Math-Net.Ru

Общероссийский математический портал

С. Б. Гашков, Арифметическая сложность преобразований Стирлинга, *Дискрет. матем.*, 2014, том 26, выпуск 4, 23–35

DOI: <https://doi.org/10.4213/dm1301>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.224.187.69

26 апреля 2023 г., 13:19:53

Арифметическая сложность преобразований Стирлинга

© 2014 г. С. Б. Гашков*

Получены близкие к точным оценки сложности вычисления известных в комбинаторике линейных преобразований Стирлинга обоих родов векторными аддитивными цепочками и неветвящимися программами, составленными из арифметических операций над действительными числами. Рассмотрена связь рассматриваемых задач с интерполяцией по Лагранжу и Ньютону.

Работа выполнена при финансовой поддержке РФФИ, проекты 14-01-00598 и 14-01-00671а.

Ключевые слова: преобразования Стирлинга 1 и 2 го рода, аддитивные векторные цепочки, схемы в арифметических базисах, лагранжева и ньютонова интерполяция, матрицы Вандермонда, q -биномиальные коэффициенты Гаусса.

1. Введение

В статье рассматриваются встречающиеся в различных вопросах конечномерные линейные преобразования над полями действительных и комплексных чисел и оценивается сложность их вычисления с помощью различных наборов элементарных арифметических операций в этих полях. Эти множества элементарных операций называются *базисами*. *Сложность вычисления* — это число выполненных операций (далее будет дано формальное определение).

Широко известно (см., например, [1, 2, 3, 4]) *биномиальное преобразование* $\mathbb{R}^n \rightarrow \mathbb{R}^n$, и обратное к нему, определяемые формулами

$$y_k = \sum_{m=0}^k \binom{k}{m} x_m, \quad x_k = \sum_{m=0}^k (-1)^{k-m} \binom{k}{m} y_m, \quad k = 0, \dots, n-1,$$

где $\binom{k}{m} = k!/(m!(k-m)!)$ — биномиальные коэффициенты. Матрица $C_n = \|c_{k,m}\|$, $c_{k,m} = \binom{k}{m}$, этого преобразования треугольная, ее ненулевые элементы образуют *треугольник Паскаля* порядка n (при $m > k$ полагаем $\binom{k}{m} = 0$). Обратную матрицу обозначаем $C_n^{-1} = \|(-1)^{k-m} \binom{k}{m}\|$ (нумерация строк и столбцов во всех матрицах здесь и далее начинается с нуля).

В комбинаторике ([1, 2, 3, 4]) и в теории интерполяции [5] используются также преобразования Стирлинга 1-го рода $y = s(x)$ и 2-го рода $x = S(y)$, взаимно обратные друг к другу:

$$y_k = \sum_{m=0}^k s_m^k x_m, \quad x_k = \sum_{m=0}^k S_m^k y_m, \quad k = 0, \dots, n-1,$$

*Место работы: МГУ им. М.В. Ломоносова, e-mail: sbgashkov@gmail.com

где s_m^k, S_m^k — числа Стирлинга 1-го и 2-го рода соответственно (используем обозначения из [3]). Матрицы этих преобразований $S_n = \|S_{k,m}\|, S_{k,m} = S_m^k$ и $s_n = \|s_{k,m}\|, s_{k,m} = s_m^k$ являются треугольными матрицами, ненулевые элементы которых образуют *треугольники Стирлинга* порядка n соответственно 1-го и 2-го рода (при $m > k$ полагаем $S_m^k = s_m^k = 0$).

Замечание 1. Эти треугольники определяются *тождествами Стирлинга*:

$$s_k^n = s_{k-1}^{n-1} - (n-1)s_k^{n-1}, S_k^n = kS_{k-1}^{n-1} + S_{k-1}^{n-1}, S_0^n = s_0^n = 0, n > 0, S_0^0 = s_0^0 = 1.$$

Числа Стирлинга встречаются и в других тождествах, например

$$x^n = \sum_{k=0}^n S_k^n(x)_k, (x)_k = x(x-1)\dots(x-k+1), (x)_0 = 1, (x)_n = \sum_{k=0}^n s_k^n x^k,$$

$$(x)^n = x(x+1)\dots(x+n-1) = \sum_{k=0}^n |s_k^n| x^k, k^n = \sum_{i=0}^n S_i^n(k)_i = \sum_{i=0}^n i! S_i^n \binom{k}{i},$$

и имеют комбинаторные интерпретации [1, 2, 3, 4]: S_k^n равно числу разбиений n -элементного множества на k непустых подмножеств, а $|s_k^n|$ равно числу подстановок n -го порядка, разлагающихся в произведение k непересекающихся циклов.

Монотонное преобразование Стирлинга 1-го рода определяется равенством $y_k = \sum_{m=1}^k |s_m^k| x_m, k = 1, \dots, n$. Число $\mathbb{B}_n$ всех разбиений n -элементного множества на непустые подмножества называется *числом Белла*: $\mathbb{B}_n = \sum_{k=1}^n S_k^n$.

Применяя тождество Паскаля, можно вычислить все числа, образующие треугольник Паскаля, с помощью $(n-1)(n-2)/2$ операций сложения¹. Этот алгоритм оптимален в случае *аддитивного базиса* $B_+ = \{x+y, 1\}$. Используемая при этом модель вычислений известна под названием *аддитивные цепочки* [6]. Аддитивной цепочкой длины l называется последовательность натуральных чисел $a_0, \dots, a_l$, в которой $a_0 = 1$, и каждое число a_i равно сумме каких-то двух предшествующих чисел $a_j, a_k, j \leq k < i$. Аддитивная цепочка может начинаться не с единицы, а с переменной x . Тогда она вычисляет функцию nx . Для функции $l(n)$, равной длине кратчайшей аддитивной цепочки, заканчивающейся числом n , известна асимптотически точная формула $\log_2 n$. Легко получить оценку $l(n) = O(\log n)$.

Известно обобщение аддитивных цепочек, называемое *векторными аддитивными цепочками* (см., например, [6, 7]). Эти цепочки состоят из векторов заданной размерности m и начинаются с m векторов вида $(0, \dots, 1, 0, \dots, 0)$. Их можно рассматривать как *схемы с входами* $x_1, \dots, x_m$ в базисе B_+ .

Замечание 2. Приведем определение *схемы (неветвящейся программы)* для произвольного базиса B , состоящего из операций вида $\omega_i(x_1, \dots, x_{n_i}) : E^{n_i} \rightarrow E$, где E заданное множество, например, множество действительных чисел, $n_i \geq 0$. Схемой с входами $x_1, \dots, x_n$ называется последовательность функций $f_i(x_1, \dots, x_n) : E^n \rightarrow E, i = 1, \dots, l$, начинающаяся с селекторных функций $e_i(x_1, \dots, x_n) = x_i, i = 1, \dots, n$ (т. е. попросту с переменных x_i), в которой каждая функция $f_i(x_1, \dots, x_n)$ вычисляется с помощью некоторой базисной операции $\omega_j(y_1, \dots, y_{n_j})$ и каких-то предыдущих функций $f_{k_r}, k_r < i, r = 1, \dots, n_j$:

$$f_i(x_1, \dots, x_n) = \omega_j(f_{k_1}(x_1, \dots, x_n), \dots, f_{k_{n_j}}(x_1, \dots, x_n)).$$

¹В силу его симметрии фактическое число сложений почти в два раза меньше.

Сложностью схемы называется число всех функций, встречающихся в ней (кроме селекторных функций, которые рассматриваются как *входы схемы*). Схема *реализует (вычисляет)* функцию $g(x_1, \dots, x_n)$, если эта функция встречается в последовательности $f_i(x_1, \dots, x_n)$, $i = 1, \dots, l$. Аналогичное определение дается и для вектор-функций. *Сложностью вычисления функции g (или вектор-функции $g_1, \dots, g_m$) схемами в заданном базисе B* называется минимальная сложность схемы в этом базисе, реализующей данную функцию (вектор-функцию).

Кроме аддитивного базиса B_+ , далее используются *конечный арифметический базис* $B_{+,*./} = \{x + y, x - y, xy, 1/x, 1\}$, *конечный полиномиальный базис* $B_{+,*} = \{x + y, xy, -1\}$, *континуальный линейный базис* $B_l = \{ax + by : a, b \in \mathbb{R}\}$ и его подмножества: *базис с вычитанием* $B_{+,-}$, *счетный линейный базис* $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$, *линейный базис с ограниченными константами* $B_{l,C} = \{x + y\} \cup \{ax : |a| \leq C\}$. В [8] было доказано, что сложность вычисления биномиального преобразования C_n в базисах $B_{l,C}, B_+, B_{+,-}$ равна¹ $\Theta(n^2)$. В [8] также было доказано, что сложность вычисления C_n в базисе B_l равна $O(n \log n)$, а в базисе $B_{+,*./}$ равна $O(n \log n \log \log n)$. Такая же оценка справедлива и для базиса $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$. Для сложности преобразования s_n (и монотонного преобразования Стирлинга 1-го рода) в базисах $B_{l,C}, B_{+,-}$, в [8] было доказано равенство $\Theta(n^2 \log_2 n)$. Для преобразования S_n в [9] (секция 7) показано, что сложность его вычисления в базисе $B_{+,*./}$ равна $O(n^{3/2}) \log^{3/2} n \log \log n$. Далее эта оценка будет улучшена.

2. Сложность вычисления преобразования Вандермонда и лагранжева интерполяция

Преобразованием Вандермонда назовем линейное преобразование с матрицей Вандермонда $V_n(a_0, \dots, a_{n-1}) = \|a_n^k\|$, где $a_0, \dots, a_{n-1}$ — заданные различные числа. В частности, матрицу $V_n(0, 1, \dots, n-1) = \|m^k\|$ обозначаем просто V_n (в ней $0^0 = 1$, поэтому в нулевом столбце стоят везде единицы, а в нулевой строке — везде нули, кроме единицы в левом верхнем углу). Известно, что

$$\det V_n(a_0, \dots, a_{n-1}) = \prod_{i>j} (a_i - a_j), \quad \det V_n = \prod_{k=1}^{n-1} k! = 2^{\Theta(n^2 \log n)}.$$

Если X_n — вектор-столбец, образованный переменными $x_1, \dots, x_n$, то вектор-столбец

$$(y_1, \dots, y_n)^T = Y_n = V_n(a_1, \dots, a_n) \cdot X_n$$

состоит из значений многочлена $f(t) = x_1 + x_2 t + \dots + x_n t^{n-1}$ при $t = a_i$, т. е. $y_i = f(a_i), i = 1, \dots, n$.

Теорема 1. Для базисов $B_+, B_{+,-}, B_{l,C}$ сложность вычисления преобразования V_n равна $\Theta(n^2 \log n)$. Для базиса B_l сложность вычисления преобразований $V_n(a_0, \dots, a_{n-1})$ и обратного к нему равна $O(n \log^2 n)$. Для базиса $B_{+,*}$ сложность вычисления преобразования $V_n(a_0, \dots, a_{n-1})$ при целых a_i равна $O(n \log^2 n \log \log n + n \log \max_i |a_i|)$. Для базиса $B_{+,*./}$ сложность вычисления преобразования $V_n(a_0, \dots, a_{n-1})$ и обратного к нему при целых a_i равна

¹Здесь и далее $a(n) = \Theta(b(n))$ означает равенство по порядку $a(n) \leq C_1 b(n)$, $b(n) \leq C_2 a(n)$, C_i не зависят от n .

$O(n \log^2 n \log \log n + n \log \max_i |a_i|)$. Для базиса $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ сложность вычисления преобразования $V_n(a_0, \dots, a_{n-1})$ и обратного к нему при целых a_i равна $O(n \log^2 n \log \log n)$.

Доказательство верхней оценки первого утверждения достаточно провести для базиса B_+ . Для этого каждое значение $y_i = f(i) = x_1 + x_2 i + \dots + x_n i^{n-1}$ вычисляем с помощью схемы Горнера, а умножение на i реализуем аддитивной цепочкой длины $O(\log i) = O(\log n)$. Доказательство нижней оценки проводится так же, как в [8] с помощью следующей леммы из [10] (доказательство этой леммы тоже можно найти в [8]).

Лемма 1 [10]. Для любой последовательности линейных форм $l_i(x_1, \dots, x_n)$, такой что $l_i = x_i$, $i = 1, \dots, n$, и для всякого $j > n$ либо $l_j = al_m$, $|a| \leq C$, $m < j$, либо $l_j = l_s + l_m$, $s, m < j$, рассмотрим $(L \times n)$ -матрицу A , строки которой образованы коэффициентами рассматриваемой последовательности линейных форм l_j , $j = 1, \dots, L$. Для любого $k \leq n$ обозначим Δ_k максимум модулей миноров порядка k (т.е. определителей подматриц порядка k) этой матрицы. Тогда $\Delta_k \leq \max(C, 2)^{L-n}$.

При применении леммы используем формулу для определителя $\det V_n = 2^{\Theta(n^2 \log n)}$.

Замечание 3. Для полноты приведем в сокращенном изложении доказательство и второго утверждения. Применяем метод, описанный, например, в [11, 12], в котором вычисление значений данного многочлена $f(t)$ в данных точках $a_0, \dots, a_{n-1}$ сводится к вычислению $f(t) \bmod (t - a_i)$, $i = 0, \dots, n-1$, а это вычисление проводится путем деления многочлена $f(t)$ с остатком на многочлены

$$\prod_{i < n/2} (t - a_i), \quad \prod_{i \geq n/2} (t - a_i),$$

потом деления полученных остатков соответственно на пары многочленов

$$\prod_{i < n/4} (t - a_i), \quad \prod_{n/2 > i \geq n/4} (t - a_i),$$

и

$$\prod_{3n/4 > i \geq n/2} (t - a_i), \quad \prod_{i \geq 3n/4} (t - a_i),$$

и так далее. Так как делители имеют фиксированные коэффициенты (т.е. не зависящие от $x_1, \dots, x_n$), то коэффициенты получающихся остатков от деления всегда выражаются через $x_1, \dots, x_n$ линейными функциями. Деление на многочлен с фиксированными коэффициентами сводится (как и в [11, 12]) к трехкратному умножению на заранее вычисленные многочлены с фиксированными коэффициентами (см. также, например, [13]), поэтому (как и в [11, 12]) сложность одновременного вычисления значений $f(a_i)$ оценивается $O(M(n) \log_2 n)$, где $M(n)$ — сложность умножения произвольного многочлена степени $n-1$ на фиксированный многочлен степени $n-1$ (в естественном предположении, что $M(2n) \geq 2M(n)$.) В случае базиса B_l для этого можно использовать метод [11, 12], основанный на применении быстрого преобразования Фурье [14], и получить оценку $O(n \log^2 n)$. В случае базисов $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ и $B_{+,*}$ для умножения многочленов применяется метод [15, 16] (см. также его изложение в [12]). Оценка для $M(n)$

в нем $M(n) = O(n \log n \log \log n)$. В случае базиса $B_{+,*}$ необходимо предварительно вычислить константы a_i со сложностью $O(n \log \max_i |a_i|)$, применяя аддитивные цепочки.

Для доказательства теоремы в отношении обратных преобразований заметим, что их вычисление равносильно задаче нахождения коэффициентов многочлена $f(t) = x_1 + x_2 t + \dots + x_n t^{n-1}$ по его значениям $y_i = f(a_i), i = 0, \dots, n-1$, в данных n точках. Эта задача есть в точности задача интерполяции по Лагранжу. Для ее решения известен алгоритм сложности $O(M(n) \log_2 n)$ (см. [11, 12]). Этот алгоритм также сводится к выполнению нескольких умножений произвольного многочлена на фиксированные, и для него получаем тем же путем те же оценки. Действительно, многочлен $f(t)$ можно интерполировать по формуле Лагранжа

$$f(t) = p(t) \sum_{i=0}^{n-1} \frac{f(a_i)/p'(a_i)}{t - a_i}, \quad p(t) = \prod_{i=0}^{n-1} (t - a_i).$$

Многочлен $p(t)$ вычисляется со сложностью $O(M(n)) \log_2 n$ с помощью умножения методом «деления пополам». Производная $p'(t)$ вычисляется со сложностью $O(n)$. Вычисление значений $f(a_i), p'(a_i)$ выполняется со сложностью $O(M(n)) \log_2 n$ методом, указанным выше. Далее складываются дроби методом «деления пополам». Числитель полученной суммы дробей равен многочлену $f(t)$. Сложность указанных вычислений равна $O(M(n)) \log_2 n$ (напомним, что во всех этих операциях вычисляются коэффициенты многочленов, а не их значения). Подробности можно найти также в [17]. Заметим, что для вычисления прямого преобразования $V_n(a_0, \dots, a_{n-1})$ при целых a_i достаточно использовать базис без деления $B_{+,*}$. Деление требуется только при вычислении обратного преобразования для нахождения $f(a_i)/p'(a_i)$. Для выполнения операций деления на многочлены с фиксированными коэффициентами числовое деление не требуется, так как старшие коэффициенты у упомянутых многочленов всегда равны 1. Для умножения многочленов методом Кантора–Калтофена числовое деление также не требуется (см., например, [18]).

3. Сложность вычисления преобразований Стирлинга

Далее понадобятся несколько лемм. Первая из них является матричной записью известного комбинаторного тождества.

Лемма 2. *Справедливо матричное равенство $V_n = C_n D_n S_n^T$, где D_n — диагональная матрица, в которой на i -м месте диагонали стоит $i!$ (нумерация строк и столбцов начинается с нуля, $0! = 1$), T означает операцию транспонирования.*

Доказательство. Выше указывалось известное тождество для чисел Стирлинга 2-го рода:

$$m^k = \sum_{i=0}^k i! S_i^k \binom{m}{i} = \sum_{i=0}^{n-1} i! S_i^k \binom{m}{i}, \quad m, k < n.$$

Так как в m -й строке матрицы C_n стоят числа $\binom{m}{i}, i = 0, \dots, n-1$, а в транспонированной матрице S_n^T в k -м столбце стоят числа $S_i^k, i = 0, \dots, n-1$, то в матрице $D_n S_n^T$ в k -м столбце стоят числа $S_i^k i!, i = 0, \dots, n-1$ (потому что при умножении

на диагональную матрицу слева i -я строка умножается на $i!$), значит, при умножении матриц C_n и $D_n S_n^T$ на пересечении m -й строки и k -го столбца результата будет стоять число

$$\sum_{i=0}^{n-1} i! S_i^k \binom{m}{i} = m^k.$$

Значит, получится матрица V_n .

Известна также простая, но полезная лемма.

Лемма 3. Для любых линейных преобразований $F_n \cdot X_n$ и $G_n \cdot X_n$ для их композиции $F_n \cdot G_n \cdot X_n$ для произвольного базиса B справедливо неравенство $L_B(F_n G_n) \leq L_B(F_n) + L_B(G_n)$, где $L_B(A_n)$ — сложность вычисления линейного преобразования $A_n \cdot X_n$.

Доказательство. Пусть S_F — схема в базисе B , вычисляющая F_n , и S_G — схема в базисе B , вычисляющая G_n , тогда композиция схем S_G и S_F , при которой выходы схемы S_G подаются на входы схемы S_F , вычисляет преобразование $F_n \cdot G_n$.

Следующая лемма многократно переоткрывалась¹.

Лемма 4 (Лемма о транспонировании). Для любых таких линейных преобразований $F_n \cdot X_n$, что в матрице F_n все строки и столбцы ненулевые, справедливо равенство $L_B(F_n) = L_B(F_n^T)$, где $L_B(A_n)$ — сложность вычисления линейного преобразования $A_n \cdot X_n$ в базисах $B = B_l, B_{l,C}, B_+, B_{+,-}$.

Доказательство леммы можно найти в [6, 7, 9].

Теорема 2. Для базиса $B_{l,C}$ сложность вычисления преобразований S_n, S_n^T, s_n, s_n^T равна $O(n^2 \log n)$. Для базиса B_l сложность вычисления преобразований S_n, S_n^T, s_n, s_n^T равна $O(n \log^2 n)$. Для базисов $B_{+,*}, /$ и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ сложность вычисления преобразований S_n, S_n^T, s_n, s_n^T равна $O(n \log^2 n \log \log n)$. Утверждения теоремы справедливы и для монотонного преобразования Стирлинга 1 рода.

Доказательство. Из леммы 2 следует, что $S_n^T = D_n^{-1} C_n^{-1} V_n$. Так как D_n^{-1} — диагональная матрица с числами $1/i!$ на диагонали, то для базисов $B_{+,*}, /$, B_l и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$, очевидно, $L_B(D_n^{-1}) = O(n)$. В [8] доказано, что $L_B(C_n^{-1}) = O(n \log n \log \log n)$ и даже $L_B(C_n^{-1}) = O(n \log n)$ при $B = B_l$. В теореме 1 доказано, что $L_B(V_n) = O(n \log^2 n \log \log n)$ для $B_{+,*}, /$ и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ и $L_B(V_n) = O(n \log^2 n)$ для B_l . Поэтому из леммы 3 следуют второе и третье утверждения теоремы для преобразования S_n^T . Согласно лемме 4 для линейных базисов B_l и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ справедливо равенство $L_B(S_n^T) = L_B(S_n)$.

К базису $B_{+,*}, /$ эту лемму непосредственно применить нельзя. Но можно заметить, что при вычислении преобразований C_n, C_n^{-1} в [8] и V_n, V_n^{-1} в теореме 1 кроме сложений-вычитаний выполнялись только умножения многочленов методом Кантора–Калтофена [12, 15], причем один из сомножителей всегда имел фиксированные коэффициенты, а так как алгоритм Кантора–Калтофена обладает свойством билинейности, то в нем, кроме сложений-вычитаний, выполнялись только скалярные умножения, т. е. умножения на некоторую константу. Поэтому схему для вычисления, например, преобразования S_n^T можно представить как композицию схемы,

¹В том числе и автором этой статьи в середине 80-х годов.

вычисляющей все необходимые константы (а они являются рациональными числами), и схемы в линейном базисе $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$, используя эти константы. Но последнюю схему можно, согласно лемме 4, преобразовать в схему той же сложности в том же базисе, вычисляющую преобразование S_n . Используемые для скалярных умножений константы при этом не изменятся, значит, для их вычисления можно использовать ту же схему, что и при вычислении S_n^T . Поэтому сложности схем для вычисления S_n^T и S_n равны. Аналогично доказывается равенство и для преобразований s_n^T и s_n .

Из матричного равенства $S_n^T = D_n^{-1}C_n^{-1}V_n$ следует равенство $s_n^T = (S_n^T)^{-1} = V_n^{-1}C_nD_n$. Из этого равенства так же, как и выше, следуют остальные утверждения теоремы для линейных базисов.

Из равенства $S_n^T = D_n^{-1}C_n^{-1}V_n$ для базиса $B_{l,C}$ следует, что

$$L_B(S_n) = L_B(S_n^T) \leq L_B(D_n^{-1}) + L_B(C_n^{-1}) + L_B(V_n) = O(n^2 \log n),$$

поскольку последнее слагаемое можно оценить также, как в теореме 1, второе слагаемое оценивается $O(n^2)$ в [8], а первое слагаемое, очевидно, равно $O(n)$.

Соотношения $L_{B_l}(s_n) = \Theta(n^2 \log n)$ доказаны в [8]. Там же доказано аналогичное соотношение для монотонного преобразования Стирлинга 1-го рода. Остальные утверждения для монотонного преобразования Стирлинга следуют из аналогичных утверждений для обычного преобразования Стирлинга 1-го рода с помощью замены переменных путем смены знака у некоторых из них. Аналогичные равенства для s_n^T следуют из предыдущих равенств и леммы 4. Теорема доказана.

Замечание 4. Соотношения $L_B(s_n) = \Theta(n^2 \log n)$ доказаны в [8] и при $B = B_{l,C}, B_{+,-}$. Там же доказаны аналогичные соотношения равенство для монотонного преобразования Стирлинга для базисов $B = B_{l,C}, B_{+,-}, B_+$. В силу леммы 4 такие же равенства верны и для транспонированных преобразований Стирлинга.

Из матричного равенства $D_n S_n^T = C_n^{-1}V_n$ для базисов $B_{l,C}, B_{+,-}$ следует, что

$$L_B(S_n D_n) = L_B(D_n S_n^T) \leq L_B(V_n) + O(n^2) = O(n^2 \log n),$$

а из матричного равенства $C_n D_n S_n^T = V_n$ для базисов $B_{l,C}, B_{+,-}, B_+$ следует, что¹

$$\Omega(n^2 \log n) \leq L_B(V_n) \leq L_B(S_n D_n) + O(n^2),$$

откуда следует, что $L_B(S_n D_n)$ и $L_B(V_n)$ для базиса $B_{l,C}$ имеют одинаковые асимптотики.

Одинаковые нижние оценки для базисов $B_{l,C}, B_{+,-}, B_+$ следуют также из леммы 2 и равенства $\det V_n = \det S_n D_n$, вытекающего из равенства $\det C_n = 1$ и формулы для умножения определителей. Поэтому и для базиса $B_{+,-}$ величины $L_B(S_n D_n)$ и $L_B(V_n)$ имеют одинаковые асимптотики.

Из теоремы 2 следует, что сложность вычисления последовательности чисел Белла $B_1, \dots, B_n$ в базисе $B_{+,*}/$ равна $O(n \log^2 n \log \log n)$. В [9] (секция 7) получена оценка $O(n \log n \log \log n)$.

Преобразование Лаха можно определить как композицию преобразований Стирлинга 2-го рода и монотонного преобразования Стирлинга 1-го рода (см., например, [3, 8]). Из теоремы 2 следует, что его сложность в базисе B_l оценивается как

¹Здесь и далее запись $\Omega(f(n)) < g(n)$ означает при $f(n) > 0, g(n) > 0$, как обычно, что $Cf(n) < g(n)$ при некотором > 0 , не зависящем от n .

$O(n \log^2 n)$, а в базисах $B_{+,*,/}$ и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ — как $O(n \log^2 n \log \log n)$. Но в [8] для преобразования Лаха получены лучшие оценки: $O(n \log n)$ и $O(n \log n \log \log n)$. Известно (см. [3]), что последовательность многочленов Чебышёва–Лагерра выражается через коэффициенты Лаха:

$$l_n(x) = x e^x \left(\frac{d}{dx} \right)^n (e^{-x} x^{n-1}) = \sum_{k=0}^n (-1)^k L_k^n x^k, \quad L_k^n = \binom{n-1}{k-1} \frac{n!}{k!},$$

а значит, и через преобразование Лаха:

$$(l_0(x), \dots, l_n(x))^T = L_n \times (1, -x, x^2, \dots, (-1)^n x^n)^T, \quad L_n = \|L_m^n\|.$$

Поэтому из полученной в [8] оценки следует, что последовательность $l_0(x), \dots, l_n(x)$ вычисляется в базисе $B_{+,*,/}$ со сложностью $O(n \log n \log \log n)$, а в базисе $\{x + y, xy, 1/x\} \cup \mathbb{R}$ — со сложностью $O(n \log n)$.

Замечание 5. Для сложности вычисления в базисе $\{x + y, xy, 1/x, e^x\} \cup \mathbb{R}$ последовательности многочленов

$$x e^x \left(\frac{d}{dx} \right)^k (e^{-x} x^{n-1}), \quad k = 0, \dots, n,$$

последний член которой совпадает с многочленом $l_n(x)$, из теоремы 2 [9] легко вывести оценку $O(n \log^2 n)$. Действительно, сложность вычисления $e^{-x} x^{n-1}$ равна $O(\log n)$. Теорема 2 [9] применима, конечно, и в более общем случае.

4. Сложность ньютоновой интерполяции и еще одного преобразования Вандермонда

Известно, что интерполяция по Ньютону также решает задачу восстановления многочлена $f(t) = x_1 + x_2 t + \dots + x_n t^{n-1}$ по его значениям $y_i = f(a_i), i = 0, \dots, n-1$, в данных n точках, если многочлен $f(t)$ представляется в виде

$$c_0 + c_1(x - a_0) + c_2(x - a_0)(x - a_1) + \dots + c_{n-1}(x - a_0)(x - a_1) \dots (x - a_{n-1}).$$

Для вычисления коэффициентов c_i при этом рекурсивно применяются формулы:

$$F_0(a_i) = f(a_i), \quad F_{k+1}(a_i) = \frac{F_k(a_i) - F_k(a_k)}{a_i - a_k}, \quad i \neq k, \quad c_i = F_i(a_i), \quad i = 0, \dots, n-1.$$

Сложность вычисления в базисе $\{x - y, x/y\} \cup \mathbb{R}$ по этим формулам равна $3n(n+1)/2$. Далее будет показано, что в частных случаях $a_i = i$ и $a_i = q^i, q > 1$, эти вычисления можно выполнить быстрее. В случае $a_i = i$ имеем дело с представлением многочленов степени n в виде

$$\sum_{k=0}^n c_k(x)_k, \quad (x)_k = x(x-1) \dots (x-k+1),$$

а в случае $a_i = q^i$ — в виде

$$\sum_{k=0}^n c_k g_{k,q}(x), \quad g_{k,q}(x) = (x-1)(x-q) \dots (x-q^{k-1}).$$

Рассмотрим еще одну матрицу типа Вандермонда $W_n(a_0, \dots, a_{n-1}) = \|(a_m)_k\|$, где $a_0, \dots, a_{n-1}$ — заданные различные числа и

$$(t)_k = t(t-1)\dots(t-k+1), (t)_0 = 1.$$

Еще определим матрицу W_n равенством $W_n = W_n(0, 1, \dots, n-1) = \|(m)_k\|$. Известно, что

$$\det W_n(a_0, \dots, a_{n-1}) = \det V_n(a_0, \dots, a_{n-1}) = \prod_{i>j} (a_i - a_j),$$

$$\det W_n = \det V_n = \prod_{k=1}^{n-1} k! = 2^{\Theta(n^2 \log n)}.$$

Задача вычисления значений многочлена $f(t) = \sum_{k=0}^{n-1} x_k(t)_k$ в произвольных данных точках a_i , $i = 0, \dots, n-1$, равносильна вычислению линейного преобразования

$$(y_0, \dots, y_{n-1})^T = W_n(a_0, \dots, a_{n-1}) \cdot (x_0, \dots, x_{n-1})^T.$$

В частности, при $a_i = i$, $i = 0, \dots, n-1$, она сводится к вычислению преобразования $W_n \cdot X_n$. Вычисление преобразования $W_n(a_0, \dots, a_{n-1})$ сводится к вычислению преобразования $V_n(a_0, \dots, a_{n-1})$, если преобразовать многочлен $f(t) = \sum_{k=0}^{n-1} x_k(t)_k$ к стандартному виду $f(t) = \sum_{k=0}^{n-1} z_k t^k$. Подставляя вместо $(t)_k$ выражения $\sum_{m=0}^k s_m^k t^m$ в сумму $\sum_{k=0}^{n-1} x_k(t)_k$ и приводя подобные члены, получаем, что

$$z_m = \sum_{k=0}^m x_k s_m^k, m = 0, \dots, n-1,$$

т. е. коэффициенты z_i вычисляются по коэффициентам x_i с помощью преобразования s_n^T , транспонированного к преобразованию Стирлинга 1-го рода, откуда следует матричное равенство $W_n(a_0, \dots, a_{n-1}) = V_n(a_0, \dots, a_{n-1}) \cdot s_n^T$, а из него — равенства $W_n(a_0, \dots, a_{n-1})^{-1} = s_n^T \cdot V_n(a_0, \dots, a_{n-1})^{-1}$, $W_n = V_n \cdot s_n^T$, $W_n^{-1} = s_n^T \cdot V_n^{-1}$, $V_n = W_n \cdot s_n^T$. Из последнего равенства и леммы 2 следует равенство $W_n = C_n D_n$, где D_n — определенная в ней диагональная матрица.

Теорема 3. Для базиса B_l сложность вычисления преобразования $W_n(a_0, \dots, a_{n-1})$ и обратного к нему равна $O(n \log^2 n)$. Для базиса $B_{+,*}$ сложность вычисления преобразования $W_n(a_0, \dots, a_{n-1})$ и обратного к нему при целых a_i равна $O(n \log^2 n \log \log n + n \log \max_i |a_i|)$. Для базиса $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ сложность вычисления преобразования $W_n(a_0, \dots, a_{n-1})$ и обратного к нему при целых a_i равна $O(n \log^2 n \log \log n)$. Для базисов $B_+, B_{+,-}, B_{l,C}$ сложность вычисления преобразования W_n равна $\Theta(n^2 \log n)$. Для базиса B_l сложность вычисления преобразования W_n и обратного к нему равна $O(n \log n)$. Для базисов $B_{+,*}$ и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$ сложность вычисления преобразования W_n и обратного к нему равна $O(n \log n \log \log n)$.

Для доказательства оценок для преобразования $W_n(a_1, \dots, a_n)$ и обратного к нему достаточно воспользоваться указанными выше матричными равенствами, теоремами 1, 2 и леммой 3.

Докажем оценки сложности для преобразования W_n . Доказательство верхней оценки первого из этих утверждений достаточно провести для базиса B_+ . Для этого каждое значение $y_i = f(i) = x_0 + x_1 i + \dots + x_{n-1} (i)_{n-1}$ вычисляем с помощью схемы Горнера, а умножение на i реализуем аддитивной цепочкой длины

$O(\log i) = O(\log n)$. Доказательство нижней оценки проводится так же, как в [8] с помощью леммы 1. Применяя выведенное выше из леммы 2 равенство $W_n = C_n D_n$, лемму 3 и теорему 2 [8], заканчиваем доказательство теоремы 3.

Замечание 6. Из теоремы 3 следует, что ньютонова интерполяция при выборе равномерных узлов $h, 2h, \dots, nh$ выполняется со сложностью $O(n \log n)$ в базисе $\{x \pm y, xy, 1/x\} \cup \mathbb{R}$, так как в этом базисе выражается базис B_l , для которого это утверждение о сложности верно. При $h = 1$ целочисленная интерполяция выполняется со сложностью $O(n \log n \log \log n)$ в базисе $B_{+,*./}$.

Вычисление значений многочлена, заданного в ньютоновой форме

$$f(x) = \sum_{k=0}^n c_k x(x-1) \dots (x-k+1),$$

в произвольных целых узлах a_i , $i = 0, \dots, n$, а также восстановление его коэффициентов c_k по его значениям в данных целых узлах a_i выполняется со сложностью $O(n \log^2 n \log \log n + n \log \max_i |a_i|)$ в базисе $B_{+,*./}$.

В случае произвольных узлов a_i и базисов B_l и $\{x \pm y, xy, 1/x\} \cup \mathbb{R}$ указанная выше задача решается со сложностью $O(n \log^2 n)$.

В реальных вычислениях нет смысла сводить задачу к вычислению биномиального преобразования, а нужно сразу свести ее к вычислению свертки, т. е. к умножению многочленов.

Для полноты приведем еще одну теорему.

Теорема 4. Сложность ньютоновой интерполяции с узлами $1, q, \dots, q^n$ равна $O(n \log n)$ в базисе B_l . Сложность вычисления значений многочлена степени n , представленного в ньютоновой форме

$$f(x) = \sum_{k=0}^n c_k g_{k,q}(x), \quad g_{k,q}(x) = (x-1)(x-q) \dots (x-q^{k-1}),$$

в узлах $1, q, \dots, q^n$ равна $O(n \log n)$ в базисе B_l . Для базиса $B_{+,*./}$ и натурального q в обоих случаях оценки заменяются на $O(n \log n \log \log n + \log q)$. Те же оценки справедливы и в случае лагранжевой интерполяции в узлах $1, q, \dots, q^n$.

Доказательство. Для вычисления значений многочлена $f(x) = \sum_{k=0}^n c_k g_{k,q}(x)$ в узлах $1, q, \dots, q^n$ воспользуемся формулами

$$\begin{aligned} f(q^k) &= y_k = \sum_{m=0}^k c_m g_{m,q}(q^k) = \sum_{m=0}^k c_m (q^k - 1)(q^k - q) \dots (q^k - q^{m-1}) = \\ &= \sum_{m=0}^k c_m q^{\binom{m}{2}} (q^k - 1)(q^{k-1} - 1) \dots (q^{k-m+1} - 1) = \\ &= \sum_{m=0}^k c_m q^{\binom{m}{2}} (q - 1) \dots (q^m - 1) \binom{k}{m}_q, \quad k = 0, \dots, n, \end{aligned}$$

где

$$\binom{k}{m}_q = \frac{(q^k - 1)(q^{k-1} - 1) \dots (q^{k-m+1} - 1)}{(q - 1) \dots (q^m - 1)}$$

— q -биномиальный коэффициент Гаусса. Вычислив числа $c_m q^{\binom{m}{2}}$, $m = 0, \dots, n$ (это можно сделать со сложностью $O(n + \log q)$ по формулам $q, q^2, \dots, q^n, q_{i+1} = q^{\binom{i+1}{2}} = q_i q^i$), сводим задачу к вычислению преобразования Гаусса; применяя теорему 8 [8], получаем оценки сложности $O(n \log n)$ и $O(n \log n \log \log n) + \log q$ для базисов B_l и $B_{+,*,/}$ соответственно (в последнем случае предполагаем, что q — натуральное число). Аналогично оценивается и сложность обратного преобразования.

Если многочлен задан в стандартной форме $f(x)$, то для преобразования его в ньютонову форму с узлами $1, q, \dots, q^n$ воспользуемся формулами

$$f(x) = \sum_{k=0}^n a_k x^k = \sum_{k=0}^n a_k \sum_{m=0}^k \binom{k}{m}_q g_{m,q}(x).$$

Приводя подобные члены, получим, что

$$f(x) = \sum_{k=0}^n c_k g_{k,q}(x), \quad c_k = \sum_{k=0}^m a_k \binom{m}{k}_q, \quad k = 0, \dots, n,$$

т. е. ньютоновы коэффициенты c_k можно вычислить по обычным коэффициентам a_k с помощью транспонированного преобразования Гаусса G_n^T . Используя теорему 8 [8] и рассуждения, аналогичные доказательству теоремы 2, а также доказанное выше утверждение о сложности вычисления многочлена в ньютоновой форме в узлах q^k , получаем что сложность вычисления многочлена в стандартной форме в тех же узлах будет такой, как указано в теореме. Для решения обратной задачи, т. е. лагранжевой интерполяции в узлах $1, q, \dots, q^n$, аналогичным способом сведем ее к ньютоновой интерполяции, для которой нужные оценки уже получены. Для преобразования ньютоновых коэффициентов в стандартные используем обратное к указанному выше преобразованию. Очевидно, оно будет транспонированным к обратному гауссову преобразованию, т. е. равно $(G_n^{-1})^T$. Оценка для него получается аналогично данной выше.

Разумеется, при реальном выполнении вычислений нет смысла сводить задачу к вычислению гауссова преобразования, а нужно сразу свести ее к вычислению свертки, т. е. к умножению многочленов. Возможность лагранжевой интерполяции в узлах $1, q, \dots, q^n$ со сложностью $O(n \log n)$ отмечалась в [11] (утверждение было сформулировано в виде задачи и, вероятно, ее решение предполагалось получить с помощью преобразования Блустейна).

Замечание 7. Каждому линейному преобразованию $A_n \cdot X_n$ можно сопоставить последовательность многочленов $a_m(x) = \sum_k a_{m,k} x^k$. Если базис B содержит умножение, то $L_B(a_0(x), \dots, a_{n-1}(x)) \leq L_B(A_n) + n$. В случае рассмотренных выше преобразований иногда указанную оценку можно улучшить. Например, очевидно, что если $A_n = C_n, C_n^{-1}, s_n$ то для базиса $B_{+,*,/}$ сложность соответствующей системы многочленов равна $\Theta(n)$. Если $A_n = S_n$, то соответствующие многочлены $\sum_{k=0}^n S_k^n x^k$ называются в [3] экспоненциальными и в [9] (раздел 7) для сложности этой системы многочленов получена оценка $O(n \log n \log \log n)$ в случае $B_{+,*,/}$.

Если $A_n = G_n^{-1}$, где $G_n = \|\binom{k}{m}_q\|$ — q -преобразование Гаусса, то согласно следствию q -биномиальной формулы

$$\sum_{m=0}^k (-1)^{k-m} q^{\binom{k-m}{2}} \binom{k}{m}_q x^m = g_{q,m}(x) = (x-1)(x-q) \dots (x-q^{m-1}),$$

откуда видно, что при фиксированном натуральном q сложность вычисления указанной системы многочленов равна $\Theta(n)$ в базисе $B_{+,*./}$. Для преобразования $V_n(a_0, \dots, a_{n-1})$ при целых a_i можно показать, что для базиса $B_{+,*}$ сложность соответствующей системы многочленов равна $O(n \log(n \max_i |a_i|))$, так как сложность вычисления многочлена $1 + x + \dots + x^n$ равна $\Theta(\log^2 n)$.

Для преобразования Лаха L_n соответствующая последовательность многочленов есть $l_k(-x)$, $k = 0, \dots, n-1$, где $l_n(x)$ — многочлены Чебышева-Лагеррера. Сложность вычисления этой системы равна $\Theta(n)$.

После того, как этот текст был написан, автор узнал, что утверждения теорем 3 и 4, касающиеся базиса $B_{+,*./}$ (и фактически базисов B_l и $\{x \pm y\} \cup \{nx : n \in \mathbb{N}\} \cup \{x/n : n \in \mathbb{N}\}$) содержатся, среди прочего, в статье [19]. Доказательства в ней похожи на приведенные выше, но некоторые из них сложнее, так как в этих случаях в [19] получались оценки с более точными мультипликативными константами. Недавно появилась статья [20], в которой для сложности умножения многочленов получена оценка $M(n) < n \log n \Psi(n)$, где $\Psi(n)$ — чрезвычайно медленно растущая последовательность (медленнее любой итерации логарифма). Благодаря этому результату все оценки вида $O(n \log n \log \log n)$ в нашей статье можно заменить на $O(n \log n \Psi(n))$, а оценки вида $O(n \log^2 n \log \log n)$ — на $O(n \log^2 n \Psi(n))$.

Автор благодарен И. С. Сергееву за информацию об обеих этих публикациях.

Список литературы

1. Риордан Дж., *Введение в комбинаторный анализ*, ИЛ, Москва, 1963.
2. Гульден Я. Джексон Д., *Перечислительная комбинаторика*, Наука, Москва, 1990.
3. Айгнер М., *Комбинаторная теория*, Мир, Москва, 1982.
4. Риордан Дж., *Комбинаторные тождества*, Наука, Москва, 1982.
5. Стефенсен И. П., *Теория интерполяции*, ОНТИ, М., 1935.
6. Кнут Д., *Искусство программирования*, т. 2, Вильямс, Киев, Москва, С.Петербург, 2000.
7. Gashkov S., Kochergin V., "On addition chains of vectors, gate circuits, and the complexity of computation of power", *Siberian Adv. in Math.*, **4:4** (1994), 1–16.
8. Гашков С.Б., "Об арифметической сложности вычисления некоторых линейных преобразований", *Вестн. Моск. ун-та. Серия 1. Математика. Механика*, 2014.
9. Гашков С.Б., Гашков И.Б., "О сложности вычисления дифференциалов и градиентов", *Дискретная математика*, **17:2** (2005), 45–67.
10. Morgenstern J., "Note on lower bound of the linear complexity of the fast Fourier transform", *J.ACM*, **20:2** (1973), 305–306.
11. Ахо А., Хопкрофт Дж., Ульман Дж., *Проектирование и анализ вычислительных алгоритмов*, Мир, Москва, 1979.
12. von zur Gathen J., Gerhard J., *Modern Computer Algebra*, Cambridge Univ. Press, Cambridge, 1999.
13. Гашков С.Б., Чубариков В.Н., *Арифметика. Алгоритмы. Сложность вычислений*, Высшая школа, М., 2000.
14. Cooley J., Tukey J., "An algorithm for the machine calculation of complex Fourier series", *Math. Comp.*, **19** (1965), 297–301.
15. Cantor D., Kaltofen E., "On fast multiplication of polynomials over arbitrary algebras", *Acta Informatica*, **28** (1991), 693–701.
16. Schönhage A., "Schnelle multiplikation von polynomen über körpern der charakteristik 2", *Acta Informatica*, **7** (1977), 395–398.

17. Гашков С.Б., “О сложности интегрирования рациональных дробей”, *Труды Матем. института им. В. А. Стеклова*, **218** (1997), 122–133.
18. Гашков С.Б., Сергеев И.С., “О сложности и глубине булевых схем для умножения и инвертирования в конечных полях характеристики два”, *Дискретная математика*, **25(1)** (2013), 3–32.
19. Bostan A., Schost E., “Polynomial evaluation and interpolation on special sets of points”, *J. Complexity*, **21(4)** (2005), 420–446.
20. Harvey D., van der Hoeven J., Lecerf G., “Faster polynomial multiplication over finite fields”, 12 Jul 2014, [arXiv.org: 1407.3361](https://arxiv.org/abs/1407.3361).

Статья поступила 01.06.2014.