Fast multiplication

D. J. Bernstein
University of Illinois at Chicago

Part 1: polynomial multiplication

Commutative ring A.
Given coefficients of $f, g \in A[x]$.
Want coefficients of $h = fg$.

e.g. $f = f_0 + f_1 x$, $g = g_0 + g_1 x$:
$h = h_0 + h_1 x + h_2 x^2$ where
$h_0 = f_0 g_0$,
$h_1 = f_0 g_1 + f_1 g_0$,
$h_2 = f_1 g_1$.

4 mults in A. 1 add in A.

Or: $h_0 = f_0 g_0$, $h_2 = f_1 g_1$,
$h_1 = (f_0 + f_1)(g_0 + g_1) - h_0 - h_2$.
3 mults, 2 adds, 2 subs.

Proof of the formula for h_1:
$h_0 + h_1 + h_2 = h(1)$
$= f(1)g(1) = (f_0 + f_1)(g_0 + g_1)$.

$p \mapsto p(1)$ is a ring morphism
$A[x] \to A$.

Algebraic algorithm:
Start from $f_0, \ldots, f_d, g_0, \ldots, g_d$
and some constants in A.
Obtain new elements of A
by a constant sequence of
adds, subs, mults.
Eventually obtain $h_0, \ldots, h_{2d}$.

Total A-complexity:
number of adds, subs, mults.

Karatsuba's method

Assume $\deg f < 2n$, $\deg g < 2n$.

Write f as $p_0 + p_1 x^n$
with $\deg p_0 < n$, $\deg p_1 < n$.
Similarly g as $q_0 + q_1 x^n$.

Then $h = (p_0 + p_1)(q_0 + q_1)x^n$
$+ (p_0 q_0 - p_1 q_1 x^n)(1 - x^n)$.

(Karatsuba 1963)

$y \mapsto x^n$ is a ring morphism $A[x][y] \to A[x]$.

$p_0 + p_1 y \mapsto f$ and
$q_0 + q_1 y \mapsto g$ so
$(p_0 + p_1 y)(q_0 + q_1 y) \mapsto h$.

Multiply $p_0 + p_1 y$ by $q_0 + q_1 y$ in $A[x][y]$.
Substitute $y \mapsto x^n$ to get h.

Complexity of Karatsuba's method:

Three products with $\deg < n$.

$7n - 3$ extra adds/subs.

10111100110101100000101100011110 00111010111101001010100100011101

1011110011010110 0011101011110100 0000101100011110 1010100100011101 1011011111001000 1001001111101001

10111100 00111010 11010110 11110100 01101010 11001110 00001011 10101001 00011110 00011101 00010101 10110100 10110111 10010011 11001000 11101001 01111111 01111010

1011 0011 1100 1010 0111 1001 1101 1111 0110 0100 1011 1011 0110 1100 1010 1110 1100 0010 0000 1010 1011 1001 1011 0011 0001 0001 1110 1101 1111 1100 0001 1011 0101 0100 0100 1111 1011 1001 0111 0011 1100 1010 1100 1110 1000 1001 0100 0111 0111 0111 1111 1010 1000 1101

For $n = 2^k$, $k \geq 2$: Complexity
$(103/18) \cdot 3^k - 7 \cdot 2^k + 3/2$
if $\deg f < n$, $\deg g < n$.

$3^k = n^{\lg 3} < n^{1.585}$
where $\lg = \log_2$.

The fast Fourier transform

To multiply in $\mathbf{C}[x]/(x^{64}-1)$:

$\mathbf{C}[x]/(x^{64}-1)$
$\to \mathbf{C}[x]/(x^{32}-1) \times \mathbf{C}[x]/(x^{32}+1)$.

$\mathbf{C}[x]/(x^{32}+1)$
$\to \mathbf{C}[x]/(x^{16}-i) \times \mathbf{C}[x]/(x^{16}+i)$.

Continue to $\mathbf{C} \times \mathbf{C} \times \cdots \times \mathbf{C}$.

(Gauss 1805)

≤ 3 operations in $\mathbf{C}$ for
$ax^j + bx^{n+j}$
$\mapsto (a + b\zeta)x^j, (a - b\zeta)x^j$
under $\mathbf{C}[x]/(x^{2n} - \zeta^2)$
$\to \mathbf{C}[x]/(x^n - \zeta) \times \mathbf{C}[x]/(x^n + \zeta)$.

$\mathbf{C}$-complexity $(3/2)n \lg n - n + 1$
for $\mathbf{C}[x]/(x^n - 1) \to \mathbf{C}^n$
when $n = 2^k$, $k \geq 0$.

$\mathbf{C}$-complexity $(9/2)n \lg n - n + 3$
to multiply in $\mathbf{C}[x]/(x^n - 1)$.

Represent $\mathbf{C}$ as $\mathbf{R}[i]/(i^2+1)$.

$(a, b) \mapsto (a + b\zeta, a - b\zeta)$
takes 10 operations in $\mathbf{R}$.
Only 4 operations if $\zeta^2 = -1$.
Only 8 operations if $\zeta^4 = -1$.

$\mathbf{R}$-complexity $15n \lg n - 22n + 48$
to multiply in $\mathbf{C}[x]/(x^n - 1)$
when $n = 2^k$, $k \geq 3$.

Split-radix FFT

$$\mathbf{C}[x]/(x^{64}-1)$$
$$\to \mathbf{C}[x]/(x^{32}-1) \times \mathbf{C}[x]/(x^{32}+1)$$
$$\to \mathbf{C}[x]/(x^{32}-1) \times$$
$$\quad \mathbf{C}[x]/(x^{16}-i) \times \mathbf{C}[x]/(x^{16}+i)$$
$$\to \mathbf{C}[x]/(x^{32}-1) \times$$
$$\quad \mathbf{C}[y]/(y^{16}-1) \times \mathbf{C}[z]/(z^{16}-1)$$

by $x \mapsto \zeta y$, $x \mapsto \zeta^{-1} z$
where $\zeta^{16} = i$.

$\mathbf{R}$-complexity $12n \lg n - 10n + 24$ to multiply in $\mathbf{C}[x]/(x^n - 1)$ when $n = 2^k$, $k \geq 3$. (Yavne 1968; Duhamel; Hollmann; Martens; Stasinski; Vetterli; Nussbaumer)

Arbitrary n: $(12 + o(1))n \lg n$. (reduction to power-of-2 case: Gauss 1805; Good 1951; better reduction: Crandall, Fagin 1994)

Real FFT

$\mathbf{R}[x]/(x^{64} - 1)$
$\to \mathbf{R}[x]/(x^{32} - 1) \times \mathbf{R}[x]/(x^{32} + 1)$
$\to \mathbf{R}[x]/(x^{32} - 1) \times \mathbf{C}[x]/(x^{16} - i)$
(Gauss 1805; Bergland 1968)

R-complexity $(12 + o(1))n \lg n$
to multiply in $\mathbf{R}[x]/(x^{2n} - 1)$;
e.g. to multiply $f, g \in \mathbf{R}[x]$
if $\deg fg < 2n$.

Part 2: integer multiplication

Given $f, g \in \mathbf{Z}$. Want $h = fg$.

f represented as $(f_0, f_1, f_2, \ldots)$
with $f_j \in \mathbf{Z}$, $|f_j| \le 2^{53}$,
$f = f_0 + 2^{48} f_1 + 2^{96} f_2 + \cdots$.
Not unique. Similarly g, h.

Use **floating-point algorithms**.
Try to minimize number of
floating-point operations.

A **floating-point number**

is a real number $2^a b$

with $a, b \in \mathbf{Z}$ and $|b| \le 2^{53}$.

Floating-point operations:

$u, v \mapsto \mathrm{fp}_{53}(u + v)$

$u, v \mapsto \mathrm{fp}_{53}(u - v)$

$u, v \mapsto \mathrm{fp}_{53}\, uv$

For each $z \in \mathbf{R}$:

$\mathrm{fp}_{53}\, z$ is a floating-point number.

$|z - \mathrm{fp}_{53}\, z| \le 2^{a-1}$ if $|z| \le 2^{a+53}$.

If u is a floating-point number and $|u| \leq 2^{75}$:

Define $\alpha = 3 \cdot 2^{75}$,
$u_1 = \text{fp}_{53}(\text{fp}_{53}(u + \alpha) - \alpha)$,
$u_0 = \text{fp}_{53}(u - u_1)$.

Then $u_1 \in 2^{24}\mathbf{Z}$, $|u_0| \leq 2^{23}$, and $u = u_0 + u_1$.

(Kahan 1965; et al.)

Can build big-integer arithmetic using floating-point operations. (Veltkamp 1968; Dekker 1971)

Carry $f = f_0 + 2^{48} f_1 + \cdots$
into $f = s_0 + 2^{24} s_1 + 2^{48} s_2 + \cdots$
with $s_j \in \mathbf{Z}$, $|s_j| \le 2^{23}$.
Similarly $g = t_0 + 2^{24} t_1 + \cdots$.

Then $s_0 t_1 + s_1 t_0$
$= \mathrm{fp}_{53}(\mathrm{fp}_{53}\, s_0 t_1 + \mathrm{fp}_{53}\, s_1 t_0)$,
etc. Be careful past 127.

The Schönhage-Strassen method

Define $A = \mathbf{Z}/(2^{1536} + 1)$.
A has a 1024th root of -1,
namely $\zeta = 2^{1153} - 2^{385}$.

Can multiply in $A[x]/(x^{2048} - 1)$
using FFT over A.

Easy to multiply by powers of ζ.
Powers of $\zeta^{32} = 2^{48}$ are easiest.
Can eliminate most other powers
as in split-radix FFT.

$x \mapsto 2^{768}$ is a ring morphism
$\mathbf{Z}[x]/(x^{2048} - 1)$
$\to \mathbf{Z}/(2^{1572864} - 1)$.

Lift elements of $\mathbf{Z}/(2^{1572864} - 1)$
to elements of $\mathbf{Z}[x]/(x^{2048} - 1)$
with coefficients under 2^{768}.

Product in $\mathbf{Z}[x]/(x^{2048} - 1)$
is determined by images
in $A[x]/(x^{2048} - 1)$
and $(\mathbf{Z}/2^{11})[x]/(x^{2048} - 1)$.

Can multiply $f, g \in \mathbf{Z}$ with
a circuit of size $O(n \lg n \lg \lg n)$
if $|f| < 2^n$, $|g| < 2^n$.
(Schönhage, Strassen 1971)

For any ring A:
Can multiply $f, g \in A[x]$ with
$O(n \lg n \lg \lg n)$ operations in A
if $\deg f < n$, $\deg g < n$.
(Cantor, Kaltofen 1991)