{Red}(w)$ не сбалансированно. Из предложения 3, существует такое подслово v, что ava и bvb – подслова $\mathrm{Red}(w)$. Так как bvb – подслово $\mathrm{Red}(w)$, то $a^{n+1} h^n(v) b a^{n+1}$ – подслово w. Теперь заметим, что если ava – суффикс $\mathrm{Red}(w)$, то существует целое l, $(0 \le l \le n)$ такое, что $h^n(ava) b a^l$ – суффикс w. Тогда $ba^n h^n(v) b a^n b$ является подсловом w, и, значит, w не сбалансированно.

3.4 Расширение сбалансированных слов.

В этой части мы покажем, что любое конечное сбалансированное слово может быть продолжено до более длинного сбалансированного слова. Последовательность сбалансированных слов, каждое из которых является префиксом последующего, задает бесконечное (вправо) сбалансированное слово. Имеет место следующее

Предложение 3.5 *Для любого сбалансированного слова над алфавитом $A = \{a, b\}$ существует не менее двух символов x и y из A таких, что wx и yw сбалансированы.*

Доказательство.

Предположим, что wa и wb не сбалансированы, а w – сбалансированно. Тогда существуют такие подслова v и v', что

ava – суффикс wa, $bv'b$ – суффикс wb, а bvb и $av'a$ – подслова w. Ясно также, что $|v| \neq |v'|$. Если $|v| < |v'|$, то $v' = uav$ для некоторого слова u. Тогда ava и bvb являются подсловами w и w не сбалансированно: противоречие. Аналогично доказывается, что в случае $|v| > |v'|$, $av'a$ и $bv'b$ – подслова w.

Тепреь, если w сбалансированно, то и обратное слово $\widetilde{w}$ тоже сбалансированно. Для него существует $y \in A$, что $\widetilde{w}y$ – сбалансированно, значит, и yw – сбалансированно.

□

Для непериодических сбалансированных слов над бинарным алфавитом существует наглядная геометрическая интерпретация: координатную плоскость и произвольную прямую с иррациональным тангенсом наклона. Рассмотрим такие точки на этой прямой, у которых хотя бы одна координата – целое число (отметим, что в силу иррациональности тангенса наклона,точка с двумя целыми координатами может быть только одна). Точкам с целой координатой по оси X сопоставим символ a, точкам с целой координатой по Y сопоставим b. Если точка имеет две целые координаты – ей сопоставляется a, если прямая имеет угол наклона меньше $\pi/4$, и b в противном случае. Слово, которое можно прочитать вдоль прямой будет словом Штурма и сбалансированным непериодическим.

Эту конструкцию можно обобщить в n-мерный аналог. В этом случае образуются классы т.н. кубических (3-мерный аналог) и гиперкубические (n-мерный аналог) *биллиардные* слова.

Еще одной важной характеристикой бесконечных слов является *плотность*:

Определение 3.6 *Пусть $W = (w_n)_{n\in\mathbb{Z}}$ – бесконечное слово над алфавитом $A = \{a_1, a_2, \ldots, a_n\}$. Плотностью символа a_i называется предел (если он существует):*

$$\rho(a_i) = \lim_{n\to\infty} |w_{-n}\cdots w_{-1}w_0w_1\cdots w_n|_{a_i}/2n$$

В случае бесконечного вправо слова:

$$\rho(a_i) = \lim_{n\to\infty} |w_0w_1\cdots w_n|_{a_i}/n$$

Если слово W – слово Штурма, то оно порождается сдвигом окружности. Сдвиг на иррациональную величину является равномерным (см. [67]), поэтому плотность каждого символа совпадает с мерой Хаара характеристического множества. То же верно и для любого слова, порождаемого сдвигом окружности.

Случай периодических сбалансированных слов был изучен Тайдеманом ([63]). Он показал, что любое периодическое сбалансированное слово обладает плотностью по каждому символу.

3.5 Слова, порождаемые повороотом окружности.

В этой части мы сформулируем и докажем несколько полезных утверждений о сдвиге окружности и словах, порождаемых сдвигом окружности на иррациональную величину.

Предложение 3.7 1. *Траектория произвольной точки при иррациональном α всюду плотна (лемма Кронекера)*

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

Предложение 3.8 *Пусть $W = (w_n)$ — слово, которое порождается динамикой $(\mathbb{S}^1, T_\alpha, U, x_0)$ и пусть $\beta = \alpha/q + r/q$, $U_q = \{x | qx \in U\}$, $y_0 = x_0/q$, где q, r — целые. Тогда эволюция точки x_0, порожденная 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 $T_\beta^{(n)}(y_0) \in U_q$. Такую операцию над динамикой мы будем называть *q-размножением*. Ясно также, что если U – симметричное подмножество, то есть при сдвиге на величину $1/q$ (q-целое) переходит в себя, то аналогично можно рассмотреть множество $U^q = \{qx | x \in U\}$ и сдвиг, в q раз больший. Такую операцию мы будем называть *q-сжатием*.

Предложение 3.9 *Пусть два слова W_1 и W_2 порожденные динамиками $(\mathbb{S}^1, T_\alpha, U, x_0)$ и $(\mathbb{S}^1, T_\beta, V, y_0)$ совпадают. Тогда существуют 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 $m\alpha + n\beta = k$). Действительно, предположим противное. Пусть α и β линейно независимы с единицей над $\mathbb{Q}$. Рассмотрим сдвиг на двумерном торе $\mathbb{T}^2$ h : $(x, y) \to (x + \alpha, y + \beta) \bmod 1$. Согласно лемме Вейля, траектория любой точки при таком сдвиге всюду плотна. Тогда существует k такое, что $h^k(x_0, y_0) \in U \times (\mathbb{S}^1/V)$, но это означает, что в k-ом символе последовательности различаются. Противоречие. Значит $\alpha = p/q\beta + r/s$, где p, q, r, s–целые. Покажем, что данные p и q — искомые. Положим $\alpha_p = \alpha/p = \beta/q + r/sp$, $\beta_q = \beta/q$, тогда $ps\alpha_p = ps\beta_q = \gamma(\mathrm{mod}1)$. 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 $U_p \triangle V_q$ (здесь

$A \triangle B$ обозначает симметрическую разность множеств A и B), т.е. слова в данной позиции различаются.

□

Это предложение можно усилить и потребовать не полное совпадение слов, а их совпадение на множестве позиций, имеющем положительную меру в $\mathbb{Z}$, то есть потребовать ненулевую плотность рассогласования.

Теперь, пусть окружность разбита на k дуг:

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

Для произвольного символа $a \in A$ мы также 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 3.10 *Конечное слово $u = x_0 x_1 x_2 \cdots x_k$ является подсловом W тогда и только тогда, когда*

$$\bigcap_{i=0}^{k} T_\alpha^{-i}(I_{x_i}) \neq \emptyset$$

Доказательство. Это утверждение является следствием леммы Кронекера. Действительно, если пересечение не пусто, то из всюдуплотности траектории следует, что траектория попадет в пересечение (конечное пересечение полуинтервалов – интервал или отрезок или полуинтервал). Кусок эволюции, начиная с момента попадания траектории в пересечение, с длиной, равной $k+1$ будет совпадать слову u.

□

3.6 Сбалансированные слова над алфавитом из n символов.

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
$A_1 = \{a_1, \bar{a_1}\}, A_2 = \{a_2, \bar{a_2}\}, \ldots, A_n = \{a_n, \bar{a_n}\}$
следующим образом:

$$W_i = (w_n^i)_{n\in\mathbb{Z}},$$

$$w_n{}^i = w_n = \begin{cases} a_i, w_n = a_i \\ \bar{a_i}, w_n \neq a_i \end{cases}$$

Предложение 3.11 *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 $u = w_j w_{j+1} \cdots w_k \subset W$ и $u' = w_j^i w_{j+1}^i \cdots w_k^i$ – соответствующее подслово W^i, то $|u|_{a_i} = |u'|_{a_i}$. Сбалансированность по одному символу для слов над бинарным алфавитом влечет сбалансированность в целом. □

Согласно теореме эквивалентности 2.34, 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 3.12 $1, \alpha_i, \alpha_j$ *линейно зависимы над* $\mathbb{Q}$ *для любых* i *и* j.

Доказательство. Предположим противное и 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 линейно независимы с 1 над $\mathbb{Q}$. Тогда плотна и траектория точки (x_i, x_j), следовательно, на каком-то шаге она попадет в множество $U = \{(x, y | x \in \Delta_i, y \in \Delta_j)\}$. Это означает, что на n-м месте в слове W одновременно стоят символы a_i и a_j, что невозможно. Противоречие. □

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

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

где $T_\gamma : \mathbb{T}^n \to \mathbb{T}^n$ – сдвиг на n-мерном торе на вектор $\gamma = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, а $U_i = \{(y_1, y_2, \ldots, y_n) \in \mathbb{T}^n | y_i \in \Delta_i\}$.

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

Доказательство. Так как все α_i попарно линейно зависимы с единицей над $\mathbb{Q}$, то существует такое α (например, α_1) такое, что

$$\begin{aligned}
\alpha_1 &= r_1/s_1\alpha + p_1/q_1 \\
\alpha_2 &= r_2/s_2\alpha + p_2/q_2 \\
&\vdots \\
\alpha_1 &= r_n/s_n\alpha + p_n/q_n,
\end{aligned}$$

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

То есть $\gamma = \alpha(r_1/s_1, \ldots, r_n/s_n) + (p_1/q_1, \ldots, p_n/q_n) = \alpha\vec{\rho_1} + \vec{\rho_2}$, где $\vec{\rho_1}, \vec{\rho_2}$ – рациональные векторы.

Отсюда видно, что, так как сдвиги между собой коммутируют, то замыкание траектории есть прямая направленная вдоль вектора $\vec{\rho_1}$ и все ее образы при сдвиге на вектора вида $m\vec{\rho_2}$ ($m \in \mathbb{Z}$). Поскольку $\vec{\rho_1}$ – рациональный вектор, то прямая на торе будет замкнута, то есть гомеоморфна окружности, а поскольку $\vec{\rho_2}$ – рационален, то 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:

$$f : (x, k) \to (x + \alpha, k + 1 \bmod N).$$

Будем теперь понимать под U_i характеристическое множество для символа a_i, которое лежит на замыкании траектории, и пусть также U_i^k обозначает часть характеристического множества, которое лежит на k-ой компоненте связности (окружности) M.

Из всюдуплотности траектории на M следует, что каждая точка принадлежит какому-то множеству U_i^k. Для составления общей картины нам необходимо описание характеристических множеств.

Предложение 3.14 *Каждое множество U_i^l является объединением k_i дуг, периодически разбивающих окружность (то есть инвариантное относительно сдвига на $1/k_i$), причем k_i не зависит от номера компоненты l.*

Доказательство. Рассмотрим слово, порождаемое на l-ой компоненте и характеристическом множестве U_i^l которое мы построим следующим образом: рассмотрим произвольную точку траектории x_0, которая попадает на l-ю компоненту, через N шагов преобразования f она снова попадет 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 со некоторым словом вида $\{v_k = w^i_{kN}\}$ составленного из N-х символов слова W_i, то есть совпадает со словом, порожденным динамикой $(\mathbb{S}^1, T_N\alpha_i, \Delta)$. Из предложения и следствия к нему 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 3.15 *Для каждого i выполняется $\alpha = \lambda_i + l_i/k_i$ или $\alpha = -\lambda_i + l_i/k_i$*

Доказательство. Предложение также является следствием предложения 3.9. □

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

Предложение 3.16 *Все дуги одного цвета имеют одинаковую длину.*

Доказательство. Покажем для красного цвета. Так как $\alpha = \lambda_i + l_i/k_i = \lambda_j + l_j/k_j$, то $\lambda_i - \lambda_j \in \mathbb{Q}$, что возможно только когда $\lambda_i = \lambda_j$. Для синего цвета действуем аналогично. □

Предложение 3.17 *На одной компоненте дуги, имеющие общую границу, имеют разный цвет.*

Доказательство. Предположим противное. Пусть какие-то две 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 грасного цвета, то $x_1 + \alpha$ является концом дуги того же цвета, что и (x_1, x_2), а $x_2 + \alpha$ – концом 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, $x_1 + \alpha$ должен быть левым концом дуги этого же цвета, значит, x_1 одновременно является левым концом двух разных дуг. Противоречие.$\square$

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

1. Каждое характеристическое множество U_i представляет собой объединение дуг одной длины, на каждой компоненте середины дуг образуют правильный k_i-угольник (то есть 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-ой к $(i+1)$-ой компоненте. Пусть m – количество красных (а, значит, и синих дуг на одной компоненте)

Предложение 3.18 *При переходе от i-ой к $(i+1)$-ой компоненте дуги красного цвета остаются на месте, дуги синего цвета поворачиваются на $1/m$. Количество компонент N кратно m.*

3.7 Основная теорема о непериодических сбалансированных словах над произвольным алфавитом

Для более наглядного представления основной теоремы будем считать, что окружность на, которой реализуется динамика имеет длину m

Теорема 3.19 *Пусть W– сбалансированное непериодическое слово над алфавитом A. Тогда для W существует динамическая система (M, f), удовлетворяющая следующим условиям:*

1. *$M = \mathbb{S} \times \mathbb{Z}_m$ как топологическое пространство.*

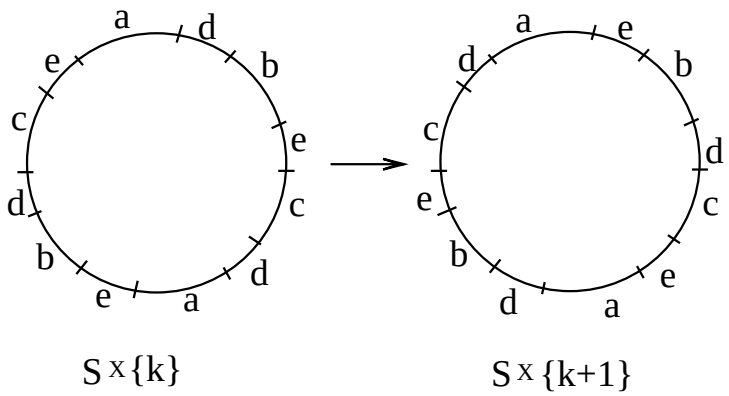
2. *$f : M \to M$ есть композиция поворота на α в $\mathbb{S}$ и сдвига на 1 в $\mathbb{Z}_m$. Длина $\mathbb{S}^1$ равна m.*

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, $k+l = |A|$ – 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 точки).*

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.*

Пример.



Переход от k-ой компоненты к $(k+1)$-ой: три красных "двуугольника" (символы a, b и c) и два синих треугольника (символы d и e).

Замечания.

1. Из вида динамической системы следует, что при "склейке" всех оттенков красного в один цвет, а всех оттенков синего – в другой, получится бинарное слово Штурма.

2. 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 о покрытии множества целых арифметическими прогрессиями, найдутся два одинаковых многоугольника.

3.8 Замечания о периодических сбалансированных словах над произвольным алфавитом

В рассуждениях о непериодических сбалансированных словах мы использовали всюдуплотность траектории для каждого слова W_i. В периодическом случае эти рассуждения не проходят. Однако из имеющейся динамической системы можно сконструировать периодическое сбалансированное слово.

Предложение 3.20 *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 видеть, что в этом случае система будет эквивалентна системе, подобной системе, порождающей слова Штурма, но с рациональной величиной сдвига. По предложению 3.7 это слово является периодическим сбалансированным. □

Возникает естествнный

Вопрос: *любое ли периодическое слово может быть получено таким путем?*

Ответ на этот вопрос отрицательный. Действительно, как следует из замечания к основной теореме, в случае алфавитов более, чем из двух символов, для сбалансированных слов, порождающихя такой динамической системой, обязательно существуют символы, имеющие одинаковую плотность.

Но уже для алфавита из трех символов существует слово с попарно различными плотностями символов:

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

Несложно проверить, что W является сбалансированным. Эту конструкцию можно расширить на алфавиты из большего числа символов рекуррентным образом:

$$W_k = (U_{k-1}a_kU_{k-1}),\ U_3 = a_1a_2a_1a_3a_1a_2a_1 \tag{3}$$

В связи с этим сформулируем гипотезу Френкеля:

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

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

4 Слова с минимальной функцией роста.

В этой главе изучается другое естественное обобщение слов Штурма – слова с минимальной функцией роста, то есть с функцией роста, удовлетворяющей соотношению $F_W(n+1) - F_W(n) = 1$ при всех достаточно больших n.

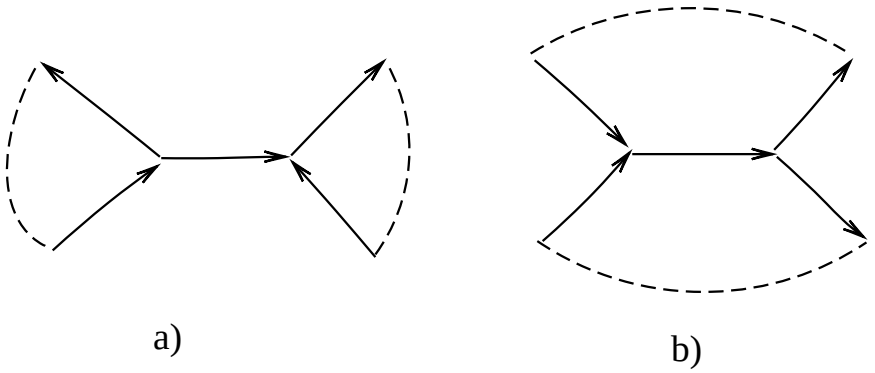
4.1 Постановка задачи.

Другим обобщением слов Штурма являются слова с минимальной функцией сложности. Наша цель – построить динамическую систему, которая аналогичным образом могла бы порождать слова медленного роста. В этой части мы будем рассматривать слова над произвольным алфавитом $A = \{a_1, a_2, \ldots, a_n\}$.

Пусть слово W будет словом медленного роста, то есть, начиная с некоторого N верно $T_W(k+1) = T_W(k) + 1, k \geq N$.

Рассмотрим k-графы этого слова для $k \geq N$. Так как ребра графа соответствуют $(k+1)$-словам, то в этом графе m вершин и $m+1$ ребро и, соответственно, имеет одну входящую и одну выходящую развилку, которые, возможно, совпадают (если данная вершина соответствует биспециальному слову).

Имеется два типа таких графов:



Два типа графов с одной входящей и одной входящей развилкой.

Легко видеть, что граф типа a) не является сильно связным, а в случае b) – является. Слова с k-графами типа b) имеют вид $u^\infty V v^\infty$ и не являются равномерно-рекуррентными

Нас интересует случай a).

Назовем путь, ведущий из входящей развилки в выходящую *перегородкой*, а два пути из выходящей развилки будем называть *дугами*.

Предложение 4.1 *Пусть k-граф слова W имеет перегородку длины $l \geq 1$ и дуги длины r, s. Тогда последователь* $\mathrm{Fol}(G)$ *имеет перегородку длины $l-1$, дуги длины $r+1$ и $s+1$ и совпадает с $(k+1)$-графом слова W.*

Рассмотрим теперь предельный случай, когда перегородка вырождается, то есть входящая развилка совпадает с выходящей. В этом случае развилка соответствует биспециальному слову.

Предложение 4.2 *Последователь графа с вырожденной перегородкой имеет две входящие и две выходящие развилки*

Пусть u – биспециальное подслово W, то есть при некоторых $a_i, a_j, a_r, a_s \in A$ a_iu, a_ju, ua_r, ua_s – тоже подслова W. Тогда развилками (входящими и выходящими) в $\mathrm{Fol}(G)$ будут вершины соответствующие этим словам.

Таким образом, для $(k+1)$-графа слова W имеется 4 возможности для удаления одного ребра из $\mathrm{Fol}(G)$, соответствующего минимальному запрещенному $(k+2)$-слову: $a_iua_r, a_iua_s, a_jua_s, a_jua_r$, В двух случаях мы получаем сильно связный граф, а в двух – не сильно связный.

Непосредственной проверкой доказывается

Предложение 4.3 *Пусть в графе G перегородка вырождена, а дуги имеют длину r, s соответственно. Тогда $(k+1)$-граф имеет вид $(s-1, 1, r+1)$ или $(r-1, 1, s+1)$.*

Предложение 4.4 *Любой граф G с одной входящей и одной выходящей развилкой имеет предшественника, причем только одного.*

Доказательство. Пусть граф имеет вид (l, r, s). Заметим, что графа с $r = s = 1$ не бывает. В случае, если $r, s > 1$ граф предшественника имеет вид $(l+1, r-1, s-1)$, если имеет вид $(l, 1, s), s > 1$, то предшественник – $(0, l+1, s-1)$, если $(l, r, 1), r > 1$ – то $(0, s+1, l-1)$. Граф вида $(0, 1, k)$ имеет предшественника $(0, 1, k-1)$. □

Ясно, что любой граф может быть получен из графа типа $(0, 1, 2)$ или $(0, 2, 1)$. Из предложения 4.4 следует, что для слова с минимальной функцией роста W существует такое слово Штурма V и такие натуральные n и l, что k-графы для слова W совпадают с $(k+l)$-графами слова Штурма V для $k \geq n$.

4.2 Свойства слов с минимальной функцией роста

Предложение 4.5 *Пусть W – слово с минимальной функцией роста, тогда существует натуральное k, что для любого подслова $v \subset W$ такого, что $|v| \geq k$, существует ровно два возвращаемых слова.*

Доказательство. Рассмотрим слово Штурма, такое что эволюция k-графов Рози слова W и графов слова V совпадают, начиная с некоторого n. Тогда k-словам слова W биективно соответствуют $(k+l)$-словам слова V. Поскольку для подслов слова Штурма существует ровно 2 возвращаемых слова, то это же верно и для подслов слова W. □

Из этих же соображений доказывается

Предложение 4.6 *рекуррентное слово с минимальной функцией роста равномерно рекуррентно.*

□

4.3 Эволюции k-графов слов медленного роста

В этой части мы рассматриваем эволюцию k-графов бесконечных слов с минимальным ростом, то есть таких слов W, у которых функция сложности растет минимально: $T_W(n+1) - T_W(n) = 1$, начиная с какого–то n.

В случае, когда $T_W(n+1) - T_W(n) = 1$ для всех $n > 0$, мы получим эволюцию k-графов слов Штурма.

Так как ребра графа соответствуют $(k+1)$–словам, то в таком графе m вершин и $m+1$ ребро и, соответственно, имеет одну входящую и одну выходящую развилку, которые, возможно, совпадают (если данная вершина соответствует биспециальному слову).

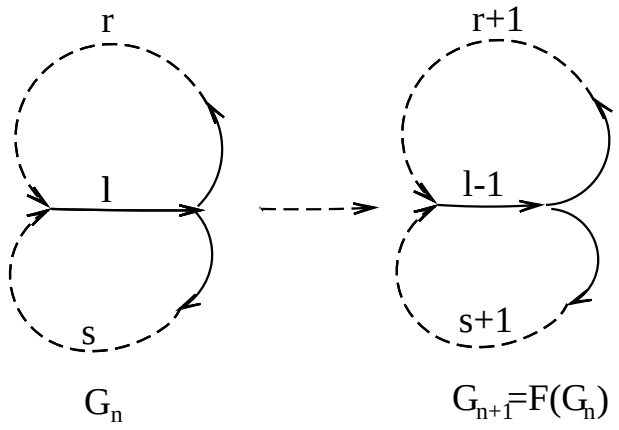
Имеется два типа таких графов: в одном случае граф является сильно связным, в другом – нет.

Слова с k-графами типа b) имеют вид $u^\infty V v^\infty$. Для нас интересным является случай a).

Назовем путь, ведущий из входящей развилки в выходящую *перегородкой*, а два пути из выходящей развилки будем называть *дугами*.

Предложение 4.7 *Пусть k-граф слова W имеет перегородку длины $l \geq 1$ и дуги длины r, s. Тогда последователь* $\mathrm{Fol}(G)$ *имеет перегородку длины $l-1$, дуги длины $r+1$ и $s+1$ и совпадает с $(k+1)$-графом слова W.*

Доказательство. Непосредственно проверяется. □



Переход от G_n к G_{n+1} для невырожденной развилки.

Рассмотрим теперь предельный случай, когда перегородка вырождается, то есть входящая развилка совпадает с выходящей. В этом случае развилка соответствует биспециальному слову.

Предложение 4.8 *Последователь графа с вырожденной перегородкой имеет 2 входящие и две выходящие развилки*

Пусть u – биспециальное подслово W, то есть при некоторых $a_i, a_j, a_r, a_s \in A$ a_iu, a_ju, ua_r, ua_s – тоже подслова W. Тогда развилками (входящими и выходящими) в $\mathrm{Fol}(G)$ будут вершины соответствующие этим словам.

Таким образом, для $(k+1)$–графа слова W имеется четыре возможности для удаления одного ребра из $\mathrm{Fol}(G)$, соответствующего минимальному запрещенному $(k+2)$-слову: $a_iua_r, a_iua_s, a_jua_s, a_jua_r$, В двух случаях мы получаем сильно связный граф, а в двух – не сильно связный.

Предложение 4.9 *Пусть в графе G перегородка вырождена, а дуги имеют длину r, s соответственно. Тогда $(k+1)$-граф имеет вид $(s-1, 1, r+1)$ или $(r-1, 1, s+1)$.*

Предложение 4.10 *Любой граф G с одной входящей и одной выходящей развилкой имеет предшественника, причем только одного.*

Доказательство. Пусть граф имеет вид (l, r, s). Заметим, что графа с $r = s = 1$ не бывает. В случае, если $r, s > 1$ граф предшественника имеет вид $(l+1, r-1, s-1)$, если имеет вид $(l, 1, s), s > 1$, то предшественник есть $(0, l+1, s-1)$, если $(l, r, 1), r > 1$ – то $(0, s+1, l-1)$.

Граф вида $(0,1,k)$ имеет предшественника $(0,1,k-1)$.

Ясно, что любой граф может быть получен из графа типа $(0,1,2)$ или $(0,2,1)$.

4.4 Частоты подслов в словах медленного роста

Напомним, что *частотой* какого-либо подслова $w \subset W$ бесконечного слова W называется следующий предел (если он существует):

$$\mu(w) = \lim_{n\to\infty} \mu_n(w), \tag{4}$$

где $\mu_n(w)$ – количество вхождений подслова w в префикс W длины n.

В случае слов Штурма предел существует всегда и равен длине характеристического интервала для данного подслова, что следует из равномерной распределенности последовательностей вида $\{\alpha n + \beta\}$ для иррациональных α. Докажем полезное

Предложение 4.11 *Пусть W – слово с минимальной функцией роста и пусть (v, a_i, u) – ребро в его k–графе, где $k \geq N$. Тогда:*

1. *Любое конечное подслово слова W обладает частотой, то есть существует предел (4).*
2. *Если v не правое специальное подслово, а u – не левое специальное, то $\mu(u) = \mu(v)$.*

Доказательство. Из определения k графов следует, что существует символ $a_j \in A$ такой, что $ua_i = a_j v \in F_{n+1}(W)$. Так как $ua_k, a_l v \notin F_W(n+1)$ для любых $a_k \neq a_i$ и $a_l \neq a_j$, то $\mu(u) = \mu(ua_i) = \mu(a_j v) = \mu(v)$. □

Следующее утверждение является обобщением т.н. *теоремы о трех расстояниях* (three distance theorem).

Теорема 4.12 *Пусть W – слово медленного роста над произвольным алфавитом A с функцией роста $T_W(n)$ как в предложении 4.11. Тогда для любого $n \geq N$ подслова длины n могут принимать не более трех значений. Если они принимают три значения, то одно из них является суммой двух других.*

Эволюция графов с одной развилкой может быть обобщена на случай слов с функцией сложности $T_W(n)$, удовлетворяющей соотношению

$$\lim_{n\to\infty} T_W(n)/n = 1$$

.

4.5 Конструкция динамической системы

Теперь мы покажем, как по соответствующему графу слов построить динамическую систему, которая порождала бы данное слово.

Построение динамической системы мы осуществим двумя способами. Первый способ опирается на уже известный результат о словах Штурма. Второй способ – более конструктивный и обобщающийся на случай нескольких развилок.

Итак, пусть слово W имеет минимальный рост, т.е., с некоторого k его k-графы слова W имеют вид а).

Поскольку каждый граф имеет единственного предшественника, то существует последовательность графов типа а) $G'_1, G'_2, \ldots, G'_n$ таких, что G'_{l+1} является предшественником G'_l, а граф G'_n совпадает с k-графом слова W.

Иными словами, существует эволюция k-графов типа а), начиная с $k = 1$, что n-ый граф в эволюции совпадает с k-графом слова W, $(n+1)$-ый совпадает с $(k+1)$-графом слова W и т.д.

Последовательность графов такого типа однозначно соответствует некоторому слову Штурма V.

Значит, существует взаимнооднозначное соответствие между k-словами слова W и n-словами слова Штурма V, при этом соответствие продолжается на $(k+1)$-слова W и $(n+1)$-слова V и т.д.

Разберем сначала случай, когда $k = 1$, то есть уже 1-граф слова W имеет вид а). Символам алфавита A взаимнооднозначно соответствуют n-слова некоторого слова Штурма V. По теореме эквивалентности, слово V порождается сдвигом окружности T_α. Как показано в части 2, n - словам в динамике соответствуют интервалы разбиения: слову $w = w_1 w_2 \ldots w_n$ $(w_i \in \{a, b\})$ соответствует интервал

$$I_w = T_\alpha^{(-n+1)}(I_{w_1}) \cup T_\alpha^{(-n+2)}(I_{w_2}) \cup \ldots \cup I_{w_n},$$

где I_{w_i} – характеристичесий интервал для w_i.

Для сдвига окружности интервалы I_w, соответствующие n-словам, будут иметь вид $I_w = (n_i\alpha, n_{i+1}\alpha)$, где n_i – некоторые целые числа.

Тогда характеристическим множеством для каждого символа из алфавита A будет являться интервал I_w, такой, что слово w соответствует данному символу. Ясно тогда, что слово W будет порождаться тем же сдвигом T_α и характеристическими множествами I_w.

Теперь разберем общий случай. Пусть k-словам слова W соответствуют n-слова слова Штурма V. Точно также мы можем построить соот-

ветствие между интервалами разбиения для n-слов слова Штурма I_w и k-словами слова V. Построим характеристическое множество для каждого символа из A.

А, именно, символу a_i поставим в соответствие все интервалы, которые соответствуют словам, начинающимся на a_i. В этом случае характеристическое множество для произвольного символа может быть несвязным и представлять собой объединение нескольких интервалов.

Покажем, что при таком выборе характеристических множеств найдется точка, чья эволюция будет совпадать с W.

Действительно, пусть слово W начинается с некоторого k-слова $w = w_1 w_2 \cdots w_k$. Ему мы поставим в соответствие интервал I_w. Рассмотрим следующий $k+1$ символ w_{k+1}; этому $(k+1)$-слову мы поставим в соответствие интервал $I_w \cup I_{w'}$, где $w' = w_2 w_3 \cdots w_{k+1}$, и т.д.

Бесконечное пересечение интервалов, соответствующих префиксам данного слова будет содержать точку, эволюция которой совпадает с W.

Таким образом, доказана следующая

Теорема 4.13 *Пусть W – рекуррентное слово над произвольным конечным алфавитом A. Тогда следующие условия на слово W эквивалентны:*

1. *Существует такое натуральное N, что функция сложности слова W равна $T_W(n) = n + K$, для $n \geq N$ и некоторого постоянного натурального K.*

2. *Существуют такое иррациональное α и целые $n_1, n_2, \ldots, n_m$, что слово W порождается динамической системой*

$$(\mathbb{S}^1, T_\alpha, I_{a_1}, I_{a_2}, \ldots, I_{a_n}, x),$$

где $T\alpha$ – сдвиг окружности на иррациональную величину α, I_{a_i} – объединение дуг вида $(n_j\alpha, n_{j+1}\alpha)$.

5 Перекладывания отрезков.

5.1 Введение и постановка задачи.

Данная глава посвящена изучению слов, получаемых из перекладывания отрезков. Каждое такое слово W удовлетворяет сотношению:

$$F_W(n+1) - F_W(n) = k,\ k \geq N,$$

которое означает, что количество n-подслов для достаточно больших n увеличивается не постоянную величину k. Мы ограничиваемся случаем, когда W рекуррентно. Наша цель – получить комбинаторный критерий порождаемости слова W перекладыванием отрезков.

Напомним (см. 2.10), что если перекладывание k отрезков является регулярным, то эволюция любой точки является словом с линейной функцией сложности: $T_{U(x)}(n) = n(k-1)+1$.

5.2 Необходимые условия для порождаемости слова перекладыванием отрезков

Пусть слово W является эволюцией точки $x \in [0,1]$ при перекладывании k отрезков и характеристичеких множествах $U_1, U_2, \ldots, U_n$. Каждое характеристическое множество U_i является объединением нескольких непересекающихся интервалов или полуинтервалов.

Как было показано в части 2.9, подслова длины k взаимнооднозначно соответствуют k-разбиениям характеристических множеств. Поскольку точка границы одномерного множества может являться границей только для двух множеств, то k-подслово слова W может иметь максимум два продолжения. Мы получили первое необходимое условие для того, чтобы слово порождалось перекладыванием отрезков:

Предложение 5.1 *Пусть слово W порождается перекладыванием отрезков. Тогда для некоторого N все специальные слова подслова длины, не меньше N должны иметь валентность, равную* 2.

Это условие аналогично тому, что начиная с какого-то N все k-графы слова W ($k \geq N$) должны иметь входящие и исходящие развилки степени 2.

Следующие условия мы получим на языке графов Рози. Пусть рекуррентное слово W имеет функцию роста $F_W(n) = Kn + L$ для $n > N$ и порождается перекладыванием отрезка. Рассмотрим эволюцию k-графов

Рози слова W, начиная с $k \geq N + 1$. Как было показано выше, все входящие и выходящие развилки имеют степень 2, поэтому у всех k-графов есть ровно K входящих и K исходящих развилок.

В минимальном случае, описанном в предыдущей главе, когда $F_W(n) = n + L$, графы Рози имеют ровно одну входящую и одну исходящую развилку. Если входящая и выходящая развилка совпадают, то в этом случае выбор удаляемого ребра в последователе $D(G)$ определяется однозначно из условия сильной связанности графа.

Более интересным случаем являются графы слов, в которых есть более одной развилки. Рассмотрим его подробно. При переходе от графа G_n к G_{n+1} возможны следующие варианты:

1. В графе G_n нет сцепленных циклов (то есть, нет входящих развилок, являющихся одновременно исходящими). В этом случае граф G_{n+1} совпадает с последователем $D(G_n)$.

2. В графе G_n одна развилка является одновременно входящей и выходящей. Граф последователя $D(G_n)$ в этом случае имеет три развилки, так как одна развилка размножилась. Следовательно, граф G_{n+1} получается из последователя $D(G_n)$ путем удаления одного ребра, соответствующего минимальному невстречающемуся слову.

3. В графе есть две развилки или более развилок, являющихся одновременно входящими и выходящими. Граф G_{n+1} получается из $D(G_n)$ удаления двух или более ребер, соответствующих минимальным невстречающимся словам.

Поскольку слово W рекуррентно, то из предложения 2.13 следует, что при удалении ребер должна сохраняться сильная связность, то есть из любой вершины можно по стрелкам перейти в любую другую.

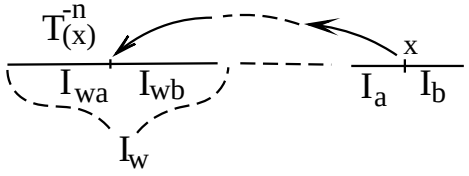
Рассмотрим подробнее второй случай. Пусть в G_k есть одна двойная развилка. Это означает, что в W есть ровно одно биспециальное подслово w длины k. Значит, существуют такие $a_i, a_j, a_k, a_l \in A$, что a_iw, a_jw, wa_k, wa_l – подслова W. Тогда $(k+1)$-граф G_{k+1} получается из последователя путем удаления какого-то ребра, соответствующего одному из четырех слов: $a_iwa_k, a_iwa_l, a_jwa_k, a_jwa_l$. Рассмотрим интервал, являющийся характеристическим множеством для слова $I_w = [x_w, y_w]$.

Так как w – правое специальное слово, то $I_w \subset T^{-1}(I_{a_k} \cup I_{a_l})$, так как w – левое специальное, то $I_w \subset T(I_{a_i} \cup I_{a_j})$.

Пусть точка $A \in [0,1]$ делит I_w на два интервала, образы которых лежат в I_{a_k} и I_{a_l} соответственно, а точка $B \in [0,1]$ – делит на интервалы, прообразы которых лежат в I_{a_i} и I_{a_j} соответственно.

Выбор минимального невстречающегося слова, а , значит, удаляемого ребра, определяется взаиморасположением точек A и B, а также сохранением или сменой ориентации отображения на этих множествах. Итого, имеется 8 вариантов, которые разбиваются на четыре пары, соответствующие одинаковым наборам слов:

1. $B < A$, $T^{-1}([x_w, B]) \subset I_{a_i}$, $T([x_w, A] \subset I_{a_k})$
2. $B < A$, $T^{-1}([x_w, B]) \subset I_{a_j}$, $T([x_w, A] \subset I_{a_k})$
3. $B < A$, $T^{-1}([x_w, B]) \subset I_{a_i}$, $T([x_w, A] \subset I_{a_l})$
4. $B < A$, $T^{-1}([x_w, B]) \subset I_{a_j}$, $T([x_w, A] \subset I_{a_l})$
5. $B > A$, $T^{-1}([x_w, B]) \subset I_{a_i}$, $T([x_w, A] \subset I_{a_k})$
6. $B > A$, $T^{-1}([x_w, B]) \subset I_{a_j}$, $T([x_w, A] \subset I_{a_k})$
7. $B > A$, $T^{-1}([x_w, B]) \subset I_{a_i}$, $T([x_w, A] \subset I_{a_l})$
8. $B > A$, $T^{-1}([x_w, B]) \subset I_{a_j}$, $T([x_w, A] \subset I_{a_l})$



Разбиение характеристического множества для слова w на n-м шаге.

Варианты 1 и 5 соответствуют запрещенному слову $a_i w a_k$.

Варианты 2 и 6 соответствуют запрещенному слову $a_i w a_l$.

Варианты 3 и 7 запрещенному слову соответствуют $a_j w a_k$.

Варианты 4 и 8 соответствуют запрещенному слову $a_j w a_l$.

Две пары вариантов соответствуют одновременной смене или сохранению ориентации отображения на характеристических интервалах, а две пары – различным ориентациям.

В случае, когда перекладывание отрезков сохраняет ориентацию, у нас остается только два возможных варианта для удаляемого ребра.

В случае, когда мы разрешаем отобажению менять ориентацию отрезков при перекладывании, возможны все четыре варианта.

Введем понятие *размеченного графа Рози*. Граф Рози называется *размеченным*, если

1. Ребра каждой развилки помечены символами l (“left”) и r (“right”)
2. Некоторые вершины помечены символом “–”.

Последователем размеченного графа Рози назовем ориентированный граф, являющийся его последователем как графа Рози, разметка ребер которого определяется по правилу:

1. Ребра, входящие в развилку должны быть помечены теми же символами, как и ребра, входящие в любого левого потомка этой вершины;
2. Ребра, выходящие из развилки должны быть помечены теми же символами, как и ребра, выходящие из любого правого потомка этой вершины;
3. Если вершина помечена знаком “–”, то все ее правые потомки также должны быть помечены знаком “–”.

Замечание. Поясним смысл разметки графа. Пусть ребра входящей развилки соответствуют a_i и a_j, символы l и r соответствуют левому и правому множеству в паре $(T(I_{a_i}), T(I_{a_j}))$. Если символы a_k и a_l соответствуют ребрам исходящей развилки, то символы l и r ставятся в соответствии с порядком “лево-право” в паре (I_{a_k}, I_{a_l}). Знак “–” ставится в вершине, если характеристическое множество, ей соответствующее, принадлежит интервалу перекладывания, на котором меняется ориентация.

Условие для перехода от графа G_n к G_{n+1}:

Предложение 5.2 *1. Если в графе нет двойных развилок, соответствующих биспециальным подсловам, то при переходе от G_n к G_{n+1} имеем $G_{n+1} = D(G_n)$;*

2. Если вершина, соответствующая биспециальному слову не помечена знаком “–”, то ребра, соответствующие запрещенным словам выбираются из пар lr и rl

3. Если вершина помечена знаком “–”, то удаляемые ребра должны выбираться из пары ll или rr.

Назовем эволюцию размеченных графов Рози *правильной*, если **правила 1** и **2** выполняются для всей цепочки эволюции графов, начиная с G_1, назовем эволюцию *асимптотически правильной*, если **правила 1** и **2** выполняются, начиная с некоторого G_n. Будем говорить, что эволюция размеченных графов Рози *ориентированна*, если в k-графах нет вершин, помеченных знаком "–".

С помощью введенных определений мы можем сформулировать следующие условия на слово, порождающееся перекладыванием отрезков:

Теорема 5.3 *Пусть рекуррентное слово W*

1. *порождается перекладыванием отрезков. Тогда это слово обеспечивается асимптотически правильной эволюцией размеченных графов Рози.*

2. *порождается перекладыванием отрезков с сохранением ориентации. Тогда это слово обеспечивается асимптотически правильной ориентированной эволюцией размеченных графов Рози.*

5.3 Построение динамической системы

Покажем, что условия теоремы 5.3 являются достаточными для того, чтобы слово порождалось перекладыванием отрезков. Сначала мы покажем, что слово W, удовлетворяющее условиям теоремы 5.3 может являться эволюцией некоторой точки при кусочно-непрерывном преобразовании отрезка $T : I \to I$ следующего вида:

1. $I = [x_0, x_1] \cup [x_1, x_2] \cup \ldots \cup [x_{n-1}, x_n]$, $x_0 = 0$, $x_n = 1$.

2. $I = [y_0, y_1] \cup [y_1, y_2] \cup \ldots \cup [y_{n-1}, y_n]$, $y_0 = 0$, $y_n = 1$.

3. $\sigma \in S_n$ – некоторая перестановка из n элементов.

4. T отображает (x_i, x_{i+1}) на $(y_{\sigma(i)}, y_{\sigma(i)+1})$ непрерывно и взаимноодно-значно.

Затем будет показано, как случай кусочно-непрерывного преобразования можно свести к перекладыванию отрезков.

Строить кусочно-непрерывное преобразование T мы будем поэтапно. При первой итерации мы разбиваем отрезок произвольным образом на

отрезки, которые соответствуют соответствующим символам. Для построение отображения нам достаточно определить траекторию этих точек и затем из соображений рекуррентности продолжить по непрерывности на весь отрезок.

Замечание. Из непрерывности и биективности следует, что отображения на интервалах являются монотонными функциями. Мы рассмотрим два случая:

1. Все отображения интервалов являются одновременно возрастающими функциями. Назовем такое преобразование *сохраняющим ориентацию.*

2. Во втором случае встречаются как возрастающие, так и убывающие отображения интервалов. Этот случай назовем *не сохраняющим ориентацию.*

Разобьем отрезок $I = [0, 1]$ на $n = \operatorname{Card} A$ интервалов произвольным образом, которые будут являться характеристическими множествами для символов алфавита A: $[0, 1] = I_{a_1} \cup I_{a_2} \cup \ldots \cup I_{a_n}$. Соответствие интервалов символам алфавита будет определено в соответствии с предложением 5.5

Будем считать, что отображение T непрерывно на каждом множестве I_{a_i}, то есть точками разрыва отображения T могут быть только конечные точки характеристических множеств. Интервалы характеристических множеств или их образы, имеющие общую граничную точку, будем называть *соседними.*

Замечание. Мы всегда можем расширить алфавит таким образом, чтобы в новом алфавите характеристические множества были устроены ровно так, как описано выше. Графы k-слов для старого алфавита будут соответствовать 1-графам нового алфавита, а эволюции графов, начиная с этого момента, будут совпадать.

Непосредственно проверяется следующее

Предложение 5.4 *Преобразования* $, T^{(-1)}$ *переводит два соседних множества либо в соседние множества, либо их образы не могут целиком покрывать интервал.*

Предложение 5.5 *Интервалы разбиения могут быть поставлены в соответствие символам алфавита таким образом, что если* w *– специальное правое* 1*-слово и* wa_i, wa_j *– подслова, то* I_{a_i} *и* I_{a_j} *– соседние*

множества. Аналогично, если w – специальное левое 1-слово и $a_k w, a_l w$ – подслова, то I_{a_k} и I_{a_l} – также соседние множества.

Доказательство. Рассмотрим произвольную развилку в 1-графе. Пусть из нее стрелки ведут в вершины, соответствующие символам a_i и a_j. Тогда положим $I_{a_i} = [x_0, x_1]$, $I_{a_j} = [x_1, x_2]$,

Так как характеристические множества являются интервалами, то на них можно естественным образом ввести отношение порядка (то есть $I_{a_i} < I_{a_j}$, если $x_i < x_j$), точно также можно ввести отношение порядка и на их образах. Если при преобразовании T для какой-то пары характеристических множеств порядок меняется, будем говорить, что на этой паре преобразование меняет ориентацию.

Рассмотрим образы интервалов при отображении T^{-1}. Ясно, что если символ a_i не является правым специальным 1-словом и за ним однозначно следует символ a_j, то $I_{a_i} \subset T^{-1}(I_{a_j})$, а если не является левым специальным и ему всегда предшествует символ a_k, то $T^{-1}(I_{a_i}) \subset I_{a_k}$.

В случае, если символ a_i является правым специальным 1-словом и за ним могут идти символы a_j и a_k, то $I_{a_i} \subset T^{-1}(I_{a_j} \cup I_{a_k})$, а если левым специальным, то $T^{-1}(I_{a_i}) \subset I_{a_j} \cup I_{a_k}$.

Обозначим через I^n множество образов концов интервалов при отображении T^{-n}, I^0 – множество концов интервалов характеристических множеств, то есть $I^0 = \{x_0, x_1, \ldots, x_n\}$.

Как следует из рассуждений в части 2.9, множество, отвечающие слову $w = w_1 w_2 \cdots w_n$ есть $I_w = I_{w_n} \cap T^{-1}(I_{w_{n-1}}) \cap \ldots \cap T^{(-n+1)}(I_{w_1})$, соответственно, множество граничных точек множеств, соответствующих словам длины n есть в точности $I^0 \cup I^1 \cup \ldots \cup I^{(n-1)}$.

Если правое специальное слово не является биспециальным (то есть не является одновременно и левым специальным), то положение точки, которая делит данное характеристическое множество, несущественно и его можно выбирать произвольно.

Для случая сохраняющего ориентацию преобразования удаления должно соответствовать сохраняющим ориентацию правилам.

В случае не сохраняющего ориентацию преобразования нам просто необходимо, чтобы внутри отображаемых отрезков при эволюции происходило конечное число “переломов”.

Таким образом мы можем определить образы точек на некотором подмножестве $N \subset I$. Из построения следует, что существуют такие интервалы $I_k = (x_k, x_{k+1})$, внутри которых наше преобразование монотонно. Мы всегда можем продолжить его по непрерывности до отображения

отрезка в себя. Построенное кусочно-непрерывное преобразование и есть искомое. Обозначим его T. Отметим, что начальная точка, эволюцией которой является искомое слово $W = \{w_n\}$, получается как предельная последовательность множества вложенных отрезков, соответствующих префиксам $w_0, w_0w_1, w_0w_1w_2, \ldots$ и т.д.

Нами доказана

Теорема 5.6 *Чтобы рекуррентное слово W порождалось кусочно-непрерывным преобразованием с сохранением ориентации, необходимо и достаточно чтобы слово обеспечивалось асимптотически правильной ориентированной эволюцией размеченных графов Рози.*

В случае меняющего ориентацию преобразования получается

Теорема 5.7 *Чтобы рекуррентное слово W порождалось кусочно-непрерывным преобразованием с сохранением ориентации, необходимо и достаточно чтобы слово обеспечивалось асимптотически правильной эволюцией размеченных графов Рози.*

5.4 Эквивалентность множества р.р слов, порождаемых кусочно-непрерывным преобразованием, множеству слов, порождаемых перекладыванием отрезков.

Покажем вначале, как можно перейти к динамике, в которой почти все точки (в смысле меры Лебега) имеют различные существенные эволюции. Рассмотрим существенные эволюции точек при преобразовании T.

Рассмотрим кусочно-непрерыное преобразование отрезков. Из теоремы о существовании инвариантной меры (см. [25]) следует, что любое отображение компакта имеет инвариантную вероятностную меру. Значит, отображение T имеет инвариантную меру μ. Следовательно, на отрезке мы можем ввести новую метрику $d(x_1, x_2) = \mu((x_1, x_2))$. Отметим, что из того, что $\mu(x_1, x_2) > 0$ не следует, что точки имеют различную существенную эволюцию, поскольку построенное отображение может быть разрывным. Кроме того, нам потребуется подобрать подходящую меру.

Назовем разбиение отрезка на характеристические множества *хорошим*, если для каждого символа его характеристическое множество является выпуклым (то есть представляет собой отрезок, интервал или полуинтервал) и если между двумя точками одноцветного множетсва есть точка разрыва – то их цвета их образов различны.

Для произвольного разбиения $U_1, U_2, \ldots, U_n$ назовем его *доразбиением* любое разбиение $V_1, V_2, \ldots, V_m$ такое, что каждое характеристическое множество U_i является объединением нескольких множеств V_j: $U_i = V_{i_1} \cup V_{i_2} \cup V_{i_j}$.

Эволюция доразбиения есть слово над алфавитом из большего числа букв и эволюция разбиения получается из него путем склейки букв. Ясно, что любое разбиение имеет хорошие доразбиения. Следующее предложение мы сформулируем для хороших разбиений.

Предложение 5.8 *Пусть точки $x_1 < x_2$ имеют одинаковые эволюции и преобразование T непрерывно на интервале (x_1, x_2). Тогда любая точка интервала (x_1, x_2) имеет такую же эволюцию.*

Доказательство. Из того, что x_1, x_2 имеют одинаковую эволюцию следует, что $T(x_1), T(x_2)$ тоже имеют одинаковую эволюцию и образом интервала (x_1, x_2) является интервал $(T(x_1), T(x_2))$. □

Замечание. Рассмотрим топологию, склеивающую точки с одинаковой существенной эволюцией и построим соответствующую фактор-динамику. Нашем случае целые отрезки окажутся склеенными. Однако таких отрезков не более, чем счетное множество. В то же время, если W – непериодическое равномерно-рекуррентное слово, то количество эквивалентных (то есть имеющих одинаковый набор конечных подслов) равномерно-рекуррентных слов имеет мощность континуума (см. [7]). Это означает, что существует несчетное множество точек вне этой системы отрезков и сингулярность меры на эволюцию этих точек влияние не оказывает. Таким образом, найдется эквивалентное слово W' W, отвечающее перекладыванию отрезков. Но тогда в этом перекладывании встретятся *все* слова , эквивалентные W', в том числе и слово W (в смысле существенной эволюции).

Пусть W – непериодическое р.р. слово, наблюдающееся при кусочном гомеоморфизме системы отрезков. Рассмотрим пространство последовательностей с оператором сдвига. В этом пространстве есть минимальное замкнутое инвариантное множество N_W, отвечающее W [7]. Каждой точке N_W отвечает множество точек системы отрезков с данной существенной эволюцией. Далее, на пространстве N_W имеется вероятностная инвариантая относительно оператора сдвига мера, которая индуцирует на системе отрезков естественным образом вероятностную инвариантую меру.

Рассмотрим топологию, склеивающую точки с нулевым расстоянием и построим соответствующую фактор-динамику. Полученное отображение и будет кусочной изометрией, т.е. перекладыванием отрезков.

Доказана следующая

Теорема 5.9 *Множество слов, порождаемых кусочно-непрерывным преобразованием отрезка эквивалентно множеству слов, порождаемых преобразованием перекладывания отрезка.*

Как следствие, мы получаем следующие теоремы:

Теорема 5.10 *Чтобы равномерно-рекуррентное слово W порождалось перекладыванием отрезка с сохранением ориентации, необходимо и достаточно чтобы слово обеспечивалось асимптотически правильной ориентированной эволюцией размеченных графов Рози.*

В случае меняющего ориентацию перекладывания доказана

Теорема 5.11 *Чтобы равномеро-рекуррентное слово W порождалось перекладыванием отрезков, необходимо и достаточно чтобы слово обеспечивалось асимптотически правильной эволюцией размеченных графов Рози.*

Список литературы

[1] С.И. Адян. Проблема Бернсайда и тождества в группах // М., Наука, 1975.

[2] Адян С. И., Разборов А. А. Периодические группы и алгебры Ли. Успехи мат. наук, 1987, т. 42, no 2, стр. 2–68.

[3] С.В. Алешин, О свободной группе конечного автомата.//Вестник Моск. Унив. Сер 1. Мат. и Мех.1983, №4, 12–14.

[4] С.В. Алешин, О суперпозициях автоматных отображений // Кибернетика, Киев, 1975, №1, 29–34.

[5] С.В. Алешин, Конечные автоматы и проблема Бернсайда для периодических групп // Мат. Заметки 11 (1972), 319–328.

[6] А.Я.Белов, Классификация слабо нетеровых мономиальных алгеб // Фунд. и прикл. мат. – 1995.–1.,б4, –С. 1085-1089.

[7] А.Я.Белов, В.В.Борисенко, В.Н.Латышев, Мономиальные алгебры // Итоги науки и техники. Совр. Мат. Прил. Тем. Обзоры т. 26 (алг. 4), М. 2002. 35-214.

[8] А.Белов, Г.Кондаков, Обратные задачи символической динамики // Фундаментальная и прикладная математика, Т1, №1, 1995

[9] Голод Е. С., Шафаревич И. Р. О башне полей классов. Изв. АН СССР. Сер. мат., 964, т. 28, no 2, стр. 261–272.

[10] Голод Е. С. О нильалгебрах и финитно-аппроксимируемых p-группах. Изв. АН СССР. Сер. мат., 1964, т. 28, no 2, стр. 273–276.

[11] Григорчук Р. И. К проблеме Милнора о групповом росте. Докл. АН СССР, 1983, т. 271, no 1, стр. 53–54.

[12] Григорчук Р. И. Степени роста конечно порожденных групп и теория инвариантных средних. Изв. АН СССР. Сер. мат., 1984, т. 48, no 5, стр. 939–985.

[13] Григорчук Р. И. О рядах Гильберта-Пуанкаре градуированных алгебр, ассоциированных с группами. Мат. сб., 1989, т. 180, no 2, стр. 207–225.

[14] Каргаполов М.И. Мерзляков Ю.И. Основы теории групп, (3е изд., Наука, 1982) 288стр.

[15] А.Т. Колотов, Апериодические последовательности и функции роста в алгебрах // Алгебра и логика, 20 (1981), 138-154, 250.

[16] А.Т. Колотов, Алгебры и группы с периодической функцией роста // Алгебра и логика, 19 (1980), 659-668, 745.

[17] Кострикин А. И. Вокруг Бернсайда. — М.: Наука, 1986, 232 стр.

[18] В. Б. Кудрявцев, С. В. Алешин, А. С. Подколзин, Введение в теорию автоматов // Москва, "Наука", 1985. 320 стр.

[19] Р.Линдон ,П.Шупп, Комбинаторная теория групп // М., Мир, 1980.

[20] Михалев А. А. Лемма о слиянии и проблема равенства для цветных супералгебр Ли. Вестн. МГУ. Серия 1, 1989, num 5, стр. 88–91.

[21] Михалев А. А. Техника композиции А. И. Ширшова в супералгебрах Ли (некоммутативные базисы Гребнера). Тр. семинара им. И. Г. Петровского, 1994, т. 18,

[22] Мищенко С. П., Вариант теоремы о высоте для алгебр Ли. Мат. заметки, 1990, ;7, no 4, стр. 83–89.

[23] Ольшанский А. Ю. Геометрия определяющих соотношений в группах. сер. Соврем.алгебра. М.: Наука, 1989, 447 стр.

[24] Саломаа А. Жемчужины формальных языков. — М.: Мир, 1986, 159 стр.

[25] Я.Г. Синай, Введение в эргодическую теорию // М.: ФАЗИС, 1996. 144 с.

[26] Уфнаровский В.А. Комбинат. и асимпт. методы в алгебре. // Итоги науки и тех. Серю Совр. Пробл. Мат. Фунд. направл. М. ВИНИТИ. 1990- 57, стр 5-177. (РЖМат, 1990).

[27] Ширшов А. И. О некоторых неассоциативных ниль-кольцах и алгебраических алгебрах. Мат. сб., 1957, т. 41, no 3, 381–394.

[28] Ширшов А. И. О кольцах с тождественными соотношениями. Мат. сб., 1957, т. 43, no 2, стр. 277–283.

[29] Ширшов А. И. О свободных кольцах Ли. Мат. сб., 1958, т. 45, no 2, стр. 113–122.

[30] Ширшов А. И. О базах свободной алгебры Ли. Алгебра и логика, 1962, т. 1, no 1, стр. 14–19.

[31] V.I. Arnold: Small denominators and problems of stability of motion in classical and celestial me- chanics, Usp. Math. Nauk. 18 (1963), p. 91-192, (in Russian) translated in Russian Math. Surveys 18 (1963), p. 86-194.

[32] A.B. Katok, A.M. Stepin: Approximations in ergodic theory, Usp. Math. Nauk. 22 (1967), p. 81-106 (in Russian), translated in Russian Math. Surveys 22 (1967), p. 76-102.

[33] V.I. Oseledec, The spectrum of ergodic automorphisms, Dokl. Akad. Nauk. SSSR 168 (1966), p. 1009-1011 (in Russian).

[34] A.Aberkane, Words whose complexity satisfies $\lim p(n)/n = 1$ // Theor. Comp. Sci., 307, (2003), 31-46.

[35] P. Arnoux and G. Rauzy [1991], Representation geometrique des suites the complexite $2n + 1$ // Bull. Soc. Math. France 119, 199-215.

[36] A.Belov, T. Gateva-Ivanova, Radicals of monomial algebras // Proceedings of Taiwan-Moscow Algebra Workshop, – C. 159-169.

[37] A. Belov. Some estimations for nilpotence of nil-algebras over a field of an arbitrary characteristic and height theorem // *Communications in algebra*, 20 (10):2919-2922, 1992.

[38] G.M. Bergman, A note on growth functions of algebras and semigroups // Research Note, University of California, Berkeley, 1978.

[39] V. Berthé, S. Ferenczi, L.Zamboni: Interactions between dynamics, arithmetics and combina- torics: the good, the bad, and the ugly, AMS Contemporary Math. 385 (2005), p. 333-364.

[40] J. Berstel, P. Séébold, Sturmian words, in: M. Lothaire (Ed.) // Algebraic Combinatorics on Words, Encyclopedia of Mathematics and Its Applications, Vol. 90, Cambridge University Press, Cambridge, 2002 (Chap. 2).

[41] J.Berstel, Resent results on Sturmian words // Developments in language theory II, 13-24, World Scientific, 1996.

[42] X. Droubay, G. Pirillo, Palindromes and sturmian word// Theroret. Comput. Sci., 223:73-85, 1999.

[43] X.Droubay, J.Justin, G.Pirillo, Episturmian words and some construction of de Luca and Rauzy // Theoret. Comp. Sci.,2000, to appear.

[44] H. Furstenberg. Poincaré reccurence and number theory // Bull. Amer. Math. Soc., 5:211-234, 1981.

[45] R. L. Graham, Covering the Positive Integers by disjoints sets of the form $\{[n\alpha + \beta] : n = 1, 2, \ldots\}$ // J. Combin. Theory Ser A15 (1973) 354-358.

[46] P. Hubert, Well balanced sequences // Theoret. Comput. Sci. 242 (2000) 91–108.

[47] S. Ferenczi, C. Holton, L. Zamboni, The structure of three-interval exchange transformations II: a combinatorial description of the trajectories// J. Anal. Math. 89 (2003), p. 239-276.

[48] Ferenczi, L. Zamboni, A new induction for symmetric k-interval exchange transformations and distances theorem, submitted, http://iml.univ-mrs.fr/ ferenczi/fz1.pdf

[49] Examples of 4-interval exchange transformations, preprint (2006), http://iml.univ-mrs.fr/ ferenczi/fz2.pdf

[50] Ferenczi, L. Zamboni, Languages of k-interval exchange transformations, submitted, http://iml.univ-mrs.fr/ ferenczi/fz3.pdf

[51] Krause, G.; Lenagan, T.H.: Growth of algebra and Gelfand-Kirillov dimension // Research Notes in Math., Pitman, London, 1985.

[52] A. de Luca, Sturmian words: structure, combinatorics and their arithmetics // Theoret. Comp. Sci., 183, (1997), 45-82.

[53] A. de Luca and S. Varricchio. Combinatorial properties of uniformly recurrent words and an application to semigroups. // *Inter. J. Algebra Comput.*, 1(2):227–246, 1991.MR 92h:20084.

[54] M. Lothaire, Combinatorics on Words // Encyclopedia of Mathematics and its Applications, Addison-Wesley, Reading, MA, 1983, Vol. 17.

[55] M. Lothaire, Algebraic combinatorics on words // A collective work by Jean Berstel, Dominique Perrin, Patrice Seebold, Julien Cassaigne, Aldo De Luca, Steffano Varricchio, Alain Lascoux, Bernard Leclerc, Jean-Yves Thibon, Veronique Bruyere, Christiane Frougny, Filippo Mignosi, Antonio Restivo, Christophe Reutenauer, Dominique Foata, Guo-Niu Han, Jacques Desarmenien, Volker Diekert, Tero Harju, Juhani Karhumaki and Wojciech Plandowski. With a preface by Berstel and Perrin. Encyclopedia of Mathematics and its Applications, 90. Cambridge University Press, Cambridge, 2002, 504 pp.

[56] M. Morse and G. A. Hedlund [1940], Symbolic dynamics II. Sturmian trajectories, // Amer. J. Math. 62, 1-42.

[57] G. Rauzy, Mots infnis en arithmetique, in: Automata on Infnite Words // Ecole de Printemps d'Informatique Thfeorique, Le Mont Dore, May 1984, ed. M. Nivat and D. Perrin, Lecture Notes in Computer Science, vol. 192, Springer-Verlag, Berlin etc., pp. 165-171, 1985

[58] G. Rauzy, Suites á termes dans un alphabet fini // In *Sémin. Théorie des Nombres,* p. 25-01-25-16, 1982-1983, Exposé No 25.

[59] G. Rauzy, Échanges d'intervalles et transformations induites, (in French), Acta Arith. 34 (1979), p. 315-328.

[60] G. Rozenberg and A. Salomaa // The Mathematical Theory of L Systems, Academic Press, New York etc., 1980

[61] G.Rote, Sequences with subword complexity $2n$ // J. Number Theory 46 (1994) 196–213.

[62] Arto Salomaa.*Jewels of Formal Language Theory* // Computer Science Press, 1981.

[63] R.Tijdeman, Decomposition of the integers as a direct sum of two subsets // in: Number Theory, ed. by S. David, Number Theory Seminar Paris 1992-93, Cambridge University Press, (1995), 261-276

[64] R.Tijdeman, Fraenkel's conjecture for six sequences // Discrete Mathematics, Volume 222, Issue 1-3, 223 - 234, 2000

[65] L. Vuillon, Balanced words // Bull. Belg. Math. Soc. Simon Stevin 10 (2003), no. 5, 787-805

[66] L.Vuillon, A characterization of Sturmian word by return words // European Journal of Combinatorics (2001) 22, 263-275.

[67] H.Weyl, Über der gleichverteilung von zahlen mod 1, // Math. Ann., v. 77, 313-352, 1916.