Д.Ю.Григорьев

НИЖНИЕ ОЦЕНКИ В АЛГЕБРАИЧЕСКОЙ СЛОЖНОСТИ ВЫЧИСЛЕНИЙ

(Результаты доложены 21 и 23 января 1981 г.)

Содержание

Введение. I. Основные понятия. Глава I. Алгебро-геометрический подход к получению нижних оценок сложности вычисления многочленов. 2. Вычисление многочлена с "общими" коэффициентами. 3. Сложность вычисления индивидуальных многочленов. 4. Метод степени и его обобщения (случай бесконечного основного поля). 5. Метод степени (случай конечного основного поля). 6. Аддитивная сложность и вещественные корни. Глава II. Нижние оценки мультипликативной сложности в задачах линейной алгебры. 7. Мультипликативная сложность и ранг. 8. Ранг пары билинейных форм. 9. Мультипликативная сложность билинейной формы над коммутативным кольцом. I0. Оценки ранга алгебр. II. Линеаризованная мультипликативная сложность. Глава III. Сложность в неветвящихся программах нестандартных типов. I2. Иррациональная сложность вычисления алгебраических функций. I3. Монотонные вычисления. I4. Нижние оценки для произведения времени и памяти. I5. Методы теории графов в алгебраической сложности. I6. Аддитивная сложность в треугольных и направленных вычислениях и разложение Брюа.

Введение

Проблема нижних оценок является одной из наиболее трудных в теории сложности вычислений, и без преувеличения можно сказать, что их получение и составляет собственно основной предмет теории сложности, так как установление верхних оценок – т.е. построение достаточно быстрых алгоритмов – это, скорее, прерогатива других математических наук, из которых происходят конкретные вычислительные задачи. Несмотря на то, что проблема получения нетривиальных нижних оценок (т.е. доказательства невозможности достаточно быстрых алгоритмов для данных вычислительных задач, и тем самым проникновение в тайну быстрых алгоритмов) далека пока что от решения в полном объеме, в ней имеются некоторые интересные продвижения, особенно в той части теории сложности, которая от-

носится к задачам, восходящим к алгебре, и получившей название алгебраической сложности (по поводу нижних оценок в некоторых других разделах теории сложности см.обзор А.П.Бельтюкова в настоящем сборнике).

Алгебраическая сложность – одна из самых старых ветвей теории сложности (но и одна из наиболее интенсивно развиваемых в настоящее время); ей около 25 лет, однако на русском языке не было пока что посвященного ей достаточно полного обзора. Из иностранных изданий следует отметить в первую очередь книгу [27], а также [I], [I3], но в последних фактически отсутствуют нижние оценки, а в [27] не вошли достижения последних лет.

Автор ни в коей мере не претендует на полноту изложения всех результатов в области получения нижних оценок в алгебраической сложности вычислений; скорее, настоящий текст – это обзор избранных методов и достижений, цель которого заполнить имеющийся пробел в литературе на русском языке. Методы, получившие уже широкое распространение, а также не имеющие пока достаточно сильных приложений, изложены менее подробно. Число доказательств, приведенных в данном обзоре, сравнительно невелико; достаточная полнота ссылок позволяет при необходимости обратиться к соответствующей литературе.

Автор старался уделять основное внимание тем методам установления нижних оценок, которые связаны с нетривиальными алгебраическими методами. Глубокие связи с классической алгеброй, а также постановки задач, являющиеся необычными для традиционной алгебры – вообще характерная черта алгебраической сложности, которая может сделать ее привлекательной и для алгебраистов.

Обзор методов в алгебраической сложности оставляет несколько мозаичное впечатление. Это объясняется, по-видимому, тем, что этот раздел математики еще достаточно молод, в нем пока что не сформировалось объединяющих идей, задачи трудны и требуют индивидуального подхода. Поэтому различные главы формально мало между собой связаны (кроме § I, в котором приводятся определения, необходимые для понимания дальнейшего). В каждом параграфе содержится, по-существу, описание отдельного метода; в то же время, порядок расположения материала не случаен и имеет определенные исторические и методические причины (если угодно, онтогенез и филогенез алгебраической сложности). Отметим, что содержание §§ 6,II,I6 и частично § I5 публикуется впервые.

Нумерация параграфов и формул сквозная, теоремы, леммы и следствия нумеруются двумя числами, первое из которых – номер соответствующего параграфа.

§ I. Основные понятия

Основной вычислительной моделью, используемой в алгебраической сложности, является неветвящаяся программа (*straight-line computation*) – см., например, [I,27], которую мы сейчас опишем. Пусть заданы

I) набор входных переменных $x_1,\dots,x_n$;

2) кольцо K (обычно это поле, которое будет обозначаться через F), будем в дальнейшем называть его основным;

3) множество P базисных операций (обычно $P=\{+,\times,/\}\cup\{\times\gamma\}_{\gamma\in K}$, где $+,\times,/$ – двухместные арифметические операции, $\times\gamma$ – одноместная операция – умножение на γ).

Переменные $x_1,\dots,x_n$ могут предполагаться или не предполагаться попарно коммутирующими, часто это ясно из существа рассматриваемой задачи.

4) Собственно неветвящаяся программа (НП) представляет собой последовательность строк (команд), ℓ-ая из которых имеет следующий вид:

$$z_\ell = f_\ell(z_{i_1},\dots,z_{i_u},x_{j_1},\dots,x_{j_v}),$$

где $i_1<\ell,\dots,i_u<\ell$ и $f_\ell\in P$.

Индукцией по ℓ естественно определяется терм от переменных $x_1,\dots,x_n$, соответствующий рабочей переменной z_ℓ и называемый значением z_ℓ. Скажем, что некоторое семейство термов (или функций) от $x_1,\dots,x_n$ вычисляется данной НП, если термы из рассматриваемого семейства встречаются среди значений рабочих переменных z_ℓ этой НП.

Фиксирована также целозначная функция $\lambda=\lambda_\odot : P\longrightarrow N\cup\{0\}$, и число $\lambda(f)$ для $f\in P$ называется сложностью операции f. Сложностью НП β (обозначим ее $C(\beta)=C_\odot(\beta)$) называется сумма всех $\lambda(f_\ell)$ по всем строкам этой НП. Наконец, сложностью набора термов (или функций) $g_1,\dots,g_k$ называется наименьшая сложность НП, вычисляющих это семейство (обозначим ее $C(g_1,\dots,g_k)=C_\odot(g_1,\dots,g_k)$). Иногда вместо слова сложность будем употреблять термин мера сложности. Приведем один пример сложности. Пусть $\lambda_t(f)=1$ для всякой $f\in P$. Тогда соответствующая сложность C_t "считает" число всех операций НП и называется тотальной сложностью.

Ниже будем ссылаться на обозначения, принятые в настоящем

параграфе, конкретизируя каждый раз K (или F), P, $\lambda = \lambda_\odot$.

ГЛАВА I. АЛГЕБРО-ГЕОМЕТРИЧЕСКИЙ ПОДХОД К ПОЛУЧЕНИЮ НИЖНИХ ОЦЕНОК СЛОЖНОСТИ ВЫЧИСЛЕНИЯ МНОГОЧЛЕНОВ

§ 2. Вычисление многочлена с "общими" коэффициентами

Одной из первых задач, рассматривавшихся в алгебраической сложности, было вычисление многочлена с "общими" коэффициентами в одной точке (см. [I8] и имеющиеся там ссылки на более раннюю литературу). Иными словами: F – алгебраически замкнутое поле, $P = \{+, \times, /\} \cup \{x\gamma\}_{\gamma \in F}$. Обозначим $\lambda_+(+) = 1$, $\lambda_+(\times) = \lambda_+(/) = \lambda_+(x\gamma) = 0$ (соответствующая сложность C_+ называется иногда а д д и т и в н о й); $\lambda_{\times/}(+) = 0$, $\lambda_{\times/}(\times) = \lambda_{\times/}(/) = \lambda_{\times/}(x\gamma) = 1$. Требуется оценить $C_\odot(a_0 + a_1 x + \ldots + a_n x^n)$, где $\odot = +$ или $\odot = \times/$; в данном случае $\{x, a_0, \ldots, a_n\}$ – набор входных попарно коммутирующих переменных, причем $x, a_0, \ldots, a_n$ алгебраически независимы над F, поэтому коэффициенты называются "общими" (в этом случае рассматриваемые НП называются в [I8] схемами без предварительной обработки коэффициентов).

ТЕОРЕМА 2.I. ([I8]). $C_+(a_0 + \ldots + a_n x^n) = C_{\times/}(a_0 + \ldots + a_n x^n) = n$.

Верхние оценки в обоих случаях достигаются, нетрудно видеть, с помощью схемы Горнера.

Идея получения нижней оценки, неформально говоря, следующая. Значением всякой рабочей переменной из НП – схемы без предварительной обработки коэффициентов – является какая-то рациональная функция $g \in F(x, a_0, \ldots, a_n)$, которую можно некоторым образом записать в виде дроби $g = (b_0 + \ldots + b_p x^p)/(c_0 + \ldots + c_q x^q)$, где $b_i, c_i \in F(a_0, \ldots, a_n)$. Тогда индукцией по k можно показать, что если $C_\odot(g) \leqslant k$, то степень трансцендентности поля $F(b_0, \ldots, b_p, c_0, \ldots, c_q)$ над F не превосходит $k+1$ при подходящем выборе записи g в виде дроби, откуда уже следует теорема ($\odot$ может обозначать либо $+$ либо $\times/$).

Отметим, что из доказанного непосредственно вытекает справедливость теоремы 2.I для случая произвольного бесконечного поля.

Другой рассматривавшийся класс НП, называемый в [I8] схемами с предварительной обработкой коэффициентов, в терминах § I определяется так: $F = \mathbb{Q}(a_0, \ldots, a_n)$ – поле алгебраических функций от алгебраически независимых над $\mathbb{Q}$ переменных, P и $\lambda_\odot$ те же, что и выше. Требуется оценить $C_\odot(a_0 + \ldots + a_n x^n)$,

причем здесь множеством входных переменных является $\{x\}$. Интерпретация следующая: если нужно много раз вычислять значение одного и того же многочлена в разных точках, то имеет смысл некоторые вспомогательные алгебраические функции от коэффициентов вычислить заранее – многократное вычисление значений многочлена в различных точках компенсирует затраты на вычисление этих алгебраических функций. Обозначим $g = a_0 + \ldots + a_n x^n$.

ТЕОРЕМА 2.2. ([I8,4I]). I) $C_+(g) = n$;
2) $C_{x/}(g) = n/2 + 1$ при четном n;
3) $C_{x/}(g) = (n+1)/2$ при нечетном $n \geqslant 9$;
4) $C_{x/}(g) = (n+3)/2$ при нечетном $n \leqslant 7$.

Доказательства нижних оценок сходны с доказательством нижних оценок в теореме 2.I. Верхние оценки в случаях 2),3),4) (которые доказаны только для алгебраически замкнутых полей F нулевой характеристики и вещественно-замкнутых полей) требуют нетривиальных конструкций. В [4I] приведены также оценки на совместное поведение $C_+(\beta)$ и $C_{x/}(\beta)$ для НП β, вычисляющей g.

Теорема 2.2 обобщается на случай нескольких многочленов от одной общей для них всех переменной. Именно, пусть $g_i = a_0^{(i)} + a_1^{(i)} x + \ldots + a_{n_i}^{(i)} x^{n_i}$ $(1 \leqslant i \leqslant k)$ и $\{a_j^{(i)}\}$ алгебраически независимы над $\mathbb{Q}$. Положим $\sum_{1 \leqslant i \leqslant k} n_i = N$, тогда

ТЕОРЕМА 2.3 ([I8]). I) $C_+(g_1, \ldots, g_k) = N$;
2) $C_{x/}(g_1, \ldots, g_k) = N/2 + 1$ в случае четного N;
3) $(N-1)/2 + 1 \leqslant C_{x/}(g_1, \ldots, g_k) \leqslant (N-1)/2 + 2$ в случае нечетного N.

В [I8] построены НП, для которых C_+ и $C_{x/}$ одновременно близки к нижним оценкам из теорем 2.2 и 2.3.

Последний тип НП, который мы рассмотрим в этом параграфе, определяется следующим образом: $F = \overline{\mathbb{Q}(a_0, \ldots, a_n)}$; $P = \{+, \times, /\} \cup \{x\gamma\}_{\gamma \in F}$; $\lambda_m(+) = \lambda_m(x\gamma) = 0$, $\lambda_m(\times) = \lambda_m(/) = 1$; $\{x\}$ – множество входных переменных. Определяемую тем самым сложность C_m будем называть м у л ь т и п л и к а т и в н о й сложностью (т.е. C_m "считает" количество нелинейных операций). Ниже $\asymp$ и $\preccurlyeq$ означают, соответственно, равенство и неравенство с точностью до мультипликативной константы.

ТЕОРЕМА 2.4 ([42]) $C_m(g) \asymp \sqrt{n}$.

Верхняя оценка получается на основе следующего равенства (без ограничения общности считаем, что $n = k^2$): $g = (a_0 + a_1 x + \ldots + a_k x^k) + (a_{k+1} x + a_{k+2} x^2 + \ldots + a_{2k} x^k) x^k + \ldots + (a_{n-k+1} x + \ldots + a_n x^k) x^{n-k}$. Доказательство нижней оценки аналогично доказательству нижних оценок в предыдущих теоремах.

§ 3. Сложность вычисления индивидуальных многочленов

В предыдущем параграфе рассматривались НП для вычисления многочлена $g = a_0 + a_1 x + \dots + a_n x^n$, коэффициенты которого $a_0, \dots, a_n$ алгебраически независимы над F. Гораздо больший интерес представляет случай, когда коэффициенты $a_0, \dots, a_n$ "устроены просто", например, они суть целые или алгебраические числа. Иначе говоря, пусть $F = \overline{\mathbb{Q}}$; $P = \{+, \times, /\} \cup \{x\gamma\}_{\gamma \in F}$; $\lambda_+, \lambda_\times, \lambda_m$ имеют тот же смысл, что и выше; $\{x\}$ — множество входных переменных. Задача состоит в оценке $C_\odot(g)$ для различных $g \in F[x]$.

Доказательства теорем 2.2 и 2.4 позволяют заодно доказать следующее утверждение: размерность многообразия (лежащего в F^{n+1}) векторов коэффициентов $(a_0, \dots, a_n)$ многочленов g, для которых или

$$C_+(g) \leqslant k, \text{ или } C_\times(g) \leqslant k - n + \left[\frac{n}{2}\right] + 1, \text{ либо } C_m(g) \leqslant \sqrt{k},$$

не превосходит $k+1$. В частности, почти все (в смысле алгебраической геометрии, т.е. векторы коэффициентов которых принадлежат всюду плотному множеству в топологии Зарисского) многочлены g удовлетворяют неравенствам

$$C_+(g) = n, \quad C_\times(g) \geqslant \left[\frac{n}{2}\right] + 1, \quad C_m(g) \geqslant \sqrt{n}. \qquad (I)$$

Несмотря на это обстоятельство, не удается удовлетворительно решить даже следующий ослабленный вариант сформулированной выше задачи об оценке $C_\odot(g)$ (хотя в последние годы имеются значительные продвижения в ней, о которых пойдет речь ниже в настоящем параграфе): "явно указать" многочлен g с "хорошими" коэффициентами, удовлетворяющий неравенствам (I) или хотя бы одному из этих неравенств. Это пролило бы свет на тайну сложности. Слова, стоящие в кавычках, требуют уточнения, но из дальнейшего контекста будет ясна разумная постановка задачи (например, слова – "хорошие" коэффициенты – могут означать коэффициенты из множества $\{0, 1\}$).

Отклоняясь несколько в сторону, отметим, что подобная несколько необычная для классической математики ситуация, когда трудно "явно указать" хоть один просто устроенный конкретный элемент из довольно естественного всюду плотного множества (в данном случае множества трудновычислимых многочленов, т.е. многочленов, удовлетворяющих неравенствам (1) или даже более слабым

неравенствам), является весьма распространенной в алгебраической сложности и вызывающей большой интерес (см.также следующие параграфы). Эту не очень четко поставленную проблему "явного указания" трудновычислимых функций (в главе I многочленов или семейств многочленов), мы будем несколько неточно называть проблемой получения нижней оценки.

Перейдем теперь к изложению некоторых продвижений в этой проблеме, происшедших в последние годы. В ряде работ (напр., [55,50,51]), которые были первыми на эту тему, строились явно многочлены, которые удовлетворяют несколько ослабленным неравенствам (I) или какой-то их части, или их дизъюнкции. Методы из этих работ в идейном плане весьма сходны, отличаются рядом технических деталей, и они слабее одного изящного метода Хайнца и Зивикинга ([36]), который мы изложим в настоящем параграфе несколько далее. Поэтому для полноты картины набросаем кратко идею этих методов, следуя первой работе [55] в этом цикле работ. Итак, пусть многочлен $g = a_0 + \ldots + a_d x^d$ вычисляется с помощью НП β, для которой выполнены неравенства $C_+(\beta) \le u$, $C_\times(\beta) \le v_1$, $C_/(\beta) \le v_2$ (здесь, естественно, $\lambda_\times(+) = \lambda_\times(/) = 0$, $\lambda_\times(\times) = \lambda_\times(\times\gamma) = 1$; $\lambda_/(+) = \lambda_/(\times) = \lambda_/(\times\gamma) = 0$, $\lambda_/(/) = 1$). Положим $m = \min\{u, 2(v_1+v_2)\}$. Максимум модулей коэффициентов многочлена назовем его весом.

ТЕОРЕМА 3.I ([55]). Пусть натуральное число $h > d\,\frac{(u+v_1+v_2)^2}{d-m-2}$. Тогда существует нетривиальная форма $H \in \mathbb{Z}[y_0, \ldots, y_d]$, $\deg H \le h$ с весом не больше трех, обладающая тем свойством, что $H(a_0, \ldots, a_d) = 0$.

Набросок доказательства. Коэффициенты рациональной функции, являющейся значением рабочей переменной z_ℓ из НП β (см.обозначения из § I), можно представить как рациональные функции от параметров, которые вводятся в β как константы из F (т.е. каждую вновь введенную в β константу из F считаем параметром) Хотя "выписать явно" эти рациональные функции не представляется возможным, индукцией по ℓ нетрудно явно оценить через ℓ (или в конечном счете через u, v_1, v_2) степени и веса этих рациональных функций. Отсюда следует (здесь мы неявно используем теорему об устранении делений в НП, вычисляющих многочлены [54] – речь о ней в ослабленной форме пойдет ниже в § 7), что вектор коэффициентов многочлена g является значением некоторого вектора многочленов $(P_0, \ldots, P_d)$ с целыми коэффициентами, с ограниченными сверху априори (через u, v_1, v_2) степенями и весами. Тогда лемма Дирихле – Зигеля (см.напр., [55]) дает

верхнюю оценку на степень нетривиальной формы H веса три с целыми коэффициентами, такой что $H(p_0,\dots,p_d)\equiv 0$ (прямое вычисление дает оценку для $\deg H$, указанную в теореме).

В качестве следствия из теоремы 3.I получаем, что если коэффициенты многочлена g степени d не удовлетворяют равенству нулю н и к а к о й формы степени $\leqslant h$ с целыми коэффициентами и веса не больше трех, то многочлен g нельзя вычислить на НП с параметрами u, v_1, v_2. Приведем некоторые приложения указанных и близких соображений.

$$C_m\Big(\sum_{0\leqslant k\leqslant d} 2^{2^{kd^3}} x^k\Big) > \sqrt{d} - 3 \qquad ([55]);$$

$$C_{\mathbb{X}/}\Big(\sum_{1\leqslant k\leqslant d} \exp(2\pi i/2^{kd^2})\, x^k\Big) > \frac{d}{2} - 1 \qquad ([5I]);$$

$$C_+\Big(\sum_{1\leqslant k\leqslant d} \exp(2\pi i/2^{kd^2})\, x^k\Big) \geqslant d - 2 \qquad ([5I]).$$

(ср. с теоремами 2.2 и 2.4).

Как уже упоминалось, более мощный метод для установления нижних оценок сложности вычисления полиномов был предложен в [36], к его изложению сейчас и переходим. Предварительно приведем некоторый дайджест из алгебраической геометрии (все это можно найти, например, в [I5]), необходимый в настоящем и следующем параграфе.

Если X – неприводимое алгебраическое многообразие (над некоторым алгебраически замкнутым полем F), $X\subset \mathbb{P}^k$, где $\mathbb{P}^k$ – проективное k-мерное пространство, то почти все (в смысле топологии Зарисского) линейные пространства $\pi\subset\mathbb{P}^k$ размерности $\dim\pi = k - \dim X$ в пересечении $X\cap\pi$ имеют одно и то же конечное число точек, называемое с т е п е н ь ю $\deg X$ многообразия X. Отметим попутно, что если $X\cap\pi$ состоит из конечного числа s точек, то $s\leqslant \deg X$. Всякое многообразие Y распадается на неприводимые компоненты – $Y = X_1\cup\dots\cup X_t$, тогда $\deg Y = \sum_i \deg X_i$. Понятие степени $\deg X$ инвариантно, т.е. не зависит от вложения X в проективное пространство. Отметим полезное в дальнейшем неравенство Безу:

$$\deg(Y\cap Z)\leqslant \deg Y\cdot\deg Z. \qquad (2)$$

Введем, следуя [36], еще одну меру сложности λ_H (в обозначениях § I), где $H \subset F$ – некоторое бесконечное подполе поля F. Положим $\lambda_H(+) = \lambda_H(\times\gamma) = 0$ для $\gamma \in H$ и $\lambda_H(\times\gamma) = \lambda_H(\times) = \lambda_H(/) = 1$ для $\gamma \in F \setminus H$. Иными словами, бесплатно допускаются умножения на элементы поля H. (типичный пример приложения – см. ниже, когда $F = \overline{\mathbb{Q}}$, $H = \mathbb{Q}$).

Пусть $v, d_1, \ldots, d_s$ – натуральные числа. Далее, пусть C' – одна из сложностных мер C_+, C_H, C_m (обозначения см. в §2). Положим $m = 2v$ в случае рассмотрения C_+ или C_H, и $m = v^2 + 2v$ в случае рассмотрения C_m. Тогда

ТЕОРЕМА 3.2 ([50,5I]). Для всяких $v, d_1, \ldots, d_s$ существуют многочлены $Q_{k,j} \in H[y_1, \ldots, y_m]$, где $d_k \geqslant j \geqslant 0$, $1 \leqslant k \leqslant s$, $\deg Q_{k,j} \leqslant 3jv + 2$ такие, что, если $g_1, \ldots, g_s \in F[x]$, $\deg g_k \leqslant d_k$ $(1 \leqslant k \leqslant s)$ и $C'(g_1, \ldots, g_s) \leqslant v$, то для всех $\xi \in H$, кроме конечного числа, найдутся такие $\eta_1, \ldots, \eta_m \in F$, что

$$g_k = \sum_{0 \leqslant j \leqslant d_k} Q_{k,j}(\eta_1, \ldots, \eta_m)(x - \xi)^j \quad (1 \leqslant k \leqslant s). \tag{3}$$

Идея доказательства сходна с идеей доказательства теоремы 3.I.

Следуя [36], рассмотрим морфизм Ψ аффинных пространств $F^m \xrightarrow{\Psi} F^{d_1} \times \ldots \times F^{d_s} = F^d$ $(d = d_1 + \ldots + d_s)$, определяемый вектором многочленов $(\underbrace{Q_{1,1}, \ldots, Q_{1,d_1}}, \ldots, \underbrace{Q_{s,1}, \ldots, Q_{s,d_s}})$ из теоремы 3.2. Обозначим $W = \overline{\mathrm{Im}\,\Psi}$ – замыкание образа Ψ (в топологии Зарисского). Многообразие $\mathrm{Im}\,\Psi$, а тем самым и W определены над полем H ([I5]). Ниже $\log$ обозначает двоичный логарифм, $|M|$ – мощность множества M.

ЛЕММА 3.3 ([36]). $\dfrac{1}{6} \dfrac{\log \deg W}{\log d} \leqslant m$.

ДОКАЗАТЕЛЬСТВО. Пусть $\Theta_1, \ldots, \Theta_{\dim W}$ – гиперплоскости, такие что $|\mathrm{Im}\,\Psi \cap \Theta_1 \cap \ldots \cap \Theta_{\dim W}| = \deg W$ (из определения степени). Тогда $\operatorname{codim} \Psi^{-1}(\Theta_i) = 1$ и $\deg \Psi^{-1}(\Theta_i) \leqslant 3dv + 2$ (последнее – из теоремы 3.2). Пусть t – число компонент в многообразии $U = \Psi^{-1}(\Theta_1) \cap \ldots \cap \Psi^{-1}(\Theta_{\dim W})$. Тогда по неравенству Безу (2) получаем $t \leqslant \deg U \leqslant \prod_{1 \leqslant i \leqslant \dim W} \deg \Psi^{-1}(\Theta_i) \leqslant (3dv+2)^{\dim W}$. Так как $\Psi(U) = W \cap \Theta_1 \cap \ldots \cap \Theta_{\dim W}$, то $\deg W \leqslant t$, и наконец, $\deg W \leqslant (3dv+2)^{\dim W}$. Это завершает доказательство леммы, так как $\dim W \leqslant m$, и можно считать, что $v \leqslant d$.

Обозначим комбинированную сложность $J = \min\{C_+, C_H, C_m^2\}$.

Пусть $g_k = \sum_{0 \le j \le d_k} g_{k,j} x^j$ $(1 \le k \le s)$. Рассмотрим точку $x = (g_{1,1}, \dots, g_{1,d_1}, \dots, g_{s,d_s}) \in F^d$, и пусть B – замыкание над полем H в F^d этой точки (т.е. B – наименьшее замкнутое определенное над H многообразие, содержащее x – если F алгебраично над H, то B состоит из конечного числа точек). Пусть $\mathcal{D}$ – многообразие, определенное как множество общих корней некоторых полиномов $P_1, \dots, P_r$, где $\deg P_i \le l$ $(1 \le i \le r)$ для некоторого l, и $B \underset{comp}{\subset} \mathcal{D}$ (последнее отношение означает, что все компоненты многообразия B являются также компонентами многообразия $\mathcal{D}$).

ТЕОРЕМА 3.4 ([36]). $\mathcal{J}(g_1, \dots, g_s) \ge \frac{1}{24} \frac{\log \deg B}{\log(dl)}$.

ДОКАЗАТЕЛЬСТВО. Используем теорему 3.2 и принятые в ней обозначения. Допустим, что в равенствах (3) можно положить $\xi = 0$ (если нельзя, то взяв произвольное допустимое значение $\xi = \xi_0$, сведем все к рассматриваемому случаю, сделав замену переменной $x' = x - \xi_0$). Тогда $x \in Im\,\psi \subset W$, и так как W – замкнуто и определено над H, то $B \subset W$, отсюда $B \underset{comp}{\subset} W \cap \mathcal{D}$.

Найдется такая нетривиальная линейная комбинация $P^{(1)} = \sum_{1 \le i \le r} \alpha_i^{(1)} P_i$ $(\alpha_i^{(1)} \in F)$, что многообразие ее нулей (обозначим его $\{P^{(1)} = 0\}$) не содержит (а тем самым, пересекает собственно – см. [15]) никакой компоненты многообразия W, которая не содержится в $\mathcal{D}$. Проделав аналогично $\dim W$ шагов, найдем линейные комбинации $P^{(j)} = \sum_{1 \le i \le r} \alpha_i^{(j)} P_i$ $(1 \le j \le \dim W)$ такие, что $\mathcal{D} \cap W \underset{comp}{\subset} \{P_1 = \dots = P_{\dim W} = 0\} \cap \dim W = E$.

Отсюда, по неравенству Безу (2) имеем

$$\deg B \le \deg(\mathcal{D} \cap W) \le \deg E \le \deg W \cdot l^{\dim W} \le \deg W \cdot l^m.$$

Используя последнее неравенство и лемму 3.3, получаем $m \ge (1/6) \log \deg B / \log(dl)$, откуда вытекает теорема.

Укажем на некоторые приложения теоремы, которые не могут быть получены на основе теоремы 3.1 и метода из [50, 51] – ниже в следствии $F = \mathbb{Q}$, $H = \mathbb{Q}$.

СЛЕДСТВИЕ 3.5 ([36]). $\mathcal{J}(\sum_{1 \le j \le d} \exp(2\pi i / k_j) x^j) \ge$

$\ge \log НОК(k_1, \dots, k_d) / \log(d \cdot \max\{k_1, \dots, k_d\})$, где k_j – натуральные $(1 \le j \le d)$, НОК – наименьшее общее кратное.

В условиях теоремы 3.4 положим $\mathcal{D}$ – многообразие общих нулей многочленов $\{y_1^{k_1}-1=0,\dots,y_d^{k_d}-1=0\}$, и тем самым $l=\max\{k_1,\dots,k_d\}$. Например, $\mathcal{J}(\sum_{1\le j\le d}\exp(2\pi i/j)x^j)\gtrsim d/\log d$ ([36]) и $\mathcal{J}(\sum_{1\le j\le d}\exp(2\pi i/p_j)x^j)\gtrsim d$ ([36]), где p_j – j-ое простое число.

Используя свой метод, Хайнц и Зивикинг ([36]) предложили метод оценки снизу сложности $C_H(\{\sum_{1\le j\le d}\alpha_{ij}x_j\}_{1\le i\le d})$ для $d\times d$ семейства линейных форм. Пусть аналогично тому, как выше, B – замыкание над полем H точки $x=(\alpha_{1,1},\dots,\alpha_{d,d})\in F^{d^2}$, многообразие $\mathcal{D}$ – как и выше. Тогда

УТВЕРЖДЕНИЕ 3.6 ([36]). $C_H(\{\sum_{1\le j\le d}\alpha_{ij}x_j\})\gtrsim\dfrac{\log\deg B}{\log(dl)}$

НАБРОСОК ДОКАЗАТЕЛЬСТВА. Используя [54, 64], можно считать (увеличивая сложность C_H не более, чем в два раза), что значения всех рабочих переменных НП, вычисляющей семейство линейных форм, суть линейные функции. Отсюда заключаем, (рассуждая как при доказательстве теоремы 3.1), что найдутся такие полиномы $Q_{1,1},\dots,Q_{d,d}\in H[y_1,\dots,y_{2v}]$; $\deg Q_{ij}\le 2v$ $(1\le i,j\le d)$, что $x\in\operatorname{Im}\psi$, где $\psi=(Q_{1,1},\dots,Q_{d,d}):F^{2v}\to F^{d^2}$. Пусть $W=\overline{\operatorname{Im}\psi}$, тогда рассуждая аналогично доказательству леммы 3.3, получаем $\deg W\le(2v)^{2v}$. Далее следуем доказательству теоремы 3.4.

Таким образом, построены относительно "простые" по структуре полиномы с коэффициентами из $\mathbb{Q}$, комбинированная сложность $\mathcal{J}$ которых близка к максимально возможной $(O(d))$. Это частично отвечает на вопрос, поставленный в начале этого параграфа.

Упомянем также, что доказано существование (неэффективно) трудновычислимых полиномов с коэффициентами из множества $\{0,1\}$ (см., напр., [50]). Именно, доказано, существование полиномов степени d с коэффициентами из множества $\{0,1\}$ с а) тотальной (см.§ I) сложностью порядка $d/\log d$ (эта оценка точна, как следует из метода [46]); в) мультипликативной сложностью не меньше $\sqrt{d/\log d}$ по порядку (как следует из теоремы 2.4, эта оценка близка к точной); с) аддитивной сложностью не меньше $\sqrt{d}/\log d$ по порядку. Остается нерешенной задача получения более точных оценок для аддитивной сложности полиномов степени d с рациональными коэффициентами. Для нее известны лишь нижняя (т.е. построены примеры с указанной нижней оценкой) и верхняя оценки, соответственно, $\sqrt{d}$ и d по порядку.

§ 4. Метод степени и его обобщения (случай бесконечного основного поля)

На основе методов, изложенных в предыдущем параграфе, не удается доказывать нижние оценки для сложности естественных полиномов или семейств полиномов, так как основным инструментом при рассуждениях является установление некоторой верхней оценки на степень расширения (над примитивным полем) поля, порожденного коэффициентами вычисляемого полинома, через его сложность. В явном виде это имеется в методе Штрассена (теорема 3.I), в более завуалированном (оценка степени множества B) – в методе Хайнца – Зивикинга (лемма 3.3 и теорема 3.4).

В настоящем параграфе будут изложены методы, в основе которых лежит использование понятия и свойств степени многообразия (см. выше § 3 или [15]), и дающие нелинейные (относительно числа переменных) нижние оценки на мультипликативную сложность для некоторых естественных семейств многочленов многих переменных (см. ставшую уже классической работу [23]) и для индивидуальных многочленов (см. [25,50]).

Итак, в терминах § I, F – алгебраически замкнутое поле; $P=\{+,\times,/\}\cup\{x\gamma\}_{\gamma\in F}$; $\lambda=\lambda_m$. Требуется оценить мультипликативную сложность $C_m(g_1,\dots,g_k)$ семейства рациональных функций от входных попарно коммутирующих переменных $x_1,\dots,x_n$. Функции $g_1,\dots,g_k$ задают рациональное отображение $F^n \xrightarrow{G=(g_1,\dots,g_k)} F^k$. Рассмотрим его график $Graph(G)\subset F^{n+k}$ и обозначим через $W=\overline{Graph(G)}\subset \mathbb{P}^{n+k}$ его проективное замыкание. Заметим, что $Graph(G)$ – открытое подмножество в неприводимом замкнутом многообразии W, поэтому $\deg W = \deg Graph(G)$.

ТЕОРЕМА 4.I ([23]). $C_m(g_1,\dots,g_k)\geqslant \log_2 \deg W$.

Теорема эта достаточно широко известна, поэтому по поводу ее доказательства ограничимся лишь замечанием, что оно происходит индукцией по $C_m(g_1,\dots,g_k)$ и использует неравенство Безу (2).

Укажем на некоторые приложения теоремы Штрассена

$$C_m\Big(\sum_{1\leqslant i\leqslant d} a_i x_1^i,\dots,\sum_{1\leqslant i\leqslant d} a_i x_d^i\Big)\asymp d\log d$$

для всякого полинома степени d (т.е. $a_d\neq 0$) над F (вычисление конкретного полинома в точности степени d в d

точках, т.е. множеством входных переменных – см. § I – здесь является $\{x_1,\dots,x_d\}$). Пусть $\sigma_i=\sum_{1\le j_1<\dots<j_i\le d} x_{j_1}\dots x_{j_i}$ – элементарная симметрическая функция степени i, тогда $C_m(\sigma_1,\dots,\sigma_d)\asymp d\log d$. Далее, задача интерполяции многочлена степени d, т.е. восстановление его коэффициентов по значениям в $(d+1)$ различной точке, также имеет мультипликативную сложность по порядку $d\log d$. Отметим, что все оценки из упомянутых приложений верны также и для произвольного бесконечного основного поля F, так как НП, вычисляющее семейство полиномов над бесконечным полем, вычисляет это семейство над любым его расширением.

Теорема 4.I, к сожалению, не дает нетривиальной оценки для мультипликативной сложности отдельного многочлена, так как $\deg W\le\prod_{1\le i\le k}\deg g_i$, где $\deg(g_1/g_2)=\max\{\deg g_1,(\deg g_2)+1\}$. Этот недостаток был устранен впервые Шнорром ([30]), к изложению метода которого мы и переходим (по-прежнему считаем, что поле F алгебраически замкнуто).

Пусть $p=a_0(x_1,\dots,x_n)+a_1(x_1,\dots,x_n)y+\dots+a_d(x_1,\dots,x_d)y^d\in F[y,x_1,\dots,x_n]$ и $C_m(p)=v$. Рассмотрим НП β, вычисляющую p, такую что $C_m(\beta)=v$. Хотелось бы, неформально говоря, преобразовать β в некоторое НП $\tilde\beta$, вычисляющую коэффициенты многочлена p, т.е. многочлены $a_0,\dots,a_d\in F[x_1,\dots,x_n]$. Но если в β есть деления, то сделать это было бы затруднительно, так как естественный путь такого преобразования – вычислять все коэффициенты при степенях (не превосходящих d) переменной y для всех значений (которые можно рассматривать как степенные ряды от y) рабочих переменных z_ℓ НП β (см.§ I), что невозможно, если требуется, например, разделить на ряд от y с нулевым свободным членом. Этот недостаток можно в принципе устранить, рассматривая степенные ряды от новой переменной $(y-\eta)$ (вместо y) для некоторого $\eta\in F$ (даже для почти всех $\eta\in F$). Поэтому можно ввести η в НП как новую входную переменную, т.е. рациональную функцию $f_\ell\in F(y,x_1,\dots,x_n)$ являющуюся значением некоторой рабочей переменной z_{m_ℓ} из НП β (через m_ℓ мы обозначили номер строки в β, в которой выполняется ℓ-ое двуместное умножение или деление – см.пункт 4) определения из § I), запишем в виде $f_\ell=\sum_{i\ge0}b_{i,\ell}(\eta,x_1,\dots,x_n)(y-\eta)^i$, где $b_{i,\ell}$ – рациональные функции.

Используя не более v операций $\times$ и $/$, вычислим индукцией по ℓ все $\{b_{0,\ell}\}_{1\le\ell\le v}$, т.е. свободные члены в степенных

рядах от переменной $(y-\eta)$, соответствующих функциям f_ℓ. Затем, индукцией по ℓ можно показать, что всякий коэффициент $b_{i,\ell}$ можно представить как некоторый многочлен $Q_{i,\ell}$ степени не больше $2i\ell$ (ср.теорему 3.2) от параметров $\eta, x_1, \ldots, x_n$ и $\{b_{0,\ell}\}_{1\le\ell\le v}$. Обратим внимание, что мы не вычисляем на самом деле коэффициентов $b_{i,\ell}$ (в противовес неформальному изложению идеи метода Шнорра – см.выше), так как это достаточно трудоемко, а строим некоторое подходящее для них представление.

Пусть $p=\tilde a_0(\eta,x_1,\ldots,x_n)+\tilde a_1(\eta,x_1,\ldots,x_n)(y-\eta)+\ldots+\tilde a_d(\eta,x_1,\ldots,x_n)(y-\eta)^d$, рассмотрим рациональное отображение $F^{n+1}\xrightarrow{A=(\tilde a_0,\ldots,\tilde a_d)}F^{d+1}$, которое можно разложить в композицию двух рациональных отображений

$$F^{n+1}\xrightarrow{\varphi=(x_1,\ldots,x_n,\eta,b_{0,1},\ldots,b_{0,v})}F^{n+1+v}\xrightarrow{\psi=(\tilde a_0,\ldots,\tilde a_d)}F^{d+1}.$$

Согласно доказанному выше и теореме 4.1, $\deg Graph(\varphi)\le 2^v$. Далее, так как $\tilde a_i=Q_i(\{\eta,\{x_i\},\{b_{0,\ell}\})$, где $Q_i=\sum_{1\le j\le v}\alpha_{ij}Q_{i,j}$ – некоторая линейная $(\alpha_{ij}\in F)$ комбинация упомянутых выше многочленов, поэтому $\deg Q_i\le 2vi$, откуда $\deg Graph(\psi)\le(2vd)^d$. Нетрудно проверить, что $\deg Graph(\psi\circ\varphi)\le\deg Graph(\psi)\cdot\deg Graph(\varphi)$, следовательно, $\deg Graph(A)\le 2^v(2vd)^d$. С другой стороны, $Graph(a_0,\ldots,a_d)=Graph(\tilde a_0,\ldots,\tilde a_d)\cap\pi\subset F^{n+d+2}$, где π – гиперплоскость с уравнением $\eta=0$, поэтому согласно неравенству Безу (2) $\deg Graph(a_0,\ldots,a_d)\le\deg Graph(\tilde a_0,\ldots,\tilde a_d)\le 2^v(2vd)^d$. Итак,

ТЕОРЕМА 4.2 ([50]). Пусть $p=a_0(x_1,\ldots,x_n)+a_1(x_1,\ldots,x_n)y+\ldots+a_d(x_1,\ldots,x_n)y^d$. Тогда $C_m(p)>v$ для v, такого что $\deg Graph(a_0,\ldots,a_d)>2^v(2vd)^d$.

Некоторые приложения теоремы ([50]):

$$C_m\Big(\sum_{0\le i\le d}x_i^k y^i\Big)\gtrsim d\log k,\qquad\text{если } d^4\prec k;$$

$$C_m\Big(\sum_{0\le i\le d}(x_1+\ldots+x_i)^k y^i\Big)\gtrsim d\log k,\qquad\text{если } d^4\prec k.$$

Отметим, что в указанных приложениях как и выше в аналогичной ситуации, при обсуждении следствий из теоремы 4.1, можно считать, что F – произвольное бесконечное поле.

Более элегантный метод, позволяющий, помимо всего прочего, получать нелинейные нижние оценки сложности для индивидуальных

полиномов, был предложен Бауэром и Штрассеном ([25]) и основывался на следующей оценке для сложности вычисления рациональной функции $f \in F(x_1, \dots, x_n)$ и всех ее первых частных производных (здесь F – любое поле).

ТЕОРЕМА 4.3 ([25]). I) $C_m\left(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) \leqslant 3\, C_m(f)$;

2) $C_t\left(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) \leqslant 5\, C_t(f)$. (см.§ I).

ДОКАЗАТЕЛЬСТВО приведем только для пункта I). Рассмотрим НП β, вычисляющую f с $C_m(\beta) = C_m(f) = r$. Пусть $g_1, \dots, g_r$ значения рабочих переменных из НП β, в строках, содержащих двухместные умножения или деления (см.пункт 4) из определения в § I). Тогда для всякого i $g_i = u_i \odot v_i$ ($\odot = \times$ или $\odot = /$) и

$$u_i = \sum_{1 \leqslant j \leqslant i-1} \beta_{ij} g_j + p_i; \quad v_i = \sum_{1 \leqslant j \leqslant i-1} \gamma_{ij} g_j + q_i;$$

$$f = \sum_{1 \leqslant j \leqslant r} \alpha_j g_j + m,$$

где $\alpha_i, \beta_{ij}, \gamma_{ij} \in F$; $\max\{\deg p_i, \deg q_i, \deg m\} \leqslant 1$.

ЛЕММА 4.4 ([25]). Пусть $0 \neq y_i \in F$ $(1 \leqslant i \leqslant s)$; $\alpha_{ij} \in F$ $(1 \leqslant j < i \leqslant s)$; $w_1, \dots, w_s$ – переменные. Определим h_i $(1 \leqslant i \leqslant s)$ индукцией по i: $h_1 = y_1 w_1, \dots, h_i = y_i \left(\sum_{1 \leqslant j \leqslant i-1} \alpha_{ij} h_j + w_i\right), \dots$.

Обозначим $h_s = \sum_{1 \leqslant \sigma \leqslant s} d_\sigma w_\sigma$, $d_\sigma \in F$. Тогда

$$d_s = y_s; \quad d_j = \left(\sum_{j+1 \leqslant \sigma \leqslant s} d_\sigma \alpha_{\sigma j}\right) y_j \quad \text{для} \quad 1 \leqslant j < s.$$

ДОКАЗАТЕЛЬСТВО ЛЕММЫ. Обозначим $h_i = \sum_{1 \leqslant \sigma \leqslant s} d_{i\sigma} w_\sigma$. Рассмотрим нижнетреугольные матрицы

$$\mathbb{D} = \begin{pmatrix} d_{11} & & 0 \\ \vdots & \ddots & \\ d_{s1} & \dots & d_{ss} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & & 0 \\ \alpha_{21} & 0 & \\ \alpha_{s1} & \dots & \alpha_{s,s-1}\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 & & 0 \\ & \ddots & \\ 0 & & y_s \end{pmatrix}.$$

Тогда условие леммы равносильно равенству $\mathbb{D} = Y(A\mathbb{D} + E)$, где E – единичная матрица. Умножив это равенство слева на матрицу $\mathbb{D}Y^{-1}$, справа на $\mathbb{D}^{-1}Y$, получаем $\mathbb{D} = (\mathbb{D}A + E)Y$, откуда вытекает лемма.

Вернемся к доказательству теоремы. Положим $s = 3r+1$;

$$\alpha_{3i-2,l}=\begin{cases}\beta_{ij}, \text{ если } l=3j\\ 0, \text{ иначе}\end{cases};\quad \alpha_{3i-1,l}=\begin{cases}\gamma_{ij}, \text{ если } l=3j\\ 0, \text{ иначе}\end{cases};$$

$$\alpha_{3i,l}=\begin{cases}1 \text{ для } l=3i-2 \text{ или } l=3i-1\\ 0, \text{ иначе}\end{cases} \quad \text{для } 1\le l<3i\le s \text{ и}$$

$$\alpha_{s,l}=\begin{cases}a_j, \text{ если } l=3j\\ 0, \text{ иначе}\end{cases}$$

Далее, если $g_i=u_i\cdot v_i$, то положим $y_{3i-2}=v_i$, $y_{3i-1}=u_i$, $y_{3i}=1$. Если же $g_i=u_i/v_i$, то $y_{3i-2}=1$, $y_{3i-1}=-u_i/v_i$, $y_{3i}=1/v_i$.

Наконец, положим $\xi_{3i-2,\nu}=\dfrac{\partial p_i}{\partial x_\nu}$,

$$\xi_{3i-1,\nu}=\frac{\partial q_i}{\partial x_\nu},\ \xi_{3i,\nu}=0\ (1\le i\le r);\quad \xi_{s,\nu}=\frac{\partial m}{\partial x_\nu}.$$

Тогда индукцией по i легко проверяется, что если $h_1,\dots,h_s$ определены как в лемме 4.4 для указанных параметров α_{ij}, y_j, то $\dfrac{\partial g_i}{\partial x_\nu}=h_{3i}(\xi_{1,\nu},\dots,\xi_{s,\nu})$. Отсюда получаем $\dfrac{\partial f}{\partial x_\nu}=h_s(\xi_{1,\nu},\dots,\xi_{s,\nu})=$ $=\sum_{1\le\sigma\le s} d_\sigma\xi_{\sigma,\nu}$. Поэтому $C_m\left(\dfrac{\partial f}{\partial x_1},\dots,\dfrac{\partial f}{\partial x_n}\right)\le C_m(d_1,\dots,d_s)$, так как $\xi_{\sigma,\nu}\in F$ для всех σ,ν. По лемме 4.4 семейство $\{d_1,\dots,d_s\}$ вычисляется (если уже вычислены $g_1,\dots,g_r$) с использованием $(s-1)$ двухместных умножений на $y_{s-1},\dots,y_1$ последовательно. Так как среди $\{y_1,\dots,y_{s-1}\}$ по крайней мере r равны единице, то для вычисления $\dfrac{\partial f}{\partial x_1},\dots,\dfrac{\partial f}{\partial x_n}$ достаточно использовать не более $(s-1)-r=2r$ двухместных умножений и делений, что доказывает теорему.

Упомянем некоторые приложения теоремы 4.3.

СЛЕДСТВИЕ 4.4 ([25]) 1) $\frac{1}{3}n\log(m-1)\le C_m\left(\sum_{1\le i\le n}x_i^m\right)\le 2n\log m$;

2) $\frac{1}{3}\max\{q,n-q+1\}\log\min\{q-1,n-q\}\le C_m(\sigma_q)\le n\log\min\{q,n-q\}+2n$,

где $\sigma_1,\dots,\sigma_n$ – элементарные симметрические функции от n переменных (ср. приложения теоремы 4.1);

3) $\frac{1}{3}n\log n\le C_m\left(\prod_{i,j}(x_i-y_j)\right)\le n(2\log n+1)$;

4) $\frac{1}{6}n\log n-\frac{2}{3}n\le C_m\left(\prod_{i\ne j}(x_i-x_j)\right)=O(n\log n)$;

5) Пусть матрица $(b_{ij})=(a_{ij})^{-1}$ $(1\le i,j\le n)$. Тогда $C_m(\{b_{ij}\})\le 3C_m(\det(a_{ij}))+n^2$.

Доказательство нижних оценок в пунктах I)–4) происходит путем применения теоремы 4.I к набору из частных производных вычисляемой функции. Доказательство пункта 5) опирается на представление $b_{\ell,\kappa}=A_{\kappa,\ell}/\det(a_{ij})$, где A_{ij} — (i,j) – минор. С другой стороны, по правилу Крамера $\frac{\partial\det(a_{ij})}{\partial a_{\kappa\ell}}=\pm A_{\kappa\ell}$.

Таким образом, из пункта 5) и известных ранее результатов получаем совпадение (с точностью до мультипликативной константы) сложности умножения матриц, обращения матриц и вычисления определителя (см.также обзор В.И.Солодовникова в настоящем сборнике).

Обратим внимание, что теорема 4.3 не обобщается прямо на вычисление вторых частных производных (и тем самым, на семейства нескольких функций – $\{f_1,\dots,f_k\}$ – в отличие от одной функции $\{f\}$). В качестве контрпримера рассмотрим многочлен (моном) $f=x_1\cdots x_n$. Тогда $\frac{\partial^2 f}{\partial x_i\partial x_j}=\frac{f}{x_i x_j}$ при $i\ne j$ и

$$C_m\left(\left\{\frac{f}{x_i x_j}\right\}\right)\ge\dim\left(\left\{\frac{f}{x_i x_j}\right\}_{i>j}\right)=\frac{n(n-1)}{2}.$$

В заключение настоящего параграфа приведем одно приложение, принадлежащее Штрассену, метода степени для вычисления евклидова представления (или непрерывной дроби) рациональной функции на вычислительной модели, несколько отличной от НП ([57]). Если $A_1,A_2\in F[x]$, $\deg A_1=n\ge\deg A_2=m\ge 0$, то применим к A_1/A_2 алгоритм Евклида, получим при этом последовательность равенств

$$A_1=Q_1A_2+A_3,\ A_2=Q_2A_3+A_4,\dots,A_{t-1}=Q_{t-1}A_t.$$

Вектор полиномов $(Q_1,Q_2,\dots,Q_{t-1},A_t)$ называется е в к л и д о в ы м п р е д с т а в л е н и е м дроби A_1/A_2. Вектор степеней $(n_1,\dots,n_t)=(\deg Q_1,\dots,\deg Q_{t-1},\deg A_t)$ назовем ф о р м а т о м дроби A_1/A_2 (или пары (A_1,A_2))

Очевидно, $m=\sum_{2\le i\le t} n_i$, $n=\sum_{1\le i\le t} n_i$. Через $\mathcal{D}(n_1,\dots,n_t)$ обозначим множество пар (A_1,A_2) , имеющих формат $(n_1,\dots,n_t)$.

Ясно, что НП – неподходящая модель для вычисления евклидова представления, так как разные дроби даже при одних и тех же значениях n,m могут иметь разный формат, т.е. разный вид ответа.

Поэтому была введена ([57]) следующая вычислительная модель адекватная для данной задачи и названная ветвящимся деревом вычислений (ВДВ). ВДВ содержит дерево T, ориентированное от корня к листьям. Любая вершина дерева T имеет одного или двух сыновей. Во всякой вершине, имеющей двух сыновей (назовем ее вершиной ветвления) стоит произвольный полином, во всякой вершине, имеющей одного сына (назовем ее вершиной вычисления) стоит некоторая базисная операция (из P — см. § I). Аргументами как упомянутого полинома, так и базисной операции являются результаты каких-то вершин вычисления, расположенных на единственной ветви от корня к рассматриваемой вершине. В листе выдается ответ, состоящий из результатов расположенных над ним вершин вычисления.

Функционирование ВДВ естественное. В корень дерева T подается вход, и вычисление происходит вдоль некоторой однозначно определенной ветви: во всякой вершине вычисления вычисляется соответствующая базисная операция; после вершины ветвления, в которой стоит полином P, вычисление идет дальше по одной из ветвей в зависимости от того, равно нулю или нет значение полинома P.

Если ВДВ с деревом T вычисляет евклидово представление, то всякие два входа, на которых вычисление происходит по одной и той же ветви дерева T, имеют одинаковый формат, и тем самым, всякой ветви (или соответствующему листу) можно приписать формат. Если задан некоторый вес $\lambda_\odot$ на базисных операциях (см. § I), то вес ветви — сумма весов базисных операций, стоящих в вершинах вдоль этой ветви. Определим сложность ВДВ как функцию от формата: $C_\odot(n_1,\dots,n_t)$ положим равным максимуму весов ветвей с форматом $(n_1,\dots,n_t)$. Нормированная энтропия ([3]) определяется как $-\frac{1}{n}\sum_{1\le i\le t} n_i \log\left(\frac{n_i}{n}\right) = H(n_1,\dots,n_t)$.

ТЕОРЕМА 4.6 ([57]). I) (Д.Кнут и А.Шенхаге). Можно построить ВДВ для вычисления евклидова представления с оценкой

$$C_m(n_1,\dots,n_t) \le 30n(H(n_1,\dots,n_t)+6{,}5).$$

2) Всякое ВДВ, вычисляющее евклидово представление над алгебраически замкнутым основным полем F, имеет мультипликативную сложность $C_m(n_1,\dots,n_t) \ge n(H(n_1,\dots,n_t)-2)$, т.е. для всяких $(n_1,\dots,n_t)$ существует открытое в $\mathcal{D}(n_1,\dots,n_t)$ множество U, на каждом элементе которого ВДВ имеет мультипликативную сложность не меньше указанной величины.

Отметим (см. [57]), что мультипликативная сложность вычис-

ления произведения t полиномов от одной (для всех) переменной степеней $n_1,\dots,n_t$ соответственно, над бесконечным основным полем F имеет порядок $nH(n_1,\dots,n_t)$. Доказательство пункта 2) теоремы 4.6 и нижней оценки в последнем замечании опирается на теорему 4.I.

§ 5. Метод степени (случай конечного основного поля)

Методы, изложенные в предыдущем параграфе, работают лишь в случае бесконечного основного поля, так как существенно используется то обстоятельство, что если некоторое НП β вычисляет многочлен (или семейство многочленов) над бесконечным полем, то β вычисляет этот же многочлен (или семейство) и над любым его расширением, в частности, над его алгебраическим замыканием, для которого уже применяется алгебро-геометрическая техника, связанная со степенью многообразия. Для случая конечного основного поля F Штрассен ([56]) предложил иной метод, использующий, впрочем, также теорему 4.I (см.§ 4). Его изложению и посвящен настоящий параграф.

Итак, пусть F – конечное поле, $P=\{+,\times,/\}\cup\{x\gamma\}_{\gamma\in F}$; $\lambda=\lambda_m$. Задача состоит в оценке $v=C_m(g_1,\dots,g_k)$ для $g_1,\dots,g_k\in F[x_1,\dots,x_n]$, где $\{x_1,\dots,x_n\}$ – попарно коммутирующие входные переменные (см. § I). Пусть НП β вычисляет $\{g_1,\dots,g_k\}$ и $C_m(\beta)=v$. Обозначим $S=Graph(F^n \xrightarrow{G=(g_1,\dots,g_k)} F^k)\subset F^{n+k}$.

Будем рассматривать НП β как НП $\bar\beta$ над алгебраическим замыканием $\bar F$. НП $\bar\beta$ вычисляет тогда некоторые многочлены $\bar g_1,\dots,\bar g_k\in \bar F[x_1,\dots,x_n]$, такие что сужение $\bar g_i|_{F^n}=g_i$ $(1\le i\le k)$, и $\bar g_i$ возможно не совпадает с g_i на всем $\bar F^n$ (этот момент и отличает случай конечного основного поля от бесконечного). Очевидно, $C_m(\bar g_1,\dots,\bar g_k)\le v$. Рассмотрим теперь неприводимое многообразие $W=Graph(\bar F^n \xrightarrow{\bar G=(\bar g_1,\dots,\bar g_k)} \bar F^k)\subset \bar F^{n+k}$. Выполнено $S\subset W$ (считаем, что $F^{n+k}\subset \bar F^{n+k}$ – естественное вложение). Идея метода Штрассена состоит в том, чтобы найти эффективные достаточные условия на конечное множество точек S_1 при выполнении которых, всякое неприводимое многообразие $W_1\supset S_1$ имело бы достаточно большую степень (при фиксированной размерности $\dim W_1$), а затем применить теорему 4.I.

Скажем (см.[56]), что конечное подмножество $S_1\subset \bar F^N$ является t-м н о ж е с т в о м (t – некоторое натуральное число), если для всех $0\le d\le N$ и для всякого неприводимого

замкнутого подмножества $W_1 \subset \overline{F}^N$ выполнено $\deg W_1 \geq$ $\geq |S_1 \cap W_1| / t^{\dim W_1}$. Следующая лемма представляет отдельный интерес, по-видимому, и для специалистов по алгебраической геометрии.

ЛЕММА 5.I ([56]). Пусть $S_1 \subset \overline{F}^N$, t - натуральное число, $l_1, \ldots, l_q$ - линейные формы на $\overline{F}^N$, такие что

а) для всякого $1 \leq i \leq q$ и любых $c_1, \ldots, c_{i-1} \in \overline{F}$ линейная форма l_i принимает не более t значений на множестве

$$S_1 \cap \{ y \in \overline{F}^N : l_1(y) = c_1, \ldots, l_{i-1}(y) = c_{i-1} \};$$

в) если $l_i(y) = l_i(u)$, для некоторых $y, u \in S_1$ и для всех $1 \leq i \leq q$, то $y = u$.

Тогда S_1 является t-множеством.

Лемма доказывается индукцией по q.

ТЕОРЕМА 5.2 ([56]). Пусть $g_1, \ldots, g_K \in F[x_1, \ldots, x_n]$ и множество $S_1 = S = Graph(g_1, \ldots, g_K) \subset F^{n+K} \subset \overline{F}^{n+K}$ удовлетворяет условиям леммы 5.I. Тогда $C_m(g_1, \ldots, g_K) \geq \log(|S| / t^n)$.

Для доказательства теоремы достаточно заметить, что неприводимое замкнутое n-мерное многообразие $\overline{W} = Graph(\overline{G})$ содержит t-множество S, поэтому $\deg W \geq |S| / t^n$, и применяя теорему 4.I, получаем

$$v \geq C_m(\overline{g}_1, \ldots, \overline{g}_K) \geq \log \deg W \geq \log(|S| / t^n).$$

В качестве приложения теоремы 5.2 (см.[56]), получаем, что

$$C_m\Big(\sum_{1 \leq i \leq n} a_i x_1^i, \ldots, \sum_{1 \leq i \leq n} a_i x_n^i\Big) \asymp O(n \log(\min\{n, |F|\})).$$

($a_n \neq 0$). Для задачи интерполяции полинома степени n по значениям в $(n+1)$ точке (задача осмыслена, если $|F| > n$), ее мультипликативная сложность также равна по порядку $n \log n$ (напомним, что для этих двух задач те же оценки были в случае бесконечного поля - см. приложения теоремы 4.I). Иначе обстоит дело для элементарных симметрических функций: в [I7] показано, что даже тотальная сложность $C_t(\sigma_1, \ldots, \sigma_n)$ линейна по n над конечным полем F (ср. § 4).

Интересно отметить, что бывает в некоторых естественных случаях и обратная ситуация, т.е. сложность вычисления над конечным полем семейства многочленов может быть больше сложности вычисления этого семейства над бесконечным полем. Например, мультиплика-

тивная сложность умножения двух полиномов степени n над бесконечным полем F равна $2n+1$ (см., например, [31]); в случае поля F из двух элементов в [28] для мультипликативной сложности для этой задачи доказана (см.также ниже конец § 10) нижняя оценка $3{,}52n$ (наилучшая известная на настоящий момент для нее верхняя оценка – $n\cdot g(n)$, где g – некоторая функция, растущая медленнее любой фиксированной итерации логарифма – см. [8,9,32, 33]).

§ 6. Аддитивная сложность и вещественные корни

В предыдущих двух параграфах установлены нижние оценки для мультипликативной сложности $C_m(g_1,\dots,g_k)$ семейства многочленов через степень графика $W=Graph(F^n \xrightarrow{G=(g_1,\dots,g_k)} F^k)$.

Так как $\deg W$ не меньше числа N_1 дискретных корней системы $\overline{g}_1=\dots=\overline{g}_k=0$, над полем $\overline{F}$, то $C_m(g_1,\dots,g_k)\geq \log N_1$ (ср.теорему 4.1).

В настоящем параграфе на основе работы Хованского [21] будет указана нижняя оценка для аддитивной сложности $C_+(g_1,\dots,g_n)$ через число вещественных корней системы $g_1=\dots=g_n=0$. В этом параграфе в обозначениях § 1 $F=\mathbb{R}$; $P=\{+,\times,/\}\cup\{x\gamma\}_{\gamma\in\mathbb{R}}$; $\lambda=\lambda_+$ (см.начало § 2); $g_1,\dots,g_n\in\mathbb{R}[x_1,\dots,x_n]$. Ниже все многочлены, если не оговорено противного, предполагаются вещественными. Наличие упомянутой оценки предполагалось давно и основывалось это предположение на правиле Декарта: число неотрицательных корней полинома от одной переменной не превосходит числа его мономов. Для одного многочлена g от одной переменной более слабая оценка, чем установлена ниже, была получена в [26] (по-видимому, самостоятельный интерес представляет метод доказательства основной теоремы в этой работе). Через $\mathbb{R}^*$ обозначим множество ненулевых вещественных чисел.

ТЕОРЕМА 6.1 ([21]). Система уравнений $g_1=\dots=g_n=0 (g_1,\dots,g_n\in\mathbb{R}[x_1,\dots,x_n])$, имеет не более $2^n(n+2)^k 2^{k(k+1)/2}$ дискретных корней в $(\mathbb{R}^*)^n$, где k – общее число мономов во всех многочленах $g_1,\dots,g_n$.

СЛЕДСТВИЕ 6.2. Если система $g_1=\dots=g_n=0$ имеет N простых (т.е. с единичной кратностью) дискретных корней в $(\mathbb{R}^*)^n$, то $C_+(g_1,\dots,g_n)\geq \frac{\sqrt{\log N}-2n}{3}$.

ДОКАЗАТЕЛЬСТВО СЛЕДСТВИЯ. Пусть $y_1,\dots,y_N$ – указанные простые корни системы $g_1=\dots=g_n=0$ (корни из $(\mathbb{R}^*)^n$ будем называть нетривиальными). Для вывода следствия из теоремы воспользуемся известной канонической формой для НП (см., например, [26]), содержащих не более $v=C_+(g_1,\dots,g_n)$ сложений:

$$\vdots$$

$$T_{l+1}=T_1^{i_1^{(l+1)}}\dots T_l^{i_l^{(l+1)}}x_1^{u_1^{(l+1)}}\dots x_n^{u_n^{(l+1)}}+T_1^{j_1^{(l+1)}}\dots T_l^{j_l^{(l+1)}}x_1^{w_1^{(l+1)}}\dots x_n^{w_n^{(l+1)}}$$

$$\vdots$$

$$G_1=T_1^{p_1^{(1)}}\dots T_v^{p_v^{(1)}}x_1^{q_1^{(1)}}\dots x_n^{q_n^{(1)}} \qquad (4)$$

$$\vdots$$

$$G_n=T_1^{p_1^{(n)}}\dots T_v^{p_v^{(n)}}x_1^{q_1^{(n)}}\dots x_n^{q_n^{(n)}},$$

где i,j,u,w,p,q с индексами – целые числа; T_l – рабочая переменная НП (см. § I) в строке, в которой происходит l-ая по счету операция сложения; значение рабочей переменной G_i равно g_i $(1\le l\le v,\ 1\le i\le n)$.

Покажем, что можно так модифицировать систему (4), заменив $G_1,\dots,G_n$ на отличные от нуля достаточно маленькие по модулю вещественные числа $\varepsilon_1,\dots,\varepsilon_n$, соответственно, чтобы модифицированная система из $(v+n)$ уравнений от $(v+n)$ неизвестных $x_1,\dots,x_n,T_1,\dots,T_v$ имела бы не меньше N нетривиальных корней. Так как $y_1,\dots,y_N$ – простые корни, то по теореме о неявной функции отображение $\mathbb{R}^n \xrightarrow{G=(g_1,\dots,g_n)} \mathbb{R}^n$ биективно для каждого $1\le i\le N$ в некоторой окрестности $Y_i\ni y_i$ – можно считать, сузив окрестности Y_i, что окрестности нуля $G(Y_i)=Q$ совпадают и $Y_i\subset(\mathbb{R}^*)^n$ для всех $1\le i\le N$. Рассмотрим отображение $Y_i \xrightarrow{T_{l+1}^{(i)}=T_{l+1}} \mathbb{R}$ (здесь мы отождествляем рабочую переменную T_{l+1} с ее значением). Оно не является тождественным нулем (в противном случае, $(l+1)$-ая строка НП лишняя и ее можно вычеркнуть), поэтому прообраз нуля $(T_{l+1}^{(i)})^{-1}(0)=M_{i,l+1}\subset Y_i$ является вещественным алгебраическим многообразием размерности меньше n. Рассмотрим $M=\bigcup_{i,l} G(M_{i,l+1})$ – подмногообразие размерности меньше n (так как G/Y_i – биективный морфизм) окрестности Q нуля в $\mathbb{R}^n$. Теперь в качестве $(\varepsilon_1,\dots,\varepsilon_n)$ возьмем произвольную точку в Q вне M и координатных гипер-

плоскостей. В силу выбора $(\varepsilon_1,\dots,\varepsilon_n)$ модифицированная система имеет не меньше N нетривиальных корней (в окрестностях $\{Y_i\}_{1\le i\le N}$ точек $\{y_i\}_{1\le i\le N}$).

Модифицированная система содержит $k=(2n+3v)$ мономов. Подставив это значение k в теорему 6.1, и заметив, что $2^n(n+2)^k < 2^{k(k+1)/2}$ в нашем случае, завершаем доказательство следствия.

В качестве приложения рассмотрим многочлен $P_d(x)=(x-1)\dots(x-d)$ и семейство многочленов $\{P_d(x_1),\dots,P_d(x_n)\}$ от переменных $x_1,\dots,x_n$ при $d \succ 2^{9n}$, тогда $C_+(P_d(x_1),\dots,P_d(x_n)) \succ \sqrt{n\log d}$ (при указанных ограничениях эта оценка нелинейна по n).

Кушниренко принадлежит гипотеза (см. [21]) о том, что в условиях теоремы 6.1 верна более сильная верхняя оценка для числа нетривиальных корней: именно, $2^n(k_1-1)\dots(k_n-1)$, где k_i – число мономов в g_i. Гипотеза остается пока недоказанной при $n>1$. Справедливость гипотезы дала бы оценку $C_+(g_1,\dots,g_n)\succeq \log N$, что явилось бы элегантным аналогом теоремы 4.1 Штрассена.

Видоизменив доказательство теоремы 4.2, можно получить нижнюю оценку аддитивной сложности для индивидуальных многочленов. Пусть $g=\sum_{1\le i\le n} g_i y^i$, где $g_1,\dots,g_n\in\mathbb{R}[x_1,\dots,x_n]$.

ТЕОРЕМА 6.3. Пусть N – число простых нетривиальных корней системы уравнений $g_1=\dots=g_n=0$. Тогда $C_+(g)\succeq(\log N)^{1/2(n+2)}-n$.

Пусть НП β таково, что $v=C_+(\beta)=C_+(g)$ и β вычисляет g. Предположим, что значения всех рабочих переменных z_ℓ НП β, которые суть рациональные функции из $\mathbb{R}(x_1,\dots,x_n,y)$, определены в точке $y_0=\eta\in\mathbb{R}$. Тогда сделаем замену переменной $y_1=y-\eta$. Каждое значение рабочей переменной z_{a_ℓ}, где a_ℓ – номер строки НП β (см. § I и доказательство следствия 6.2), содержащей ℓ-ую по счету операцию сложения, будем рассматривать как степенной ряд от y_1: т.е. $T_\ell=\sum_{i\ge 0} b_i^{(\ell)} y_1^i$ (см.(4)); кроме того, положим $T_0=y=y_1+\eta$.

Перестроим НП β в НП β_1, вычисляющую рекурсией по ℓ набор $\{b_i^{(\ell)}\}_{0\le i\le n}$ (этим настоящее доказательство отличается от доказательства теоремы 4.2, где набор коэффициентов при степенях y_1 фактически не вычислялся). Сначала как и при доказательстве теоремы 4.2, строим НП, содержащую v сложений и вычисляющую $\{b_0^{(\ell)}\}_{1\le\ell\le v}$. Пусть (см.(4)) $T_{\ell+1}=$

$T_{l+1,1}/T_{l+1,2}+T_{l+1,3}/T_{l+1,4}$, где $T_{l+1,j}=T_1^{\beta_{1,j}^{(l+1)}}\dots T_l^{\beta_{l,j}^{(l+1)}} x_1^{\gamma_{1,j}^{(l+1)}}\dots x_n^{\gamma_{n,j}^{(l+1)}}$, причем $\beta_{p,j}^{(l+1)}$, $\gamma_{p,j}^{(l+1)}$ $(1\le j\le 4)$ – целые неотрицательные. Обозначим $T_{l+1,j}=\sum_{i\ge 0} b_{i,j}^{(l+1)} y_1^i$. Тогда

$$b_{i,j}^{(l+1)}=\Big(\sum_{i_1+\dots+i_l=i} A_{i_1,\dots,i_l,j}\, b_{i_1}^{(1)}\dots b_{i_l}^{(l)}\Big) x_1^{\gamma_{1,j}^{(l+1)}}\dots x_n^{\gamma_{n,j}^{(l+1)}},$$

где $A_{i_1,\dots,i_l,j}$ – некоторое натуральное число. Количество слагаемых в последней сумме не превосходит $\binom{i+l-1}{i}\le\binom{n+v}{v}$. Вычисление $b_i^{(l+1)}$ через $b_{i,j}^{(l+1)}$ и $b_0^{(l+1)}$ $(1\le j\le 4)$ требует не более n^2 сложений по порядку. В конце НП β_1 вернемся к вычислению $g_1,\dots,g_n$, имея вычисленными $\overline{g}_0,\dots,\overline{g}_n$, такие что $\sum_{0\le i\le n}\overline{g}_i(y-\eta)^i=\sum_{1\le i\le n} g_i y^i$, что потребует не более n^2 сложений по порядку.

В результате $C_+(\beta_1)\le n(v+1)\binom{n+v}{n}$. Отсюда по следствию 6.2 получаем $(n+v)^{n+2}\ge n(v+1)\binom{n+v}{n}\ge C_+(g_1,\dots,g_n)\ge\frac{\sqrt{\log N}-2n}{3}$, откуда следует теорема.

Упомянем здесь, что аналог теоремы 4.3 Штрассена неверен для C_+. Например, пусть $f=x_1\cdots x_n+x_1x_2^2\cdots x_n^n$, тогда $C_+(f)=1$. Можно доказать, что $C_+\left(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\right)=n$, так как если $i\ne j$, то $\mathrm{HOД}\left(\frac{\partial f}{\partial x_i},\frac{\partial f}{\partial x_j}\right)=(x_1\cdots x_n)/(x_ix_j)$, и далее для доказательства следует воспользоваться представлением (4).

Для полноты картины отметим, что аналог следствия 6.2 для вычислений над основным комплексным полем $F=\mathbb{C}$ неверен. В [63] для всякого n построен пример многочлена $f\in\mathbb{R}[x]$, имеющего n различных вещественных корней и при этом $C_+^{(\mathbb{C})}(f)\le 3$. В этой же работе [63] отмечено, что аддитивная (над $\mathbb{R}$) сложность многочлена Чебышева степени 3^k с 3^k различными вещественными корнями не превосходит k. Тем самым, оценка в следствии 6.2 точна с точностью до извлечения квадратного корня, а из справедливости вышеупомянутой гипотезы Кушниренко вытекала бы точная оценка по порядку.

В заключение параграфа автор хотел бы обратить внимание на

то, что многие результаты Арнольда и его учеников (см.,например, [2, 21]), касающиеся оценок некоторых других топологических характеристик вещественных алгебраических многообразий (помимо использовавшихся нами оценок для числа нульмерных компонент),например, чисел Бетти, эйлеровой характеристики и т.п., через число K мономов, входящих в многочлены, определяющие данное многообразие, могут по-видимому, дать другие интересные приложения к оценкам аддитивной сложности, которая тесно связана с K, как было видно при доказательстве следствия 6.2. Более глубокие оценки аддитивной сложности имеются, по-видимому, через многогранники Ньютона вычисляемых (вещественных) полиномов. Арнольду принадлежит гипотетический тезис о том, что все "разумные" инварианты многочленов выражаются через их многогранники Ньютона - см., например, [2], где доказано, что число корней общей системы совпадает со смешанным объемом Минковского многогранников Ньютона многочленов из этой системы. Остается выяснить, является ли сложность (например, аддитивная) "разумным" инвариантом?

В заключение главы I отметим, что в главе были изложены методы получения нелинейных нижних оценок сложности для полиномов (и семейств полиномов) относительно больших степеней (сравнительно с числом переменных). Одной из нерешенных и наиболее интересных задач в этой области является получение нелинейных нижних оценок для полиномов маленьких (например, постоянных, т.е. не зависящих от числа переменных) степеней. Решение этой задачи потребует, по-видимому, развития принципиально новой техники.

ГЛАВА II. НИЖНИЕ ОЦЕНКИ МУЛЬТИПЛИКАТИВНОЙ СЛОЖНОСТИ В ЗАДАЧАХ ЛИНЕЙНОЙ АЛГЕБРЫ

§ 7. Мультипликативная сложность и ранг

На протяжении всей главы (за исключением § II) будем иметь дело со следующей ситуацией (в обозначениях § I): $P =$

$= \{+, x\} \cup \{ x y^{\gamma} \}_{\gamma \in K}$, или $P = \{+, x, /\} \cup \{x y^{\gamma}\}_{\gamma \in F}$,

если $K = F$ - поле (в §§ 8, 10 $K = F$); $\lambda = \lambda_m$ - мультипликативная сложность; входные переменные $\{x_1, \dots, x_n\} \cup \{y_1, \dots, y_m\}$ не предполагаются коммутирующими.

Задача, которую будем рассматривать в главе II, состоит в оценке $C_m(A_1, \dots, A_p)$, где $A_\ell = \sum_{i,j} a_{ij\ell} x_i y_j$ $(1 \le \ell \le p)$ -

билинейные формы ($a_{ijl} \in K$). Той же буквой A_l будем обозначать $n \times m$ матрицу коэффициентов $(a_{ijl})_{1 \le i \le n, 1 \le j \le m}$. Плодотворным оказалось следующее понятие (одной из первых работ, в которой оно явно появилось, была [64]) ранга семейства билинейных форм (матриц):

$$Rg_K(A_1,\dots,A_p) = \min\{N : \{A_1,\dots,A_p\} \subset KC_1K + \dots + KC_NK,$$

где $C_i = u_i v_i$ $(1 \le i \le N)$ для некоторых $n \times 1$ столбцов $u_1,\dots,u_N$ и $1 \times m$ строк $v_1,\dots,v_N$ над $K\}$.

Пусть тензор $\tau \in K^n \otimes K^m \otimes K^p$, определим его ранг

$$Rg_K(\tau) = \min\{N : \tau = u_1 \otimes v_1 \otimes w_1 + \dots + u_N \otimes v_N \otimes w_N\}.$$

Для матриц $A_1,\dots,A_p$ положим $n \times m \times p$ тензор $\tau = (a_{ijl})$, тогда нетрудно проверить, что $Rg_K(\tau) = Rg_K(A_1,\dots,A_p)$. Аналогично можно определить $Rg_K(\tau)$ для любого $\tau \in M_1 \otimes_K M_2 \otimes_K \dots \otimes_K M_s$, где M_i – K – модуль $(1 \le i \le s)$, но столь общее определение нам здесь не понадобится. Роль введенного понятия ранга проявляется в следующей теореме, алгебраизирующей мультипликативную сложность в рассматриваемой ситуации.

ТЕОРЕМА 7.1 ([54]) $C_m(A_1,\dots,A_p) = Rg_K(A_1,\dots,A_p)$.

ДОКАЗАТЕЛЬСТВО. Пусть сначала $K = F$ – поле. Устраним деление с помощью метода работы [54] (кратко он описан также в [9]). Сходный прием уже применялся при доказательствах теорем 4.2 и 6.3. Пусть НП β вычисляет $A_1,\dots,A_p$. С помощью замен переменных типа $x_i \to x_i - \eta_i = \bar{x}_i$ добьемся того, чтобы свободные члены значений всех рабочих переменных z_l в НП β были отличны от нуля. Значение всякой переменной z_l можно представить как ряд $\sum_{i \ge 0} b_{i,l}$, где $b_{i,l}$ – форма степени i (от некоммутирующих переменных $\bar{x}_1,\dots,\bar{x}_n,\bar{y}_1,\dots,\bar{y}_m$). Перестроим β в НП $\bar{\beta}$, вычисляющее рекурсией по l формы $b_{0,l}, b_{1,l}, b_{2,l}$ (в общем случае, когда β вычисляет семейство многочленов $\{g_1,\dots,g_p\}$, произвольных степеней, НП $\bar{\beta}$ должно вычислять $b_{i,l}$ для всех $i \le \max\{\deg g_1,\dots,\deg g_p\}$). Нетрудно видеть, что указанная перестройка в рассматриваемом случае не увеличивает мультипликативной сложности, т.е. $C_m(\bar{\beta}) \le C_m(\beta)$.

Рассмотрим теперь (см., например, [64]) НП β без деления

(тем самым, охватывается случай произвольного кольца K), вычисляющую $A_1,\dots,A_p$. Значение каждой рабочей переменной z_ℓ НП β можно представить как $b_0^{(\ell)}+b_x^{(\ell)}+b_y^{(\ell)}+b_{xx}^{(\ell)}+b_{xy}^{(\ell)}+b_{yx}^{(\ell)}+b_{yy}^{(\ell)}+b_3^{(\ell)}$, где, например, $b_{xy}^{(\ell)}$ – сумма мономов (с коэффициентами из K) вида $x_i y_j$ и т.п., $b_3^{(\ell)}$ – сумма мономов степеней не меньше трех. Перестроим НП β в НП β_1, вычисляющую $b_0^{(\ell)}, b_x^{(\ell)}, b_y^{(\ell)}, b_{xy}^{(\ell)}$ рекурсией по ℓ для всех ℓ. Если, например, ℓ-ая строка (см.§ I) НП β имела вид $z_\ell = z_s z_t$ (где $s,t<\ell$), то $b_0^{(\ell)} = b_0^{(s)} b_0^{(t)}$;

$$b_x^{(\ell)} = b_x^{(s)} b_0^{(t)} + b_0^{(s)} b_x^{(t)};\quad b_y^{(\ell)} = b_y^{(s)} b_0^{(t)} + b_0^{(s)} b_y^{(t)};$$

$$b_{xy}^{(\ell)} = b_x^{(s)} b_y^{(t)} + b_{xy}^{(s)} b_0^{(t)} + b_0^{(s)} b_{xy}^{(t)}.$$

Это показывает, что $C_m(\beta_1) \leqslant C_m(\beta)$, причем β_1 содержит лишь нелинейные операции умножения $\times$ вида $(\sum_i \alpha_i x_i) \times (\sum_j \gamma_j y_j)$ – одно такое умножение соответствует произведению столбца на строку в матричных терминах. Отсюда следует неравенство $C_m(A_1,\dots,A_p) \geqslant Rg_K(A_1,\dots,A_p)$. Это завершает доказательство теоремы, так как обратное неравенство очевидно.

Итак, изучение мультипликативной сложности семейства билинейных форм сводится к оценке ранга $Rg_K(A_1,\dots,A_p)$ семейства матриц. Если $K = F$ – поле, то $Rg_F(A)$ – обычный ранг матрицы A, и он не зависит от выбора F, что неверно для $p>1$ (т.е. $Rg_F(A_1,\dots,A_p) \geqslant Rg_H(A_1,\dots,A_p)$ при расширении поля $F \subseteq H$, и неравенство может быть строгим – см.например, [9]). Исследование ранга при $p>1$ оказалось очень трудной задачей (результаты его изучения при $p=2$, полученные автором, сформулированы в § 8). В настоящей главе приведем некоторые оценки в этом направлении, которые удалось получить. В [30] содержится более полный обзор по рангу (на голландском языке).

В конце этого вводного параграфа ограничимся некоторыми общими замечаниями и свойствами ранга. Очевидно, $Rg_K(A_1,\dots,A_p) \leqslant \sum_{1\leqslant i\leqslant p} Rg_K(A_i)$. В отличие от ранга одной матрицы, $Rg_F(A_1,\dots,A_p)$ не является полунепрерывной сверху функцией от $A_1,\dots,A_p$ при $p>1$ (см., например, [9]). Если F алгебраически замкнуто, то нетрудно видеть, что $Rg_F(A_1,\dots,A_p)$ "почти всюду" (в смысле топологии Зарисского) равен некоторому числу $r_q(m,n,p)$, зависящему только от m,n,p (и не меняется при любой их перестановке) и характеристики q поля F. В [9] приведены неко-

торые оценки на $r_q(m,n,p)$, именно, $mnp/(m+n+p-2) \le$ $\le r_q(m,n,p) \le \lceil m/2 \rceil \min\{2n,p\}$ при $n \le p$ (верхняя оценка вытекает из следствия 8.3 ниже). При $m=n=p$ порядок роста $r_q(n,n,n)$: между $n^2/3$ и $n^2/2$ (более точные оценки автору неизвестны). Далее, в [9] показано, что для некоторых $n \times n$ матриц $A_1,\dots,A_n$ с коэффициентами из множества $\{0,1\}$ порядок роста $Rg_F(A_1,\dots,A_n)$ отличается от n^2 не более, чем в мультипликативную константу (в отличие от ситуации с многочленами – см. конец § 3 – с коэффициентами из множества $\{0,1\}$, порядок роста сложности которых меньше максимального по крайней мере в мультипликативный логарифм). В плане общего изучения ранга отметим, что в [9] вычислена также группа линейных преобразований, сохраняющих ранг тензора.

§ 8. Ранг пары билинейных форм

В настоящем параграфе приведем сводку результатов, принадлежащих автору, по оценкам $Rg_F(A,B)$, где F – поле, следуя [9, 32, 33]. Ниже считаем, что все матрицы определены над F.

Определим отношение $C \precsim D$ между матрицами, если $Rg_F(C,D) = Rg\, D$. Введем, далее, относительный ранг $Rg_F(A/B) = \min\limits_{C \precsim B} Rg(A-C)$, где минимум берется по всем $C \precsim B$.

ТЕОРЕМА 8.1 ([9]) $Rg_F(A,B) = RgB + Rg_F(A/B)$.

Пусть теперь до конца параграфа F алгебраически замкнуто. Приведем явную формулу для $Rg_F(A,B)$ в терминах канонической формы пары (A,B) относительно преобразований $(A,B) \to (CAD, CBD)$, где C,D – неособые, которая называется канонической формой Вейерштрасса–Кронекера пучка матриц $\lambda A + \mu B$ (см., например, [5] – оттуда же заимствована употребляемая в следующей теореме терминология).

ТЕОРЕМА 8.2 ([8,9,32,33]). Пусть у пучка $\lambda A + \mu B$ ненулевые минимальные индексы для столбцов равны $a_1,\dots,a_\ell$, для строк – $b_1,\dots,b_k$. Пусть, далее, регулярное $p \times p$ "ядро" $\lambda A_0 + \mu B_0$ пучка $\lambda A + \mu B$ для каждого $\gamma \in F \cup \{\infty\}$ имеет d_γ элементарных делителей вида $(\alpha\lambda + \beta\mu)^\delta$, где $\alpha/\beta = \gamma$. Положим $d = \max d_\gamma$. Тогда

$$Rg_F(A,B) = \sum_i (a_i+1) + \sum_j (b_j+1) + p + d.$$

Доказательство теоремы опирается на теорему 8.I. Независимо, близкий к теореме 8.2 результат получен в [38]. Далее, в формулировках $m \leqslant n$.

СЛЕДСТВИЕ 8.3 ([9]). Для $m \times n$ матриц

I) $Rg_F(A,B)$ почти всюду равен $\min\{n,2m\} = r_q(2,n,m)$

(в обозначениях § 7);

2) $\max\limits_{A,B} Rg_F(A,B) = \min\{m+[n/2], 2m\}$.

Из этого следствия видно, в частности, что $Rg_F(A,B)$ не является полунепрерывной сверху функцией от A, B.

§ 9. Мультипликативная сложность билинейной формы над коммутативным кольцом

Если в случае, когда $K = F$ – основное поле, трудности вызвала оценка ранга $Rg_F(A,B)$ пары матриц (см. § 8), то над произвольным коммутативным кольцом K трудности вызывает уже оценка $Rg_K(A)$ (последняя величина по теореме 7.I совпадает с мультипликативной сложностью билинейной формы A над K). Итак, пусть K – нетерово коммутативное кольцо с единицей, A – $m \times n$ матрица над K. В настоящем параграфе приведем результаты автора об оценках $Rg_K(A)$ (см. [I0, 34]). Обозначим через $rg\,A$ обычный ранг матрицы A, равный наибольшему размеру миноров в ней отличных от нуля. Очевидно, $rg\,A \leqslant Rg_K(A)$. Опишем явно такие кольца K (назовем их Rg – к о л ь ц а – м и), для которых равенство $Rg_K(A) = rg\,A$ выполнено при любой A над K. Используемые ниже понятия из гомологической алгебры можно найти в [I4]. Через $gldh(K)$ обозначается глобальная гомологическая размерность кольца K.

ТЕОРЕМА 9.I ([I0,34]). Кольцо K в том и только том случае является Rg-кольцом, если $K = K_1 \oplus \ldots \oplus K_s$ для некоторых однозначно определенных целостных колец $K_1, \ldots, K_s$, таких что

I) $gldh(K_i) \leqslant 2$ $(1 \leqslant i \leqslant s)$;

2) всякий проективный K_i-модуль свободен $(1 \leqslant i \leqslant s)$.

СЛЕДСТВИЕ 9.2 ([I0]) $F[z_1, z_2]$ – Rg-кольцо.

Как ведет себя $Rg_K(A)$ для колец многочленов $K = K_d = F[z_1, \ldots, z_d]$ при $d \geqslant 3$? Этот вопрос удалось практически полностью решить в случае матриц вида $A = z_1 A_1 + \ldots + z_d A_d$,

где $A_i\ (1 \leqslant i \leqslant d)$ – матрица над полем F (такие матрицы назовем с в о б о д н ы м и о т к в а д р а т о в и сохраним за ними до конца настоящего параграфа за исключением последнего в нем абзаца обозначение A с индексами или без). Обозначим $R_d(r) = \sup\limits_{rgA=r} Rg_{K_d}(A)$ и $R(r) = \sup\limits_{d} R_d(r)$.

ТЕОРЕМА 9.3 ([10,34]). 1) $R(r_1+r_2) \geqslant R(r_1)+R(r_2)$;

2) $R(r) < 2r$;

3) $\lim\limits_{r\to\infty} \frac{R(r)}{r} = 2$;

4) $R_3(r) = [3/2\ r]$.

Пример семейства матриц $\{A_i\}$, на котором последовательность $Rg_{K_d}(A_i)/rg(A_i) \xrightarrow[i\to\infty]{} 2$, строится следующим образом. Рассмотрим комплекс Кошуля (см.[14,15]) кольца $K = K_d$ относительно системы элементов $\{z_1,\dots,z_d\}$:

$$0 \to K^1 \xrightarrow{A_{1,d}} K^d \dots K^{\binom{d}{i}} \xrightarrow{A_{i,d+1-i}} K^{\binom{d}{i+1}} \dots K^d \xrightarrow{A_{d,1}} K^1 \longrightarrow 0.$$

Можно показать ([10]), что

$$Rg_K(A_{i,d+1-i}) = \min\left\{\binom{d}{i},\binom{d}{i+1}\right\};\ rg(A_{i,d+1-i}) = \binom{d-1}{i}.$$

В качестве последовательности $\{A_i\}$ можно взять средние члены комплексов Кошуля, т.е. $A_i = A_{i,i}$.

В заключение параграфа упомянем, что для $Rg_K(A)$ аддитивность не выполняется для всех колец K относительно прямой суммы матриц, определяемой как $A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ (ср. с гипотезой Штрассена об аддитивности ранга, упомянутой в следующем параграфе). Например, пусть $K = \mathbb{Z}[\sqrt{5}]$ и $A = \begin{vmatrix} \frac{\sqrt{5}-1}{2} & 2 \\ & \sqrt{5}+1 \end{vmatrix}$. Тогда $Rg_K(A) = 2$, но $Rg_K(A \oplus \dots \oplus A) \leqslant p+1$, где p – число слагаемых в указанной прямой сумме. В то же время аддитивность $Rg_K(A)$ имеет место над кольцом многочленов и для матриц свободных от квадратов ([10]). Автор позволяет себе высказать две гипотезы: а) аддитивность ранга выполнена над кольцом

многочленов для любых матриц; б) для всякого регулярного кольца K (т.е. $gldh(K)<\infty$) найдется такое число c_K, что $Rg_K(A)\leqslant c_K\,rg\,A$ для произвольной матрицы A над K.

§ 10. Оценки ранга алгебр

К оценкам ранга некоторых алгебр приводят различные сложностные задачи линейной алгебры – например, умножение матриц или полиномов. Ранг алгебры $\mathcal{A}$ над полем F (обозначим его $Rg_F(\mathcal{A})$) определяется как ранг ее структурного тензора в некотором базисе этой алгебры, и не зависит от выбора базиса (см. [54]). Ранг $Rg_F(\mathcal{A})$ интерпретируется как мультипликативная сложность умножения двух элементов алгебры $\mathcal{A}$ – т.е. по разложению сомножителей по базису найти разложение их произведения. Пусть $\mathcal{M}_n$ – алгебра $n\times n$ матриц. Тогда $Rg_F(\mathcal{M}_n)$ равен мультипликативной сложности умножения $n\times n$ матриц, т.е. $C_m\left(\left\{\sum_{1\leqslant i\leqslant n} x_{ki}y_{il}\right\}_{1\leqslant k,l\leqslant n}\right)$. Следующая теорема принадлежит Альдеру и Штрассену.

ТЕОРЕМА 10.1 ([24]) $Rg_F(\mathcal{A})\geqslant 2\dim_F(\mathcal{A})-\kappa$, где κ равен количеству максимальных идеалов в $\mathcal{A}$.

Доказательство разбивается на две следующие леммы.

ЛЕММА 10.2 ([24]). $Rg(\mathcal{A}\oplus\mathcal{B})\geqslant Rg((\mathcal{A}/rad\,\mathcal{A})\oplus\mathcal{B})+2\dim rad\,\mathcal{A}$ ($rad\,\mathcal{A}$ – радикал алгебры $\mathcal{A}$).

ЛЕММА 10.3 ([24]). Если $\mathcal{A}$ – простая алгебра, то

$$Rg(\mathcal{A}\oplus\mathcal{B})\geqslant 2\dim\mathcal{A}-1+Rg(\mathcal{B})$$

(напомним, что алгебра называется простой, если единственный ее идеал – нулевой).

Так как $\mathcal{M}_n$ – простая алгебра, то получаем

СЛЕДСТВИЕ 10.4 ([24]). $Rg_F(\mathcal{M}_n)\geqslant 2n^2-1$.

Отметим, что теорема 10.1 (как и следующие за ней три утверждения) доказаны в [24] на самом деле в более сильной форме – для мультипликативной сложности C'_m умножения в алгебре в предположении коммутирования входных переменных $x_iy_j=y_jx_i$ билинейных форм (ср.начало § 7). Легко видеть (см.,например, [64]), что $C'_m\leqslant Rg=C_m\leqslant 2C'_m$ для семейств билинейных форм.

Упомянем еще два полезных неравенства для ранга тензоров:

$$Rg_K(\tau_1\oplus\tau_2)\leqslant Rg_K(\tau_1)+Rg_K(\tau_2) \qquad (5)$$

$$Rg_K(\tau_1 \otimes_K \tau_2) \leq Rg_K(\tau_1) Rg_K(\tau_2). \qquad (6)$$

Неравенство (6) применяется в следующем виде при получении верхних оценок для $Rg_F(\mathcal{M}_n)$ (см., например, [52,54]): если $Rg_K(\mathcal{M}_{n_0}) \leq N_0$ для некоторых n_0, N_0, то $Rg_K(\mathcal{M}_n) = O(n^{\log_{n_0} N_0})$ (неравенство (6) используется здесь в форме $Rg_K(\mathcal{M}_{n_1 n_2}) \leq Rg_K(\mathcal{M}_{n_1}) Rg_K(\mathcal{M}_{n_2})$, учитывая, что $\mathcal{M}_{n_1 n_2} = \mathcal{M}_{n_1} \otimes \mathcal{M}_{n_2}$). Здесь уместно упомянуть асимптотически лучшую известную на момент написания настоящего обзора оценку для сложности умножения матриц:

ТЕОРЕМА 10.5 ([29]) $Rg_F(\mathcal{M}_n) = O(n^{2,49555\ldots})$.

Доказательство этой оценки опирается на следующее интересное само по себе понятие граничного ранга $\overline{Rg}(\tau)$ (см., например, [29, 52]). Именно, определим

$$\overline{Rg}(\tau) = \min\left\{ N : \varepsilon^K(\tau + \varepsilon\sigma) = u_1 \otimes v_1 \otimes w_1 + \cdots + u_N \otimes v_N \otimes w_N \right\},$$

где K – некоторое натуральное число; u_i, v_i, w_i – какие-то векторы над кольцом $F[\varepsilon]$; σ – тензор над кольцом $F[\varepsilon]$. Неформально говоря, неравенство $\overline{Rg}(\tau) \leq p$ означает, что "сколь угодно близко" (в смысле топологии Зарисского) к тензору τ имеются тензоры ранга, не превосходящего p. Основной инструмент в использовании $\overline{Rg}$ – граничный аналог неравенства (6) ([52]): если $\overline{Rg}(\mathcal{M}_{n_0}) < N_0$ для некоторых n_0, N_0, то $Rg(\mathcal{M}_n) = O(n^{\log_{n_0} N_0})$. В [29,52] построен ряд примеров такого сорта оценок для подходящих n_0, N_0.

Штрассену ([54]) принадлежит гипотеза о том, что $Rg_F(\tau_1 \oplus \tau_2) = Rg_F(\tau_1) + Rg_F(\tau_2)$ (ср. неравенство (5)). Для граничного ранга она неверна (см. [52]).

Обозначим через $A_K = \sum_{0 \leq i \leq K} x_i y_{K-i}$ билинейную форму, выражающую коэффициент (через коэффициенты сомножителей) при K-ой степени в произведении двух полиномов степени n, т.е. $Rg_F(A_0, \ldots, A_{2n})$ – мультипликативная сложность умножения двух полиномов. Результаты, упомянутые в конце § 5, на языке, принятом в настоящем параграфе, можно переформулировать следующим образом:

ТЕОРЕМА 10.6. 1) ([31]) $Rg_F(F[z]/(f)) = 2n-k$, где F – бесконечное поле, $\deg f = n$ и $f = f_1^{m_1} \dots f_k^{m_k}$, причем $f_i \in F[z]$ – неприводим над F $(1 \le i \le k)$ и $\{f_i\}$ попарно взаимно просты.

2) ([8,9,32,33]) $Rg_F(A_0,\dots,A_{2n}) \le ng(n)$ для конечного поля F и некоторой функции g, растущей медленнее любой фиксированной итерации логарифма.

3) ([28]) $Rg_F(A_0,\dots,A_{2n}) \ge (3{,}52)n$ для поля F из двух элементов. Очевидно, $Rg_F(F[z]/(f)) \le Rg_F(A_0,\dots,A_{2n}) \le Rg_F(F[z]/(h))$, где $\deg h = 2n$. Из [22] вытекала оценка $Rg_F(A_0,\dots,A_{2n}) = O(n\log n)$ для любого поля F. Верхняя оценка в пункте 1) вытекает из неравенства (5) и изоморфизма $F[z]/(f) = \bigoplus_{1 \le i \le k} F[z]/(f_i^{m_i})$; нижняя оценка следует из теоремы 10.1.

Приведем теперь набросок доказательства пункта 3). Зафиксируем некоторое $0 \le j \le n$. Очевидно, $Rg_F(A_0,\dots,A_{2n}) \ge Rg_F(A_{j-1},\dots,A_{2n-j+1}) = p_j$. По теореме 7.1 $p = p_j$ равно наименьшему количеству $C_1,\dots,C_p$ билинейных форм ранга 1, линейная над F оболочка которых содержит линейную оболочку $\mathcal{L} = \mathcal{L}(A_{j-1},\dots,A_{2n-j+1})$ рассматриваемых билинейных форм. С другой стороны, для любой билинейной формы $0 \ne A \in \mathcal{L}$ выполнено неравенство $rg\,A \ge j$, поэтому $\mathcal{L}$ можно рассматривать как линейный код в линейной оболочке $\mathcal{L}(C_1,\dots,C_p)$ (с базисом $\{C_1,\dots,C_p\}$), причем кодовое расстояние (по поводу употребляемых понятий из теории кодирования см., например, [19]) этого кода не меньше j. Поэтому для оценки p снизу можно применить границы Варшамова – Гилберта (см. [19], глава 4), что приводит к неравенству $p \ge (3{,}52)n$ при подходящем выборе j.

Усиление оценок из пунктов 2), 3) теоремы – по-видимому, тонкая теоретико-числовая задача.

§ 11. Линеаризованная мультипликативная сложность

В предыдущих параграфах главы II рассматривался ранг элементов тензорного произведения векторных пространств (точнее, произведения трех пространств, но можно в принципе и большего числа – см. замечание в § 7). В настоящем параграфе мы рассмотрим ана-

лог ранга для элементов симметрического произведения (более точно, симметрической d_1-ой степени векторного пространства размерности n над полем F – иными словами, пространства однородных форм степени d_1 в кольце многочленов $F[x_1,\dots,x_n]$). Именно, для всякой формы $f\in F[x_1,\dots,x_n]$ степени d_1 ее линеаризованной мультипликативной сложностью $C_\ell(f)$ назовем

$$\min\Big\{N: f=\sum_{1\le i\le N} \ell_1^{(i)}\dots\ell_{d_1}^{(i)}\ ,\ \text{где}\ \ \ell_j^{(i)}-\text{линейная форма}\Big\}.$$

Изложим метод получения нелинейных нижних оценок на $C_\ell(f)$ – метод дает нелинейные оценки для $d_1\ge 4$. Считаем ниже, что степень $d_1=2d$ четна – это предположение сделано для удобства обозначений (без большой потери общности вследствие характера предлагаемых в этом параграфе оценок).

Будем рассматривать вспомогательные матрицы A размера $n^d\times n^d$. Их строки и столбцы занумеруем всевозможными векторами $I=(i_1,\dots,i_d)$, где $1\le i_1,\dots,i_d\le n$, в этих обозначениях $A=(a_{I,J})$. Через x^I обозначим моном $x_{i_1}\dots x_{i_d}$ степени d, через $|I|$ обозначим число $\kappa_1!\dots\kappa_n!$, где κ_i – количество вхождений числа i в вектор I.

Определим на пространстве $n^d\times n^d$ матриц линейный оператор L (назовем его оператором коммутации), положив $L(A)=B$, где $b_{P,Q}=\Big(\sum_{x^Ix^J=x^Px^Q=x^K} a_{I,J}\Big)|K|$, т.е. $x^Px^Q=x^K$ и суммирование происходит по всем парам векторов I,J, таких что $x^Ix^J=x^K$. Всякой матрице A можно поставить в соответствие форму $g_A=\sum_{I,J} a_{I,J}x^Ix^J$ степени $2d$, где суммирование происходит по всем парам векторов I,J. Обратно, всякой форме g можно поставить в соответствие матрицу A, такую что $g_A=g$, но уже неоднозначно, именно

ЛЕММА II.1. Равенство форм $g_A=g_B$ равносильно тому, что $L(A)=L(B)$.

Всякому (упорядоченному) семейству линейных форм $\{\ell_1,\dots,\ell_{2d}\}$ поставим в соответствие $n^d\times n^d$ матрицу $A=A_{\ell_1,\dots,\ell_{2d}}$ ранга I, положив ее элемент $a_{I,J}=\ell_{1,i_1}\dots\ell_{d,i_d}\ell_{d+1,j_1}\dots\ell_{2d,j_d}$, где форма $\ell_i=\sum_{1\le j\le n}\ell_{i,j}x_j$ и $I=(i_1,\dots,i_d)$, $J=(j_1,\dots,j_d)$.

Очевидно, форма $g_{A_{\ell_1,\dots,\ell_{2d}}} = \ell_1 \dots \ell_{2d}$ – произведение линейных форм. Выясним действие оператора коммутации на построенную матрицу $A_{\ell_1,\dots,\ell_{2d}}$.

ЛЕММА II.2. $L(A_{\ell_1,\dots,\ell_{2d}}) = \sum_{\pi \in S_{2d}} A_{\ell_{\pi(1)},\dots,\ell_{\pi(2d)}}$,

где суммирование происходит по всем перестановкам π.

Следовательно, $rg\,(L(A_{\ell_1,\dots,\ell_{2d}})) \leq (2d)!$

Пусть теперь $C_\ell(f) \leq N$ и $f = \sum_{1 \leq i \leq N} \ell_1^{(i)} \dots \ell_{2d}^{(i)}$. Обозначим через A_i матрицу $A_{\ell_1^{(i)},\dots,\ell_{2d}^{(i)}}$. Тогда $f = g_{A_1} + \dots + g_{A_N} = g_{A_1 + \dots + A_N}$. По лемме II.1 матрица $L(A_1 + \dots + A_N) = L(f)$ – инвариант формы f. Наконец, по лемме II.2

$$rg(L(A_1 + \dots + A_N)) \leq rg\,L(A_1) + \dots + rg\,L(A_N) \leq N(2d)!,$$

и в итоге получаем

ТЕОРЕМА II.3. Для всякой формы f степени $2d$

$$C_\ell(f) \geq \frac{rg\,L(f)}{(2d)!}.$$

Приведем пример приложения теоремы, демонстрирующий наличие большого разрыва между линеаризованной мультипликативной сложностью и тотальной сложностью C_t (см.§ I). Пусть $f = (x_1 \overline{x_1} + \dots + x_n \overline{x_n})^d$ – форма степени $2d$ от $2n$ переменных $x_1, \dots, x_n, \overline{x_1}, \dots, \overline{x_n}$. Для всяких векторов I, J обозначим $x^I = x_{i_1} \dots x_{i_d}$, $\overline{x}^J = \overline{x_{j_1}} \dots \overline{x_{j_d}}$. Выделим в матрице $L(f)$ подматрицу B размерами $n^d \times n^d$, натянутую на строки вида I и столбцы вида J (по всем I, J указанного выше вида). Выделим теперь в матрице B подматрицу C, натянутую на некоторый любой такой максимальный набор строк $\{I\}$, для которых все мономы x^I попарно различны; столбцы выбираем с соответствующими индексами $\{\overline{I}\}$, что и у строк. Тогда квадратная матрица C имеет сторону $\binom{n+d-1}{n-1}$, она диагональна с ненулевыми элементами на диагонали, поэтому $rg\,C = \binom{n+d-1}{n-1}$. Согласно теореме II.3 $C_\ell(f) \geq \binom{n+d-1}{n-1} / (2d)!$

С другой стороны, $C_t(f) \preccurlyeq 2n + \log d$. Ясно, что при достаточно малом d по сравнению с n, именно, здесь можно взять $n \succ d^6$, величина $C_\ell(f) \succ (n/d^5)\,d$ не оценивается сверху никаким, например, полиномом от $C_t(f)$, т.е. от n.

Отметим здесь для полноты картины, что в [37] показано, что мультипликативная сложность $C_m(f)$ формы f (здесь можно рассматривать оба случая – с коммутирующими и некоммутирующими входными переменными) степени d_1 оценивается снизу через следующую величину (определение которой близко по духу к определениям ранга и $C_\ell(f)$, рассматриваемым в настоящей главе):

$\min\left\{N : f = \sum_{1 \le i \le N} g_i h_i\right.$, где g_i, h_i – формы, причем $\left.\frac{d_1}{3} \le \deg g_i, \deg h_i \le \frac{2d_1}{3}\right\}$. В [37] предложен прием для оценивания снизу этой величины (и тем самым мультипликативной сложности) в случае, когда входные переменные коммутирующими не предполагаются. Как получать нижние оценки для этой величины в случае коммутирующих переменных, пока неясно.

В заключении главы II скажем несколько слов о том, что ранг семейства билинейных форм таит в себе пока что много загадок. Одной из наиболее интересных нерешенных проблем является получение нелинейных (от числа переменных и числа форм) нижних оценок для ранга каких-либо естественных семейств форм (сюда можно отнести фактически все сказанное в начале § 3 о сложности вычисления многочленов), например, для вызывающих большой интерес задач умножения полиномов (над конечным полем) или умножения матриц (см. § 10). Автор питает надежду, что для ранга прорыв в получении нелинейных нижних оценок произойдет скорее, нежели в других направлениях алгебраической сложности, для чего по-требуется, по-видимому, дальнейшая алгебраизация понятия ранга семейства билинейных форм.

ГЛАВА III. СЛОЖНОСТЬ В НЕВЕТВЯЩИХСЯ ПРОГРАММАХ НЕСТАНДАРТНЫХ ТИПОВ

В последней главе параграфы менее связаны между собой, чем в предыдущих двух главах. Объединяет их, пожалуй, то, что в них рассматриваются НП и сложностные меры нестандартных типов, удов-

летворяющие (за исключением § I2) каким-то ограничениям (в разных параграфах различным) в сравнении с НП достаточно общего вида, изучавшимся в главах I–II. Принятие именно этих ограничений имеет иногда практические основания, а также объясняется возможностью получить нижние оценки, которые не удается пока получить для НП общего вида. Часто для нестандартных (ограниченных) вычислительных моделей удается выявить связь сложности с классическими для математики объектами (именно на этой основе автор выбирал материал для настоящей главы). Кроме того, ограничения позволяют иногда проследить влияние отдельных факторов на тотальную сложность (см. § I), что полезно для проникновения в тайну нижних оценок сложности.

§ I2. Иррациональная сложность вычисления алгебраических функций

В настоящем параграфе рассматриваем НП следующего типа (в терминологии § I): F – основное поле характеристики отличной от двух; $P=\{+,\times,/,\sqrt{}\}\cup\{x\gamma\}_{\gamma\in F}$; $\lambda=\lambda_{\sqrt{}}$, $\lambda(+)=\lambda(\times)=\lambda(/)=\lambda(x\gamma)=0$, $\lambda(\sqrt{})=1$;

множество входных переменных пусто (несколько модифицируя определение из § I, считаем, что в качестве аргументов базисной операции в пункте 4) определения могут стоять константы из поля F). Здесь знак $\sqrt{}$ означает одноместную операцию извлечения квадратного корня; рациональные операции допускаются бесплатно, поэтому м е р ы с л о ж н о с т и указанного вида естественно назвать и р р а ц и о н а л ь н ы м и.

Если $f\in F$, то элемент $\sqrt{f}$ расширения поля F называется простым радикалом. Пусть $\{g_1,\dots,g_k\}\subset F(\sqrt{f_1},\dots,\sqrt{f_\ell})$, где $f_1,\dots,f_\ell\in F$; задача, рассматриваемая в настоящем параграфе, состоит в оценке иррациональной сложности $C_{\sqrt{}}(g_1,\dots,g_k)$. Расширение полей $F\subseteq F(g_1,\dots,g_k)$ – абелево расширение Галуа степени 2^c (для некоторого c) с группой Галуа $Z_2\times\dots\times Z_2$ (прямое произведение c экземпляров циклической группы Z_2 порядка два). Нетрудно видеть, что $c=C_{\sqrt{}}(g_1,\dots,g_k)$. Неравенство $c\leqslant C_{\sqrt{}}(g_1,\dots,g_k)$ следует из того, что присоединение одного нового простого радикала увеличивает степень расширения не более, чем в два раза; с другой стороны, каждый этаж (степени два) башни полей $F=F_0\subset F_1\subset\dots\subset F_c=F(g_1,\dots,g_k)$ получа-

ется из предыдущего присоединением одного нового радикала, откуда вытекает обратное неравенство.

Приведем один пример, пришедший из древности. Пусть $F = \mathbb{R}(y_1, y_2)$, и тогда $C_{\sqrt{\ }}(g_1, \ldots, g_k)$ равна в точности необходимому числу применений циркуля для построения $g_1, \ldots, g_k$ с помощью циркуля и линейки.

Трудность состоит в том, что зачастую в интересных примерах неясно как априори оценить снизу степень расширения полей. Поэтому в [53] был предложен метод оценки $C_{\sqrt{\ }}$ для случая, когда $F = F_0(x_1, \ldots, x_n)$, где $F_0 \subseteq \mathbb{C}$, и было показано, что $C_{\sqrt{\ }}(g) \geqslant \log N$, где N – число листов римановой поверхности аналитической функции g. Полное строгое доказательство этого результата требует громоздкой техники (не проведенной в [53] до конца), и мы обратимся здесь к более элементарному и удобному для приложений методу, предложенному в [45].

Пусть $r = \{r_1, \ldots, r_k\}$ – множество простых радикалов, обозначим $r_I = \prod_{i \in I} r_i$ для $\emptyset \neq I \subseteq \{1, \ldots, k\}$ $(r_\emptyset = 1)$. Скажем, что множество r радикально независимо, если $r_I \notin F$ для всякого $I \neq \emptyset$. Через $R(r)$ обозначим радикальную размерность множества r, т.е. наибольшее число радикально независимых элементов среди $r_1, \ldots, r_k$. Нетрудно показать ([45]), что в любом максимальном по включению радикально независимом подмножестве множества содержится $R(r)$ элементов, т.е. радикальная размерность обладает теми же общими (матроидными) свойствами, что и обычная размерность векторного пространства, степень трансцендентности расширения полей и т.п. В связи с этим $R(r)$ удобна для вычисления.

Если r – радикально независимое множество, то элементы $\{r_I\}_{I \subseteq \{1, \ldots, k\}}$ составляют базис расширения $F \subset F(r)$, степень которого тем самым равна 2^k ([45]). Пусть $g = \{g_1, \ldots, g_k\} \subset F(r)$, тогда носителем $S(g, r)$ семейства g назовем множество таких r_I, что для некоторого $1 \leqslant i \leqslant k$ в разложении (единственном) $g_i = \sum_{J \subseteq \{1, \ldots, k\}} c_{i,J} r_J$, где $c_{i,J} \in F$, коэффициент $c_{i,I}$ при r_I отличен от нуля.

ЛЕММА 12.1 ([45]). Если r – радикально независимое множество и $g \subset F(r)$, то $F(g) = F(S(g, r))$.

ЛЕММА 12.2 ([45]). Если s – некоторое множество простых радикалов, то $C_{\sqrt{\ }}(s) = R(s) = \log [F(s) : F]$.

В качестве следствия получаем

ТЕОРЕМА I2.3 ([45]). Пусть τ – радикально независимое множество и $g \subset F(\tau)$, тогда $C_{\sqrt{\ }}(g) = R(S(g,\tau))$.

Так как в приложениях часто $g \subset F(\mathfrak{z})$, где $\mathfrak{z}$ – некоторое данное множество простых радикалов, то теорема дает следующий путь для вычисления $C_{\sqrt{\ }}(g)$, который иногда оказывается эффективным. Из $\mathfrak{z}$ нужно выделить произвольное максимальное по включению радикально независимое множество $\tau \subset \mathfrak{z}$; далее, разложив g_i по базису из $\{\tau_I\}$, найти $S(g,\tau)$, которое суть некоторое множество простых радикалов, и наконец, выделить в нем максимальное по включению радикально независимое подмножество, его мощность и равна $C_{\sqrt{\ }}(g)$.

Приведем одно конкретное приложение теоремы ([45]). Пусть $F = F_0(\{y_i, z_i\}_{1 \le i \le N})$, тогда

$$C_{\sqrt{\ }}\left(g = \sum_{1 \le i < j \le N} \sqrt{(y_i - y_j)^2 + (z_i - z_j)^2}\right) = \binom{N}{2}.$$

В данном случае множество простых радикалов $\mathfrak{z} = \tau = \left\{\sqrt{(y_i - y_j)^2 + (z_i - z_j)^2}\right\}_{1 \le i < j \le N}$ радикально независимо – это следует из того, что функции $\{(y_i - y_j)^2 + (z_i - z_j)^2\}_{1 \le i < j \le N}$ попарно взаимно простые. В результате $\tau = S(g,\tau)$ и $R(\tau) = \binom{N}{2}$.

§ I3. Монотонные вычисления

Монотонные вычисления – довольно узкий, но достаточно естественный класс НП, для которых удалось получить нижние оценки сложности и даже в некоторых случаях явно ее вычислить. Это проделывалось для разных задач в большой серии работ. Здесь мы приведем теорему Шнорра [49], которая обобщает соображения, содержащиеся во многих из этих работ.

Итак, м о н о т о н н о й НП называется НП (используем обозначения из § I), у которой роль основного кольца K играет некоторое полукольцо $S \subset F \setminus \{0\}$ для некоторого поля F (полукольцо образует моноид по сложению и моноид по умножению); $P = \{+, \times\} \cup \{+\gamma, \times\gamma\}_{\gamma \in S}$; $\lambda = \lambda_+$ (см. § 2); т.е. $\lambda(+) = \lambda(+\gamma) = 1$, $\lambda(\times) = \lambda(\times\gamma) = 0$; $\{x_1, \dots, x_n\}$ – некоторое множество входных переменных; соответствующую меру сложности (монотонную) обозначим через C_{mon}. Например, если $F = \mathbb{R}$, то в качестве S можно взять все положительные числа.

Для всякого полинома g через $Mon(g)$ обозначим множество мономов, входящих в g с ненулевыми коэффициентами. Подмножество мономов $B \subset Mon(g)$ назовем о т д е л я ю щ и м, если для всяких $s, t \in B$ и $r \in Mon(g)$, если выполнено $r \mid st$ (вертикальная черта обозначает отношение делимости), то либо $r = s$, либо $r = t$. Шнорр показал, что $C_{mon}(g) \geq |B| - 1$ (здесь и далее, полагаем $C_{mon}(g) = \infty$, если она не определена). Более того, для целей работы [49] доказано следующее усиление этого результата. Пусть σ - некоторое отображение из множества переменных в множество мономов, тогда для всякого многочлена g через g^{σ} обозначим результат замены в g всякой переменной x_j на моном $\sigma(x_j)$.

ТЕОРЕМА I3.I ([49]). Для всякого полинома g $C_{mon}(g) \geq |B| - 1$ для любого отделяющего множества B многочлена g^{σ}.

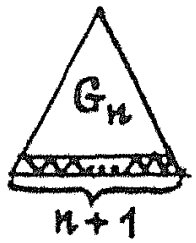
В качестве приложений в [49] показано, что для вычисления произведения полиномов степени n выполнено

$$C_{mon}\left(\left\{\sum_{k+l=s} x_k y_l\right\}_{0 \leq s \leq 2n}\right) = (n+1)^2,$$ для умножения $n \times n$ матриц $$C_{mon}\left(\left\{\sum_{1 \leq i \leq n} x_{ki} y_{il}\right\}_{1 \leq k, l \leq n}\right) = n^3 - n^2$$ (ср. с § I0). Рассмотрим, далее, многочлен

$$CL_{n,k} = \sum_{1 \leq v_1 < \ldots < v_k \leq n} \prod_{1 \leq i < j \leq k} x_{v_i, v_j}$$

степени $\binom{k}{2}$ от $\binom{n}{2}$ переменных $\{x_{i,j}\}_{1 \leq i < j \leq n}$. Из теоремы I3.I следует, что $C_{mon}(CL_{n,k}) = \binom{n}{k} - 1$. С другой стороны, вопрос о наличии полиномиальной верхней оценки для тотальной сложности $C_t(CL_{n,k})$ (см. § I) тесно связан с $P = NP$ проблемой, так как многочлен $CL_{n,k}$ соответствует NP-универсальной задаче о существовании k-клики в n-вершинном графе (см.например,[I], гл.I0).

Несмотря на то, что было получено много различных нижних оценок для C_{mon}, некоторое время оставался открытым вопрос, насколько большим может быть разрыв между $C_t(g)$ и $C_{mon}(g)$ для многочлена g ? В [62] показано, что разрыв этот может быть экспоненциальным. Именно, построим семейство плоских графов $\{G_n\}$ индукцией по n :

$G_1 = \Delta, \ldots, G_{n+1} =$  $, \ldots,$

т.е. G_n образует правильный треугольник со стороной n, разбитый паркетом из правильных треугольников со стороной I.

Совершенным паросочетанием графа G_n называется всякое множество его ребер, не имеющих попарно общих вершин и покрывающих в совокупности все вершины (в частности, число вершин $(n+1)(n+2)/2$ должно быть четным). Припишем каждому ребру графа G_n свою переменную, и для каждого множества I ребер через h_I обозначим моном, равный произведению переменных, приписанных ребрам из I. Определим многочлен g_n как сумму $\sum_I h_I$ по всем совершенным паросочетаниям I графа G_n.

ТЕОРЕМА I3.2 ([62]). Для некоторой константы $c > 1$ выполнено $C_{mon}(g_n) > c^n$ для всех n.

С другой стороны, согласно одному результату Кастелейна (см., например, [62]) можно построить эффективно (конструкция годится для всякого плоского графа) по всякому n такую кососимметрическую матрицу A_n, что $g_n = Pfaffian(A_n) = \sqrt{\det A_n}$, что доказывает, что $C_{\mathbb{C}}(g_n)$ оценивается сверху полиномом от n (см. [I], гл.6). В сочетании с теоремой Вальянта I3.2 это дает ответ на сформулированный выше вопрос о разрыве между тотальной и монотонной сложностью.

Для полноты картины отметим, что в булевском случае, т.е. когда поле F состоит из двух элементов, монотонные вычисления изучались очень интенсивно (см., например, [7] и имеющуюся там библиографию).

§ I4. Нижние оценки для произведения времени и памяти

В настоящем параграфе мы несколько отклонимся от принятого в § I понятия НП, для того чтобы ввести понятие НП с памятью S (см. [6], § 2). Пусть F – основное поле, $P = \{+, \times, /\} \cup \{x^\gamma\}_{\gamma \in F}$; $\{x_i\}_{1 \le i \le n}$ – набор входных переменных. Каждая команда НП β с памятью S (где S – некоторое натуральное число) имеет вид

$$z_{s_0} = f(z_{i_1}, \ldots, z_{i_u}, x_{j_1}, \ldots, x_{j_v}),$$

где $f \in P$, $1 \leqslant s_0, i_1, \ldots, i_u \leqslant S$ (важно отметить, что индексы $i_1, \ldots, i_u$ могут быть и больше s_0 в отличие от НШ, определенных в § I).

Пусть НШ β с памятью S (далее в настоящем параграфе эти последние слова будем иногда опускать) состоит из T команд, и пусть выписанная команда является t_0-ой по порядку в программе для некоторого $1 \leqslant t_0 \leqslant T$. Индукцией по t естественно определяется для каждого $1 \leqslant t \leqslant T$ и $1 \leqslant s \leqslant S$ рациональная функция $Z_s^{(t)} \in F(x_1, \ldots, x_n)$, называемая значением рабочей переменной z_s в момент времени t. База индукции $Z_s^{(0)} = 0$ для любого $1 \leqslant s \leqslant S$. Если $Z_s^{(t)}$ уже определено при всех $t < t_0$, то в рассматриваемом случае

$$Z_{s_0}^{(t_0)} = f(Z_{i_1}^{(t_0-1)}, \ldots, Z_{i_u}^{(t_0-1)}, x_{j_1}, \ldots, x_{j_v}) \quad \text{и}$$

$$Z_s^{(t_0)} = Z_s^{(t_0-1)}$$ для всех $s \neq s_0$. Скажем, что функции $g_1, \ldots, g_k \in F(x_1, \ldots, x_n)$ вычисляются данной β, если для всякого $1 \leqslant r \leqslant k$ найдутся такие $1 \leqslant s \leqslant S$ и $1 \leqslant t \leqslant T$, что $Z_s^{(t)} = g_r$. Число S естественно интерпретируется как память, T – как время.

Пусть F – поле из двух элементов, $P = \{+, \times\} \cup \{+1\}$. В этом случае в [6], § 2 был предложен метод для установления нижней оценки на произведение ST для НШ, вычисляющей семейство полиномов $g_1, \ldots, g_k \in F[x_1, \ldots, x_n]$, удовлетворяющее некоторому условию l-независимости ($1 \leqslant l \leqslant k$). Сформулируем следующие два его варианта I) и 2) (первый на самом деле является следствием второго). Для $g \in F[x_1, \ldots, x_n]$ через R_g обозначим разбиение n-мерного куба F^n на два множества:

$$R_g^{(0)} = \{X \in F^n : g(X) = 0\}, \quad R_g^{(1)} = \{X \in F^n : g(X) = 1\}.$$

Скажем, что семейство $\{g_1, \ldots, g_k\}$ обладает свойством l-независимости, если для всякого $1 \leqslant v \leqslant l$ и для всяких $1 \leqslant j_1 < \ldots < j_v \leqslant n$ и $1 \leqslant i_1 < \ldots < i_{l-v} \leqslant k$ выполнено одно из следующих двух условий (см. замечание выше):

I) найдется такой набор значений переменных $x_{j_1} = x_{j_1}^{(0)}, \ldots, x_{j_v} = x_{j_v}^{(0)}$, что при этих фиксированных значениях вектор функций $(g_{i_1}, \ldots, g_{i_{l-v}})$ принимает больше 2^{l-v-1} значений (при изменении остальных переменных);

2) $H(R_{g_{i_1}} \vee \ldots \vee R_{g_{i_{l-v}}} / R_{x_{j_1}} \vee \ldots \vee R_{x_{j_v}}) > l - v - 1$,

где H – условная энтропия в равномерной мере на кубе F^n, знак $\vee$ – измельчение разбиений (см. [3]).

ТЕОРЕМА I4.I ([6], § 2). Пусть семейство $\{g_1, \ldots, g_k\}$ удовлетворяет условию l-независимости. Тогда для всякой НП с памятью S, вычисляющей это семейство, выполнено $ST \geq kl/8$.

Во многих результатах настоящего параграфа встречается нижняя оценка на произведение ST, которую можно рассматривать как некоторый аналог принципа неопределенности (английский термин – time-space tradeoff) для памяти (пространства) и времени.

В качестве приложения теоремы получаем (см. [6]), что для задачи умножения полиномов степени n (ср. теорему I0.6) $ST \geq n^2/16$, а для задачи умножения матриц (ср. следствие I0.4 и теорему I0.5) $ST \geq n^3/8$.

В ряде последующих работ метод доказательства теоремы I4.I был обобщен на произвольные поля F. Для формулировки обобщения ограничимся случаем НП с памятью S, вычисляющих семейство линейных форм, т.е. $P = \{+\} \cup \{x_\gamma\}_{\gamma \in F}$ и функция $\lambda = \lambda_t$, соответствующую меру сложности (время или число команд) обозначим через T_{ad} (аббревиатура от слова аддитивный). Задача состоит в оценке $T_{ad}(A_1, \ldots, A_n)$, где $A_j = \sum_{1 \leq i \leq n} \alpha_{ij} x_i$ – линейная форма $(1 \leq j \leq n)$. Обозначим через $A = (\alpha_{ij})$ $n \times n$ матрицу коэффициентов этих форм.

ТЕОРЕМА I4.2 ([58]). Пусть все миноры (см. [5]) матрицы A отличны от нуля. Тогда для всякой НП с памятью S, вычисляющей семейство $\{A_1, \ldots, A_n\}$, выполнено $ST_{ad} \geq n^2$.

Идея доказательства восходит к работам [6,47,48,59] и опирается на понятие суперконцентратора (см., например, [44,59]), являющегося некоторым усилением понятия концентратора, введенного независимо в работах [I6,43]. Суперконцентраторы сыграли значительную роль в установлении нижних оценок сложности, поэтому рассмотрим их здесь несколько более подробно.

Пусть G – ориентированный граф без ориентированных циклов (графы этого типа в точности те, которые задают частичное упорядочение на вершинах, и мы будем называть их у п о р я д о ч е н н ы м и г р а ф а м и). В х о д н ы м и в е р ш и н а м и $w_1, \ldots, w_n$ графа G назовем вершины, в которые не входят дуги; в ы х о д н ы м и в е р ш и н а м и $v_1, \ldots, v_m$

графа G назовем вершины, из которых не выходят дуги. Пусть $m=n$. Скажем, что G является n-суперконцентратором, если для всякого $1\leqslant k\leqslant n$ и для всяких двух k-элементных множеств $I,J\subset\{1,\dots,n\}$, в графе G можно провести k ориентированных путей, попарно непересекающихся по вершинам, причем начала путей лежат в I, концы – в J. Мотивировка этого определения будет ясна из дальнейшего.

Здесь и в следующем параграфе полезной будет распространенная конструкция, ставящая в соответствие всякой НП- β (обычного типа, как определено в § I) некоторый упорядоченный граф G_β (см.например, [6,58,59]). Граф G_β имеет n входных вершин – по одной для каждой входной переменной $x_1,\dots,x_n$. Далее, каждой команде $z_\ell=f(z_{i_1},\dots,z_{i_u},x_{j_1},\dots,x_{j_v})$ (см. § I) соответствует вершина, в которую идут дуги из вершин, соответствующих рабочим переменным $z_{i_1},\dots,z_{i_u}$ и входным переменным $x_{j_1},\dots,x_{j_v}$ (таким образом, T – число невходных вершин графа G_β). Проделав тривиальные преобразования, можно считать, что выходные вершины графа G_β отвечают в точности функциям, вычисляемым с помощью НП β. Ниже в настоящем параграфе под НП будем понимать также и НП с памятью, которую можно рассматривать как частный случай обычного НП (забыв про ограничение на память).

ЛЕММА I4.3 ([59]). Если НП β вычисляет семейство линейных форм с $n\times n$ матрицей A коэффициентов, все миноры которой отличны от нуля, то G_β является n-суперконцентратором.

Отметим,что как показано в работах [I6,43,44,59] существует семейство n-суперконцентраторов с линейным по n числом ребер (в [44] имеется верхняя оценка $29n$ на число ребер),поэтому сама по себе лемма I4.3 не приводит непосредственно к нелинейным нижним оценкам на сложность.

ЛЕММА I4.4 ([58]). Если β является НП с памятью S и G_β n-суперконцентратор, то $ST\geqslant n^2$.

Теорема I4.2 теперь легко следует из этих двух лемм.

В качестве приложений теоремы (помимо нее, для доказательства нижеперечисленных результатов привлекались еще другие дополнительные соображения) приведем следующие примеры:

СЛЕДСТВИЕ I4.5. I) ([58]) для задачи умножения полиномов степени n (над любым полем) $ST\geqslant n^2$;

2) ([47,58]) для вычисления дискретного преобразования Фурье, т.е. семейства линейных форм с $n\times n$ матрицей коэффициентов $(\exp(2\pi ijk))_{1\leqslant j,k\leqslant n}$ выполнено $ST\geqslant n^2$ (для устранения возможных недоразумений отметим, что здесь $i^2=-1$);

3) ([58,39]) для умножения $n \times n$ матриц $ST \geq n^3$ и для обращения $n \times n$ матриц $ST \geq n^4$;

4) ([48]) для умножения целых чисел, не превосходящих 2^n (т.е. нахождение двоичных разрядов произведения чисел по n разрядам сомножителей), выполнено $ST \geq n^2$.

§ I5. Методы теории графов в алгебраической сложности

Применения теории графов в алгебраической сложности основаны на конструкции упорядоченного графа G_β , описанной в предыдущем параграфе. Одно из таких применений (к установлению "принципа неопределенности") было рассмотрено выше (см. § I4). В настоящем параграфе мы изложим менее абсолютные применения теории графов, чем в § I4, но, по мнению автора, интерес представляют сами методы и постановки задач.

Мы не будем сначала (как делали обычно выше) фиксировать класс НП β (т.е. параметры из определения в § I), так как основным рассматриваемым объектом будет граф G_β . Будем лишь предполагать, что мера сложности $C = C_\lambda$ определяется через функцию λ равную единице (ср. тотальную сложность - см.§ I) для каждой базисной операции из $\mathfrak{P}$ (т.е. $C(\beta)$ равна количеству невходных вершин графа G_β). Количество входных переменных НП обозначим через n . Отметим здесь одно важное отличие конструкции графа $G = G_\beta$ от приведенной в § I4. Именно (это оправдано сформулированным ниже ограничением (7) на рассматриваемый класс НП), следует считать, что в графе G входных вершин может быть больше n , и одна входная переменная может соответствовать нескольким входным вершинам (это предположение лишь расширяет класс допустимых НП).

В первом приложении, о котором пойдет речь, граф $G = G_\beta$ будет удовлетворять следующему ограничению (см.[6], § I). Через T_v обозначим подграф графа G , содержащий вершины, расположенные в нем над вершиной v (и содержащий все дуги из G , оба конца которых лежат в этих вершинах), т.е. вершины, из которых имеется ориентированный путь в v (на языке упорядочений T_v содержит вершины меньшие или равные v). Ограничение состоит в том, что

T_v является деревом для всякой вершины v (7)

Скажем, что семейство функций $\{g_1, \ldots, g_m\}$ (в соответствие с тем, что говорилось выше, природа функций не существенна)

является (θ,η) – о т д е л и м ы м , если для всякой НП β, удовлетворяющей ограничению (7), и вычисляющей семейство $\{g_1,\dots,g_m\}$, при вычеркивании из графа G_β любых θ вершин, найдутся η различных пар вершин, таких что, один член каждой пары – некоторая входная вершина графа G_β, другой – выходная, и между вершинами – членами всякой пары – можно провести в G_β путь (такой путь согласно (7) единственен), не проходящий через вычеркнутые вершины (неформально говоря, даже если "бесплатно" пользоваться произвольными θ функциями, то для многих индексов $i\,(1\leqslant i\leqslant m)$ для вычисления выхода g_i требуется обратиться ко многим входам). Сформулированное понятие, по-существу, содержалось уже в § I [6] и очень близко к независимо предложенному в [60] понятию $(n,\eta(\theta))$–*grate*.

Условие отделимости (так же как и понятие *grate*) описано на постороннем по отношению к вычисляемым функциям языке достаточно обширного класса НП, что, несомненно, усложняет проверку его выполнения. В [60] приведено некоторое более внутреннее свойство семейства линейных форм $\{A_1,\dots,A_m\}$, при выполнении которого это семейство является *grate* (а также удовлетворяет условию отделимости) для подходящих значений параметров. Именно, обозначим через A – матрицу коэффициентов семейства $\{A_1,\dots,A_m\}$, и пусть для всякой матрицы $\mathcal{D}$, такой что $\operatorname{rg}\mathcal{D}\leqslant\theta$, в матрице $(A+\mathcal{D})$ не менее η ненулевых элементов. Тогда семейство $\{A_1,\dots,A_m\}$ является (θ,η)–отделимым. В [60] поставлено два открытых вопроса, суть которых сводится к следующему (их можно в совокупности рассматривать как некоторый план на пути получения нижних оценок):

I) построить (явно – ср.начало § 3) пример семейства функций $\{g_1,\dots,g_m\}$, удовлетворяющего условию (θ,η)–отделимости для "нетривиальных" значений θ и η (например, $\theta\times m, \eta\times mn$);

2) показать, что сложность $C(\beta)$ нелинейна по $\max\{m,n\}$ при некоторых θ и η (считаем, что β вычисляет семейство функций, удовлетворяющих условию (θ,η)–отделимости).

Частичным ответом на вопрос 2) Вэльянта служит следующая принадлежащая автору теорема, метод доказательства которой содержится в § I [6] (там она доказана для случая $\theta\sim\frac{m}{2}$, $\eta\times mn$ и сформулирована в менее общем виде).

ТЕОРЕМА I5.I. Если семейство $\{g_1,\dots,g_m\}$ является (θ,η)–отделимым, то для всякой НП β, вычисляющей это семейство, и граф G_β которой удовлетворяет ограничению (7), выполнено $C(\beta)\geqslant M$, где M – единственное положительное решение

уравнения

$$M=(\eta/m)^{\left(1+\frac{\theta}{Mn2}\right)}.$$

Отметим, что при $\theta \asymp m \asymp n$ и $\eta \asymp m^2$ выполнено $M \asymp m \log m / \log\log m$. Метод доказательства теоремы был использован также в [7] для получения нижней оценки (уже без ограничений типа (7)) для монотонной сложности семейства дизъюнкций (ср.§ I3).

В § I [6] предложено приложение теоремы к оценке аддитивной сложности вычисления семейства линейных форм над $\mathbb{R}$ (по-прежнему в предположении (7)), причем $P=\{+\}\cup\{x\gamma\}_{|\gamma|\le 1}$; $\lambda_{11}=\lambda_t$ – тотальная сложность (см.§ I). Пусть семейство векторов $a_1,\dots,a_m\in\mathbb{R}^n$, таково что

$$\forall I\,(\varnothing\neq I\subsetneq\{1,\dots,m\}\Rightarrow\rho_1(Conv\{a_i\}_{i\in I},\ Conv\{a_j\}_{j\notin I})\ge c),$$

где $Conv$ обозначает выпуклую оболочку, ρ_1 – метрика, отвечающая норме $l_1(\alpha_1,\dots,\alpha_n)=\sum\limits_{1\le i\le n}|\alpha_i|$ (в § I [6] такое семейство векторов названо (m,c)-системой). На основании теоремы Радона (см.[I2]) в § I [6] показано, что семейство $a_1,\dots,a_m$ (обладающее выписанным свойством) удовлетворяет условию $(m/2,\ mc/4)$-отделимости в рассматриваемой ситуации. Отсюда, по теореме I5.I получаем, что $C_{11}(a_1,\dots,a_m)\ge M$, где M взято из теоремы для параметров $\theta=m/2$, $\eta=mc/4$. Интересно было бы попытаться условие, аналогичное сформулированному, использовать в ситуациях, где нет теорем типа теоремы Хелли ([I2]) и даже нет прямого аналога выпуклости (в рассмотренном случае теорема Радона позволила фактически дать ответ на поставленный выше вопрос I) Вэльянта).

НП над $\mathbb{R}$ (или $\mathbb{C}$) с таким же P, как и в предыдущем абзаце (про условие (7) здесь и ниже можно забыть), рассматривались в [40], где замечено, что в случае $m=n$ сложность $C_{11}(a_1,\dots,a_n)\ge\log\det(a_1,\dots,a_n)$. Оценка из предыдущего абзаца слабее в интересных случаях оценки Маргенштерна, но возможно, самостоятельный интерес представляет (помимо теоремы I5.I) проделанная реализация пути, предложенного в вопросах I), 2) Вэльянтом.

Другой частичный ответ на сформулированный выше вопрос 2) дал Вэльянт в [61]. Глубиной $d(G)$ упорядоченного графа G назовем наибольшую длину ориентированных путей в графе G.

ТЕОРЕМА I5.2 ([61]). Пусть для какого-то $\varepsilon > 0$ некоторое (θ,η)-отделимое семейство, где $\eta \geq \theta^{1+\varepsilon}$, вычисляется с помощью НП β, для которой глубина $d(G_\beta) = O(\log\theta)$. Тогда

$$C_t(\beta) \geq \theta \log\log\theta / \log\log\log\theta .$$

§ I6. Аддитивная сложность в треугольных и направленных вычислениях и разложение Брюа

В последнем параграфе мы введем два класса НП (треугольные и направленные), вычисляющих семейства линейных форм (см. [35]). Для треугольных НП ниже предлагается метод установления нелинейных нижних оценок сложности. Для направленных вычислений помимо нижних оценок предложена, более того, явная формула сложности на основе развитого автором аппарата в теории групп Шевалле (необходимые здесь сведения изложены так, что никакой предварительной информации о группах Шевалле не предполагается).

Итак, рассматриваем следующие несколько модифицированные НП. В обозначениях § I, F – основное поле, $\{x_1,\dots,x_n\}$ – входные переменные, $P = \{a \to a + \alpha b\}_{\alpha\in F} \cup \{x\alpha\}_{\alpha\in F}$. В матричных терминах этим командам естественно соответствуют элементарные операции. Имеется также N переменных $z_1,\dots,z_N$ (которые можно рассматривать как память) среди которых выделены n переменных $z_{i_1},\dots,z_{i_n}$ $(1 \leq i_1 < \dots < i_n \leq N)$, которые назовем о с н о в н ы м и (остальные – в с п о м о г а т е л ь н ы м и). Сама НП β является последовательностью строк, и для каждых $1 \leq t \leq T$ и $1 \leq j \leq N$ естественно (аналогично тому, как это делалось в § I4), индукцией по t определяется линейная форма $Z_j^{(t)}$ от переменных $x_1,\dots,x_n$ с коэффициентами из F. В начальный момент $(t=0)$ полагаем $Z_{i_j}^{(0)} = x_j$ $(1 \leq j \leq n)$ для основных переменных и $Z_s^{(0)} = 0$ $(s \neq i_1,\dots,i_n)$ для вспомогательных. Считаем по определению, что β вычисляет семейство n линейных форм $Z_{i_1}^{(T)},\dots,Z_{i_n}^{(T)}$. Это ограничение, т.е. считывание в конце вычисления выходов там, где в начальный момент были входы, не являющееся существенным для НП общего вида, в нашем случае очень важно.

Сначала обратимся к треугольным вычислениям. По определению, всякая команда т р е у г о л ь н о й НП β имеет вид $z_j = z_j + \alpha z_i$, где $i > j$ или $z_j = \alpha z_j$ $(\alpha \in F)$.

Эти команды (на матричном языке) отвечают верхнетреугольным элементарным преобразованиям. Положим функцию $\lambda = \lambda_\Delta$ (см. § I) равной единице на командах первого типа и нулю на командах второго типа. Порождаемую при этом меру сложности C_Δ (см. § I) назовем треугольной. Если A – $n \times n$ матрица коэффициентов семейства линейных форм $a_1, \ldots, a_n$, то обозначим $C_\Delta(A) = C_\Delta(a_1, \ldots, a_n)$. Отметим, что в оправдание своего названия, треугольная сложность $C_\Delta(A)$ определена только для верхнетреугольных матриц A (т.е. матриц с нулями ниже диагонали).

ТЕОРЕМА I6.I ([35]). Пусть верхнетреугольная матрица A представлена в виде $A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$, где A_1, A_2 – верхнетреугольные. Тогда

$$C_\Delta(A) \geq C_\Delta(A_1) + C_\Delta(A_2) + rg\, B .$$

В качестве приложения теоремы рассмотрим семейство верхнетреугольных матриц $\{A_n\}$, где $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \ldots, A_{n+1} = \begin{pmatrix} A_n & E \\ 0 & A_n \end{pmatrix}, \ldots$; здесь E – единичная матрица, (т.е. A_n имеет размеры $2^n \times 2^n$). Тогда

СЛЕДСТВИЕ I6.2 ([35]) $C_\Delta(A_n) = n \cdot 2^{n-1}$.

Иными словами, рост треугольной сложности $C_\Delta(A_n)$ нелинеен по размеру матриц.

Второй класс НП, который мы здесь рассмотрим – направленные НП. Всякая команда направленного НП β имеет вид

$$z_{k+1} = z_{k+1} + \alpha z_k \quad \text{или} \quad z_j = z_j + \alpha z_i \ , \text{где } j \leq i \ (\alpha \in F).$$

Положим функцию $\lambda = \lambda_d$ равной единице на командах первого типа и нулю на командах второго типа. Порождаемую при этом меру сложности C_d (см. § I) назовем направленной. В отличие от треугольной сложности, направленная сложность $C_d(A)$ (используем введенные выше обозначения) определена для любой квадратной $(n \times n)$ матрицы A.

Для $C_d(A)$ нетрудно получить нелинейные (по n) нижние оценки, но удалось сделать гораздо большее: получить явную формулу для $C_d(A)$. Для формулировки этого результата потребуются некоторые предварительные сведения, к изложению которых мы и переходим.

Пусть S_n – симметрическая группа (т.е. группа всех перестановок n-элементного множества), которую одновременно будем

рассматривать как подгруппу группы GL_n неособых матриц (все встречающиеся здесь и далее матрицы имеют размеры $n \times n$). Обозначим через $\mathcal{T}$ многообразие верхнетреугольных матриц, через $\mathcal{B} = \mathcal{T} \cap GL_n$ подгруппу всех неособых верхнетреугольных матриц. Разложение Брюа (см. § 3 [20]) состоит в том, что для всякой матрицы $A \in GL_n$ существует и единственна перестановка $w_A \in S_n$, такая что $A \in \mathcal{B} w_A \mathcal{B}$. На группе S_n определяется (см. § I гл.4 [4]) функция длины $l(w)$, где $w \in S_n$, как наименьшее такое l, что $w = \sigma_{j_1} \dots \sigma_{j_l}$, где $\sigma_j = (j\ j+1)$ - транспозиция двух соседних индексов $(1 \leq j < n)$. Всякое разложение w в произведение транспозиций вида σ_j с наименьшим числом сомножителей равным $l = l(w)$, называется приведенным. Нетрудно видеть, что $l(w)$ совпадает с числом инверсий в w, т.е. числом пар $i < j$, для которых $w(i) > w(j)$. Обозначим $l(A) = l(w_A)$, т.е. определим функцию l на GL_n (неособых матрицах).

На группе S_n вводится (см. § 8 [20]) отношение $\preccurlyeq$ частичного порядка: $w_1 \preccurlyeq w_2$ $(w_1, w_2 \in S_n)$, если w_1 равно некоторому подпроизведению (с сохранением порядка сомножителей) какого-то приведенного разложения элемента w_2. Можно показать (см. § 8 [20]), что определение порядка не зависит от выбора приведенного разложения w_2. Порядок проявляет себя в следующей теореме Шевалле (см. § 8 [20]): $\overline{\mathcal{B} w \mathcal{B}} = \bigcup_{w_1 \geq w} \mathcal{B} w_1 \mathcal{B}$, где черта обозначает замыкание в топологии Зарисского (здесь предполагается, что поле F бесконечно). Как следует из разложения Брюа, под знаком объединения стоят попарно непересекающиеся множества.

Нетрудно видеть, что $C_d(A) \leq l(A)$ для неособых матриц A. Чтобы установить обратное неравенство, автору пришлось продолжить функцию l с GL_n на $\mathcal{M}_n$ (через $\mathcal{M}_n$ обозначим многообразие всех $n \times n$ матриц), доказать аналог разложения Брюа (см. теорему I6.3 ниже) и аналог теоремы Шевалле (см. следствие I6.4 ниже) для $\mathcal{M}_n$, затем установить некоторое свойство монотонности продолженной функции l (см. лемму I6.5 ниже), и наконец доказать равенство $C_d(A) = l(A)$ уже для всех $A \in \mathcal{M}_n$.

ТЕОРЕМА I6.3 ([II,35]). Для всякой $A \in \mathcal{M}_n$ существует и единственна перестановка $w_A \in S_n$, такая что

I) $A \in \mathcal{T} w_A \mathcal{T}$;

2) если $A \in \mathcal{T} w \mathcal{T}$ для некоторой $w \in S_n$, то $w_A \leqslant w$.

Перестановка w_A строится за время $O(n^3)$ по A. Теперь, на основе теоремы, можно продолжить функцию ℓ на $\mathcal{M}_n$, положив $\ell(A) = \ell(w_A)$.

СЛЕДСТВИЕ I6.4 ([II,35]). I) $\overline{\mathcal{T} w \mathcal{T}} = \bigcup_{w_1 \leqslant w} \mathcal{T} w_1 \mathcal{T}$;

2) функция ℓ полунепрерывна сверху на $\mathcal{M}_n$.

В отличие от неособого случая, множества, входящие в объединение в пункте I) следствия, могут пересекаться.

Скажем, что $n \times n$ матрица A является г л а в н о й подматрицей $m \times m$ матрицы D $(n \leqslant m)$, если A – подматрица матрицы D, и ее диагональ лежит на диагонали матрицы D; иными словами, подматрица A выделяется из матрицы D с помощью строк и столбцов с одними и теми же индексами.

ЛЕММА I6.5 ([II,35]). Если A – главная подматрица матрицы D, то $\ell(A) \leqslant \ell(D)$.

Наконец, доказав еще ряд утверждений для функции ℓ (см. [II,35]), и опираясь на I6.3, I6.4, I6.5, получаем в итоге обещанную явную формулу для $C_d(A)$.

ТЕОРЕМА I6.6 ([35]). $C_d(A) = \ell(A)$.

Отметим, что теоремы I6.3 и I6.6 и лемма I6.5 верны и над конечным основным полем F.

В заключение упомянем о том, что теорема I6.3 и следствие I6.4, а также частично лемма I6.5 обобщаются (в инвариантной форме) на произвольные классические группы Шевалле ([II]).

Литература

I. А х о А., Х о п к р о ф т Дж., У л ь м а н Дж. Построение и анализ вычислительных алгоритмов. – М., Мир, I979.

2. Б е р н ш т е й н Д.Н. Число корней системы уравнений. – Функц.анализ прилож., I975, т.9, № 3, с.I–4.

3. Б и л л и н г с л е й П. Эргодическая теория и информация. – М., Мир, I969.

4. Б у р б а к и Н. Группы и алгебры Ли. гл.IУ–УI. – М.,Мир, I972.

5. Г а н т м а х е р Ф.Р. Теория матриц. – М., Гостехиздат, I954.

6. Г р и г о р ь е в Д.Ю. Использование понятий отделенности и независимости для получения нижних оценок сложности схем. – Зап.научн.семинаров ЛОМИ АН, I976, т.60, с.38–48.

7. Григорьев Д.Ю. Об одной нижней оценке сложности вычисления семейства дизъюнкций в монотонном базисе. – Зап. научн.семинаров ЛОМИ АН, 1977, т.68, с.19–25.
8. Григорьев Д.Ю. Ранг пары матриц и свертки. – Успехи матем.наук, 1979, т.34, № 2, с.193–194.
9. Григорьев Д.Ю. Алгебраическая сложность вычисления семейства билинейных форм. – Ж.выч.матем. и матем.физ., 1979, т.19, № 3, с.563–580.
10. Григорьев Д.Ю. Соотношение ранга и мультипликативной сложности билинейной формы над нетеровым коммутативным кольцом. – Зап.научн.семинаров ЛОМИ АН, 1979, т.86, с.66–81.
11. Григорьев Д.Ю. Аналог разложения Брюа для замыкания конуса группы Шевалле классической серии. – Докл.АН СССР 1981, т.257, № 5, с.1040–1044.
12. Данцер Л., Грюнбаум Б., Кли В. Теорема Хелли и ее применение. – М., Мир, 1968.
13. Кнут Д. Искусство программирования для ЭВМ, т.2. – М. Мир, 1977.
14. Маклейн С. Гомология. – М., Мир, 1966.
15. Манин Ю.И. Лекции по алгебраической геометрии, ч.1. – М., изд-во МГУ, 1968.
16. Маргулис Г.А. Явные конструкции концентраторов. – Пробл.перед.информ., 1973, т.9, № 4, с.71–80.
17. Михайлюк М.В. О сложности вычисления элементарных симметрических функций в конечных полях. – Докл.АН СССР, 1979, т.244, № 5, с.1072–1076.
18. Пан В.Я. О способах вычисления значений многочленов. – Успехи матем.наук, 1966, т.21, № 1, с.103–134.
19. Питерсон У. Коды, исправляющие ошибки. – М., Мир, 1964.
20. Стейнберг Р. Лекции о группах Шевалле. – М., Мир, 1975.
21. Хованский А.Г. Об одном классе систем трансцендентных уравнений. – Докл.АН СССР, 1980, т.244, № 5, с.1072–1076.
22. Шёнхаге А., Штрассен Ф. Быстрое умножение больших чисел. – Кибернет.сб., нов.сер., 1973, вып.10, с.87–98.
23. Штрассен Ф. Сложность вычисления элементарных симметрических функций и коэффициентов интерполяционного поли-

нома. - Кибернет.сб., нов.сер., 1978, вып.15, с.5-21.

24. Alder A., Strassen V. On the algorithmic complexity of associative algebras. - Theoret.Comput.Sci., 1981, vol.15, p.201-211.
25. Baur W., Strassen V. The complexity of partial derivatives. - Prepr.Univ.Zurich, 1981.
26. Borodin A., Cook S. On the number of additions to compute specific polynomials. - SIAM J. Comput., 1976, vol.5, N 1, p.146-157.
27. Borodin A., Munro I. Computational complexity of algebraic and numeric problems. - Amer.Els., N.Y., 1975.
28. Brown M.R., Dobkin D.P. An improved lower bound on polynomial multiplication . - IEEE Trans.Comput., 1980, vol.29, N 5, p.337-340.
29. Coppersmith D., Winograd S. On the asymptotic complexity of matrix multiplication. - Prepr.Res. Centre IBM, 1981.
30. Emde Boas P. van. Berekeningskomplexiteit van Bilineaire en Kwadratische Vormen. - Prepr.Univ.van Amsterdam, 1980.
31. Fiduccia C.M., Zalcstein Y. Algebras having linear multiplicative complexity. - J.Assoc.Comput. Mach., 1977, vol.24, N 2, p.311-331.
32. Grigor'ev D.Yu. Some new bounds on tensor rank. - LOMI Prepr. E-2-78, Leningrad, 1978.
33. Grigor'ev D.Yu. Multiplicative complexity of a pair of bilinear forms and of the polynomial multiplication. - Lect.Notes.Comput.Sci., 1978, vol.64, p.250-256.
34. Grigor'ev D.Yu. Multiplicative complexity of a bilinear form over a commutative ring. - Lect.Notes Comput. Sci., 1981, vol.118, p.281-286.
35. Grigor'ev D.Yu. Additive complexity in directed computations. - Theoret.Comput.Sci., 1982, vol.19.
36. Heintz J., Sieveking M. Lower bounds for polynomials with algebraic coefficients. - Theoret.Comput. Sci., 1980, vol.11, N 3, p.321-330.
37. Hyafil L. The power of commutativity. - 18th Annu. Symp.Found.Comput.Sci., 1977, p.171-174.
38. Ja Ja J. Optimal evaluation of pair of bilinear forms. - Proc.10th Annu.ACM Symp.Theory Comput., 1978,p.173-183.

39. J a J a J. Time-space tradeoffs for some algebraic problems. - Proc.12th Annu.ACM Symp.Theory Comput., 1980, p.339-350.
40. M o r g e n s t e r n J. Note on a lower bound of the linear complexity of the fast Fourier transform. - J.Assoc. Comput.Mach., 1973, vol.20, N 2, p.305-306.
41. P a n V.Ya. Computational complexity of computing polynomials over the fields of real and complex numbers. - Proc. 10th Annu.ACM Symp.Theory Comput., 1978, p.162-172.
42. P a t e r s o n M.S., S t o c k m e y e r L.J. On the number of nonscalar multiplications necessary to evaluate polynomials. - SIAM J.Comput., 1973, vol.2, p.60-66.
43. P i n s k e r M.S. On the complexity of a concentrator. - 7th Intern.Teletraffic Congr., Stockholm, 1973.
44. P i p p i n g e r N. Superconcentrators. - Techn.Rep. RC 5937, IBM Yorktown Heights, 1976.
45. P i p p i n g e r N. Computational complexity in algebraic function fields. - 20th Annu. Symp.Found.Comput.Sci., 1979, p.61-65.
46. S a v a g e J.E. An algorithm for the computation of linear forms. - SIAM J.Comput., 1974, vol.3, N 2, p.150-158.
47. S a v a g e J.E., S w a m y S. Space-time tradeoffs on the FFT algorithm. - IEEE Trans. Inform.Th., 1978, vol.IT-24, N 5, p.563-568.
48. S a v a g e J.E., S w a m y. S. Space-time tradeoffs for oblivious sorting and integer multiplication. - Techn.Rep. CS-32, Brown Univ., 1978.
49. S c h n o r r C.P. A lower bound on the number of additions in monotone computations. - Theoret.Comput.Sci.,1976, vol.2, p.305-315.
50. S c h n o r r C.P. On the additive complexity of polynomials and some new lower bounds. - Lect.Notes Comput.Sci., 1979, vol.67, p.286-297.
51. S c h n o r r C.P. W i e l e J.P. V a n d e. On the additive complexity of polynomials. - Theoret.Comput.Sci., 1980, vol.10, p.1-18.
52. S c h ö n h a g e A. Partial and total matrix multiplication. - SIAM J. Comput., 1981, vol.10, N 3, p.434-455.
53. S h a m o s M.I. Y u v a l G. Lower bounds from complex function theory. - 17th Annu.Symp.Found.Comput.Sci., 1976, p.268-273.

54. S t r a s s e n V. Vermeidung von Divisionen. - J.reine angew.Math., 1973, B.264, S.184-202.
55. S t r a s s e n V. Polynomials with rational coefficients that are hard to compute. - SIAM J.Comput., 1974, vol.3, N 2, p.128-149.
56. S t r a s s e n V. Computational complexity over finite fields. - SIAM J.Comput., 1976, vol.5, N 2, p.324-331.
57. S t r a s s e n V. The computational complexity of continued fractions. - Proc.ACM Symp.Symb.Algebr.Comput.,1981, p.51-67.
58. T o m p a M. Time-space tradeoffs for computing functions, using connectivity property of their circuits. - Proc.10th Annu.ACM Symp. Theory Comput., 1978, p.196-204.
59. V a l i a n t L.G. On non-linear lower bounds in computational complexity. - Proc.7th Annu.ACM Symp.Theory Comput., 1975, p.45-53.
60. V a l i a n t L.G. Some conjectures relating to super-linear complexity bounds. - Techn.Rep. N 85, Univ.Leeds, 1976.
61. V a l i a n t L.G. Graph-theoretic arguments in low-level complexity. - Comput.Sci.Rep.13-77, Univ.Edinburgh, 1977.
62. V a l i a n t L.G. Negation can be exponentially powerful. - Theoret.Comput.Sci., 1980, vol.12, p.303-314.
63. W i e l e J.P.V. d e. Complexité additive et zéros des polynômes a coefficients réels et complexes. - Proc. 1st meet. AFCET-SMT Appl.Math., 1978.
64. W i n o g r a d S. On the number of multiplications necessary to compute certain functions. - Communs Pure Appl. Math., 1970, vol.23, p.165-179.

Grigor'ev D.Yu. Lower bounds in the algebraic computational complexity.

This paper is a survey on some selected methods for obtaining lower bounds in the algebraic complexity. The content is the following:

Introduction. 1. Basic notions

Chapter I. Algebraic-geometric approach to obtaining lower bounds of computational complexity of polynomials. 2. Evaluating of a polynomial with "general" coefficients. 3. Computational complexity of the individual polynomials. 4. The degree method and its generalizations (the case of an infinite ground

field). 5. The degree method (the case of a finite ground field) 6. Additive complexity and real roots.

Chapter II. Lower bounds on the multiplicative complexity for the problems in the linear algebra. 7. Multiplicative complexity and the rank. 8. Rank of a pair of bilinear forms. 9. Multiplicative complexity of a bilinear form over a commutative ring. 10. Bounds on the rank of algebras. 11. Linearized multiplicative complexity.

Chapter III. Complexity for straight-line programs of non-standard kinds. 12. Irrational computational complexity of algebraic functions. 13. Monotone programs. 14. Time-space tradeoffs. 15. Graph-theoretic methods in the algebraic complexity. 16. Additive complexity in triangular and directed computations and Bruhat decomposition.

We shall explain in more detail the author's results from §§6, 11, 15, 16 unpublished earlier.

In § 6 the best previous known bound on the additive complexity over the real numbers in the terms of the number of real roots ([26]) is essentially improved. Namely, let polynomials $g_1,\dots,g_n \in \mathbb{R}[x_1,\dots,x_n]$ and the system $g_1=\dots=g_n=0$ have N pairwise distinc discrete simple real roots. Then Corollary 6.2. The additive complexity of a set $\{g_1,\dots,g_n\}$ is greater than $(\sqrt{\log N}-2n)/3$.

The proof is based on one Hovansky theorem [21]. This bound is sharp up to taking of the square root according to [63].

In § 11 the following invariant $C_\ell(f)$ of the homogenuous form f of the degree d_1 over n variables $x_1,\dots,x_n$ is introduced:

$$C_\ell(f)=\min\Big\{N: f=\sum_{1\le i\le N} \ell_1^{(i)}\dots\ell_{d_1}^{(i)},\ \text{where } \ell_j^{(i)} \text{ is a linear form } (1\le i\le N,\ 1\le j\le d_1)\Big\}.$$

This definition is similar in spirit to the definition of the tensor rank (see e.g. [54]). We suggest a method for obtaining lower bounds for $C_\ell(f)$.

For the convenience of notations (w.l.o.g.) we assume that $d_1=2d$ is even. Consider the auxiliary matrices A of the size $n^d\times n^d$ the rows and columns of which are numbered by the vectors of the form $I=(i_1,\dots,i_d)$ for all $1\le i_1,\dots,i_d\le n$. To each

vector I there corresponds a monomial $x^I = x_{i_1} \dots x_{i_d}$ of degree d. By $|I|$ we denote the number $\kappa_1! \dots \kappa_n!$, where κ_i is the number of j such that $i_j = i$. To every matrice A there corresponds the form $g_A = \sum_{I,J} a_{I,J} x^I x^J$ of degree $d_1 = 2d$. Finally, we define a linear operator L on the matrices of the size $n^d \times n^d$ by the formula: $L(A) = B$, where an entry

$$b_{P,Q} = \Big(\sum_{x^P x^Q = x^K = x^I x^J} a_{I,J} \Big) |K| .$$

A form $g_A = 0$ iff $L(A) = 0$, so $L(f)$, which can be defined as $L(A)$ for any A such that $f = g_A$, is an invariant of the form f.

Theorem 11.3. $C_\ell(f) \geq (\operatorname{rg} L(f))/d_1!$

As a corollary we obtain for example that

$$C_\ell\Big(\Big(\sum_{1 \leq i \leq n} x_i \overline{x_i}\Big)^d\Big) \geq \binom{n+d-1}{n-1} / (2d)!$$

In § 15 we consider a class of straight-line programs β which satisfy the following restriction (7): for any vertex v of the ordered graph G_β corresponding to β in the usual manner, the subgraph T_v consisting of all the vertices situated in G_β above the vertex v is a tree. From the other hand we allow an initial variable to be represented by more than one initial vertices in G_β. We say, that a set of the functions $\{g_1, \dots, g_m\}$ is (θ, η)-s e p a r a t e d if after the deleting from the graph G_β (for any straight-line program β computing $\{g_1, \dots, g_m\}$ and satisfying the restriction (7)) of arbitrary θ vertices, there exist η pairs of vertices such that one member of each pair is an initial vertex and the other is a terminal one and between the members of each pair there is some path in G_β (this path is unique according to restriction (7) above) containing no vertex from the deleted set of θ vertices (informally speaking, even if one can utilize "free of charge" some θ auxiliary functions, a lot of the terminal functions depends essentially on many initial variables) This notion is very close to the notion of $(n, \eta(\theta))$-grate ([60]) and was actually introduced independently in [6]. The following theorem can be considered as a kind of weakened answer on the problem of Valiant posed in [60], and one can find in fact the method for proving of the theorem in [6].

Theorem 15.1. For any (θ, η)-separated set $\{g_1, \dots, g_m\}$

the complexity of every straight-line program β satisfying the restriction (7) and computing $\{g_1,\dots,g_m\}$ is greater than a number M, where M is an unique positive solution of the equation

$$M=(n/m)^{\left(1+\frac{\theta}{M\ln 2}\right)}.$$

In the last §16 we introduce two classes of straight-line linear programs (they are called triangular and directed). Every linear program can be represented informally as a sequence of elementary matrix transformations. A triangular program consists of instructions of the form

$$z_j =: z_j + \alpha z_i \quad , \text{ where } i \geq j \ , \ \alpha \in F \ .$$

The triangular complexity C_Δ counts the number of all the instructions. A method for obtaining nonlinear lower bounds on C_Δ is suggested. For example, set

$$A_1=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},\dots, A_{n+1}=\begin{pmatrix}A_n & E\\ 0 & A_n\end{pmatrix},\dots,$$ where E

is the unit matrix. Then $C_\Delta(A_n)=n\cdot 2^{n-1}$.

The second class of linear programs under consideration is a class of directed programs. A directed program consists of the instructions of the following two forms:

$$z_j =: z_j + \alpha z_i \quad , \text{ where } i \geq j \ , \alpha \in F \text{ and}$$
$$z_{k+1} =: z_{k+1} + \alpha z_k$$

The directed complexity C_d is the number of instructions only of the second form. For $C_d(A)$ the explicit effective formula is obtained in the terms of the so-called generalized Bruhat decomposition introduced by the author and based on the suitable algebraic apparatus in the Chevalley group theory ([11]). The exposition of the results of §16 in detail can be found in [35]. Also it is noticed in §6 that a direct analogy of theorem of Baur and Strassen [25] about the complexity of the partial derivatives is false for the additive complexity C_+. Namely, set $f=x_1\cdots x_n+x_1x_2^2\cdots x_n^n$, then $C_+(f)=1$ and from the other hand $C_+\left(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\right)=n$.

УДК 519.5

Нижние оценки сложности для машинных моделей вычисления. Белятыков А.П. – В кн.: Теория сложности вычислений. I (Зап. научн. семин. ЛОМИ, т. 118). Л., "Наука", 1982, с. 4–24.

Статья является текстом обзорного доклада по методам получения нижних оценок вычислительной сложности для абстрактных вычислительных машин. Кроме методов получения нижних оценок излагаются родственные им методы моделирования одних машин другими с сокращением одних мер сложности за счет увеличения других (результаты типа *trade-off*). Рассматриваются методы следов, хвостов, перекрытий и родственные им методы. В качестве работы метода иногда дается новое доказательство старого результата или доказывается новый результат. Библ. – 43 назв.

УДК 519.5

Нижние оценки в алгебраической сложности вычислений. Григорьев Д.Ю. – В кн.: Теория сложности вычислений. I (Зап. научн. семин. ЛОМИ, т. 118). Л., "Наука", 1982, с. 25–82.

Настоящая работа представляет собой обзор по избранным методам в получении нижних оценок в алгебраической сложности. Приведем оглавление.

Введение. I. Основные понятия. Глава I. Алгебро-геометрический подход к получению нижних оценок сложности вычисления многочленов. 2. Вычисление многочлена с "общими" коэффициентами. 3. Сложность вычисления индивидуальных многочленов. 4. Метод степени и его обобщения (случай бесконечного основного поля). 5. Метод степени (случай конечного основного поля). 6. Аддитивная сложность и вещественные корни. Глава II. Нижние оценки мультипликативной сложности в задачах линейной алгебры. 7. Мультипликативная сложность и ранг. 8. Ранг пары билинейных форм. 9. Мультипликативная сложность билинейной формы над коммутативным кольцом. I0. Оценки ранга алгебр. II. Линеаризованная мультипликативная сложность. Глава III. Сложность в неветвящихся программах нестандартных типов. I2. Иррациональная сложность вычисления алгебраических функций. I3. Монотонные вычисления. I4. Нижние оценки для произведения времени и памяти. I5. Методы теории графов в алгебраической сложности. I6. Аддитивная сложность в треугольных и направленных вычислениях и разложение Брюа. Библ. – 64 назв.