CSE 548: Analysis of Algorithms

Lectures 4 & 5 (Divide-and-Conquer Algorithms: Polynomial Multiplication & the Fast Fourier Transform)

> **Rezaul A. Chowdhury** Department of Computer Science SUNY Stony Brook Fall 2012

$$A(x) = \sum_{j=0}^{n-1} a_j x^j$$

= $a_0 + a_1 x + a_2 x^2 + \dots + a_{n-1} x^{n-1}$

A(x) is a polynomial of degree bound n represented as a vector $a = (a_0, a_1, \dots, a_{n-1})$ of coefficients.

The *degree* of A(x) is k provided it is the largest integer such that a_k is nonzero. Clearly, $0 \le k \le n - 1$.

Evaluating A(x) at a given point:

Takes $\Theta(n)$ time using Horner's rule:

$$A(x_0) = a_0 + a_1 x_0 + a_2 (x_0)^2 + \dots + a_{n-1} (x_0)^{n-1}$$

= $a_0 + x_0 (a_1 + x_0 (a_2 + \dots + x_0 (a_{n-2} + x_0 (a_{n-1})) \dots))$

Adding Two Polynomials:

Adding two polynomials of degree bound n takes $\Theta(n)$ time.

C(x) = A(x) + B(x)

where,
$$A(x) = \sum_{j=0}^{n-1} a_j x^j$$
 and $B(x) = \sum_{j=0}^{n-1} b_j x^j$.
Then $C(x) = \sum_{j=0}^{n-1} c_j x^j$, where, $c_j = a_j + b_j$ for $0 \le j \le n-1$.

Multiplying Two Polynomials:

The product of two polynomials of degree bound n is another polynomial of degree bound 2n - 1.

C(x) = A(x)B(x)

where,
$$A(x) = \sum_{j=0}^{n-1} a_j x^j$$
 and $B(x) = \sum_{j=0}^{n-1} b_j x^j$.
Then $C(x) = \sum_{j=0}^{2n-1} c_j x_j^j$ where, $c_j = \sum_{k=0}^j a_k b_{j-k}$ for $0 \le j \le 2n-2$.

The coefficient vector $c = (c_0, c_1, \dots, c_{2n-2})$, denoted by $c = a \otimes b$, is also called the *convolution* of vectors $a = (a_0, a_1, \dots, a_{n-1})$ and $b = (b_0, b_1, \dots, b_{n-1})$.

Clearly, straightforward evaluation of c takes $\Theta(n^2)$ time.

Multiplying Two Polynomials:

We can use Karatsuba's algorithm (assume *n* to be a power of 2):

$$A(x) = \sum_{j=0}^{n-1} a_j x^j = \sum_{j=0}^{\frac{n}{2}-1} a_j x^j + x^{\frac{n}{2}} \sum_{j=0}^{\frac{n}{2}-1} a_{\frac{n}{2}+j} x^j = A_1(x) + x^{\frac{n}{2}} A_2(x)$$
$$B(x) = \sum_{j=0}^{n-1} b_j x^j = \sum_{j=0}^{\frac{n}{2}-1} b_j x^j + x^{\frac{n}{2}} \sum_{j=0}^{\frac{n}{2}-1} b_{\frac{n}{2}+j} x^j = B_1(x) + x^{\frac{n}{2}} B_2(x)$$

Then
$$C(x) = A(x)B(x)$$

= $A_1(x)B_1(x) + x^{\frac{n}{2}}[A_1(x)B_2(x) + A_2(x)B_1(x)] + x^n A_2(x)B_2(x)$

But $A_1(x)B_2(x) + A_2(x)B_1(x)$

 $= [A_1(x) + A_2(x)][B_1(x) + B_2(x)] - A_1(x)B_1(x) - A_2(x)B_2(x)$

3 recursive multiplications of polynomials of degree bound $\frac{n}{2}$. Similar recurrence as in Karatsuba's integer multiplication algorithm leading to a complexity of $O(n^{\log_2 3}) = O(n^{1.59})$.

Point-Value Representation of Polynomials

A point-value representation of a polynomial A(x) is a set of n pointvalue pairs $\{(x_0, y_0), (x_1, y_1), \dots, (x_{n-1}, y_{n-1})\}$ such that all x_k are distinct and $y_k = A(x_k)$ for $0 \le k \le n - 1$.

A polynomial has many point-value representations.

Adding Two Polynomials:

Suppose we have point-value representations of two polynomials of degree bound *n* using the same set of *n* points.

 $A: \{(x_0, y_0^a), (x_1, y_1^a), \dots, (x_{n-1}, y_{n-1}^a)\}$ $B: \{(x_0, y_0^b), (x_1, y_1^b), \dots, (x_{n-1}, y_{n-1}^b)\}$

If C(x) = A(x) + B(x) then

$$C: \{ (x_0, y_0^a + y_0^b), (x_1, y_1^a + y_1^b), \dots, (x_{n-1}, y_{n-1}^a + y_{n-1}^b) \}$$

Thus polynomial addition takes $\Theta(n)$ time.

Point-Value Representation of Polynomials

Multiplying Two Polynomials:

Suppose we have *extended* (why?) point-value representations of two polynomials of degree bound n using the same set of 2n points.

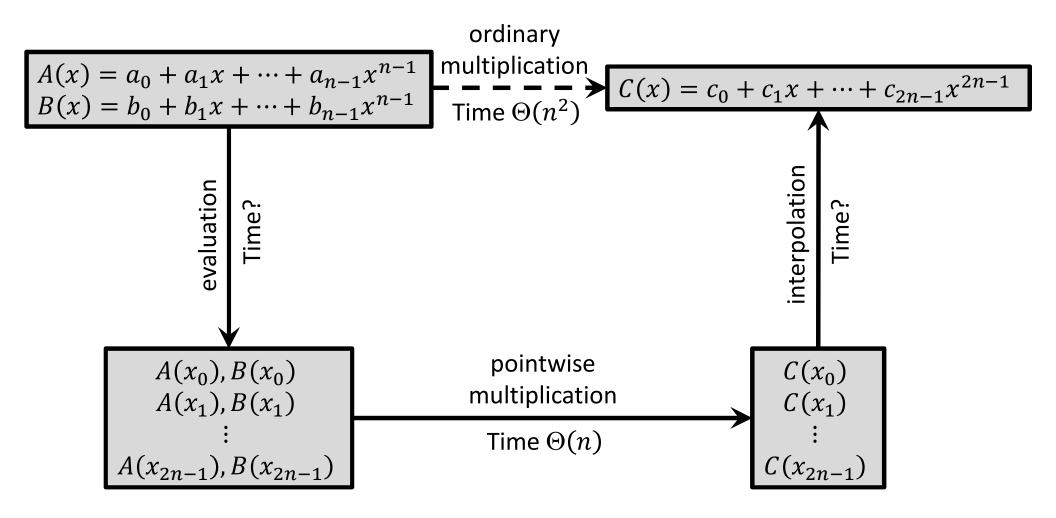
 $A: \{(x_0, y_0^a), (x_1, y_1^a), \dots, (x_{2n-1}, y_{2n-1}^a)\}$ $B: \{(x_0, y_0^b), (x_1, y_1^b), \dots, (x_{2n-1}, y_{2n-1}^b)\}$

If C(x) = A(x)B(x) then

$$C: \{ (x_0, y_0^a y_0^b), (x_1, y_1^a y_1^b), \dots, (x_{2n-1}, y_{2n-1}^a y_{2n-1}^b) \}$$

Thus polynomial multiplication also takes only $\Theta(n)$ time! (compare this with the $\Theta(n^2)$ time needed in the coefficient form)

<u>Faster Polynomial Multiplication?</u> (in Coefficient Form)



<u>Faster Polynomial Multiplication?</u> (in Coefficient Form)

Coefficient Representation \Rightarrow **Point-Value Representation**:

We select any set of n distinct points $\{x_0, x_1, \dots, x_{n-1}\}$, and evaluate $A(x_k)$ for $0 \le k \le n-1$.

Using Horner's rule this approach takes $\Theta(n^2)$ time.

Point-Value Representation ⇒ **Coefficient Representation**:

We can interpolate using Lagrange's formula:

$$A(x) = \sum_{k=0}^{n-1} \frac{\prod_{j \neq k} (x - x_j)}{\prod_{j \neq k} (x_k - x_j)} y_k$$

This again takes $\Theta(n^2)$ time.

In both cases we need to do much better!

A polynomial of degree bound n: $A(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ A set of n distinct points: $\{x_0, x_1, \dots, x_{n-1}\}$

Compute point-value form: $\{(x_0, A(x_0)), (x_1, A(x_1)), ..., (x_{n-1}, A(x_{n-1}))\}$

Using matrix notation: $\begin{bmatrix} A(x_0) \end{bmatrix}$

$$\mathsf{n:} \begin{bmatrix} A(x_0) \\ A(x_1) \\ \vdots \\ \vdots \\ A(x_{n-1}) \end{bmatrix} = \begin{bmatrix} 1 & x_0 & (x_0)^2 & \cdots & (x_0)^{n-1} \\ 1 & x_1 & (x_1)^2 & \cdots & (x_1)^{n-1} \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ 1 & x_{n-1} & (x_{n-1})^2 & \cdots & (x_{n-1})^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ a_{n-1} \end{bmatrix}$$

We want to choose the set of points in a way that simplifies the multiplication.

In the rest of the lecture on this topic we will assume:

n is a power of 2.

Let's choose $x_{n/2+j} = -x_j$ for $0 \le j \le n/2 - 1$. Then

$$\begin{bmatrix} A(x_{0}) \\ A(x_{1}) \\ \vdots \\ A(x_{n/2-1}) \\ A(x_{n/2+1}) \\ A(x_{n/2+1}) \\ \vdots \\ A(x_{n/2+(n/2-1)}) \end{bmatrix} = \begin{bmatrix} 1 & x_{0} & (x_{0})^{2} & \cdots & (x_{0})^{n-1} \\ 1 & x_{1} & (x_{1})^{2} & \cdots & (x_{1})^{n-1} \\ \vdots & \vdots & \ddots & \ddots & \cdots & \vdots \\ 1 & -x_{0} & (-x_{0})^{2} & \cdots & (-x_{0})^{n-1} \\ 1 & -x_{1} & (-x_{1})^{2} & \cdots & (-x_{1})^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -x_{n/2-1} & (-x_{n/2-1})^{2} & \cdots & (-x_{n/2-1})^{n-1} \end{bmatrix} \begin{bmatrix} a_{0} \\ a_{1} \\ \vdots \\ \vdots \\ a_{n-1} \end{bmatrix}$$

Observe that for $0 \le j \le n/2 - 1$: $(x_{n/2+j})^{k} = \begin{cases} (x_{j})^{k}, & \text{if } k = even, \\ -(x_{j})^{k}, & \text{if } k = odd. \end{cases}$

Thus we have just split the original $n \times n$ matrix into two almost similar $\frac{n}{2} \times n$ matrices!

How and how much do we save?

$$A(x) = \sum_{\substack{l=0\\n/2-1}}^{n-1} a_l x^l = \sum_{\substack{l=0\\l=0}}^{n/2-1} a_{2l} x^{2l} + \sum_{\substack{l=0\\l=0}}^{n/2-1} a_{2l+1} x^{2l+1}$$
$$= \sum_{\substack{l=0\\l=0}}^{n/2-1} a_{2l} (x^2)^l + x \sum_{\substack{l=0\\l=0}}^{n/2-1} a_{2l+1} (x^2)^l = A_{even} (x^2) + x A_{odd} (x^2),$$

where, $A_{even}(x) = \sum_{l=0}^{n/2-1} a_{2l} x^{l}$ and $A_{odd}(x) = \sum_{l=0}^{n/2-1} a_{2l+1} x^{l}$.

Observe that for $0 \le j \le n/2 - 1$: $A(x_j) = A_{even}(x_j^2) + x_j A_{odd}(x_j^2)$ $A(x_{n/2+j}) = A(-x_j) = A_{even}(x_j^2) - x_j A_{odd}(x_j^2)$

So in order to evaluate $A(x_j)$ for all $0 \le j \le n - 1$, we need:

n/2 evaluations of A_{even} and n/2 evaluations of A_{odd} n multiplications

n/2 additions and n/2 subtractions

Thus we save about half the computation!

If we can recursively evaluate A_{even} and A_{odd} using the same approach, we get the following recurrence relation for the running time of the algorithm:

$$T(n) = \begin{cases} \Theta(1), & \text{if } n = 1, \\ 2T\left(\frac{n}{2}\right) + \Theta(n), & \text{otherwise.} \end{cases}$$
$$= \Theta(n \log n)$$

Our trick was to evaluate A at x (positive) and -x (negative).

But inputs to A_{even} and A_{odd} are always of the form x^2 (positive)! How can we apply the same trick?

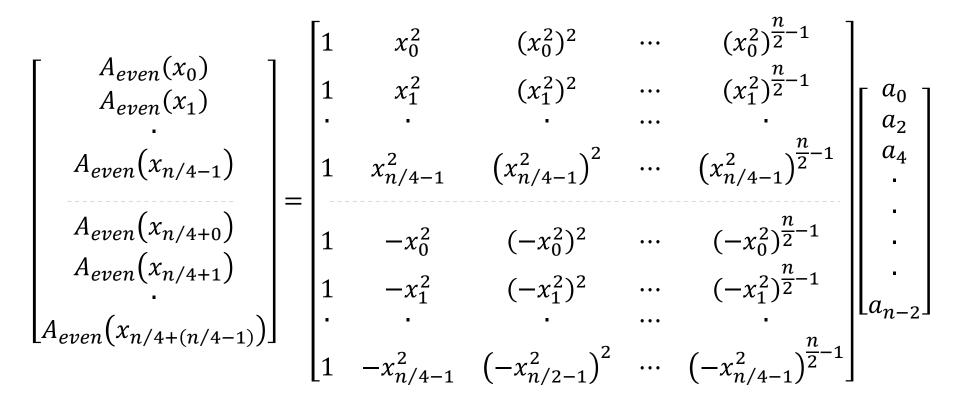
Let us consider the evaluation of $A_{even}(x_j)$ for $0 \le j \le n/2 - 1$:

$$\begin{bmatrix} A_{even}(x_0) \\ A_{even}(x_1) \\ \vdots \\ A_{even}(x_{n/2-1}) \end{bmatrix} = \begin{bmatrix} 1 & (x_0)^2 & (x_0)^4 & \cdots & (x_0)^{n-2} \\ 1 & (x_1)^2 & (x_1)^4 & \cdots & (x_1)^{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_{n/2-1})^2 & (x_{n/2-1})^4 & \cdots & (x_{n/2-1})^{n-2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_2 \\ a_4 \\ \vdots \\ a_{n-2} \end{bmatrix}$$

In order to apply the same trick on A_{even} we must set:

$$(x_{n/4+j})^2 = -(x_j)^2$$
 for $0 \le j \le n/4 - 1$

In A_{even} we set: $x_{n/4+j}^2 = -x_j^2$ for $0 \le j \le n/4 - 1$. Then



This means setting $x_{n/4+j} = ix_j$, where $i = \sqrt{-1}$ (imaginary)!

This also allows us to apply the same trick on A_{odd} .

We can apply the trick once if we set:

$$x_{n/2+j} = -x_j$$
 for $0 \le j \le n/2 - 1$

We can apply the trick (recursively) 2 times if we also set:

$$(x_{n/2^2+j})^2 = -(x_j)^2$$
 for $0 \le j \le n/2^2 - 1$

We can apply the trick (recursively) 3 times if we also set:

$$(x_{n/2^3+j})^{2^2} = -(x_j)^{2^2}$$
 for $0 \le j \le n/2^3 - 1$

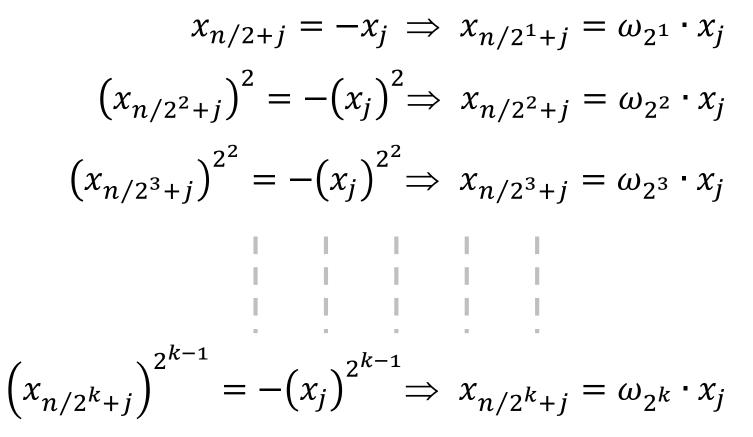
We can apply the trick (recursively) k times if we also set:

$$(x_{n/2^k+j})^{2^{k-1}} = -(x_j)^{2^{k-1}}$$
 for $0 \le j \le n/2^k - 1$

Consider the *t*th primitive root of unity:

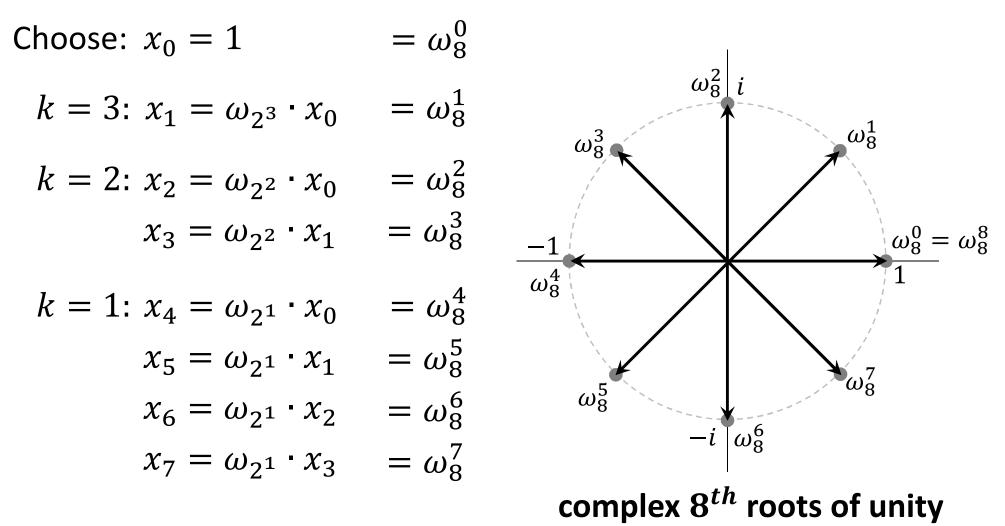
$$\omega_t = e^{\frac{2\pi i}{t}} = \cos\frac{2\pi}{t} + i \cdot \sin\frac{2\pi}{t} \quad (i = \sqrt{-1})$$

Then



If $n = 2^k$ we would like to apply the trick k times recursively. What values should we choose for $\{x_0, x_1, \dots, x_{n-1}\}$?

Example: For $n = 2^3$ we need to choose $\{x_0, x_1, \dots, x_7\}$.



For a polynomial of degree bound $n = 2^k$, we need to apply the trick recursively at most $\log n = k$ times.

We choose $x_0 = 1 = \omega_n^0$ and set $x_j = \omega_n^j$ for $1 \le j \le n - 1$.

Then we compute the following product:

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} A(1) \\ A(\omega_n) \\ A(\omega_n^2) \\ \vdots \\ A(\omega_n^{n-1}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_n & (\omega_n)^2 & \cdots & (\omega_n)^{n-1} \\ 1 & \omega_n^2 & (\omega_n^2)^2 & \cdots & (\omega_n^2)^{n-1} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 1 & \omega_n^{n-1} & (\omega_n^{n-1})^2 & \cdots & (\omega_n^{n-1})^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

The vector $y = (y_0, y_1, \dots, y_{n-1})$ is called the *discrete Fourier* transform (DFT) of $(a_0, a_1, \dots, a_{n-1})$.

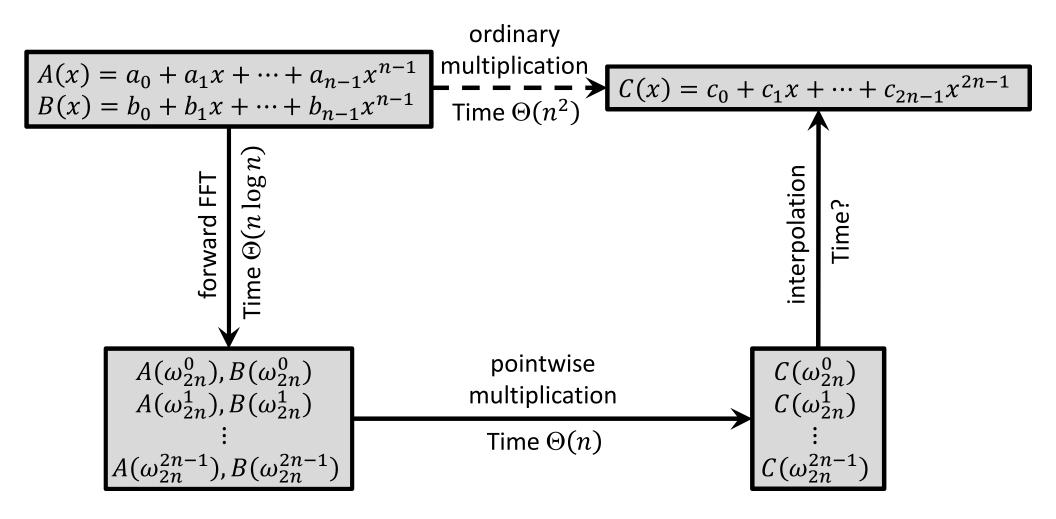
This method of computing DFT is called the *fast Fourier transform* (FFT) method.

Rec-FFT (($a_0, a_1, ..., a_{n-1}$)) { $n = 2^k$ for integer $k \ge 0$ } 1. *if n* = 1 *then* 2. return (a_0) 3. $\omega_n \leftarrow e^{2\pi i/n}$ 4. $\omega \leftarrow 1$ 5. $y^{\text{even}} \leftarrow \textit{Rec-FFT} ((a_0, a_2, ..., a_{n-2}))$ 6. $y^{\text{odd}} \leftarrow \text{Rec-FFT} ((a_1, a_3, ..., a_{n-1}))$ 7. for $j \leftarrow 0$ to n/2 - 1 do 8. $y_j \leftarrow y_j^{\text{even}} + \omega y_j^{\text{odd}}$ 9. $y_{n/2+j} \leftarrow y_j^{\text{even}} - \omega y_j^{\text{odd}}$ 10. $\omega \leftarrow \omega \omega_n$ 11. return y

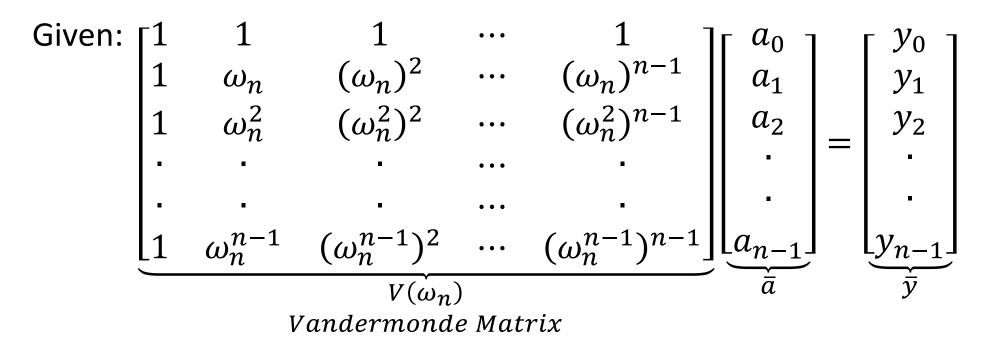
Running time:

$$T(n) = \begin{cases} \Theta(1), & \text{if } n = 1, \\ 2T\left(\frac{n}{2}\right) + \Theta(n), & \text{otherwise.} \end{cases}$$
$$= \Theta(n \log n)$$

<u>Faster Polynomial Multiplication?</u> (in Coefficient Form)



<u>Point-Value Form ⇒ Coefficient Form</u>



 $\Rightarrow V(\omega_n) \cdot \bar{a} = \bar{y}$

We want to solve: $\bar{a} = [V(\omega_n)]^{-1} \cdot \bar{y}$

It turns out that:
$$[V(\omega_n)]^{-1} = \frac{1}{n}V\left(\frac{1}{\omega_n}\right)$$

That means $[V(\omega_n)]^{-1}$ looks almost similar to $V(\omega_n)!$

<u>Point-Value Form ⇒ Coefficient Form</u>

Show that: $[V(\omega_n)]^{-1} = \frac{1}{n}V\left(\frac{1}{\omega_n}\right)$ Let $U(\omega_n) = \frac{1}{n}V\left(\frac{1}{\omega_n}\right)$

We want to show that $U(\omega_n)V(\omega_n) = I_n$, where I_n is the $n \times n$ identity matrix.

Observe that for $0 \le j, k \le n - 1$, the $(j, k)^{th}$ entries are: $[V(\omega_n)]_{jk} = \omega_n^{jk}$ and $[U(\omega_n)]_{jk} = \frac{1}{n} \omega_n^{-jk}$

Then entry (p,q) of $U(\omega_n)V(\omega_n)$,

$$[U(\omega_n)V(\omega_n)]_{pq} = \sum_{k=0}^{n-1} [U(\omega_n)]_{pk} [V(\omega_n)]_{kq} = \frac{1}{n} \sum_{k=0}^{n-1} \omega_n^{k(q-p)}$$

<u>Point-Value Form ⇒ Coefficient Form</u>

$$[U(\omega_n)V(\omega_n)]_{pq} = \frac{1}{n} \sum_{k=0}^{n-1} \omega_n^{k(q-p)}$$

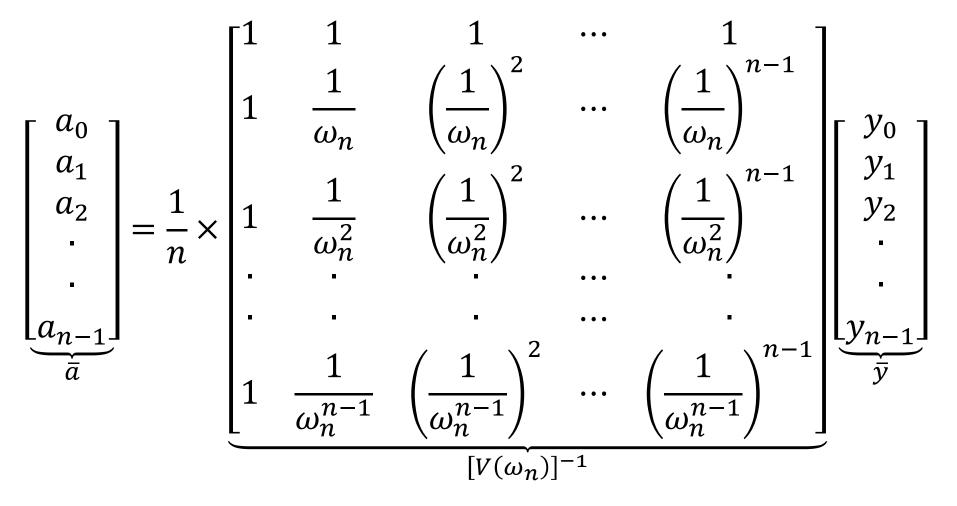
CASE
$$p = q$$
:
 $[U(\omega_n)V(\omega_n)]_{pq} = \frac{1}{n} \sum_{k=0}^{n-1} \omega_n^0 = \frac{1}{n} \sum_{k=0}^{n-1} 1 = \frac{1}{n} \times n = 1$

CASE
$$p \neq q$$
:
 $[U(\omega_n)V(\omega_n)]_{pq} = \frac{1}{n} \sum_{k=0}^{n-1} (\omega_n^{q-p})^k = \frac{1}{n} \times \frac{(\omega_n^{q-p})^n - 1}{\omega_n^{q-p} - 1}$
 $= \frac{1}{n} \times \frac{(\omega_n^n)^{q-p} - 1}{\omega_n^{q-p} - 1} = \frac{1}{n} \times \frac{(1)^{q-p} - 1}{\omega_n^{q-p} - 1} = 0$

Hence $U(\omega_n)V(\omega_n) = I_n$

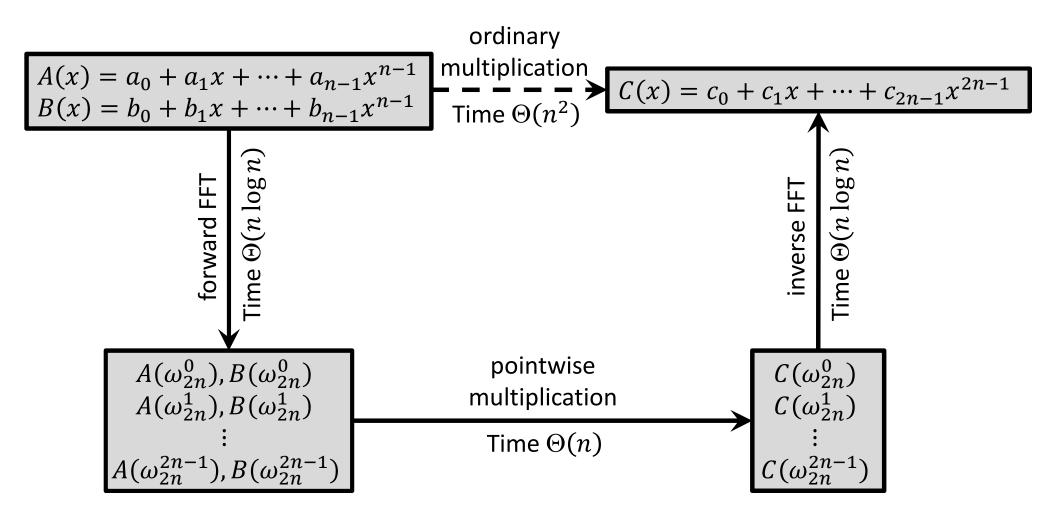
<u>Point-Value Form \Rightarrow Coefficient Form</u>

We need to compute the following matrix-vector product:



This inverse problem is almost similar to the forward problem, and can be solved in $\Theta(n \log n)$ time using the same algorithm as the forward FFT with only minor modifications!

<u>Faster Polynomial Multiplication?</u> (in Coefficient Form)



Two polynomials of degree bound n given in the coefficient form can be multiplied in $\Theta(n \log n)$ time!