Тема 4.3. Метод Фюрера

С. Б. Гашков, И. С. Сергеев

Метод Фюрера, предложенный в 2007 г., позволяет достичь наилучшей известной сегодня оценки сложности умножения n-разрядных чисел $nc^{\log^* n}\log n$, где $\log^* n$ — сверхлогарифм, очень медленно растущая функция, определяемая из соотношения:

$$\lfloor \underbrace{\log \ldots \log}_{\log^* n} n \rfloor = 1.$$

1 Выбор подходящего кольца

Рассмотрим кольцо $\mathbb{C}[x]/(x^{2^p}+1)$. В нем элемент x является примитивным корнем степени 2^{p+1} из единицы (что проверяется прямо по определению: многочлены $x^{2^p}-1$ и $x^{2^p}+1$ взаимно просты). Обозначим $\zeta = e^{\mathbf{i}\pi/2^p}$. Заметим, что

$$x^{2^p}+1 = \prod_{k=0}^{2^p-1}(x-\zeta^{2k+1}).$$

Для $q \in \mathbb{N}$ положим $\xi = e^{\mathbf{i}\pi/(2^{p+q})}$. Определим многочлен $\rho_q(x)$ степени меньше 2^p следующим образом:

$$\rho_q(\zeta^{2k+1}) = \xi^{2k+1}, \quad k = 0, \ldots, 2^p-1.$$

Лемма 1. *Многочлен $\rho_q(x)$ является примитивным корнем степени 2^{p+q+1} в $\mathbb{C}[x]/(x^{2^p}+1)$.*

Доказательство. По построению, $\rho_q^{2^q}(x) = x$ в кольце $\mathbb{C}[x]/(x^{2^p}+1)$. Действительно, при любом k

$$\rho_q^{2^q}(x) \bmod (x-\zeta^{2k+1}) = \rho_q^{2^q}(\zeta^{2k+1}) = \xi^{(2k+1)2^q} = \zeta^{2k+1}.$$

Следовательно, $\rho_q^{2^q}(x) \equiv x \mod (x^{2^p}+1)$. Значит, $\rho_q(x)$ является корнем степени 2^{p+q+1} из единицы. Примитивность следует из сравнения

$$\rho_q^{2^{p+q}}(x) - 1 \equiv x^{2^p} - 1 \equiv -2 \mod (x^{2^p}+1).$$

Лемма 2. *Модули коэффициентов многочлена* $\rho_q^m(x) \bmod (x^{2^p}+1)$ *не превосходят* 1.

Доказательство. Пусть $(\rho_q^m(x) \bmod (x^{2^p}+1)) = \sum_{k=0}^{2^p-1} r_k x^k$. Подставляя в равенство ζx вместо x, получим $(\rho_q^m(\zeta x) \bmod (x^{2^p}-1)) = \sum_{k=0}^{2^p-1} r_k \zeta^k x^k$. По определению многочлена $\rho_q(x)$:

$$\text{ДПФ}_{2^p,\zeta^2}\left(r_0\zeta^0, r_1\zeta^1, \ldots, r_{2^p-1}\zeta^{2^p-1}\right) = \left(\xi^m, \xi^{3m}, \ldots, \xi^{(2(2^p-1)+1)m}\right).$$

Тогда, применяя формулу для обратного ДПФ, искомые коэффициенты r_k могут быть вычислены как

$$r_k = \zeta^{-k} 2^{-p} \sum_{l=0}^{2^p-1} \xi^{m(2l+1)} \zeta^{-2kl}.$$

Очевидно, выражение в правой части по модулю не превосходит 1.

2 ДПФ в кольце $\mathbb{C}[x]/(x^{2^p}+1)$

Пусть $\omega = \rho_{(s-1)(p+1)}(x)$ — примитивный корень степени $2^{s(p+1)}$ в кольце $\mathbb{C}[x]/(x^{2^p}+1)$. ДПФ порядка $N = 2^{s(p+1)}$, построенное методом Кули-Тьюки с разложением порядка на множители $2^{(s-1)(p+1)}2^{p+1}$, имеет следующую структуру:

– $2^{(s-1)(p+1)}$ параллельно выполняемых преобразований $\text{ДПФ}_{2^{p+1},x}$;

– N параллельных умножений результатов предыдущего шага на степени примитивного корня ω;

– 2^{p+1} применяемых к векторам, вычисленным на предыдущем шаге, и параллельно выполняемых ДПФ порядка $2^{(s-1)(p+1)}$.

Расщепляя внешние ДПФ аналогичным образом, получаем окончательно, что при вычислении ДПФ порядка N чередуются s блоков, в которых параллельно выполняются $N/2^{p+1}$ ДПФ порядка 2^{p+1} с примитивным корнем x, и $s-1$ блоков, в которых выполняются покомпонентные умножения текущего вектора на степени ω.

При этом реализация ДПФ порядка 2^{p+1} состоит в выполнении операций сложения-вычитания и умножения на степени x в кольце — последние сводятся к циклическому сдвигу комплексных коэффициентов, если элементы кольца $\mathbb{C}[x]/(x^{2^p}+1)$ записываются многочленами по модулю $x^{2^{p+1}}-1$.

Более точно (опять пользуемся методом Кули—Тьюки), ДПФ порядка 2^{p+1} состоит из $p+1$ слоев, в которых параллельно выполняются ДПФ порядка 2, а между слоями выполняются покомпонентные умножения на степени x. ДПФ порядка 2 — это просто сложение и вычитание: $\text{ДПФ}_{2,\,x^{2^p}}(\gamma_0,\gamma_1)=(\gamma_0+\gamma_1,\,\gamma_0-\gamma_1)$.

Пусть действительные и мнимые части комплексных коэффициентов кодируются E двоичными разрядами до и после запятой (плюс один разряд под знак числа), а операции с ними производятся с погрешностью (абсолютной величиной ошибки) 2^{1-E} и с отбрасыванием разрядов старше $E-$го. Тогда сложность каждой из операций сложения-вычитания, умножения на степень x в кольце $\mathbb{C}[x]/(x^{2^p}+1)$ составляет $O(2^pE)$.

Тогда сложность умножения многочленов из кольца $(\mathbb{C}[x]/(x^{2^p}+1))[y]$ по модулю y^N-1 можно оценить как

$$3s(N/2^{p+1})(p+1)2^{p+1}O(2^pE)+(3s-2)NL^*\le O(2^pEN\log N)+3sNL^*,\ (1)$$

где через L^* обозначена сложность умножения в кольце $\mathbb{C}[x]/(x^{2^p}+1)$ с погрешностью 2^{1-E} (под погрешностью понимается максимальная ошибка в каждом коэффициенте многочлена).

Подчеркнем, что все вычисления являются приближенными, в частности, константы ω^m даны с точностью E знаков после запятой. В силу накопления погрешности (абсолютной величины ошибки) по ходу вычислений точность, с которой вычисляются коэффициенты произведения, вообще говоря, существенно меньше, чем точность исходных коэффициентов. Оценим сопутствующее алгоритму умножения изменение погрешности, считая, что не происходит переполнения в старших разрядах.

3 Точность вычислений

Далее с каждым многочленом a, возникающим в процессе вычислений, мы будем ассоциировать величину e_a, которая является верхней оценкой абсолютной величины погрешности вычисления a. Эту оценку будем называть просто погрешностью.

Погрешность $e_{a\pm b}$ результата $a \pm b$ выполнения любой из операций сложения или вычитания комплексных чисел a и b, данных с погрешностями e_a и e_b можно положить равной $e_a + e_b$, считая, что используются точные алгоритмы сложения и вычитания ($2E$-разрядных чисел).

Рассмотрим операцию умножения элементов $u, v \in \mathbb{C}[x]/(x^{2^p} + 1)$ (каждый из соответствующих многочленов имеет степень не выше $2^{p+1} - 1$). Пусть абсолютные величины коэффициентов этих многочленов ограничены числами U и V соответственно, а погрешности — соответственно e_u и e_v. Будем полагать, что погрешность алгоритма умножения в кольце $\mathbb{C}[x]/(x^{2^p} + 1)$ не превосходит 2^{1-E}.

Поскольку каждый из коэффициентов произведения uv является суммой 2^{p+1} попарных произведений коэффициентов многочленов-сомножителей, то его погрешность можно оценить как

$$e_{uv} = 2^{p+1}(e_v U + e_u V + e_v e_u + 2^{1-E}).$$

Если $U = V = \Omega(2^{-E})$ и $1 \geq e_v = e_u = \Omega(2^{-E})$, то положим $e_{uv} = c_1 2^{p+1} U e_u$ при некотором $c_1 \geq 1$.

Особо рассмотрим случай, когда многочлен v совпадает с ω^m. Согласно лемме 2, $V \leq 1$. Т.к. v — постоянный многочлен, то положим $e_v = 2^{-E}$. При выполнении условий $e_u = \Omega(2^{-E})$ и $e_u/U \geq 2^{-E}$ можно положить $e_{uv} = c_2 2^p e_u$ при некотором $c_2 \geq 1$.

В обоих случаях для каждого отдельного коэффициента a произведения uv положим $e_a = e_{uv}$.

Обе указанные оценки для погрешности зависят от величины коэффициентов многочленов-сомножителей. Рассмотрим, как изменяется величина комплексных коэффициентов при вычислении ДПФ.

Пусть комплексные коэффициенты компонент γ_j исходного вектора не превосходят A по абсолютной величине. Из предыдущего раздела следует, что под действием ДПФ порядка 2^{p+1} максимум абсолютной величины коэффициентов увеличивается не более, чем в 2^{p+1} раз. Согласно лемме 2, при умножении на ω^m этот максимум увеличивается не более, чем в 2^p раз. Таким образом, для максимума абсолютной величины коэффициентов компонент γ_j^* вектора значений ДПФ справедлива оценка $2^{s(p+1)} 2^{(s-1)p} A < N^2 A$.

Оценим величину накопленной погрешности. Пусть она не превосходит e_γ для коэффициентов компонент γ_j. Тогда коэффициент увеличения погрешности при выполнении внутреннего ДПФ порядка 2^{p+1} можно

оценить как 2^{p+1}, а коэффициент увеличения при умножении на ω^m — как $c_2 2^p$, если выполняется условие $e_u/U \geq 2^{-E}$. Таким образом, для максимума e^*_γ погрешности коэффициентов компонент γ^*_j получаем оценку $e^*_\gamma \leq N(c_2 2^p)^s < c_2^s N^2$.

Условие $e_u/U \geq 2^{-E}$ в алгоритме выполняется при всех умножениях на степени ω, если изначально выполнено $e_\gamma/A \geq 2^{-E}$, т.к. в случае каждой элементарной операции (сложение, вычитание, умножение на ω^m) отношение e_u/U не убывает.

4 Умножение в $\mathbb{C}[x]/(x^{2^p}+1)$

Осталось указать как умножаются многочлены из $\mathbb{C}[x]/(x^{2^p}+1)$ с погрешностью 2^{1-E}. Поле $\mathbb{C}$ изоморфно фактор-кольцу $\mathbb{R}[\mathbf{i}]/(\mathbf{i}^2+1)$ (элементы которого представляются вещественными многочленами степени 1). Многочлены из $\mathbb{C}[x]/(x^{2^p}+1) \cong \mathbb{R}[\mathbf{i},x]/(\mathbf{i}^2+1,\, x^{2^p}+1)$ могут быть перемножены как обычные многочлены переменных $\mathbf{i}$ и x с последующим приведением по модулям.

Умножим коэффициенты перемножаемых многочленов на 2^E (коэффициенты, таким образом, становятся целыми числами). Запишем эти многочлены как

$$a(\mathbf{i},x) = \sum_{j=0}^{1}\sum_{l=0}^{2^{p+1}-1} a_{j,l}\mathbf{i}^j x^l, \qquad b(\mathbf{i},x) = \sum_{j=0}^{1}\sum_{l=0}^{2^{p+1}-1} b_{j,l}\mathbf{i}^j x^l.$$

Коэффициенты произведения исходных многочленов получаются делением коэффициентов произведения ab на 2^{2E}. Обозначим

$$h(\mathbf{i},x) = \sum_{j=0}^{1}\sum_{l=0}^{2^{p+1}-1} 2^{2E}\mathbf{i}^j x^l, \quad \begin{array}{l} a'(\mathbf{i},x) = a(\mathbf{i},x) + h(\mathbf{i},x), \\ b'(\mathbf{i},x) = b(\mathbf{i},x) + h(\mathbf{i},x). \end{array}$$

Заметим, что коэффициенты многочленов a', b', h положительны. При этом $ab = a'b' - a'h - b'h - h^2$. Тем самым задача сведена к умножению многочленов с неотрицательными целочисленными $(2E+1)-$разрядными коэффициентами.

Такое умножение сводится к умножению $2^{p+1}3(4E+p+4)-$разрядных чисел при подстановке $2^{4E+p+4} \to \mathbf{i}$ и $2^{3(4E+p+4)} \to x$. Окончательно имеем, что сложность умножения в $\mathbb{C}[x]/(x^{2^p}+1)$ по порядку не превосходит

сложности умножения $O(2^p(E+p))-$разрядных чисел (в этой же оценке учитывается сложность $O(2^pE)$ дополнительных сложений и вычитаний при вычислении a', b', восстановлении ab и приведении по модулям $\mathbf{i}^2+1$, $x^{2^p}+1$). Алгоритм умножения с погрешностью 2^{1-E} получается отбрасыванием разрядов коэффициентов результата младше $E-$го после запятой.

5 Оценка сложности

В методе Фюрера используется выбор параметров $E=\Theta(\log N)$ и $p=\log_2\log_2 N+O(1)$. Перемножаемые числа представляются многочленами над $\mathbb{C}[x]/(x^{2^p}+1)$ степени меньше $N/2$ следующим образом: они разбиваются на блоки длины $E/3$, которые интерпретируются как целые числа, каждые 2^{p-1} подряд расположенных блоков интерпретируются как коэффициенты многочлена из кольца $\mathbb{C}[x]/(x^{2^p}+1)$.

Получившиеся многочлены (переменной y) перемножаются при помощи ДПФ порядка N, описанного выше. Число из многочлена восстанавливается за $O(2^pEN)$ операций при подстановке $x=2^{E/3}$, $y=2^{2^{p-1}E/3}$ и приведении подобных.

Параметр E выбирается, исходя из следующих рассуждений. В обозначениях из предыдущего пункта при старте алгоритма $A\le 2^{E/3}$, поэтому положим $e_\gamma=A2^{-E}$. Коэффициенты компонент вектора значений ДПФ имеют абсолютные величины не более N^2A и погрешности не более $c_2^sN^2e_\gamma$.

Коэффициенты произведения двух элементов из $\mathbb{C}[x]/(x^{2^p}+1)$, абсолютные величины коэффициентов которых не превосходят U, ограничены величиной $2^{p+1}U^2$. Погрешность после их перемножения возрастает не более, чем в $c_1 2^{p+1}U$ раз (см. предыдущий пункт). Поэтому для максимума абсолютных величин коэффициентов W и их погрешностей e_W после выполнения покомпонентных умножений векторов значений ДПФ имеем оценки

$$W\le 2^{p+1}U^2\le 2^{p+1}N^4A^2,$$
$$e_W=c_1c_2^s2^{p+1}N^2Ue_\gamma\le c_1c_2^s2^{p+1}N^4Ae_\gamma.$$

Заметим, что $e_W/W\ge c_1c_2^sN^2e_\gamma/U\ge c_1c_2^se_\gamma/A\ge 2^{-E}$.

Обратное ДПФ (с точки зрения оценок точности и сложности) отличается от прямого делением на N, которое выражается в сдвиге позиции запятой в представлении коэффициентов и ведет к уменьшению

оценок для абсолютных значений и погрешностей в N раз. Окончательно абсолютные величины и погрешности компонент вектора значений заключительного ДПФ оцениваются как $2^{p+1}N^5A^2$ и $c_1c_2^{2s}2^{p+1}N^5A^22^{-E}$ соответственно.

Требуется, чтобы первая оценка была меньше 2^{E-1}, а последняя оценка была меньше $1/4$ (первое условие следует из второго). Тогда, если в заключение алгоритма выполнить приведение всех коэффициентов при степенях y по модулю $x^{2^p}+1$ (это сводится к параллельным вычитаниям комплексных чисел), то итоговая погрешность будет меньше $1/2$, а абсолютные значения комплексных коэффициентов — меньше 2^E.

Приходим к условию $2^{E/3} \geq c_1c_2^{2s}2^{p+1}N^5 = o(N^6)$, выполнения которого можно добиться, полагая $E = c_3 \log_2 N$ при достаточно большой константе c_3.

Теорема 1. *Умножение n-разрядных целых чисел реализуется схемой сложности $L(n) = nc^{\log^* n}\log n$, где $\log^* n$ определяется из соотношения*

$$1 \leq \underbrace{\log_2 \ldots \log_2}_{\log^* n} n < 2.$$

Доказательство. Воспользуемся оценкой (1), в которой L^*, согласно предыдущему пункту, можно оценить сверху как $L(O(2^p(E+p)))$:

$$L(n) \leq O(2^pEN\log N) + 3sNL(O(2^p(E+p))),$$

а $n \leq 2^{p-2}NE/3$. Учитывая, что $N \asymp n/\log^2 n$, $E \asymp \log n$, $2^p \asymp \log n$, $s \asymp \log n/\log\log n$, получаем соотношение

$$L(n) \leq \alpha\left(n\log_2 n + \frac{n}{\log_2 n \log_2\log_2 n} L\left(\lfloor \log_2^2 n \rfloor\right)\right).$$

Пусть $L(n) \leq nc_0^{\log^*(\sqrt{n}/4)}\log_2 n$ при $2 \leq n < 256$. Положим $c = \max\{1,\, c_0,\, 3\alpha\}$ и проверим по индукции, что $L(n) \leq nc^{\log^*(\sqrt{n}/4)}\log_2 n$. Заметим, что при $n \geq 256$ справедливо $n > \log_2^2 n$ и $\log_2(\sqrt{n}/4) = \frac{1}{2}\log_2 n - 2 \geq \frac{1}{4}\log_2 n$, откуда следует, что $\log^*(\sqrt{\log_2^2 n}/4) \leq \log^*(\sqrt{n}/4) - 1$.

Применяя индуктивное предположение для $n \geq 256$ получаем

$$L(n) \leq \alpha\left(n\log_2 n + 2n\log_2 nc^{\log^*(\sqrt{n}/4)-1}\right) = n\log_2 n\left(2\alpha c^{\log^*(\sqrt{n}/4)-1} + \alpha\right),$$

откуда следует доказываемое соотношение. Теорема доказана.