

In-place fast polynomial modular remainder

Jean-Guillaume Dumas & **Bruno Grenet**



ISSAC, Raleigh, NC
July 16–19, 2024

Introduction

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ without extra space as fast as possible

Introduction

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ without extra space as fast as possible

Space complexity models

1. ro/wo: read-only inputs & write-only output
 - 👍 Classical model in complexity theory
 - 👎 $\Omega(n^2)$ lower bounds on time \times space
2. ro/rw: read-only inputs & read-write output
 - 👍 Allows parallel access to the inputs
 - 👎 Too restrictive in a sequential model
3. rw/rw: read-write inputs & read-write output, *inputs restored at the end*
 - 👍 Allows recursive calls / use as subroutines
 - 👎 Not suitable for parallel programming

Introduction

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ without extra space as fast as possible

Space complexity models

1. ro/wo: read-only inputs & write-only output

👍 Classical model in complexity theory

👎 $\Omega(n^2)$ lower bounds on time \times space

2. ro/rw: read-only inputs & read-write output

👍 Allows parallel access to the inputs

👎 Too restrictive in a sequential model

3. rw/rw: read-write inputs & read-write output, *inputs restored at the end*

👍 Allows recursive calls / use as subroutines

👎 Not suitable for parallel programming

alternative view: no distinction input/output \rightarrow full specification

▶ $C \leftarrow A \bmod B:$ $(A, B, C) \mapsto (A, B, A \bmod B)$

in place

▶ $C += AB:$ $(A, B, C) \mapsto (A, B, C + AB)$

in place with accumulation

▶ $B *= A:$ $(A, B) \mapsto (A, AB)$

over-place

Known results

$$A = BQ + R \text{ with } \deg(R) < \deg(B) \quad n = \deg(A), m = \deg(B), q = n - m, k = \max(q, m)$$

Algorithm	Model		QuoRem	Quotient only	Remainder only
Classical	ro/wo	time	$O(qm)$	same result [Monagan'93]	same result [Monagan'93]
		space	$O(1)$		
Fast [Strassen'73, ...]	ro/wo	time	$O(M(k))$	same result	same result
		space	$O(k)$		
Fast & low space [Giorgi G. Roche'20]	ro/rw	time	$O(M(k))$	$O(M(q)\log q)^*$	$O(M(k))$
		space	$O(1)$	$O(1)$	$\min(q, m) + O(1)$

* $O(M(q))$ if $M(q) = \Omega(q^{1+\epsilon})$

Known results

$$A = BQ + R \text{ with } \deg(R) < \deg(B) \quad n = \deg(A), m = \deg(B), q = n - m, k = \max(q, m)$$

Algorithm	Model		QuoRem	Quotient only	Remainder only
Classical	ro/wo	time	$O(qm)$	same result [Monagan'93]	same result [Monagan'93]
		space	$O(1)$		
Fast [Strassen'73, ...]	ro/wo	time	$O(M(k))$	same result	same result
		space	$O(k)$		
Fast & low space [Giorgi G. Roche'20]	ro/rw	time	$O(M(k))$	$O(M(q)\log q)^*$	$O(M(k))$
		space	$O(1)$	$O(1)$	$\min(q, m) + O(1)$

* $O(M(q))$ if $M(q) = \Omega(q^{1+\epsilon})$

- Open problems that “seem” impossible:
Remainder, modular inverse, factorization

Dan Roche's slides @ ISSAC'20 [Giorgi G. Roche'20]

Our results

In-place remainder

(model rw/rw but A read-only)

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ in time $O(M(k)\log m)^*$ and space $O(1)^\dagger$

$n = \deg(A)$, $m = \deg(B)$, $k = \max(n - m, m)$

* $O(M(k))$ if $M(k) = \Omega(k^{1+\varepsilon})$

† $O(\log m)$ pointers for call stack

Our results

In-place remainder

(model rw/rw but A read-only)

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ in time $O(M(k)\log m)^*$ and space $O(1)^\dagger$

$n = \deg(A)$, $m = \deg(B)$, $k = \max(n - m, m)$

* $O(M(k))$ if $M(k) = \Omega(k^{1+\varepsilon})$

† $O(\log m)$ pointers for call stack

Variants (same time & space)

- ▶ $C \ += A \bmod B$
- ▶ $R \ += AB \bmod C$
- ▶ $(A, B) \mapsto (Q|R, B)$

Our results

In-place remainder

(model rw/rw but A read-only)

Given $A, B \in \mathbb{F}[x]$, compute $A \bmod B$ in time $O(M(k)\log m)^*$ and space $O(1)^\dagger$

$n = \deg(A)$, $m = \deg(B)$, $k = \max(n - m, m)$

* $O(M(k))$ if $M(k) = \Omega(k^{1+\varepsilon})$

† $O(\log m)$ pointers for call stack

Variants (same time & space)

- ▶ $C \ += A \bmod B$
- ▶ $R \ += AB \bmod C$
- ▶ $(A, B) \mapsto (Q|R, B)$

Building blocks: in-place Toeplitz computations

- ▶ $\vec{c} \ += A \cdot \vec{b}$ where A is a Toeplitz matrix *in place with accumulation*
 - ▶ polynomial short and middle products
- ▶ $\vec{b} \leftarrow A \cdot \vec{b}$ and $\vec{b} \leftarrow A^{-1} \cdot \vec{b}$ where A is a triangular Toeplitz matrix *over place*
 - ▶ (truncated) power series multiplication and division

Contents

1. Toeplitz computations and convolutions

2. Remainder algorithms

Goals

Main goal: *Over-place* triangular Toeplitz computations

- ▶ $\vec{b} \leftarrow A \cdot \vec{b}$ where A is a lower triangular Toeplitz matrix
- ▶ $\vec{b} \leftarrow A^{-1} \cdot \vec{b}$ where A is an upper triangular Toeplitz matrix

Intermediate goal: *In-place* Toeplitz computations *with accumulation*

- ▶ $\vec{c} += A \cdot \vec{b}$ where A is
 - ▶ a triangular Toeplitz matrix
 - ▶ a square Toeplitz matrix
 - ▶ a rectangular Toeplitz matrix

The tool: *In-place* polynomial convolutions *with accumulation*

- ▶ $C += A \cdot B \bmod x^n - \lambda$ for any $\lambda \in \mathbb{F}$
 - ▶ $\vec{c} += A \cdot \vec{b}$ for a λ -circulant matrix A

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ \lambda \cdot a_{n-1} & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \lambda \cdot a_{n-2} & \lambda \cdot a_{n-1} & \cdots & a_{n-4} & a_{n-3} \\ & & \ddots & & \\ \lambda \cdot a_1 & \lambda \cdot a_2 & \cdots & \lambda \cdot a_{n-1} & a_0 \end{bmatrix}$$

Convolution

$$A = a_0 + x^{n/2}a_1$$

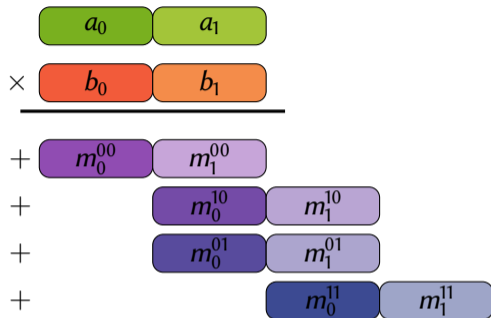
$$B = b_0 + x^{n/2}b_1$$

$$a_0b_0$$

$$a_1b_0$$

$$a_0b_1$$


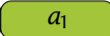
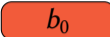
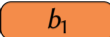
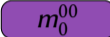
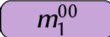
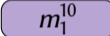
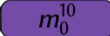
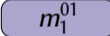
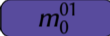
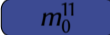
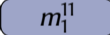
$$a_1b_1$$



Convolution

$$A = a_0 + x^{n/2}a_1$$

$$B = b_0 + x^{n/2}b_1$$

			
×			$\text{mod } x^n - 1$
	<hr/>		
	+		
	+		
	+		
	+		

$$a_0b_0$$

$$a_1b_0$$

$$a_0b_1$$

$$a_1b_1$$

Convolution

$$A = a_0 + x^{n/2}a_1$$

$$B = b_0 + x^{n/2}b_1$$

$$\begin{array}{cc} \text{green box } a_0 & \text{light green box } a_1 \\ \times & \\ \text{red box } b_0 & \text{orange box } b_1 \\ \hline \end{array}$$

$$\text{mod } x^n - \lambda$$

$$a_0b_0$$

$$+ \begin{array}{cc} \text{purple box } m_0^{00} & \text{light purple box } m_1^{00} \end{array}$$

$$a_1b_0$$

$$+ \begin{array}{cc} \text{light purple box } \lambda \cdot m_1^{10} & \text{purple box } m_0^{10} \end{array}$$

$$a_0b_1$$

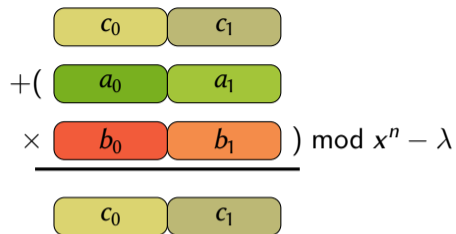
$$+ \begin{array}{cc} \text{light purple box } \lambda \cdot m_1^{01} & \text{dark purple box } m_0^{01} \end{array}$$

$$a_1b_1$$

$$+ \begin{array}{cc} \text{dark blue box } \lambda \cdot m_0^{11} & \text{light blue box } \lambda \cdot m_1^{11} \end{array}$$

Convolution: $C \ += AB \bmod x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$



Convolution: $C \ += AB \text{ mod } x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$

$$\begin{array}{r} \begin{array}{|c|c|} \hline c_0 & c_1 \\ \hline \end{array} \\ + \left(\begin{array}{|c|c|} \hline a_0 & a_1 \\ \hline \end{array} \right. \\ \times \left. \begin{array}{|c|c|} \hline b_0 & b_1 \\ \hline \end{array} \right) \text{ mod } x^n - \lambda \\ \hline \begin{array}{|c|c|} \hline c_0 & c_1 \\ \hline \end{array} \\ + \begin{array}{|c|c|} \hline m_0^{00} & m_1^{00} \\ \hline \end{array} \end{array}$$

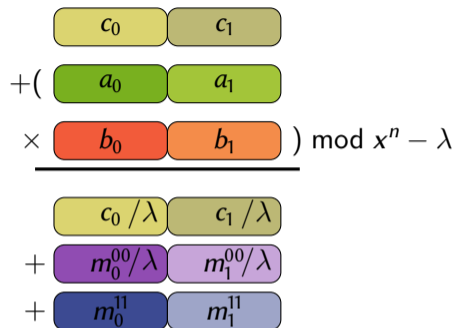
Convolution: $C \ += AB \text{ mod } x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$

$$\begin{array}{r} \begin{array}{|c|c|} \hline c_0 & c_1 \\ \hline \end{array} \\ + \left(\begin{array}{|c|c|} \hline a_0 & a_1 \\ \hline \end{array} \right. \\ \times \left. \begin{array}{|c|c|} \hline b_0 & b_1 \\ \hline \end{array} \right) \text{ mod } x^n - \lambda \\ \hline \begin{array}{|c|c|} \hline c_0 / \lambda & c_1 / \lambda \\ \hline \end{array} \\ + \begin{array}{|c|c|} \hline m_0^{00} / \lambda & m_1^{00} / \lambda \\ \hline \end{array} \end{array}$$

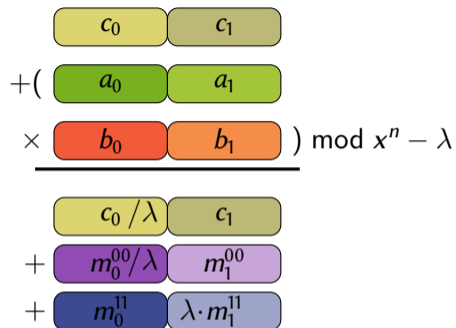
Convolution: $C \ += AB \text{ mod } x^n - \lambda, \lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$



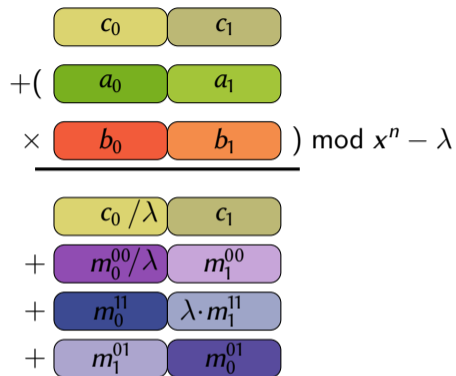
Convolution: $C \ += AB \bmod x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$



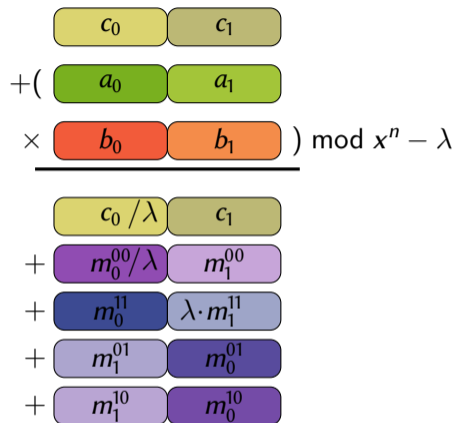
Convolution: $C \ += AB \bmod x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$



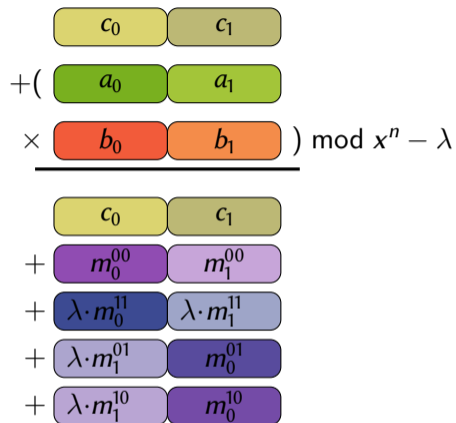
Convolution: $C \ += AB \text{ mod } x^n - \lambda, \lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$



Convolution: $C \ += AB \bmod x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) \ += a_0 \cdot b_0$
2. $C \ /= \lambda$
3. $(c_0, c_1) \ += a_1 \cdot b_1$
4. $c_1 \ *= \lambda$
5. $(c_1, c_0) \ += a_0 \cdot b_1$
6. $(c_1, c_0) \ += a_1 \cdot b_0$
7. $c_0 \ *= \lambda$

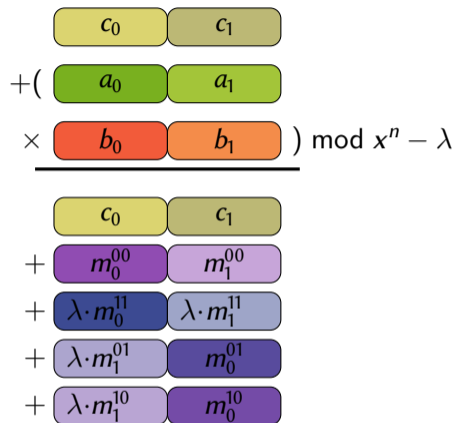


Convolution: $C += AB \bmod x^n - \lambda$, $\lambda \neq 0$

1. $(c_0, c_1) += a_0 \cdot b_0$
2. $C /= \lambda$
3. $(c_0, c_1) += a_1 \cdot b_1$
4. $c_1 *= \lambda$
5. $(c_1, c_0) += a_0 \cdot b_1$
6. $(c_1, c_0) += a_1 \cdot b_0$
7. $c_0 *= \lambda$

Remarks

- ▶ Requires $C += AB$ [D.G.'24 (previous talk)]
- ▶ Case n odd: additional technicalities
- ▶ Karatsuba-like improvement



- ▶ Time: $O(M(n))$
- ▶ Space: $O(1)$

Convolution: case $\lambda = 0$

Remark

► If $A \cdot B = x^n \times Q + R$

then $A \cdot B = x^n - \mu \times Q + R + \mu Q$

Convolution: case $\lambda = 0$

Remark

► If $A \cdot B = x^n \times Q + R$

then $A \cdot B = x^n - \mu \times Q + R + \mu Q$

► If $\alpha + (1 - \alpha) \cdot \mu = 0$ $\alpha \notin \{0, 1\}$ and $\mu = \alpha(\alpha - 1)^{-1}$

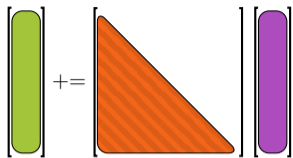
$$\alpha \underbrace{(R + Q)}_{(AB \bmod x^n - 1)} + (1 - \alpha) \underbrace{(R + \mu Q)}_{(AB \bmod x^n - \mu)} = \underbrace{R}_{(AB \bmod x^n)}$$

Algorithms: $C += A \cdot B \bmod x^n$

- If $\mathbb{F} \neq \mathbb{F}_2$: two accumulated convolutions $\bmod x^n - 1$ and $\bmod x^n - \mu$
- If $\mathbb{F} = \mathbb{F}_2$: dedicated algorithm based on a modified Toom-3 algorithm

- Time: $O(M(n))$
- Space: $O(1)$

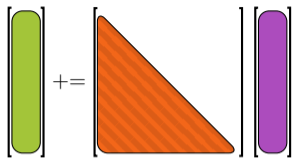
Accumulated Toeplitz matrix-vector product



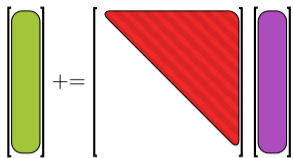
$$C += A \cdot B \pmod{x^n}$$

- ▶ Time: $O(M(n))$
- ▶ Space: $O(1)$

Accumulated Toeplitz matrix-vector product



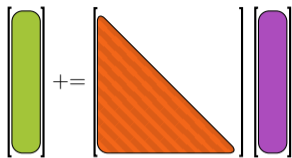
$$C += A \cdot B \bmod x^n$$



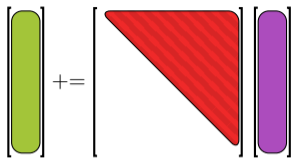
$$C += A \cdot B \operatorname{div} x^n$$

- ▶ Time: $O(M(n))$
- ▶ Space: $O(1)$

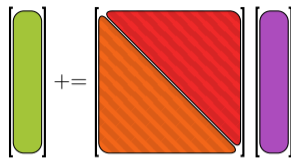
Accumulated Toeplitz matrix-vector product



$$C += A \cdot B \bmod x^n$$



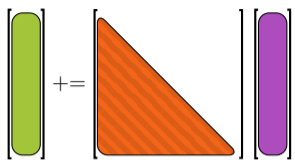
$$C += A \cdot B \operatorname{div} x^n$$



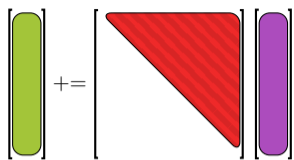
$$C += \text{MIDPROD}(A, B)$$

- ▶ Time: $O(M(n))$
- ▶ Space: $O(1)$

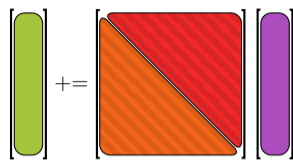
Accumulated Toeplitz matrix-vector product



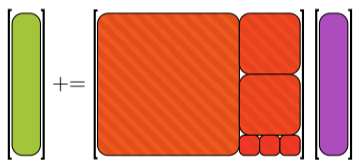
$$C += A \cdot B \bmod x^n$$



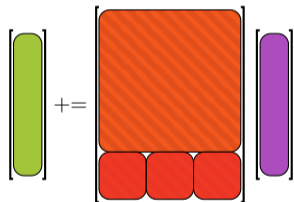
$$C += A \cdot B \operatorname{div} x^n$$



$$C += \text{MIDPROD}(A, B)$$



$$C += \text{MIDPROD}(A, B)$$

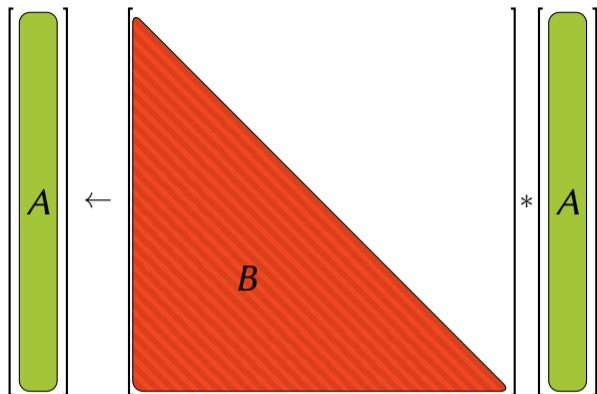


$$C += \text{MIDPROD}(A, B)$$

- ▶ Time: $O(M(n))$
- ▶ Space: $O(1)$

or $O(\frac{n}{m}M(m))$

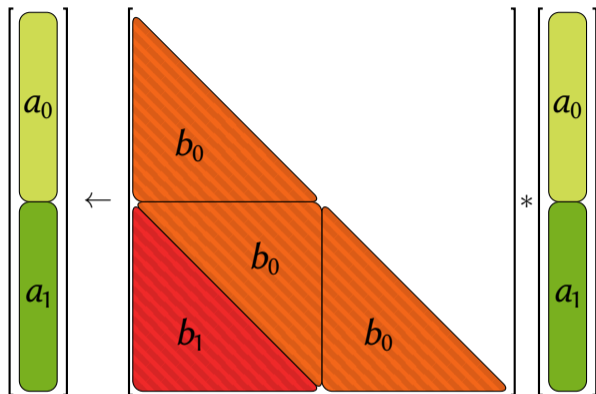
Over-place triangular Toeplitz matrix-vector product



$$A \leftarrow B \cdot A \bmod x^n$$

Inputs: A and B of degree $< n$ (even)
Output: $A \leftarrow A \cdot B \bmod x^n$

Over-place triangular Toeplitz matrix-vector product



$$A \leftarrow B \cdot A \bmod x^n$$

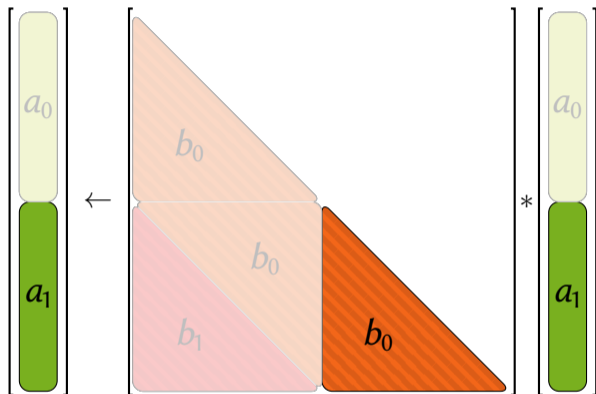
Inputs: A and B of degree $< n$ (even)

Output: $A \leftarrow A \cdot B \bmod x^n$

0. $A = a_0 + x^{n/2} a_1$

$B = b_0 + x^{n/2} b_1$

Over-place triangular Toeplitz matrix-vector product



$$A \leftarrow B \cdot A \bmod x^n$$

Inputs: A and B of degree $< n$ (even)

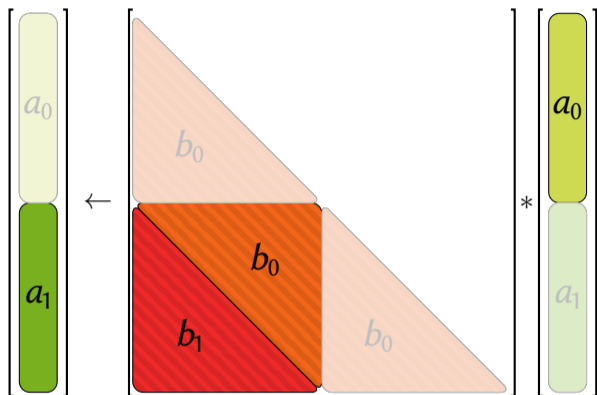
Output: $A \leftarrow A \cdot B \bmod x^n$

0. $A = a_0 + x^{n/2} a_1$

$B = b_0 + x^{n/2} b_1$

1. $a_1 *= b_0 \bmod x^{n/2}$ (recursive call)

Over-place triangular Toeplitz matrix-vector product



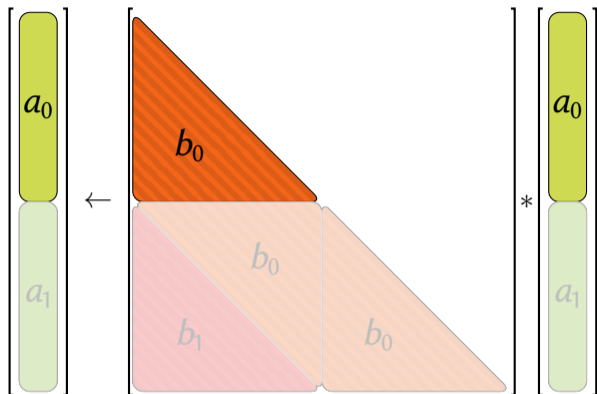
$$A \leftarrow B \cdot A \bmod x^n$$

Inputs: A and B of degree $< n$ (even)

Output: $A \leftarrow A \cdot B \bmod x^n$

0. $A = a_0 + x^{n/2} a_1$
 $B = b_0 + x^{n/2} b_1$
1. $a_1 \ast = b_0 \bmod x^{n/2}$ (recursive call)
2. $a_1 += \text{MIDPROD}(a_0, B)$

Over-place triangular Toeplitz matrix-vector product



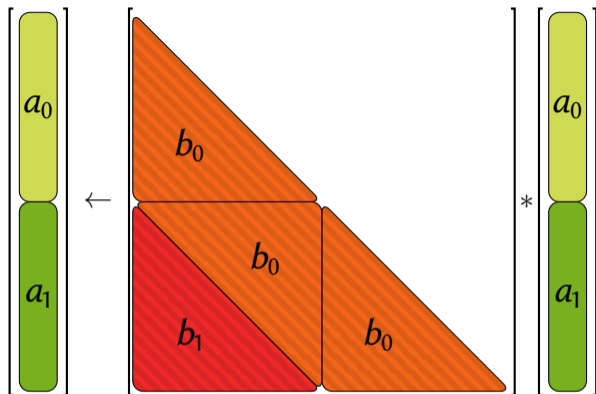
$$A \leftarrow B \cdot A \bmod x^n$$

Inputs: A and B of degree $< n$ (even)

Output: $A \leftarrow A \cdot B \bmod x^n$

0. $A = a_0 + x^{n/2} a_1$
 $B = b_0 + x^{n/2} b_1$
1. $a_1 *= b_0 \bmod x^{n/2}$ (recursive call)
2. $a_1 += \text{MIDPROD}(a_0, B)$
3. $a_0 *= b_0 \bmod x^{n/2}$ (recursive call)

Over-place triangular Toeplitz matrix-vector product



$$A \leftarrow B \cdot A \bmod x^n$$

Inputs: A and B of degree $< n$ (even)

Output: $A \leftarrow A \cdot B \bmod x^n$

0. $A = a_0 + x^{n/2} a_1$
 $B = b_0 + x^{n/2} b_1$
1. $a_1 *= b_0 \bmod x^{n/2}$ (recursive call)
2. $a_1 += \text{MIDPROD}(a_0, B)$
3. $a_0 *= b_0 \bmod x^{n/2}$ (recursive call)

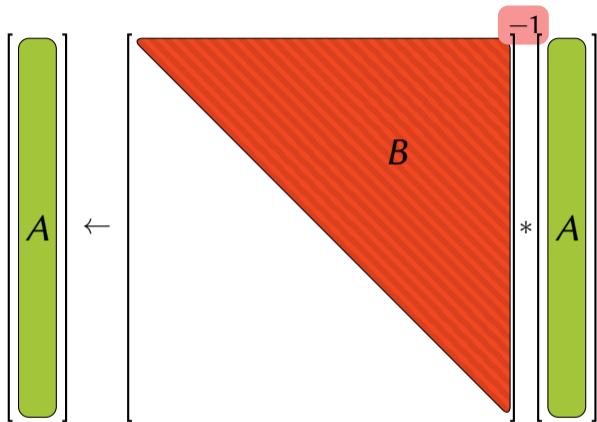
► Time: $O(M(n) \log n)^*$

► Space: $O(1)^\dagger$

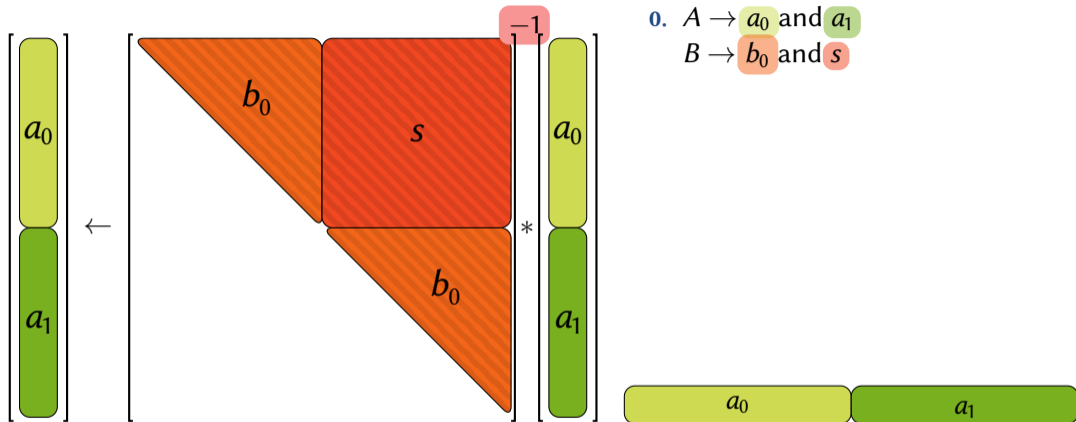
* $O(M(n))$ if $M(n) = \Omega(n^{1+\epsilon})$

† $O(\log n)$ pointers for call stack

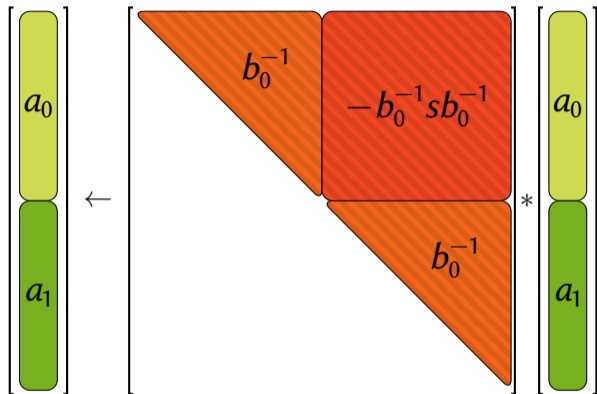
Over-place triangular Toeplitz system solving



Over-place triangular Toeplitz system solving

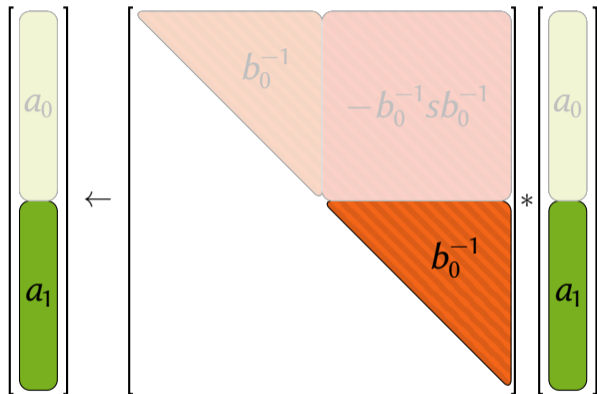


Over-place triangular Toeplitz system solving



- 0. $A \rightarrow a_0$ and a_1
 $B \rightarrow b_0$ and s

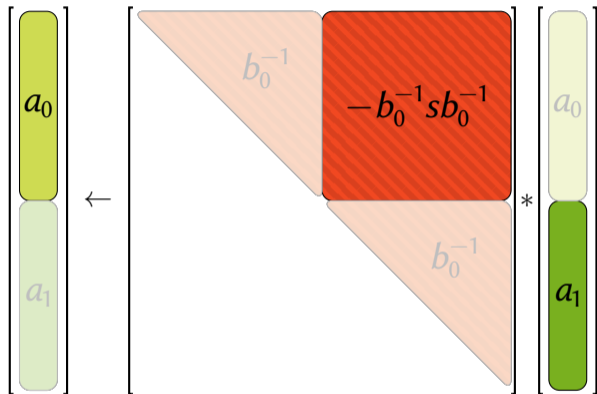
Over-place triangular Toeplitz system solving



0. $A \rightarrow a_0$ and a_1
 $B \rightarrow b_0$ and s
1. $a_1 * = b_0^{-1}$ (recursive call)



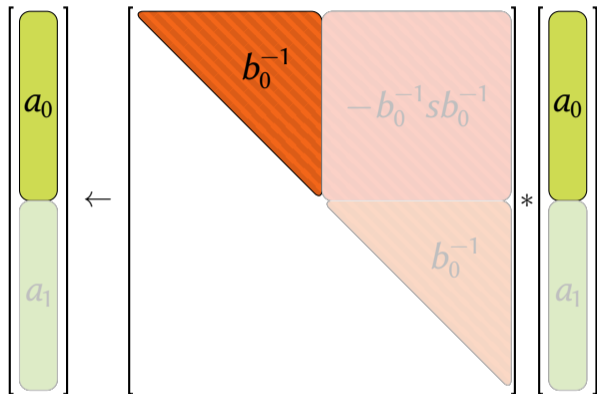
Over-place triangular Toeplitz system solving



0. $A \rightarrow a_0$ and a_1
 $B \rightarrow b_0$ and s
1. $a_1 *= b_0^{-1}$ (recursive call)
2. $a_0 -= s \cdot a_1$

$$a_0 - sb_0^{-1}a_1 \quad b_0^{-1}a_1$$

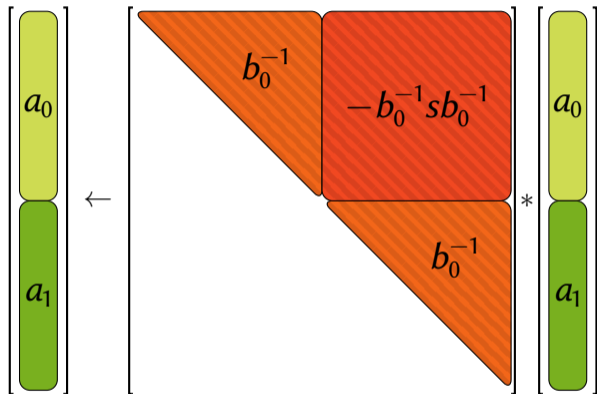
Over-place triangular Toeplitz system solving



0. $A \rightarrow a_0$ and a_1
 $B \rightarrow b_0$ and s
1. $a_1 \ast = b_0^{-1}$ (recursive call)
2. $a_0 -= s \cdot a_1$
3. $a_0 \ast = b_0^{-1}$ (recursive call)

$$b_0^{-1}(a_0 - s b_0^{-1} a_1) \quad b_0^{-1} a_1$$

Over-place triangular Toeplitz system solving



0. $A \rightarrow a_0$ and a_1
 $B \rightarrow b_0$ and s
1. $a_1 \ast = b_0^{-1}$ (recursive call)
2. $a_0 \dashv = s \cdot a_1$
3. $a_0 \ast = b_0^{-1}$ (recursive call)

$$b_0^{-1}(a_0 - s b_0^{-1} a_1)$$

$$b_0^{-1} a_1$$

- ▶ Time: $O(M(n) \log n)^\ast$
- ▶ Space: $O(1)^\dagger$

$^\ast O(M(n))$ if $M(n) = \Omega(n^{1+\epsilon})$

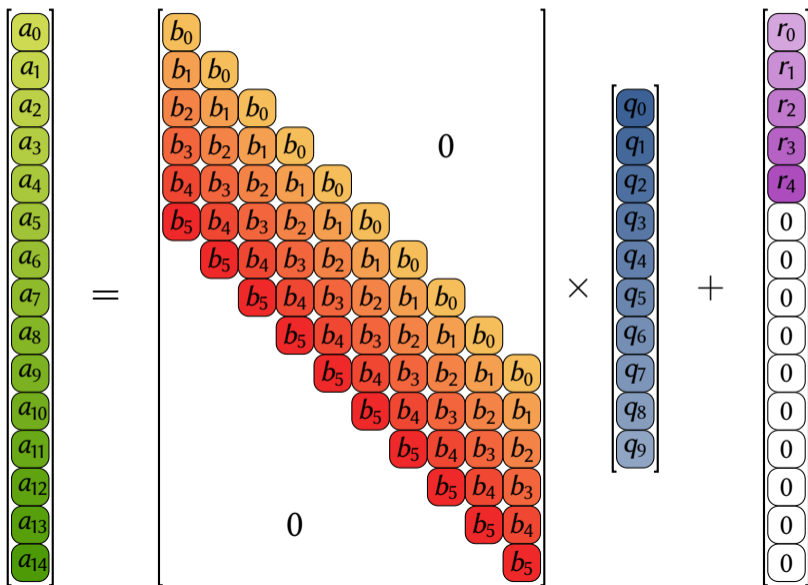
$^\dagger O(\log n)$ pointers for call stack

Contents

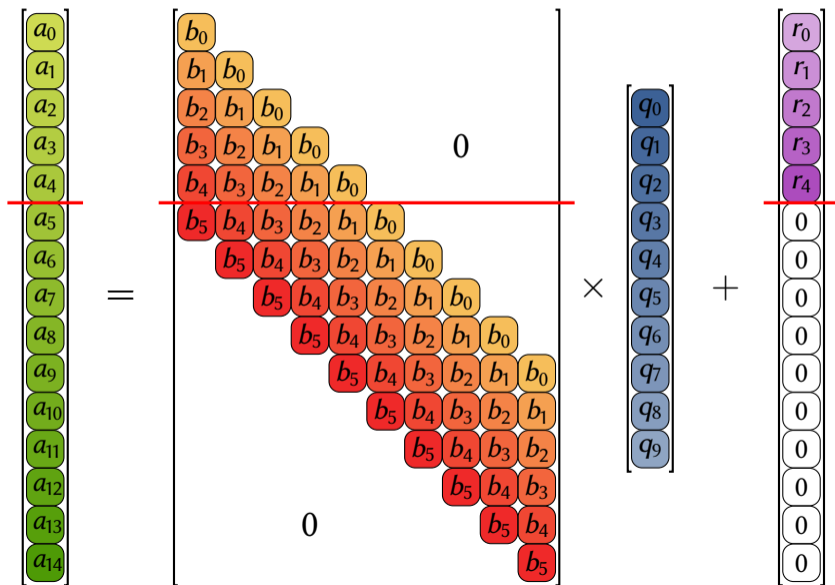
1. Toeplitz computations and convolutions

2. Remainder algorithms

$A = B \times Q + R$ as matrix-vector products



$A = B \times Q + R$ as matrix-vector products



$A = B \times Q + R$ as matrix-vector products

$$\begin{array}{c}
 \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \\
 \\
 \begin{bmatrix} a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{bmatrix}
 \end{array}
 =
 \begin{array}{c}
 \begin{bmatrix} b_0 & & & & & \\ b_1 & b_0 & & & & \\ b_2 & b_1 & b_0 & & & \\ b_3 & b_2 & b_1 & b_0 & & \\ b_4 & b_3 & b_2 & b_1 & b_0 & \\ & & & & & 0 \end{bmatrix} \\
 \\
 \begin{bmatrix} b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & & \\ & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & \\ & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & \\ & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & \\ & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & \\ & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & \\ & & & & & & b_5 & b_4 & b_3 & b_2 & \\ & & & & & & & b_5 & b_4 & b_3 & \\ & & & & & & & & b_5 & b_4 & \\ & & & & & & & & & b_5 & \\ & & & & & & & & & & 0 \end{bmatrix}
 \end{array}
 \times
 \begin{array}{c}
 \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_8 \\ q_9 \end{bmatrix} \\
 \\
 \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \\ q_8 \\ q_9 \end{bmatrix}
 \end{array}
 +
 \begin{array}{c}
 \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} \\
 \\
 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
 \end{array}$$

$A = B \times Q + R$ as matrix-vector products

$$\begin{array}{c}
 \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \\
 \\
 \begin{bmatrix} a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{bmatrix}
 \end{array}
 =
 \begin{array}{c}
 \begin{bmatrix} b_0 \\ b_1 & b_0 \\ b_2 & b_1 & b_0 \\ b_3 & b_2 & b_1 & b_0 \\ b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix} \\
 \\
 \begin{bmatrix} b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & b_5 & b_4 & b_3 & b_2 & b_1 \\ & & & & & & b_5 & b_4 & b_3 & b_2 \\ & & & & & & & b_5 & b_4 & b_3 \\ & & & & & & & & b_5 & b_4 \\ & & & & & & & & & b_5 \end{bmatrix} \\
 \\
 0
 \end{array}
 \times
 \begin{array}{c}
 \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \\
 \\
 \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \\ q_8 \\ q_9 \end{bmatrix}
 \end{array}
 +
 \begin{array}{c}
 \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} \\
 \\
 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
 \end{array}$$

$A = B \times Q + R$ as matrix-vector products

The diagram illustrates the equation $A = B \times Q + R$ as matrix-vector products in two examples.

Example 1:

- Column vector A (green boxes): a_0, a_1, a_2, a_3, a_4
- Lower triangular matrix B (orange boxes):

$$\begin{bmatrix} b_0 & & & & \\ b_1 & b_0 & & & \\ b_2 & b_1 & b_0 & & \\ b_3 & b_2 & b_1 & b_0 & \\ b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix}$$
- Column vector Q (blue boxes): q_0, q_1, q_2, q_3, q_4
- Column vector R (purple boxes): r_0, r_1, r_2, r_3, r_4

Example 2:

- Column vector A (green boxes): $a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{13}, a_{14}$
- Lower triangular matrix B (orange boxes):

$$\begin{bmatrix} b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & & \\ & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & \\ & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & \\ & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & \\ & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & \\ & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & \\ & & & & & & b_5 & b_4 & b_3 & b_2 & \\ & & & & & & & b_5 & b_4 & b_3 & \\ & & & & & & & & b_5 & b_4 & \\ & & & & & & & & & b_5 & \\ & & & & & & & & & & b_5 \end{bmatrix}$$
- Column vector Q (blue boxes): $q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9$
- Column vector R (purple boxes): $r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9$
- A zero 0 is shown below the matrix B in the second example.

A matrix formula for the remainder

$$\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} - \begin{bmatrix} b_0 & & & & & \\ b_1 & b_0 & & & & \\ b_2 & b_1 & b_0 & & & \\ b_3 & b_2 & b_1 & b_0 & & \\ b_4 & b_3 & b_2 & b_1 & b_0 & \end{bmatrix} \times \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}$$

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \\ q_8 \\ q_9 \end{bmatrix} = \begin{bmatrix} b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & & \\ & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & & \\ & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & & \\ & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & & \\ & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & \\ & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 \\ & & & & & & & b_5 & b_4 & b_3 & b_2 \\ & & & & & & & & b_5 & b_4 & b_3 \\ & & & & & & & & & b_5 & b_4 \\ & & & & & & & & & & b_5 \\ & & & & & & & & & & & 0 \end{bmatrix}^{-1} \times \begin{bmatrix} a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{bmatrix}$$

A matrix formula for the remainder

$$\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} - \begin{bmatrix} b_0 \\ b_1 & b_0 \\ b_2 & b_1 & b_0 \\ b_3 & b_2 & b_1 & b_0 \\ b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix} \times \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \times \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} \times \left[\begin{bmatrix} b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \\ & & & & & & & & & & & & & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix}^{-1} \times \begin{bmatrix} a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{bmatrix}$$

A matrix formula for the remainder

$$\begin{bmatrix} R \end{bmatrix} = \begin{bmatrix} A_0 \end{bmatrix} - \begin{bmatrix} G \end{bmatrix} \times \begin{bmatrix} 0 \\ \times \begin{bmatrix} T & G \\ & T \end{bmatrix}^{-1} \times \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \end{bmatrix}$$

A matrix formula for the remainder

$$\begin{bmatrix} R \end{bmatrix} = \begin{bmatrix} A_0 \end{bmatrix} - \begin{bmatrix} G \end{bmatrix} \times \begin{bmatrix} 0 \\ \times \begin{bmatrix} T^{-1} & -T^{-1}GT^{-1} \\ 0 & T^{-1} \end{bmatrix} \times \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \end{bmatrix}$$

A matrix formula for the remainder

$$\begin{bmatrix} R \end{bmatrix} = \begin{bmatrix} A_0 \end{bmatrix} - \begin{bmatrix} GT^{-1} & -(GT^{-1})^2 \end{bmatrix} \times \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

A matrix formula for the remainder

$$\begin{bmatrix} R \end{bmatrix} = \begin{bmatrix} A_0 \end{bmatrix} - \begin{bmatrix} GT^{-1} & -(GT^{-1})^2 \end{bmatrix} \times \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

Theorem

$$\begin{bmatrix} R \end{bmatrix} = \sum_{i=0}^{\mu} \begin{bmatrix} -GT^{-1} \end{bmatrix}^i \cdot \begin{bmatrix} A_i \end{bmatrix}$$

→ Horner-like computation

A “not-in-place” algorithm

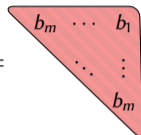
Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m

Output: $R = A \bmod B$ of degree $< m$

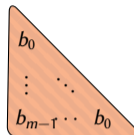
(model ro/rw)

0. $\mu = \lceil (n - m + 1) / m \rceil$

$A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



; $G =$



1. $R := A_\mu$

2. for $i = \mu - 1$ to 0:

3. $Q := T^{-1} \cdot R$

4. $R := -G \cdot Q$

5. $R += A_i$

6. return R

A “not-in-place” algorithm

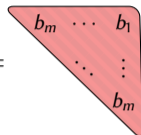
Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m

Output: $R = A \bmod B$ of degree $< m$

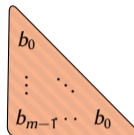
(model ro/rw)

0. $\mu = \lceil (n - m + 1) / m \rceil$

$A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



; $G =$



1. $R := A_\mu$

2. for $i = \mu - 1$ to 0 :

3. $Q := T^{-1} \cdot R$

4. $R := -G \cdot Q$

5. $R += A_i$

6. return R

▶ Time: $O(n/m \cdot M(m))$

▶ Space: $m + O(1)$

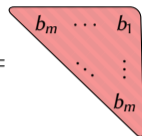
▶ Remark: similar result in [Giorgi-G.-Roche'20]

storage of Q

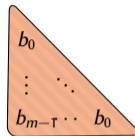
An *in-place* algorithm

Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m **A: read-only** / **B: read-write**
Output: $R = A \bmod B$ of degree $< m$

0. $\mu = \lceil (n - m + 1) / m \rceil$
 $A = \boxed{A_0} \boxed{A_1} \cdots \boxed{A_\mu}$ (blocks of size m); $T =$



; $G =$

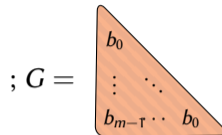
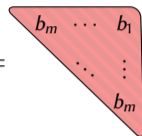


1. $R := \boxed{A_\mu}$
2. for $i = \mu - 1$ to 0 :
3. $R := \boxed{R} \cdot \boxed{T}^{-1}$
4. $R := \boxed{R} - \boxed{G}$
5. $R := \boxed{R} + \boxed{A_i}$
6. return \boxed{R}

An *in-place* algorithm

Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m **A: read-only** / B: read-write
 Output: $R = A \bmod B$ of degree $< m$

0. $\mu = \lceil (n - m + 1) / m \rceil$
 $A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



1. $R := A_\mu$
2. for $i = \mu - 1$ to 0 :
3. $R := R - T^{-1} A_i$
4. $R := R + G A_i$
5. $R := R + A_i$
6. return R

- ▶ Time: $O(n/m \cdot M(m) \log m)^*$
- ▶ Space: $O(1)^\dagger$

- * $O(M(m))$ if $M(m) = \Omega(m^{1+\epsilon})$
- † $O(\log m)$ pointers for call stack

Over-place Euclidean division

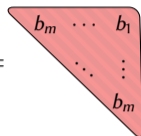
Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m

Output: $A \leftarrow R \cdot Q$ s.t. $A = BQ + R$ with R of degree $< m$

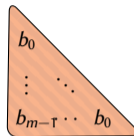
A, B : read-write

0. $\mu = \lceil (n - m + 1) / m \rceil$

$A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



; $G =$



1. for $i = \mu - 1$ to 0:

2. $A_{i+1} *= T^{-1}$

3. $A_i -= G \cdot A_{i+1}$

4. return $R = A_0$ and $Q = A_1 \cdots A_\mu$

simplification: m divides $n - 1$

Over-place Euclidean division

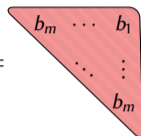
Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m

Output: $A \leftarrow RQ$ s.t. $A = BQ + R$ with R of degree $< m$

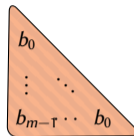
A, B : read-write

0. $\mu = \lceil (n - m + 1) / m \rceil$

$A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



; $G =$



simplification: m divides $n - 1$

1. for $i = \mu - 1$ to 0:

2. $A_{i+1} *= T^{-1}$

3. $A_i -= G \cdot A_{i+1}$

4. return $R = A_0$ and $Q = A_1 \cdots A_\mu$

▶ Time: $O(n/m \cdot M(m) \log m)^*$

▶ Space: $O(1)^\dagger$

▶ Remark: the algorithm is **reversible**

* $O(n/m \cdot M(m))$ if $M(m) = \Omega(m^{1+\epsilon})$

† $O(\log m)$ pointers for call stack

Over-place Euclidean division

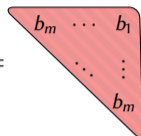
Inputs: $A \in \mathbb{F}[x]$ of degree n , $B \in \mathbb{F}[x]$ of degree m

A, B : read-write

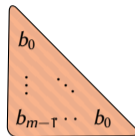
Output: $A \leftarrow R \cdot Q$ s.t. $A = BQ + R$ with R of degree $< m$

0. $\mu = \lceil (n - m + 1) / m \rceil$

$A = A_0 A_1 \cdots A_\mu$ (blocks of size m); $T =$



; $G =$



simplification: m divides $n - 1$

1. for $i = \mu - 1$ to 0:

2. $A_{i+1} *= T^{-1}$

3. $A_i -= G \cdot A_{i+1}$

4. return $R = A_0$ and $Q = A_1 \cdots A_\mu$

▶ Time: $O(n/m \cdot M(m) \log m)^*$

▶ Space: $O(1)^\dagger$

▶ Remark: the algorithm is **reversible**

$C += A \bmod B$

1. Replace A by $R|Q$

2. $C += R$

3. Restore A

* $O(n/m \cdot M(m))$ if $M(m) = \Omega(m^{1+\epsilon})$

† $O(\log m)$ pointers for call stack

Accumulated modular multiplication

Inputs: $A, B, R \in \mathbb{F}[x]$ of degrees $< m$; $C \in \mathbb{F}[x]$ of degree m
Output: $R += A \cdot B \bmod C$

Accumulated modular multiplication

Inputs: $A, B, R \in \mathbb{F}[x]$ of degrees $< m$; $C \in \mathbb{F}[x]$ of degree m
 Output: $R += A \cdot B \bmod C$

An adapted remainder formula

$$\begin{array}{c} \boxed{R} \end{array} += \underbrace{\begin{array}{c} \boxed{A_L} \\ \text{---} \\ \boxed{A_U} \end{array}}_{(AB)_0} \cdot \begin{array}{c} \boxed{B} \end{array} - \boxed{GT^{-1}} \times \underbrace{\begin{array}{c} \boxed{A_U} \\ \text{---} \\ \boxed{A_L} \end{array}}_{(AB)_1} \cdot \begin{array}{c} \boxed{B} \end{array}$$

Algorithm

1. $R += A_L \cdot B$
2. $B *= A_U$
3. $B *= T^{-1}$
4. $R -= G \cdot B$
5. $B *= T$ *undo 3.*
6. $B *= A_U^{-1}$ *undo 2.*

Accumulated modular multiplication

Inputs: $A, B, R \in \mathbb{F}[x]$ of degrees $< m$; $C \in \mathbb{F}[x]$ of degree m
 Output: $R += A \cdot B \bmod C$

An adapted remainder formula

$$\begin{array}{c} \boxed{R} \end{array} += \underbrace{\begin{array}{c} \boxed{A_L} \\ \text{---} \\ \boxed{A_U} \end{array}}_{(AB)_0} \cdot \begin{array}{c} \boxed{B} \end{array} - \boxed{GT^{-1}} \times \underbrace{\begin{array}{c} \boxed{A_U} \\ \text{---} \\ \boxed{A_L} \end{array}}_{(AB)_1} \cdot \begin{array}{c} \boxed{B} \end{array}$$

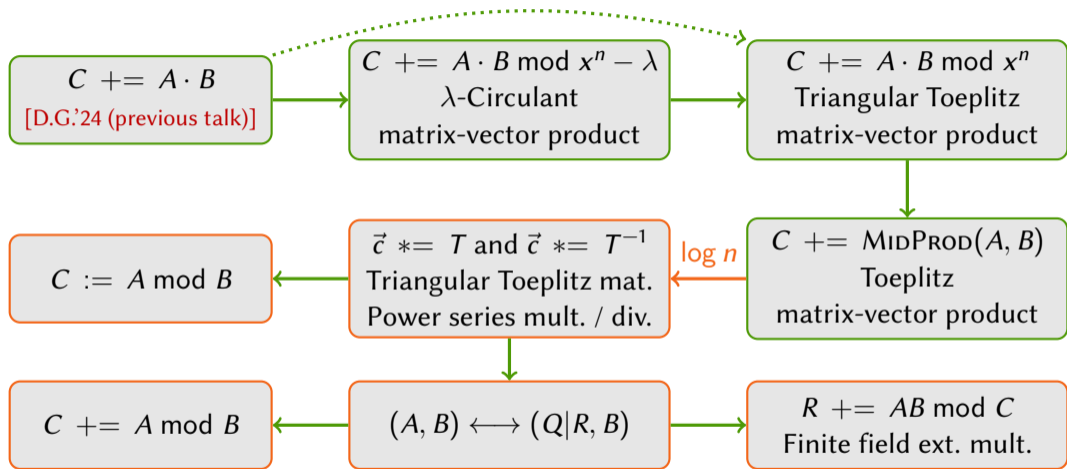
Algorithm

1. $R += A_L \cdot B$
2. $B *= A_U$
3. $B *= T^{-1}$
4. $R -= G \cdot B$
5. $B *= T$ *undo 3.*
6. $B *= A_U^{-1}$ *undo 2.*

- ▶ Time: $O(M(m) \log m)^*$
- ▶ Space: $O(1)^\dagger$
- ▶ Generalization to all degrees

* $O(M(m))$ if $M(m) = \Omega(m^{1+\epsilon})$
 † $O(\log m)$ pointers for call stack

Summary of reductions



□ Time: $O(M(n))$
Space: $O(1)$

□ Time: $O(M(n) \log n)^*$
Space: $O(1)^\dagger$

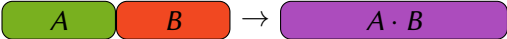
* $O(M(n))$ if $M(n) = \Omega(n^{1+\epsilon})$
† $O(\log n)$ pointers for call stack

Open problems

Non-optimal results

- ▶ $C := A \bmod B, \dots$ in $O(M(n))$
- ▶ Avoid the call stack
- ▶ Bypass the over-place computations

Other models and lower bounds

- ▶ What can be done in the model ro/rw ?
- ▶ Time-space lower bounds in models ro/rw and rw/rw ?
- ▶ Over-place computation: 

[Roche'09]

Practical efficiency

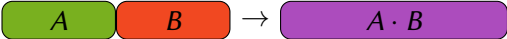
- ▶ No implementation yet
- ▶ Encouraging results in [D.G.'24 (previous talk)]

Open problems

Non-optimal results

- ▶ $C := A \bmod B, \dots$ in $O(M(n))$
- ▶ Avoid the call stack
- ▶ Bypass the over-place computations

Other models and lower bounds

- ▶ What can be done in the model ro/rw ?
- ▶ Time-space lower bounds in models ro/rw and rw/rw ?
- ▶ Over-place computation: 

[Roche'09]

Practical efficiency

- ▶ No implementation yet
- ▶ Encouraging results in [D.G.'24 (previous talk)]

Thank you!