Сибирский математический журнал
Январь—февраль, 2011. Том 52, № 1

УДК 519.115.8

О ВОЗМОЖНЫХ СКОРОСТЯХ РОСТА ЯЗЫКОВ ТЁПЛИЦА

Ж. Кассень, Ф. В. Петров, А. Э. Фрид

Аннотация. Рассматривается новое семейство факторных языков, комбинаторная сложность которых растет как $\Theta(n^\alpha)$, где α — корень некоторого трансцендентного уравнения. Асимптотический рост функции сложности исследуется с применением аналитических методов, в частности, следствия из теоремы Винера — Питта. Рассматриваемые факторные языки являются языками арифметических подслов бесконечных слов; таким образом, описывается новое семейство бесконечных слов с необычным ростом арифметической сложности.

Ключевые слова: комбинаторная сложность, арифметическая сложность, комбинаторика на словах, слова Тёплица, асимптотическая комбинаторика, аналитические методы в комбинаторике, тауберовы теоремы, теорема Винера — Питта.

1. Введение

Исследование началось как попытка построить бесконечное слово с нестандартной скоростью роста арифметической сложности. Вычисление асимптотики функции сложности на основе рекуррентного соотношения потребовало применения нетривиальных и неэлементарных аналитических методов (тауберовой теории Винера). Таким образом, помимо основного комбинаторного результата (теоремы 1) представляет интерес аналитический способ его получения. Семейство слов, рассматриваемых в данной работе, строится с помощью достаточно простой конструкции, а их арифметическая сложность растет как $\Theta(n^\alpha)$, где α — корень некоторого трансцендентного уравнения. Это устанавливается элементарными методами (лемма 4).

Для получения главного члена асимптотики применяется следствие теоремы Винера — Питта (теорема 3).

2. Определения, обзор и результаты

Мы рассматриваем конечные и бесконечные слова над алфавитом Σ мощности $d \geq 2$. Множество конечных слов называется *языком*; в частности, можно говорить о языке подслов $\mathrm{Fac}(w)$ бесконечного слова w. Такой язык всегда является *факторным*, т. е. замкнут относительно взятия подслов; *подслово* v слова u здесь определяется как всякое слово, для которого $u = svt$ для некоторых s

---

Работа второго автора выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 08–01–00379) и Совета по грантам президента РФ и государственной поддержке ведущих научных школ (код проекта НШ–2460.2008.1); работа третьего автора выполнена при финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 09–01–00244 и 10–01–00424).

и t. В общем случае можно говорить о *факторном замыкании* данного языка — множестве всех подслов его элементов.

Арифметическим подсловом данного слова $w_1 \dots w_n \dots$, где $w_i \in \Sigma$, называется всякое конечное слово вида $w_k w_{k+r} w_{k+2r} \dots w_{k+mr}$, где $r > 0$. Множество всех арифметических подслов данного слова называется его *арифметическим замыканием*. Ясно, что арифметическое замыкание всякого слова или языка — факторный язык.

Количество слов длины n в факторном языке F называется его *комбинаторной сложностью* и обозначается через $p_F(n)$; если этот факторный язык — язык подслов бесконечного слова w, то его комбинаторная сложность обозначается через $p_w(n)$. Комбинаторная сложность арифметического замыкания слова w называется его *арифметической сложностью* и обозначается через $a_w(n)$.

Ясно, что обе функции от бесконечного слова не убывают и растут не быстрее, чем d^n, где d — мощность алфавита. Обзор результатов о комбинаторной сложности можно найти в [1]; из более новых результатов хочется отметить конструкцию слова, комбинаторная сложность которого растет быстрее любого полинома и медленнее любой экспоненты [2].

Введенная в [3] функция арифметической сложности для непериодического слова может расти как экспоненциально [4], так и линейно, причем для равномерно рекуррентных слов с линейной арифметической сложностью известна характеризация [5]. Кроме того, известны примеры слов, арифметическая сложность которых растет не линейно, но субполиномиально [6], а также быстрее, чем любой полином, и медленнее, чем c^n для любого $c > 1$ [7].

Цель данной работы — построение семейства бесконечных слов с нестандартным, а именно порядка n^α, где α определяется как корень некоего трансцендентного уравнения, ростом арифметической сложности.

Для формулировки основного результата работы зафиксируем конечное множество простых чисел $Q = \{q_1, \dots, q_k\}$ мощности $k \geq 2$ и число $d > 1$ — размер используемого алфавита. Рассмотрим уравнение

$$\frac{d-1}{d} = \prod_{i=1}^{k} \left(1 - q_i^{-x}\right) \tag{1}$$

и отметим, что для любых Q и d оно имеет единственный положительный корень $x = \alpha = \alpha(Q, d)$. Действительно, функция $\Pi(x) = \prod_{i=1}^{k} \left(1 - q_i^{-x}\right)$ непрерывно возрастает с ростом $x \geq 0$, причем $\Pi(0) = 0$ и $\Pi(x) \to 1$ при $x \to \infty$.

Заметим, что при любом фиксированном d корень $\alpha(Q, d)$ уравнения (1) ни для какого Q не превосходит корня $\alpha = \alpha(d)$ уравнения

$$\frac{d-1}{d} = \prod_{q \text{ просто}} (1 - q^{-x}) = \frac{1}{\zeta(x)},$$

где ζ — дзета-функция Римана. В частности, при $d = 2$ параметр α ни для какого множества Q не превосходит корня $\alpha = 1.7286\dots$ уравнения $\zeta(x) = 2$.

Как следует из теоремы Безиковича [8], корень α уравнения (1) может быть либо целым (как в случае $d = 3$ и $Q = \{2, 3\}$, когда $\alpha = 2$), либо иррациональным. Из гипотезы Шануэля [9] следовало бы, что всякий иррациональный корень такого уравнения трансцендентен, однако, насколько нам известно, на данный момент этот факт не доказан.

Основным комбинаторным результатом данной работы является следующая

Теорема 1. *Пусть $Q = \{q_1, \dots, q_k\}$ — произвольное конечное множество простых чисел, $k \leq 2$. Тогда для любого $d > 1$ существует бесконечное слово $w = w(Q, d)$ над d-буквенным алфавитом, для арифметической сложности $a_w(n) = p_{Q,d}(n)$ которого выполняется свойство*

$$\frac{p_{Q,d}(n)}{n^{\alpha(Q,d)+1}} \to C \quad \text{при } n \to \infty,$$

где C — некоторая положительная константа.

В дальнейшем нам будет удобнее рассматривать $p_{Q,d}(n)$ не как функцию арифметической сложности последовательности $w = w(Q, d)$, а как функцию комбинаторной сложности языка $F = F(Q, d)$ (равного арифметическому замыканию последовательности w). Язык F и последовательность w строятся в следующем разделе работы. Далее в разд. 4 с помощью классической техники специальных слов выводится рекуррентная формула для первых разностей $s_{Q,d}(n) = p_{Q,d}(n+1) - p_{Q,d}(n)$ функции $p_{Q,d}(n)$. Свойства функции $s_{Q,d}(n)$ исследуются в разд. 5 (комбинаторными методами). Затем в разд. 6 доказывается свойство функции $\Pi(x)$ как функции комплексного переменного, позволяющее применить к ней следствие тауберовой теоремы — теорему 3, доказываемую в разд. 7.

Применение теоремы 3 к функции $s_{Q,d}(n)$ позволяет завершить доказательство теоремы 1 в разд. 8.

3. Конструкция языка $F(Q, d)$

Пусть Σ — конечный алфавит мощности d. В данной работе изучается семейство факторных языков, конструируемых следующим образом. Напомним, что *морфизмом* называется отображение $\varphi : \Sigma^* \to \Sigma^*$, для любых слов $x, y \in \Sigma^*$ удовлетворяющее равенству $\varphi(xy) = \varphi(x)\varphi(y)$; здесь и далее последовательная запись двух слов означает их конкатенацию.

Для всяких $a \in \Sigma$ и $m \in \mathbb{N}$ определим морфизм $\varphi_{a,m}$ равенствами

$$\varphi_{a,m}(b) = a^{m-1}b$$

для всех $b \in \Sigma$ (в частности, для $b = a$). Заметим, что морфизм $\varphi_{a,m}$ можно интерпретировать также как преобразование Тёплица, поскольку слово $\varphi_{a,m}(x)$ получается подстановкой последовательных символов слова x в «дырки» (обозначаемые ромбами) частичного слова $a^{m-1} \diamond a^{m-1} \diamond \dots$ [10, 11].

Зафиксируем конечное множество Q простых чисел, $Q = \{q_1, \dots, q_k\}$. Определим язык $L = L(\Sigma, Q)$ как замыкание алфавита Σ относительно всех морфизмов $\varphi_{a,q}$, где $a \in \Sigma$ и $q \in Q$:

$$L(\Sigma, Q) = \{\varphi_{c_m,r_m}(\varphi_{c_{m-1},r_{m-1}}(\dots(\varphi_{c_1,r_1}(c_0))\dots)) \mid c_i \in \Sigma,\ r_i \in Q\}.$$

Основной объект рассмотрения в данной работе — факторное замыкание $F = F(\Sigma, Q) = \mathrm{Fac}(L(\Sigma, Q))$. Этот факторный язык, как нетрудно видеть, не является языком подслов никакого бесконечного слова, однако может быть представлен, в частности, как язык арифметических подслов бесконечного влево слова w, являющегося пределом

$$w = \lim_{n\to\infty} \varphi_{c_1,r_1}(\varphi_{c_2,r_2}(\dots(\varphi_{c_n,r_n}(c_{n+1}))\dots)),$$

где последовательность пар $\{(c_i, r_i)\}_{i=1}^{\infty}$ содержит в качестве подслов всякой длины j все $(dk)^j$ возможных наборов из $(\Sigma \times Q)^j$. Нетрудно видеть, что предел w существует, а его арифметическое замыкание равно F: в частности, всякая арифметическая подпрогрессия с разностью $r_1 \dots r_m$ либо состоит из одинаковых символов, либо содержит в качестве подслов все слова вида $\varphi_{c_{m+1},r_{m+1}}(\varphi_{c_{m+2},r_{m+2}}(\dots(\varphi_{c_{m+n},r_{m+n}}(c_{m+n+1}))\dots))$.

Полное доказательство того, что арифметическое замыкание слова w равно F, аналогично доказательствам в работе [6], где был рассмотрен случай $Q = \{p\}$ (т. е. $|Q| = 1$): порядок роста арифметической сложности соответствующей последовательности равен $\Theta(n^{1+\log_p d})$, однако предел отношения $p_{\{p\},d}(n)$ к $n^{1+\log_p d}$ не существует. В той же статье была рассмотрена арифметическая сложность других последовательностей, порожденных конструкциями Тёплица с одной и той же длиной морфизма, поэтому здесь мы можем сконцентрироваться на случае $|Q| \geq 2$.

Пример 1. Пусть $\Sigma = \{a, b\}$, а $Q = \{3, 5\}$; рассмотрим язык $F = F(\Sigma, Q)$. Слово $u = abaabaabaabaaaaabaaba$ принадлежит языку F, поскольку является подсловом слова $\varphi_{a,3}(bbbbabbb)$; слово $bbbbabbb$ является подсловом слова $\varphi_{b,5}(aa)$, а aa — подсловом слова $\varphi_{a,3}(a)$. Таким образом, $u \in \mathrm{Fac}(\varphi_{a,3}(\varphi_{b,5}(\varphi_{a,3}(a)))) \subset F$.

Заметим, что в действительности однозначно по слову u восстанавливается только последний примененный морфизм: так, слово $bbbbabbb$ является подсловом и слова $\varphi_{b,3}(bab)$, а bab — подсловом и слова $\varphi_{b,3}(a)$, и слова $\varphi_{b,5}(a)$.

Замечание 1. Для любых $n, m > 0$ и $a \in \Sigma$ морфизмы $\varphi_{a,n}$ и $\varphi_{a,m}$ коммутируют: $\varphi_{a,n}\varphi_{a,m} = \varphi_{a,nm} = \varphi_{a,m}\varphi_{a,n}$.

Оставшаяся часть работы будет связана с нахождением асимптотики функции $p_F(n) = p_{d,Q}(n)$ (для краткости $p(n)$) комбинаторной сложности языка F. Большая часть рассуждений и вычислений будет касаться не самой функции $p(n)$, а ее первых разностей

$$s(n) = p(n+1) - p(n).$$

4. Рекуррентная формула для первых разностей

Как хорошо известно (см., например, [12]), первые разности $s(n)$ могут быть выражены в терминах специальных слов факторного языка F.

Рассмотрим слово $u \in F$ и обозначим через $L(u)$ множество символов $a \in \Sigma$ таких, что $au \in F$. Мощность множества $L(u)$ назовем (левой) *степенью специальности* слова u и обозначим через $l(u)$. Всякое слово u, для которого $l(u) \neq 1$, называется *специальным* в F; множество всех специальных слов длины n в F обозначается через $S(n)$. Известна формула, выражающая сложность языка F через его специальные слова:

$$s(n) = \sum_{u \in S(n)} (l(u) - 1). \tag{2}$$

Поэтому, чтобы найти выражение для первых разностей языка F, надо прежде всего исследовать множество его специальных слов.

Лемма 1. *Рассмотрим специальное слово $u \in S(n)$ языка F длины не меньше трех. Тогда для u верно одно из двух:*

(1) *либо $u = a^n$ для некоторого символа $a \in \Sigma$,*

(2) *либо $u = \varphi_{a,N}(v)a^{N'}$ для некоторых $a \in \Sigma$, $N \in \mathbb{N}$, причем все простые множители числа N принадлежат множеству Q, $0 \leq N' < N$ и $v \in F$; более того, v содержит символ $b \neq a$ и $L(v) = L(u)$.*

Доказательство. Рассмотрим сначала случай, когда $2 \notin Q$. В этом случае всякое слово u из F длины не меньше трех либо содержит два подряд идущих одинаковых символа, либо имеет вид $u = xyx$ для $x \neq y$, $x, y \in \Sigma$. В последнем случае u появляется в F только как подслово слов вида $\varphi_{x,q}(v)$ для некоторых $q \in Q$ и $v \in F$ и не является специальным, поскольку единственная буква, которой u может быть продолжено влево до элемента множества F, это x.

Таким образом, в каждом специальном в F слове длины не меньше трех встречаются два одинаковых символа подряд. Обозначив этот повторяющийся символ через a, видим, что u встречается в F только как подслово слов вида $\varphi_{a,q}(v)$, где $q \in Q$ и $v \in F$.

Ситуация, когда $u = a^n$, в точности соответствует случаю 1 леммы. Заметим, что $L(a^n) = \Sigma$: действительно, всякое продолжение xa^n слова a^n, где $x \in \Sigma$, принадлежит F, поскольку является фактором слова $\varphi^m_{a,q}(xx)$, а значит, и слова $\varphi^m_{a,q}\varphi_{x,q}(x)$ для всяких $q \in Q$ и достаточно большого m.

Предположим теперь, что $u \neq a^n$. В силу специальности u может быть продолжено влево некоторым символом $b \neq a$. Слово bu является фактором некоторого слова $\varphi_{a,q}(v')$, где $q \in Q$ и $v' \in F$. Следовательно, все попарные расстояния между символами слова bu, не равными a, делятся по меньшей мере на одно число $q \in Q$. В силу (взаимной) простоты всех элементов из Q в действительности все попарные расстояния делятся на некоторое произведение степеней элементов из Q: выберем его максимально возможным и обозначим через $N(u) = q_1^{l_1} \dots q_k^{l_k}$; здесь все l_i неотрицательны и по меньшей мере одно из них положительно. Таким образом, $u = \varphi_{a,N(u)}(v)a^{N'}$, и если слово v (при фиксированном $N(u)$) выбрано самым длинным из возможных, то $0 \leq N' < N(u)$. Заметим, что в силу максимальности числа $N(u)$ слово v не содержит двух последовательных вхождений символа a.

Осталось показать, что $L(v) = L(u)$. Включение $L(v) \subseteq L(u)$ практически очевидно, поскольку для любого $b \in \Sigma$ из того, что $bv \in F$, следует, что $\varphi_{a,N(u)}(bv)a^{N'} \in F$, а значит, и $bu \in F$ (поскольку $a^{N'}$ является префиксом образа под действием $\varphi_{a,N(u)}$ всякого продолжения слова bv вправо).

Докажем индукцией по $l = \sum_{i=1}^{k} l_i$ обратное включение: $L(u) \subseteq L(v)$. Предположим для начала, что $l = 1$, т. е. $N(u) = q \in Q$. Тогда для любого символа $x \in \Sigma$, $x \neq a$, слово xu может появиться в F только как суффикс некоторого слова $\varphi_{a,q}(xv)a^{N'}$, где слово v не содержит двух букв a подряд. Здесь по построению ясно, что если $x \in L(u)$, то и $x \in L(v)$. Тот же самый аргумент действует для $x = a$, если слово v содержит минимум два символа, не равных a, и число q может быть установлено по расстоянию между ними. Наконец, если v содержит только один символ c, не равный a, т. е. если v — подслово слова aca, то $a \in L(u)$ и $a \in L(v)$. База индукции доказана.

Для доказательства индукционного шага рассмотрим общий случай: $N(u) = q_1^{l_1} \dots q_k^{l_k}$; без ограничения общности предположим, что $l_1, \dots, l_i > 0$ и $l_{i+1} = \dots = l_k = 0$. Таким образом, u может появиться в языке F как результат применения любого из морфизмов φ_{a,q_j}, $j \leq i$, к слову $\varphi_{a,N(u)/q_j}(v)$ и приписывания справа слова $a^{N'}$ (которое всегда возможно). Но для любого $j \leq i$ по предположению индукции выполняется равенство $L(\varphi_{a,N(u)/q_j}(v)) = L(v)$, т. е. и $L(u) = L(v)$, что и требовалось доказать.

Заметим, что здесь мы использовали два факта: во-первых, что морфизмы

$\varphi_{a,p}$ и $\varphi_{a,q}$ коммутируют для всех p и q, а во-вторых, что элементы множества Q (взаимно) просты: это дает уверенность в том, что прообраз слова u под действием всякого морфизма φ_{a,q_j}, $j \leq i$, действительно является элементом множества F.

Для случая $2 \notin Q$ лемма доказана. Предположим теперь, что $2 \in Q$. В этом случае, как и раньше, если слово содержит две одинаковые буквы подряд, то эта буква соответствует последнему примененному к нему морфизму, а значит, к такому слову применимы все рассуждения из предыдущего случая. Единственное отличие здесь состоит в том, что в F существуют сколь угодно длинные слова, в которых двух одинаковых идущих подряд букв нет. Такое слово может быть специальным, только если имеет вид $u = av_1av_2a\ldots$ для некоторых букв $a, v_i \in \Sigma$, т. е. представимо как $u = \varphi_{a,2}(v_1v_2\ldots)a^m$, где $m \in \{0,1\}$.

Если среди символов v_i есть два различных, то очевидно, что последнему примененному морфизму всегда соответствуют не они, а символ a. Поэтому со словом u можно обращаться так же, как в предыдущем случае. Если же все v_i равны друг другу, т. е. если $u = ababa\ldots$ для некоторого символа $b \neq a$, то u является образом под действием $\varphi_{a,2}$ слова b^n для соответствующего n (или префиксом такого образа, если $m = 1$), а значит, $L(u) = L(v) = \Sigma$. Заметим, что одновременно u является подсловом слова $\varphi_{2,b}(a^{n'})$ для некоторого n', однако к множеству продолжений влево слова u это ничего не добавляет: оно и так максимально. □

Следствие 1. *Для всякого специального слова $u \in S(n)$ длины $n \geq 3$ единственным образом определены буква a и подмножество $Q(u)$ множества Q такие, что для всех $q \in Q(u)$ (и ни для каких других элементов множества Q) слово u может быть представлено в виде $u = \varphi_{a,q}(v_q)a^m$ для соответствующих слова $v_q \in F$ и длины $0 \leq m < q$. Длина слова v_q при этом равна $\lfloor n/q \rfloor$, и для всех $q \in Q(u)$ верны равенства $L(u) = L(v_q)$.*

Доказательство. Специальные слова, в которых встречаются минимум два разных символа, уже рассмотрены в предыдущей лемме (случай 2). Для слов вида a^n имеем $L(u) = \Sigma$, $Q(u) = Q$ и $v_q = a^{\lfloor n/q \rfloor}$: в частности, если длина слова u меньше, чем некоторое $q \in Q$, то v_q равно пустому слову. □

Теперь заметим, что к каждому специальному слову $v \in S(m)$ можно применить каждый из морфизмов $\varphi_{a,q}$, где $a \in \Sigma$ и $q \in Q$, а затем приписать справа любой суффикс a^r, где $r < q$. По лемме 1 полученное таким образом слово $u = \varphi_{a,q}(v)a^r$ длины $mq+r$ будет специальным с $L(u) = L(v)$, и, как было показано, все специальные слова длины начиная с трех могут быть получены таким образом. При этом одно и то же слово u может иметь несколько прообразов под действием разных морфизмов (соответствующих одной и той же букве a): $u = \varphi_{a,q_i}(v_i)a^{r_i} = \varphi_{a,q_j}(v_j)a^{r_j}$ для некоторых $i \neq j$ и соответствующих v_i, v_j, r_i, r_j. В силу взаимной простоты чисел q_i и q_j слово v_i является образом некоторого (специального) слова v_{ij} под действием морфизма φ_{a,q_j}, возможно, с приписыванием справа символов a, а v_j точно так же является образом того же самого слова v_{ij} под действием морфизма φ_{a,q_i} (опять же, возможно, с приписыванием справа нужного количества символов a). Таким образом, $u = \varphi_{a,q_iq_j}(v_{ij})a^r$ для некоторого специального слова v_{ij} с $L(v_{ij}) = L(u)$; здесь $0 \leq r < q_iq_j$, а длина слова v равна $\lfloor |u|/q_1q_2 \rfloor$.

Продолжая эти рассуждения на большее количество простых делителей из Q, получаем, что каждое слово u, которое может быть получено применением l разных морфизмов $\varphi_{a,q_{i_1}}, \ldots, \varphi_{a,q_{i_l}}$, может быть получено последовательным применением всех их (в любом порядке) к некоторому слову v длины $\lfloor |u|/q_{i_1} \ldots q_{i_l} \rfloor$, причем $L(v) = L(u)$.

Таким образом, по формуле включений-исключений вклад в $s(n)$ всех специальных слов длины $n \geq 3$, начинающихся с a, равен

$$\sum_{l=1}^{k} (-1)^{l+1} \sum_{\{q_{i_1}, \ldots, q_{i_l}\} \subseteq Q} s\left(\left\lfloor \frac{n}{q_{i_1} \ldots q_{i_l}} \right\rfloor\right).$$

Символ a при этом был выбран произвольным образом, как один из d символов алфавита Σ. В силу симметричности языка F относительно всех символов из Σ получаем рекуррентную формулу

$$s(n) = d \sum_{l=1}^{k} (-1)^{l+1} \sum_{\{q_{i_1}, \ldots, q_{i_l}\} \subseteq Q} s\left(\left\lfloor \frac{n}{q_{i_1} \ldots q_{i_l}} \right\rfloor\right) \quad \text{при } n \geq 3. \tag{3}$$

Пример 2. Если алфавит бинарный, а $Q = \{3, 5\}$, то выражение (3) имеет вид

$$s(n) = 2\left(s\left(\left\lfloor \frac{n}{3} \right\rfloor\right) + s\left(\left\lfloor \frac{n}{5} \right\rfloor\right) - s\left(\left\lfloor \frac{n}{15} \right\rfloor\right)\right)$$

и выполняется для всех $n \geq 3$.

Заметим также, что для любого множества Q значения $s(n)$ для $n = 0, 1, 2, 3$ могут быть вычислены вручную (через соответствующие значения $p(n)$ при $n = 0, \ldots, 4$) и зависят только от мощности алфавита и от того, принадлежат ли множеству Q числа 2 и 3.

Лемма 2. *Значения функции $s(n)$ для $n = 0, \ldots, 3$ описываются следующей таблицей:*

n	0	1	2	3
$2 \in Q$, $3 \in Q$	$d-1$	d^2-d	d^3-d^2	$2d^3-3d^2+d$
$2 \in Q$, $3 \notin Q$	$d-1$	d^2-d	d^3-d^2	d^3-d^2
$2 \notin Q$, $3 \in Q$	$d-1$	d^2-d	$2d^2-2d$	d^3-d^2
$2 \notin Q$, $3 \notin Q$	$d-1$	d^2-d	$2d^2-2d$	d^2-d

Доказательство проводится прямым вычислением значений функции $p(n)$ при $n = 0, \ldots, 4$. $\square$

5. Рост функции $s(n)$

В этом разделе комбинаторными методами будет доказано несколько свойств функции $s(n)$, напрямую вытекающих из формулы (3). Для удобства доопределим функцию $s(x)$ на множестве всех неотрицательных действительных чисел по формуле

$$s(x) = s(\lfloor x \rfloor). \tag{4}$$

Кроме того, для любого конечного множества $R = \{r_1, \ldots, r_m\}$ простых чисел и для всех функций $f(x)$, удовлетворяющих условию (4), определим функциональные операторы

$$D_R f(x) = \sum_{i=1}^{m} (-1)^{i+1} \sum_{\{s_1, \ldots, s_i\} \subseteq R} f\left(\frac{x}{s_1 \ldots s_i}\right)$$

и

$$E_R f(x) = f(x) - D_R f(x).$$

В этих обозначениях выражение (3) может быть переписано в виде

$$s(x) = dD_Q s(x) \quad \text{для всех } x \geq 3. \tag{5}$$

Заметим, что всякое решение этого уравнения, удовлетворяющее условию (4), полностью определяется значениями функции $s(x)$ для $x = 0, 1, 2$. С другой стороны, ясно, что решения уравнения (5), удовлетворяющие условию (4), образуют линейное пространство: для любых таких решений s_1 и s_2 функция $\alpha s_1(x) + \beta s_2(x)$ также является решением и также удовлетворяет (4).

Заметим, что для всякой функции f верны равенства $D_\varnothing f(x) = 0$, $E_\varnothing f(x) = f(x)$ и $D_{\{r\}} f(x) = f(x/r)$. Кроме того, как нетрудно проверить,

$$E_{R\cup\{r\}} f(x) = E_R f(x) - E_R f(x/r) \tag{6}$$

для любого $r \notin R$, и

$$D_{\{r_1,\ldots,r_m\}} f(x) = E_\varnothing f(x/r_1) + E_{\{r_1\}} f(x/r_2) + \cdots + E_{\{r_1,\ldots,r_{m-1}\}} f(x/r_m). \tag{7}$$

Лемма 3. *Пусть $s(x)$ — решение уравнения (5) для множества Q, удовлетворяющее условию (4). Предположим, что $s(3) \geq s(2) \geq s(1) \geq s(0) > 0$. Тогда функция $s(x)$ неубывающая для всех $x > 0$.*

Доказательство. Определим функцию $t(x) = s(x) - s(x-1)$. По условию леммы $t(x) \geq 0$ при $1 \leq x \leq 3$; наша задача — доказать, что $t(n) \geq 0$ для всех $n \geq 4$. Нам потребуются два вспомогательных утверждения.

Предложение 1. *Рассмотрим целое $n > 3$ и подмножество $Q(n)$ множества Q, состоящее в точности из элементов Q, которые делят n. Тогда*

$$t(n) = dD_{Q(n)} t(n). \tag{8}$$

Доказательство. Из определений следует, что

$$t(n) = s(n) - s(n-1) = d \sum_{m=1}^{k} (-1)^{m+1} \sum_{\{r_1,\ldots,r_m\}\subseteq Q} \left(s\left(\frac{n}{r_1\ldots r_m}\right) - s\left(\frac{n-1}{r_1\ldots r_m}\right)\right).$$

Из-за условия (4) разность $(s(n/(r_1\ldots r_m)) - s((n-1)/(r_1\ldots r_m))$ может быть не равна нулю, только если число $n/(r_1\ldots r_m)$ целое, т. е. если $r_1, \ldots, r_m \in Q(n)$ (поскольку все элементы множества Q взаимно просты). В этом случае $s((n)/(r_1\ldots r_m)) - s((n-1)/(r_1\ldots r_m)) = t((n)/(r_1\ldots r_m))$. $\square$

В частности, если $n > 3$ не делится ни на один элемент множества Q, то $t(n) = 0$.

Предложение 2. *Рассмотрим число $n > 3$, $n = q_1^{i_1}\ldots q_l^{i_l} n_0$, где $\{q_1, \ldots, q_l\} = Q(n)$, а число n_0 взаимно просто со всеми элементами множества Q. Пусть $t(n_0) \geq 0$; тогда $t(n) \geq 0$.*

Доказательство. Докажем несколько больше: что $E_{\{q_1,\ldots,q_j\}} t(n) \geq 0$ для всех $j = 0, \ldots, l$ (в частности, для $E_\varnothing t(n) = t(n)$). Воспользуемся индукцией по $i = i_1 + \cdots + i_l$; база индукции для $i = 0$, т. е. для $n = n_0$, задана условиями предложения, поскольку $Q(n_0) = \varnothing$.

Шаг индукции основан на комбинировании равенств (8) и (7), из которых

$$t(n) = E_{\varnothing} t(n) = dE_{\varnothing} t(n/q_1) + \dots + dE_{\{q_1,\dots,q_{l-1}\}} t(n/q_l)$$

и

$$E_{\{q_1,\dots,q_j\}} t(n) = (d-1)E_{\varnothing} t(n/q_1) + \dots + (d-1)E_{\{q_1,\dots,q_{j-1}\}} t(n/q_j) \\ + dE_{\{q_1,\dots,q_j\}} t(n/q_{j+1}) + \dots + dE_{\{q_1,\dots,q_{l-1}\}} t(n/q_l)$$

для всех $j = 1, \dots, l$. По определениям n/q_m для всех m является целым числом, множество $Q(n/q_m)$ равно либо $Q(n)$, либо $Q(n)\backslash\{q_m\}$, а сумма степеней сомножителей q из множества $Q(n/q_m)$ равна $i-1$. Значит, по предположению индукции предложение применимо к любому оператору E_R из правой части выписанной выше системы, и значение каждого оператора $E_{\{q_1,\dots,q_j\}} t(n)$ является суммой неотрицательных слагаемых. В частности, это верно и для $j = 0$, т. е. для самой функции $t(n)$. □

Для доказательства леммы остается заметить, что предложение 2 может быть применено ко всякому $n > 3$, где n_0 либо равно $1, 2$ или 3 (тогда $t(n_0) \geq 0$ по условию), либо является числом больше трех, взаимно простым со всеми элементами множества Q (тогда $t(n_0) = 0$ по предложению 1). □

Лемма 4. *Пусть функция $s(x)$ является решением уравнения* (5), *удовлетворяющим* (4), *причем $s(x) > 0$ для всех $x \geq 0$ и $s(0) < s(1) < s(2)$. Тогда существует положительная константа C, для которой при всех $x > 1$ верно неравенство*

$$s(x) \leq Cx^{\alpha},$$

где α — решение уравнения (1).

Доказательство. Для доказательства леммы удобно предполагать, что элементы множества Q упорядочены по возрастанию.

Заметим, что функция $s(x)$ и операторы $E_{\{q_1,\dots,q_i\}} s(x)$, где $1 \leq i \leq k$, при $1 \leq x \leq q_1 \dots q_k$ принимают лишь конечное число значений, а значит, существует положительная константа C, для которой

$$\begin{aligned} s(x) &\leq Cx^{\alpha}, \\ E_{\{q_1\}} s(x) &\leq (1 - 1/q_1^{\alpha})Cx^{\alpha}, \\ &\vdots \\ E_{\{q_1,\dots,q_{k-1}\}} s(x) &\leq (1 - 1/q_1^{\alpha}) \dots (1 - 1/q_{k-1}^{\alpha})Cx^{\alpha} \end{aligned} \tag{9}$$

при всех $1 \leq x < q_1 \dots q_k$. Эта система неравенств образует базу индукции; для индукционного шага предположим, что система (9) верна для всех $x \in [1, X)$, где $X \geq q_1 \dots q_k$, и докажем ее для $x \in [X, q_1 X)$ (заметим, что по предположению q_1 — минимальный элемент множества Q). В силу определений операторов D_Q и E_Q, а также равенств (5)–(7) для любого $x \geq 3$ имеем

$$E_{\{q_1,\dots,q_{k-1}\}} s(x) = (d-1)[s(x/q_1) + E_{\{q_1\}} s(x/q_2) + \dots + E_{\{q_1,\dots,q_{k-2}\}} s(x/q_{k-1})] \\ + dE_{\{q_1,\dots,q_{k-1}\}} s(x/q_k). \tag{10}$$

Подставив неравенства системы (9) (выполняющиеся по предположению индукции) к операторам из правой части этого равенства, получаем неравенство

$$\begin{aligned}E_{\{q_1,\ldots,q_{k-1}\}}s(x) \le Cx^\alpha(d-1)\left[\frac{1}{q_1^\alpha}+\left(1-\frac{1}{q_1^\alpha}\right)\frac{1}{q_2^\alpha}+\cdots+\frac{1}{q_{k-1}^\alpha}\prod_{i=1}^{k-2}\left(1-\frac{1}{q_i^\alpha}\right)\right]\\ + Cx^\alpha d\cdot\frac{1}{q_k^\alpha}\prod_{i=1}^{k-1}\left(1-\frac{1}{q_i^\alpha}\right)\\ = Cx^\alpha(d-1)\sum_{j=1}^{k}\frac{1}{q_j^\alpha}\prod_{i=1}^{j-1}\left(1-\frac{1}{q_i^\alpha}\right)+Cx^\alpha\cdot\frac{1}{q_k^\alpha}\prod_{i=1}^{k-1}\left(1-\frac{1}{q_i^\alpha}\right).\end{aligned}$$

Заметим, что

$$\sum_{j=1}^{k}\frac{1}{q_j^\alpha}\prod_{i=1}^{j-1}\left(1-\frac{1}{q_i^\alpha}\right)=1-\prod_{i=1}^{k}\left(1-\frac{1}{q_i^\alpha}\right).$$

Поэтому, воспользовавшись равенством (1), видим, что первое слагаемое в правой части предыдущего неравенства равно

$$Cx^\alpha(d-1)[1-(d-1)/d]=Cx^\alpha(d-1)/d.$$

В свою очередь, второе слагаемое в правой части этого неравенства можно преобразовать следующим образом:

$$\begin{aligned}Cx^\alpha\cdot\frac{1}{q_k^\alpha}\prod_{i=1}^{k-1}\left(1-\frac{1}{q_i^\alpha}\right)=Cx^\alpha\left[\prod_{i=1}^{k-1}\left(1-\frac{1}{q_i^\alpha}\right)-\prod_{i=1}^{k}\left(1-\frac{1}{q_i^\alpha}\right)\right]\\ = Cx^\alpha\left[\prod_{i=1}^{k-1}\left(1-\frac{1}{q_i^\alpha}\right)-\frac{d-1}{d}\right].\end{aligned}$$

Сумма этих двух слагаемых дает в точности правую часть последнего из неравенств системы (9), что и требовалось. Таким образом, последнее неравенство системы (9) верно и для $x\in[X,q_1X)$.

Теперь чтобы доказать для $x\in[X,q_1X)$ все предыдущие неравенства из (9), достаточно последовательно (в порядке убывания индекса $i=k-2,\ldots,0$) воспользоваться равенствами

$$E_{\{q_1,\ldots,q_i\}}s(x)=E_{\{q_1,\ldots,q_{i+1}\}}s(x)+E_{\{q_1,\ldots,q_i\}}s(x/q_{i+1}), \tag{11}$$

верными благодаря (6). Подставляя в них соответствующие, верные по предположению индукции, неравенства из (9), получаем, что неравенства (9), в частности первое из них, выполняются и для $x\in[X,q_1X)$. $\square$

6. Лемма об отсутствии корней

Лемма 5. *Для любого множества Q простых чисел мощности не меньше двух уравнение*

$$\frac{d-1}{d}=\prod_{q\in Q}(1-q^{-x})$$

не имеет корней с действительной частью, равной $\alpha=\alpha(Q,d)$, помимо корня $x=\alpha$.

Доказательство. Предположим, что существует другой корень вида $\alpha + iy$, $0 \neq y \in \mathbb{R}$. Тогда

$$\frac{d-1}{d} = \prod_{q \in Q} (1 - q^{-\alpha - iy}) = \prod_{q \in Q} |1 - q^{-\alpha - iy}|.$$

По неравенству треугольника имеем $|1 - q^{-\alpha-iy}| \geq 1 - q^{-\alpha}$, причем равенство достигается, если и только если $q^{-iy} = 1$, т. е. $y = 2\pi k / \ln q$, $k \in \mathbb{Z}$. Такое равенство возможно не более чем для одного простого q. В самом деле, если $y = 2\pi k_1 / \ln q_1 = 2\pi k_2 / \ln q_2$, то $q_2^{k_1} = q_1^{k_2}$, что невозможно. Перемножив все неравенства $|1 - q^{-\alpha-iy}| \geq 1 - q^{-\alpha}$ для $q \in Q$ (среди неравенств есть строгое, так как Q содержит хотя бы два простых числа), получаем

$$\prod_{q \in Q} |1 - q^{-\alpha - iy}| > \prod_{q \in Q} |1 - q^{-\alpha}| = \frac{d-1}{d};$$

противоречие. $\square$

Заметим, что здесь существенно наличие в Q хотя бы двух различных чисел: если $|Q| = 1$, утверждение леммы не выполняется.

7. Тауберова теорема

Обозначим через W класс Винера функций g, суммируемых на $(0, +\infty)$ и таких, что преобразование Меллина $\int\limits_0^\infty g(t) t^{ix}\, dt$ не обращается в нуль при вещественных x.

Теорема 2 (Винера — Питта [13, теорема 233]). *Пусть $g \in W$, а функция f вещественна, ограничена и медленно убывает в том смысле, что если $y > x \to +\infty$, $y/x \to 1$, то $\liminf f(y) - f(x) \geq 0$. Тогда если*

$$\frac{1}{x} \int\limits_0^\infty f(t) g(t/x) dt \to l \int\limits_0^\infty g(t)\, dt$$

при $x \to +\infty$, то $f(t) \to l$ при $t \to +\infty$.

Теорема 3. *Предположим, что вещественные числа c_1, c_2, …, c_l и m_1, m_2, …, m_l ($m_i > 1$ для всех i) таковы, что уравнение $F(\alpha) := \sum c_i m_i^{-\alpha} = 1$ имеет единственное положительное решение $\alpha = \alpha_0$, причем $F(\alpha_0 + ix) \neq 1$ при вещественном $x \neq 0$ и $F'(\alpha_0) \neq 0$.*

Предположим также, что неубывающая функция $s(x)$, $x \geq 0$, такая, что $s(x) = s(\lfloor x \rfloor)$, удовлетворяет уравнению

$$s(x) = \sum_{i=1}^{l} c_i s(x/m_i), \quad x \geq N_0 > 0,$$

и соотношению $s(x) = O(x^{\alpha_0})$ при $x \to +\infty$.

Тогда частное $s(x)/x^{\alpha_0}$ имеет предел.

Доказательство. Обозначим через $\chi(t) = \chi_{[0,1]}(t)$ характеристическую функцию отрезка $[0, 1]$ и рассмотрим ядро

$$g(t) = \frac{\chi(t) - \sum c_i m_i^{-\alpha_0} \chi(m_i t)}{t}, \quad t > 0.$$

Заметим, что $g(t)$ — кусочно постоянная финитная функция, носитель которой отделен от нуля и бесконечности. Поэтому $g \in L_1(0, +\infty)$. Вычислим преобразование Меллина функции g:

$$\int_0^\infty g(t)t^{ix} dt = \frac{1 - \sum c_i m_i^{-\alpha_0 - ix}}{ix},$$

значение при $x \neq 0$ определяется по непрерывности. Наше предположение относительно свойств функции $F(\alpha)$ можно теперь сформулировать так: преобразование Меллина функции g не имеет вещественных нулей. Таким образом, функция g принадлежит классу Винера W на $(0, +\infty)$.

Заметим, что выражение

$$I(x) = \frac{1}{x} \int_0^\infty \frac{s(t)}{t^{\alpha_0}} g(t/x)\, dt \tag{12}$$

при достаточно больших положительных x не зависит от x и равно

$$A = \sum c_i m_i^{-\alpha_0} \int_{N_0/m_i}^{N_0} \frac{s(\tau)}{\tau^{\alpha_0+1}}\, d\tau. \tag{13}$$

В самом деле, подынтегральная функция в (12) равна 0 при $t < N_0$ (если x достаточно велико), так что нижний предел интегрирования можно заменить на N_0. Тогда интеграл может быть записан как

$$I(x) = \int_{N_0}^{x} \frac{s(t)}{t^{\alpha_0+1}}\, dt - \sum c_i m_i^{-\alpha_0} \int_{N_0}^{x/m_i} \frac{s(t)}{t^{\alpha_0+1}}\, dt.$$

Распишем первый интеграл как сумму $s(t) = \sum c_i s(t/m_i)$ при $t \geq N_0$ и, заменяя в соответствующих слагаемых переменную t на $\tau = t/m_i$, получаем (13).

Заметим, что если функция $s(x)$ возрастает и $s(x) = O(x^{\alpha_0})$, то ограниченная функция $f(x) = s(x) \cdot x^{-\alpha_0}$ является также медленно убывающей. В самом деле,

$$f(y) - f(x) = s(y)y^{-\alpha_0} - s(x)x^{-\alpha_0} \geq s(x)(y^{-\alpha_0} - x^{-\alpha_0}) = (s(x)x^{-\alpha_0})((x/y)^{\alpha_0} - 1),$$

первый сомножитель ограничен, а второй стремится к нулю.

Таким образом, по тауберовой теореме Винера — Питта $s(x)/x^{\alpha} = f(x) \to A/\int g = A/(-F'(\alpha_0))$ при $x \to \infty$, что и требовалось доказать. $\square$

8. Завершение доказательства

Отметим, что теорема 3 не всегда может быть применена напрямую к функции $s(x)$ первых разностей, поскольку, как видно из леммы 2, эта функция не всегда является неубывающей. Однако во всех случаях, как замечено в начале § 5, решения уравнения (5), удовлетворяющие условию (4), образуют линейное пространство, а значит, $s(x)$ может быть представлена в виде разности $s(x) = s_1(x) - s_2(x)$, где s_1 и s_2 — решения уравнения (5), удовлетворяющие условию (4) и неубывающие при $x \leq 4$. По лемме 3 этого достаточно для того, чтобы s_1 и s_2 были неубывающими на всей положительной полуоси, и можно было применить теорему 3 (и все предыдущие утверждения, требуемые для выполнения ее условий) к каждой из них по отдельности.

Коэффициенты c_r и m_r из условия теоремы 3 определяются естественным образом с помощью уравнения (3): множество коэффициентов m_r составляют произведения $q_{i_1} \dots q_{i_l}$, а коэффициент c_r при любом таком произведении равен $d(-1)^{l+1}$. Нетрудно проверить, что уравнение (1) при таких определениях переписывается в требуемом для теоремы виде: здесь функция $F(x)$ определяется как

$$F(x) = d\left(1 - \prod_{i=1}^{k}\bigl(1 - q_i^{-x}\bigr)\right).$$

Нетрудно видеть, что ее производная в точке $x = \alpha$ строго меньше нуля; остальные условия теоремы выполняются по леммам 4 и 5. Таким образом, по теореме 3 при $x \to \infty$ существует предел отношения $s(x)/x^\alpha$ либо, строго говоря, пределы $s_1(x)/x^\alpha$ и $s_2(x)/x^\alpha$, но тогда достаточно взять их разность, чтобы получить искомый предел $s(x)/x^\alpha$.

Остается отметить, что функция $s(n)$ представляет собой первые разности интересующей нас функции арифметической сложности $a_w(n) = p_{Q,d}(n)$. Следовательно, существование положительного предела $p_{Q,d}(n)/x^{\alpha+1}$ вытекает по теореме Штольца [14]; искомый предел равен пределу для первых разностей, деленному на $\alpha + 1$. Выпишем его в явном виде.

Вычислим значение константы A из уравнения (13): его можно считать напрямую для функции s, а не для s_1 и s_2, поскольку интеграл от линейной комбинации равен линейной комбинации интегралов. Это значение зависит от того, принадлежит ли множеству Q число 2. Предположим сначала, что $2 \notin Q$. В нашем случае $N_0 = 3$, $\sum c_i = d$. Значения N_0/m_i для каждого i не превосходят единицы, так что

$$\begin{aligned} A &= \alpha^{-1} \sum c_i m_i^{-\alpha}\bigl(s(0)\bigl(3^{-\alpha} m_i^\alpha - 1\bigr) + s(1)(1 - 2^{-\alpha}) + s(2)(2^{-\alpha} - 3^{-\alpha})\bigr) \\ &= \alpha^{-1}(3^{-\alpha}(s(0)d - s(2)) + 2^{-\alpha}(s(2) - s(1)) + s(1) - s(0)). \end{aligned}$$

Остается подставить в это выражение значения $s(0) = d - 1$, $s(1) = d^2 - d$, $s(2) = 2d^2 - 2d$, чтобы получить

$$A = \frac{1}{\alpha}((d-1)^2 + (2^{-\alpha} - 3^{-\alpha})(d^2 - d)).$$

Аналогичным образом при $2 \in Q$ получаем, что

$$A = \frac{1}{\alpha}(d-1)^2.$$

Таким образом, константа, к которой стремится отношение $p_{Q,d}(n)/x^{\alpha+1}$, выражается в явном виде и равна $\frac{A}{-F'(\alpha)(\alpha+1)}$. В частности, видно, что она положительна, что завершает доказательство теоремы. □

Благодарности

Мы выражаем признательность профессорам В. С. Губе и Ф. Л. Назарову за полезные обсуждения, а также администраторам и участникам сообщества ru-math.livejournal.com, дающего редкую возможность математического общения специалистов в разных областях.

ЛИТЕРАТУРА

1. *Ferenczi S.* Complexity of sequences and dynamical systems // Discrete Math. 1999. V. 206, N 1–3. P. 145–154.
2. *Cassaigne J.* Constructing infinite words of intermediate complexity // Developments in language theory. Berlin; Heidelberg: Springer Verl., 2002. P. 173–184. (Lect. Notes Comput. Sci.; V. 2450).
3. *Avgustinovich S. V., Fon-Der-Flaass D. G., Frid A. E.* Arithmetical complexity of infinite words // Words, languages and combinatorics. III. Singapore: World Scientific, 2003. P. 51–62.
4. *Frid A. E.* Arithmetical complexity of symmetric D0L words // Theoret. Comput. Sci. 2003. V. 306, N 1–3. P. 535–542.
5. *Frid A. E.* Sequences of linear arithmetical complexity // Theoret. Comput. Sci. 2005. V. 339, N 1. P. 68–87.
6. *Frid A. E.* On possible growths of arithmetical complexity // Theoret. Informatics Appl. 2006. V. 40, N 3. P. 443–458.
7. *Salimov P. V.* Constructing infinite words of intermediate arithmetical complexity // Proceedings of the 3rd International conference on language and automata theory and applications, Berlin; Heidelberg: Springer Verl., 2009. P. 696–701. (Lect. Notes Comput. Sci.; V. 5457).
8. *Besicovitch A. S.* On the linear independence of fractional powers of integers // J. London Math. Soc. 1940. V. 15, N 1. P. 3–6.
9. *Waldschmidt M.* Open Diophantine problems // Mosc. Math. J. 2004. V. 4, N 1. P. 245–305.
10. *Koskas M.* Complexités de suites de Toeplitz // Discrete Math. 1998. V. 183, N 1–3. P. 161–183.
11. *Cassaigne J., Karhumäki J.* Toeplitz words, generalized periodicity and periodically iterated morphisms // European J. Combinatorics. 1997. V. 18, N 5. P. 497–510.
12. *Cassaigne J.* Complexité et facteurs spéciaux // Bull. Belg. Math. Soc. Simon Stevin. 1997. V. 4, N 1. P. 67–88.
13. *Харди Г.* Расходящиеся ряды. М.: Факториал Пресс, 2006.
14. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления. М.: Физматлит, 2001. Т. 1.

Статья поступила 10 марта 2010 г.

Кассень Жюльен (Julien Cassaigne)
Institut de Mathématiques de Luminy
13288 Marseille Cedex 9, France
cassaigne@iml.univ-mrs.fr

Петров Федор Владимирович
Санкт-Петербургское отделение
Математического института им. В. А. Стеклова РАН,
наб. Фонтанки, 27, Санкт-Петербург 191023
fedyapetrov@gmail.com

Фрид Анна Эдуардовна
Институт математики им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4, Новосибирск 630090
anna.e.frid@gmail.com