On Arithmetically Progressed Suffix Arrays

Jacqueline W. Daykin[1] Dominik Köppl[2] David Kübel[3]
Florian Stober[4]

[1]Department of Computer Science, Aberystwyth University, UK; Department of Information Science, Stellenbosch University, South Africa

[2]Kyushu University, Japan Society for Promotion of Science, Japan

[3]University of Bonn, Institute of Computer Science, Germany

[4]University of Stuttgart, FMI, Germany

PSC 2020

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

1	abaababa
2	baababa
3	aababa
4	ababa
5	baba
6	aba
7	ba
8	a

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

8	a
3	aababa
6	aba
1	abaababa
4	ababa
7	ba
2	baababa
5	baba

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$\mathsf{SA_T} = [\,8\,,\,3\,,\,6\,,\,1\,,\,4\,,\,7\,,\,2\,,\,5\,]$

8	`a`
3	`aababa`
6	`aba`
1	`abaababa`
4	`ababa`
7	`ba`
2	`baababa`
5	`baba`

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$\mathsf{SA_T} = [8, 3, 6, 1, 4, 7, 2, 5]$

8	a
3	aababa
6	aba
1	abaababa
4	ababa
7	ba
2	baababa
5	baba

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$$\mathsf{SA_T} = [\,8\,, \underbrace{}_{+3}\,3\,, 6\,, 1\,, 4\,, 7\,, 2\,, 5\,]$$

8	a
3	aababa
6	aba
1	abaababa
4	ababa
7	ba
2	baababa
5	baba

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$\mathsf{SA_T} = [\underbrace{8 , 3}_{+3} , \underbrace{3 , 6}_{+3} , 1 , 4 , 7 , 2 , 5]$

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.

8	a
3	aababa
6	aba
1	abaababa
4	ababa
7	ba
2	baababa
5	baba

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$$\mathsf{SA_T} = [\,8\,,\,3\,,\,6\,,\,1\,,\,4\,,\,7\,,\,2\,,\,5\,]$$

$+3\,+3$

8	`a`
3	`aababa`
6	`aba`
1	`abaababa`
4	`ababa`
7	`ba`
2	`baababa`
5	`baba`

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.
- A permutation P is arithmetically progressed if $\exists_k \forall_i : P[i + 1 \bmod n] = P[i] + k \bmod n$.

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$\mathsf{SA_T} = [\underbrace{8 , 3}_{+3} , \underbrace{3 , 6}_{+3} , 6 , 1 , 4 , 7 , 2 , 5]$

8	a
3	aababa
6	aba
1	abaababa
4	ababa
7	ba
2	baababa
5	baba

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.
- A permutation P is arithmetically progressed if $\exists_k \forall_i : P[i + 1 \bmod n] = P[i] + k \bmod n$.
- k is co-prime to n.

$\mathsf{T} = \texttt{abaababa}$
$n = |T|$

$$\mathsf{SA_T} = [\,8\,,\,3\,,\,6\,,\,1\,,\,4\,,\,7\,,\,2\,,\,5\,]$$

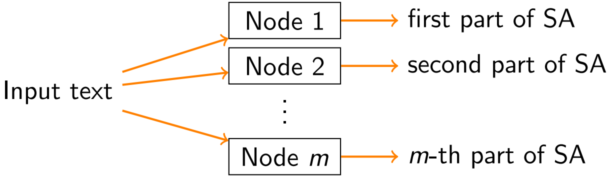
+3 +3

8	`a`
3	`aababa`
6	`aba`
1	`abaababa`
4	`ababa`
7	`ba`
2	`baababa`
5	`baba`

- $\forall_T$: $\mathsf{SA_T}$ is a permutation.
- A permutation P is arithmetically progressed if $\exists_k \forall_i : P[i+1 \bmod n] = P[i] + k \bmod n$.
- k is co-prime to n.
- $P[i] = P[1] + (i-1)k \bmod n$.

Motivation and Outline

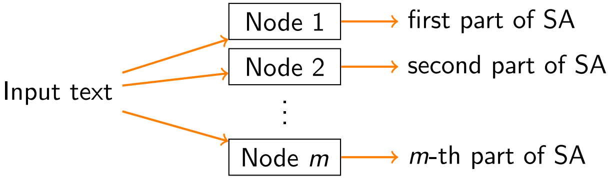
- **Scenario:** distributed suffix array construction



Test data with arithmetically progressed suffix array allows correctness of the result to be verified locally on each node.

Motivation and Outline

- **Scenario:** distributed suffix array construction

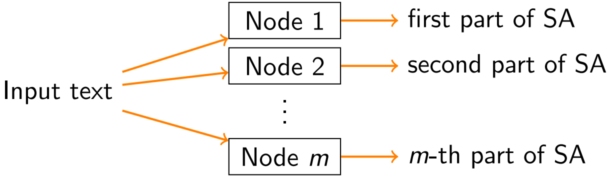


  Test data with arithmetically progressed suffix array allows correctness of the result to be verified locally on each node.

- **Outline:**
  1. Characterize all strings whose suffix arrays are arithmetically progressed.

Motivation and Outline

- **Scenario:** distributed suffix array construction



  Test data with arithmetically progressed suffix array allows correctness of the result to be verified locally on each node.

- **Outline:**
  1. Characterize all strings whose suffix arrays are arithmetically progressed.
  2. Describe the Burrows-Wheeler Transform (BWT) of those strings. Many have a simple BWT.

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \boxed{\quad p_1 \quad}\boxed{\qquad\qquad\qquad \cdots \qquad\qquad\qquad}\boxed{\quad p_n \quad}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T over a ternary *alphabet such that $\mathsf{SA_T} = P$.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1}\boxed{\cdots}\boxed{p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1}\boxed{\cdots}\boxed{n{-}k}}_{B}\underbrace{\boxed{n}\boxed{\cdots}\boxed{p_n}}_{C}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \dots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \dots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

- Example $P = [\,2\,,\,7\,,\,4\,,\,1\,,\,6\,,\,3\,,\,8\,,\,5\,]$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\,2\,,\,7\,,\,4\,|\,1\,,\,6\,,\,3\,,\,8\,,\,5\,]$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T over a ternary *alphabet such that $\mathsf{SA}_\mathsf{T} = P$.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \,|\, \cdots \,|\, p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \,|\, \cdots \,|\, n{-}k}}_{B}\underbrace{\boxed{n \,|\, \cdots \,|\, p_n}}_{C}$$

- Example $P = [\,2\,,\,7\,,\,4\,|\,1\,,\,6\,,\,3\,|\,8\,,\,5\,]$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathsf{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

- Example $P = [\underbrace{2, 7, 4}_{A} \mid \underbrace{1, 6, 3}_{B} \mid \underbrace{8, 5}_{C}]$

$$T = \underset{1}{\_}\,\underset{2}{\_}\,\underset{3}{\_}\,\underset{4}{\_}\,\underset{5}{\_}\,\underset{6}{\_}\,\underset{7}{\_}\,\underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathsf{a} & \text{if } p_i \in A, \text{ or} \\ \mathsf{b} & \text{if } p_i \in B, \text{ or} \\ \mathsf{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathsf{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A}\,|\,\underbrace{1\,,\,6\,,\,3}_{B}\,|\,\underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\_}\,\underset{2}{\mathsf{a}}\,\underset{3}{\_}\,\underset{4}{\_}\,\underset{5}{\_}\,\underset{6}{\_}\,\underset{7}{\_}\,\underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathsf{a} & \text{if } p_i \in A, \text{ or} \\ \mathsf{b} & \text{if } p_i \in B, \text{ or} \\ \mathsf{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \mid \underbrace{1\,,\,6\,,\,3}_{B} \mid \underbrace{8\,,\,5}_{C}]$

$T =$

_	a	_	_	_	_	a	_
1	2	3	4	5	6	7	8

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \,|\, \underbrace{1\,,\,6\,,\,3}_{B} \,|\, \underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\_}\ \underset{2}{\mathtt{a}}\ \underset{3}{\_}\ \underset{4}{\mathtt{a}}\ \underset{5}{\_}\ \underset{6}{\_}\ \underset{7}{\mathtt{a}}\ \underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2, 7, 4}_{A} | \underbrace{1, 6, 3}_{B} | \underbrace{8, 5}_{C}]$

$$T = \underset{1}{\mathtt{b}}\,\underset{2}{\mathtt{a}}\,\underset{3}{\_}\,\underset{4}{\mathtt{a}}\,\underset{5}{\_}\,\underset{6}{\_}\,\underset{7}{\mathtt{a}}\,\underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A}\ \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B}\ \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A}\,|\,\underbrace{1\,,\,6\,,\,3}_{B}\,|\,\underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\mathtt{b}}\,\underset{2}{\mathtt{a}}\,\underset{3}{\_}\,\underset{4}{\mathtt{a}}\,\underset{5}{\_}\,\underset{6}{\mathtt{b}}\,\underset{7}{\mathtt{a}}\,\underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \,|\, \underbrace{1\,,\,6\,,\,3}_{B} \,|\, \underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\mathtt{b}}\,\underset{2}{\mathtt{a}}\,\underset{3}{\mathtt{b}}\,\underset{4}{\mathtt{a}}\,\underset{5}{\_}\,\underset{6}{\mathtt{b}}\,\underset{7}{\mathtt{a}}\,\underset{8}{\_}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k, there exists a string T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \mid \underbrace{1\,,\,6\,,\,3}_{B} \mid \underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\mathtt{b}}\,\underset{2}{\mathtt{a}}\,\underset{3}{\mathtt{b}}\,\underset{4}{\mathtt{a}}\,\underset{5}{\_}\,\underset{6}{\mathtt{b}}\,\underset{7}{\mathtt{a}}\,\underset{8}{\mathtt{c}}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A \text{, or} \\ \mathtt{b} & \text{if } p_i \in B \text{, or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

c

Theorem (4)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k*, there exists a string* T *over a* ternary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- For $P = [n, n-1, \ldots, 1]$ we have $\mathsf{T} = \mathtt{a}^n$.
- Else split P into subarrays

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \,|\, \underbrace{1\,,\,6\,,\,3}_{B} \,|\, \underbrace{8\,,\,5}_{C}]$

$$T = \underset{1}{\mathtt{b}}\,\underset{2}{\mathtt{a}}\,\underset{3}{\mathtt{b}}\,\underset{4}{\mathtt{a}}\,\underset{5}{\mathtt{c}}\,\underset{6}{\mathtt{b}}\,\underset{7}{\mathtt{a}}\,\underset{8}{\mathtt{c}}$$

$$\mathsf{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A \text{, or} \\ \mathtt{b} & \text{if } p_i \in B \text{, or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A}\ \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B}\ \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \,|\, \underbrace{1\,,\,6\,,\,3}_{B} \,|\, \underbrace{8\,,\,5}_{C}]\ ,\ T = \texttt{b a b a c b a c}$

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A}\,|\,\underbrace{1\,,\,6\,,\,3}_{B}\,|\,\underbrace{8\,,\,5}_{C}]$, $T = \texttt{b a b a c b a c}$

2	`abacbac`
7	`ac`
4	`acbac`
1	`babacbac`
6	`bac`
3	`bacbac`
8	`c`
5	`cbac`

$$P = \underbrace{\begin{array}{|c|c|c|}\hline p_1 & \cdots & p_1{-}k{-}1 \\ \hline\end{array}}_{A}\underbrace{\begin{array}{|c|c|c|}\hline p_1{-}1 & \cdots & n{-}k \\ \hline\end{array}}_{B}\underbrace{\begin{array}{|c|c|c|}\hline n & \cdots & p_n \\ \hline\end{array}}_{C}$$

- Example $P = [\underbrace{2, 7, 4}_{A} \mid \underbrace{1, 6, 3}_{B} \mid \underbrace{8, 5}_{C}]$, $T =$ `babacbac`

Split between $n{-}k$ and n, else $T[n]$ is a prefix of $T[n{-}k..n]$. $\longrightarrow$

2	`abacbac`
7	`ac`
4	`acbac`
1	`babacbac`
6	`bac`
3	`bacbac`
8	`c`
5	`cbac`

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2, 7, 4}_{A} \mid \underbrace{1, 6, 3}_{B} \mid \underbrace{8, 5}_{C}]$, $T = \mathtt{b\,a\,b\,a\,c\,b\,a\,c}$

Split between $p_1{-}k{-}1$ and $p_1{-}1$ as $T[p_1{-}k]{=}T[p_n]{>}T[p_1]$. $\longrightarrow$

Split between $n{-}k$ and n, else $T[n]$ is a prefix of $T[n{-}k..n]$. $\longrightarrow$

2	`abacbac`
7	`ac`
4	`acbac`
1	`babacbac`
6	`bac`
3	`bacbac`
8	`c`
5	`cbac`

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

- Example $P = [\underbrace{2\,,\,7\,,\,4}_{A} \,|\, \underbrace{1\,,\,6\,,\,3}_{B} \,|\, \underbrace{8\,,\,5}_{C}]$, $T = \texttt{babacbac}$

Split between $p_1{-}k{-}1$ and $p_1{-}1$ as $T[p_1{-}k] = {\color{magenta}T[p_n]} > {\color{cyan}T[p_1]}$. $\longrightarrow$

Split between $n{-}k$ and n, else $T[n]$ is a prefix of $T[n{-}k..n]$. $\longrightarrow$

2	abacbac
7	ac
4	acbac
1	babacbac
6	bac
3	bacbac
8	c
5	cbac

- Note: If $n < p_1 - 1 \bmod n$ then

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{n{-}k}}_{A} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{B} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{p_n}}_{C}$$

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

$$\mathrm{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (8)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k *such that* $p_1 \notin \{1, k+1, n\}$, *the string* T *is unique.*

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

$$\mathrm{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A \text{, or} \\ \mathtt{b} & \text{if } p_i \in B \text{, or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (8)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k such that $p_1 \notin \{1, k+1, n\}$, the string T is unique.

- Assume that there is another string $\mathrm{S} \neq \mathrm{T}$ with $\mathrm{SA}_\mathrm{S} = P$.

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

$$\mathrm{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A \text{, or} \\ \mathtt{b} & \text{if } p_i \in B \text{, or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (8)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ with ratio k such that $p_1 \notin \{1, k+1, n\}$, the string T is unique.

- Assume that there is another string $\mathrm{S} \neq \mathrm{T}$ with $\mathrm{SA_S} = P$.
- $\mathrm{SA_S}$ with can be split into subarrays A, B and C.

$$P = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

$$\mathrm{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A \text{, or} \\ \mathtt{b} & \text{if } p_i \in B \text{, or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (8)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k *such that* $p_1 \notin \{1, k+1, n\}$*, the string* T *is unique.*

- Assume that there is another string $\mathrm{S} \neq \mathrm{T}$ with $\mathrm{SA}_\mathrm{S} = P$.
- SA_S with can be split into subarrays A, B and C.
- $\mathrm{S} \neq \mathrm{T}$ implies different splitting positions.

$$P = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$$

$$\mathrm{T}[p_i] := \begin{cases} \mathtt{a} & \text{if } p_i \in A, \text{ or} \\ \mathtt{b} & \text{if } p_i \in B, \text{ or} \\ \mathtt{c} & \text{if } p_i \in C. \end{cases}$$

Theorem (8)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ *with ratio* k *such that* $p_1 \notin \{1, k+1, n\}$, *the string* T *is unique.*

- Assume that there is another string $\mathrm{S} \neq \mathrm{T}$ with $\mathrm{SA}_\mathrm{S} = P$.
- SA_S with can be split into subarrays A, B and C.
- $\mathrm{S} \neq \mathrm{T}$ implies different splitting positions.
- $\mathrm{S}[p_1 - k - 1] = \mathrm{S}[p_1 - 1]$ or $\mathrm{S}[n - k] = \mathrm{S}[n]$ lead to the contradiction $\mathrm{SA}_\mathrm{S} \neq P$.

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n] \neq [n, n-1, \ldots, 1]$ *with ratio* k*, such that* $p_1 \in \{1, k+1, n\}$*, there exists a* unique *string* T *over a* binary *alphabet such that* $\mathsf{SA}_{\mathsf{T}} = P$*.*

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n]$ $\neq [n, n-1, \ldots, 1]$ *with ratio* k*, such that* $p_1 \in \{1, k+1, n\}$*, there exists a* unique *string* T *over a* binary *alphabet such that* $\mathsf{SA_T} = P$*.*

- $p_1 \in \{1, n\}$: One of the subarrays is empty.

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n] \neq [n, n-1, \ldots, 1]$ *with ratio* k*, such that* $p_1 \in \{1, k+1, n\}$*, there exists a* unique *string* T *over a* binary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- $p_1 \in \{1, n\}$: One of the subarrays is empty.
- $p_1 = k + 1$:

$$P = \underbrace{\begin{array}{|c|c|c|}\hline p_1{=}k{+}1 & \cdots & n{-}k \\ \hline\end{array}}_{A}\underbrace{\begin{array}{|c|}\hline n \\ \hline\end{array}}_{B}\underbrace{\begin{array}{|c|c|c|}\hline k{=}p_1{-}1 & \cdots & p_n \\ \hline\end{array}}_{C}$$

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n] \neq [n, n-1, \ldots, 1]$ *with ratio* k, *such that* $p_1 \in \{1, k+1, n\}$, *there exists a* unique *string* T *over a* binary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- $p_1 \in \{1, n\}$: One of the subarrays is empty.
- $p_1 = k + 1$:

$$P = \underbrace{\begin{array}{|c|c|c|}\hline p_1{=}k{+}1 & \cdots & n{-}k \\ \hline\end{array}}_{A}\underbrace{\begin{array}{|c|}\hline n \\ \hline\end{array}}_{B}\underbrace{\begin{array}{|c|c|c|}\hline k{=}p_1{-}1 & \cdots & p_n \\ \hline\end{array}}_{C}$$

Example:
$P = [6, 3 \mid 8 \mid 5, 2, 7, 4, 1]$
T = `ccaccacb`

6	`acb`
3	`accacb`
8	`b`
5	`cacb`
2	`caccacb`
7	`cb`
4	`ccacb`
1	`ccaccacb`

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n] \neq [n, n-1, \ldots, 1]$ *with ratio* k*, such that* $p_1 \in \{1, k+1, n\}$*, there exists a* unique *string* T *over a* binary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$*.*

- $p_1 \in \{1, n\}$: One of the subarrays is empty.
- $p_1 = k + 1$:

$$P = \underbrace{\boxed{p_1{=}k{+}1} \boxed{\cdots} \boxed{n{-}k}}_{A} \underbrace{\boxed{n}}_{B} \underbrace{\boxed{k{=}p_1{-}1} \boxed{\cdots} \boxed{p_n}}_{C}$$

Example:

$P = [6, 3 \mid 8 \mid 5, 2, 7, 4, 1]$

T $=$ `ccaccacb`

6	acb
3	accacb
8	b
5	cacb
2	caccacb
7	cb
4	ccacb
1	ccaccacb

Theorem (10,11)

Given an arithmetically progressed permutation $P := [p_1, \ldots, p_n] \neq [n, n-1, \ldots, 1]$ with ratio k, such that $p_1 \in \{1, k+1, n\}$, there exists a unique *string* T *over a* binary *alphabet such that* $\mathsf{SA}_\mathsf{T} = P$.

- $p_1 \in \{1, n\}$: One of the subarrays is empty.
- $p_1 = k+1$:

$$P = \underbrace{\boxed{p_1{=}k{+}1 \mid \cdots \mid n{-}k}}_{A}\underbrace{\boxed{n \mid k{=}p_1{-}1 \mid \cdots \mid p_n}}_{B}$$

Example:
$P = [6, 3 \mid 8, 5, 2, 7, 4, 1]$
$\mathsf{T} = \texttt{bbabbabb}$

6	abb
3	abbabb
8	b
5	babb
2	babbabb
7	bb
4	bbabb
1	bbabbabb

Relation between Permutation and Min. Alphabet Size

p_1	k	Min. Size of Σ	Properties
1		2	Lyndon word, simple BWT
$k+1$		2	period $(n-k)$*
n	$\neq (n-1)$	2	period $(n-k)$*, simple BWT
n	$= (n-1)$	1	trivially periodic*, simple BWT
$\notin \{1, k+1, n\}$		3	simple BWT

Characterization of strings over the alphabet Σ whose suffix array is an arithmetically progressed with ration k.

All words are unique over an alphabet of minimal size.

Properties marked with * only apply to the word over the minimal alphabet.

$\mathsf{T} = \texttt{babbabac}$

T = babbabac

BWT matrix

babbaba*c*
abbabac*b*
bbabacb*a*
babacba*b*
abacbab*b*
bacbabb*a*
acbabba*b*
cbabbab*a*

T = babbabac

BWT matrix

abacbab*b*
abbabac*b*
acbabba*b*
babacba*b*
babbaba*c*
bacbabb*a*
bbabacb*a*
cbabbab*a*

$T = \texttt{babbabac}$

BWT matrix

abacbab*b*
abbabac*b*
acbabba*b*
babacba*b*
babbaba*c*
bacbabb*a*
bbabacb*a*
cbabbab*a*

$\mathrm{BWT}_{\mathrm{matrix}} = \texttt{b}^4\texttt{ca}^3$

$\mathsf{T} = \mathtt{babbabac}$

BWT matrix

$\mathtt{abacbab}\mathit{b}$

$\mathtt{abbabac}\mathit{b}$

$\mathtt{acbabba}\mathit{b}$

$\mathtt{babacba}\mathit{b}$

$\mathtt{babbaba}\mathit{c}$

$\mathtt{bacbabb}\mathit{a}$

$\mathtt{bbabacb}\mathit{a}$

$\mathtt{cbabbab}\mathit{a}$

$\mathsf{SA} = [5, 2, 7, 4, 1, 6, 3, 8]$

$\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i] - 1 \bmod n]$

$\mathsf{BWT} = \mathtt{b}^4\mathtt{ca}^3$

$\mathsf{BWT}_{\mathsf{matrix}} = \mathtt{b}^4\mathtt{ca}^3$

$\mathsf{T} = \texttt{babbabac}$

BWT matrix

$\texttt{abacbab}b$
$\texttt{abbabac}b$
$\texttt{acbabba}b$
$\texttt{babacba}b$
$\texttt{babbaba}c$
$\texttt{bacbabb}a$
$\texttt{bbabacb}a$
$\texttt{cbabbab}a$

$\mathsf{SA} = [5, 2, 7, 4, 1, 6, 3, 8]$

$\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i] - 1 \bmod n]$

$\mathsf{BWT} = \texttt{b}^4\texttt{ca}^3$

$\mathsf{BWT}_{\mathsf{matrix}} = \texttt{b}^4\texttt{ca}^3$

Theorem (5)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$. *Then the BWT of* T *defined on the BWT matrix coincides with the BWT of* T *defined on the suffix array.*

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$*.*
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathsf{T}[\mathsf{SA}[1]] \cdots \mathsf{T}[\mathsf{SA}[n]]$*,*
i.e., $\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$*.*

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* SA $:= [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathrm{T}[p_1 - k - 1] \neq \mathrm{T}[p_1 - 1]$.
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathrm{T}[\mathrm{SA}[1]] \cdots \mathrm{T}[\mathrm{SA}[n]]$, *i.e.,* $\mathrm{BWT}[i] = \mathrm{T}[\mathrm{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$.

- SA =

p_1	$\cdots$	$p_t = p_1 - k - 1$	$p_1 - 1$	$\cdots$	p_n
T[SA[1]]	$\cdots$	T[SA[t]]	T[SA[t+1]]	$\cdots$	T[SA[n]]

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* SA $:= [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$.
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathsf{T}[\mathsf{SA}[1]] \cdots \mathsf{T}[\mathsf{SA}[n]]$,
i.e., $\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$.

- SA =

p_1	$\cdots$	$p_t = p_1 - k - 1$	$p_1 - 1$	$\cdots$	p_n
$\mathsf{T}[p_1]$	$\cdots$	$\mathsf{T}[p_1 - k - 1]$	$\mathsf{T}[p_1 - 1]$	$\cdots$	$\mathsf{T}[p_n]$

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$*.*
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathsf{T}[\mathsf{SA}[1]] \cdots \mathsf{T}[\mathsf{SA}[n]]$*, i.e.,* $\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$*.*

$\mathsf{SA} =$	p_1	$\cdots$	$p_t = p_1-k-1$	p_1-1	$\cdots$	p_n
	$\mathsf{T}[p_1]$	$\cdots$	$\mathsf{T}[p_1-k-1]$	$\mathsf{T}[p_1-1]$	$\cdots$	$\mathsf{T}[p_n]$
		rotation by t				
$\mathsf{BWT} =$	$\mathsf{T}[p_1-1]$	$\cdots$	$\mathsf{T}[p_n]$	$\mathsf{T}[p_1]$	$\cdots$	$\mathsf{T}[p_1-k-1]$

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* SA $:= [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathrm{T}[p_1 - k - 1] \neq \mathrm{T}[p_1 - 1]$.
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathrm{T}[\mathrm{SA}[1]] \cdots \mathrm{T}[\mathrm{SA}[n]]$, *i.e.,* $\mathrm{BWT}[i] = \mathrm{T}[\mathrm{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$.

SA =	p_1	$\cdots$	$p_t = p_1 - k - 1$	$p_1 - 1$	$\cdots$	p_n
	$\mathrm{T}[p_1]$	$\cdots$	$\mathrm{T}[p_1 - k - 1]$	$\mathrm{T}[p_1 - 1]$	$\cdots$	$\mathrm{T}[p_n]$
BWT =	$\mathrm{T}[p_1 - 1]$	$\cdots$	$\mathrm{T}[p_n]$	$\mathrm{T}[p_1]$	$\cdots$	$\mathrm{T}[p_1 - k - 1]$

  rotation by t
- Consider $P' := [p'_1, \ldots, p'_n]$ with $p'_i = p_i - 1 \bmod n$.

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$*.*
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathsf{T}[\mathsf{SA}[1]] \cdots \mathsf{T}[\mathsf{SA}[n]]$*, i.e.,* $\mathsf{BWT}[i] = \mathsf{T}[\mathsf{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$*.*

► SA =	p_1	$\cdots$	$p_t = p_1-k-1$	p_1-1	$\cdots$	p_n
	$\mathsf{T}[p_1]$	$\cdots$	$\mathsf{T}[p_1-k-1]$	$\mathsf{T}[p_1-1]$	$\cdots$	$\mathsf{T}[p_n]$
			rotation by t			
BWT =	$\mathsf{T}[p_1-1]$	$\cdots$	$\mathsf{T}[p_n]$	$\mathsf{T}[p_1]$	$\cdots$	$\mathsf{T}[p_1-k-1]$

- Consider $P' := [p'_1, \ldots, p'_n]$ with $p'_i = p_i - 1 \bmod n$.
- P' is arithmetically progressed with ratio k and $p'_1 = p_1 - 1$.

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* $\mathrm{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathrm{T}[p_1 - k - 1] \neq \mathrm{T}[p_1 - 1]$.
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathrm{T}[\mathrm{SA}[1]] \cdots \mathrm{T}[\mathrm{SA}[n]]$,
i.e., $\mathrm{BWT}[i] = \mathrm{T}[\mathrm{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$.

- SA =

SA =	p_1	$\cdots$	$p_t = p_1 - k - 1$	$p_1 - 1$	$\cdots$	p_n
	$\mathrm{T}[p_1]$	$\cdots$	$\mathrm{T}[p_1 - k - 1]$	$\mathrm{T}[p_1 - 1]$	$\cdots$	$\mathrm{T}[p_n]$

rotation by t

BWT =	$\mathrm{T}[p_1 - 1]$	$\cdots$	$\mathrm{T}[p_n]$	$\mathrm{T}[p_1]$	$\cdots$	$\mathrm{T}[p_1 - k - 1]$

- Consider $P' := [p'_1, \ldots, p'_n]$ with $p'_i = p_i - 1 \bmod n$.
- P' is arithmetically progressed with ratio k and $p'_1 = p_1 - 1$.
- P' is the t-th rotation of SA:

P' =	$p_1 - 1$	$\cdots$	p_n	p_1	$\cdots$	$p_1 - k - 1$

Lemma (6)

Let T *be a string with an arithmetically progressed suffix array* SA $:= [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathrm{T}[p_1 - k - 1] \neq \mathrm{T}[p_1 - 1]$.
Let t *be the index of* $p_1 - k - 1 \bmod n$ *in SA.*
Then the BWT *of* T *is the* t*-th rotation of* $\mathrm{T}[\mathrm{SA}[1]] \cdots \mathrm{T}[\mathrm{SA}[n]]$,
i.e., $\mathrm{BWT}[i] = \mathrm{T}[\mathrm{SA}[i + t \bmod n]]$ *for* $i \in [1..n]$.

- SA =

p_1	$\cdots$	$p_t = p_1{-}k{-}1$	$p_1{-}1$	$\cdots$	p_n
$\mathrm{T}[p_1]$	$\cdots$	$\mathrm{T}[p_1{-}k{-}1]$	$\mathrm{T}[p_1{-}1]$	$\cdots$	$\mathrm{T}[p_n]$

rotation by t

BWT = $\mathrm{T}[p_1{-}1]$ $\cdots$ $\mathrm{T}[p_n]$ $\mathrm{T}[p_1]$ $\cdots$ $\mathrm{T}[p_1{-}k{-}1]$

- Consider $P' := [p'_1, \ldots, p'_n]$ with $p'_i = p_i - 1 \bmod n$.
- P' is arithmetically progressed with ratio k and $p'_1 = p_1 - 1$.
- P' is the t-th rotation of SA:

P' =

$p_1{-}1$	$\cdots$	p_n	p_1	$\cdots$	$p_1{-}k{-}1$

- $\mathrm{BWT}_{\mathrm{T}}[i] = \mathrm{T}[p_i - 1 \bmod n] = \mathrm{T}[p'_i]$.

Corollary (7)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$*. Then the* BWT *of* T *is simple, i.e. has the minimal number of runs.*

Corollary (7)

Let T *be a string with an arithmetically progressed suffix array* $\mathsf{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathsf{T}[p_1 - k - 1] \neq \mathsf{T}[p_1 - 1]$*. Then the* BWT *of* T *is simple, i.e. has the minimal number of runs.*

Case

$$\mathsf{SA} = \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A} \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C}$$

$$P' = \underbrace{\boxed{p_1{-}1} \boxed{\cdots} \boxed{n{-}k}}_{B} \underbrace{\boxed{n} \boxed{\cdots} \boxed{p_n}}_{C} \underbrace{\boxed{p_1} \boxed{\cdots} \boxed{p_1{-}k{-}1}}_{A}$$

$\mathsf{BWT} = b^{\sigma_1} c^{\sigma_2} a^{\sigma_3}$ with $\sum \sigma_i = n$.

Corollary (7)

Let T *be a string with an arithmetically progressed suffix array* $\mathrm{SA} := [p_1, \ldots, p_n]$ *with ratio* k *and* $\mathrm{T}[p_1 - k - 1] \neq \mathrm{T}[p_1 - 1]$*. Then the* BWT *of* T *is simple, i.e. has the minimal number of runs.*

Case $\mathrm{SA} = \underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}$

$P' = \underbrace{\boxed{p_1{-}1 \mid \cdots \mid n{-}k}}_{B}\underbrace{\boxed{n \mid \cdots \mid p_n}}_{C}\underbrace{\boxed{p_1 \mid \cdots \mid p_1{-}k{-}1}}_{A}$

$\mathrm{BWT} = b^{\sigma_1} c^{\sigma_2} a^{\sigma_3}$ with $\sum \sigma_i = n$.

Case $\mathrm{SA} = \underbrace{\boxed{p_1 \mid \cdots \mid n{-}k}}_{A}\underbrace{\boxed{n \mid \cdots \mid p_1{-}k{-}1}}_{B}\underbrace{\boxed{p_1{-}1 \mid \cdots \mid p_n}}_{C}$

$P' = \underbrace{\boxed{p_1{-}1 \mid \cdots \mid p_n}}_{C}\underbrace{\boxed{p_1 \mid \cdots \mid n{-}k}}_{A}\underbrace{\boxed{n \mid \cdots \mid p_1{-}k{-}1}}_{B}$

$\mathrm{BWT} = c^{\sigma_1} a^{\sigma_2} b^{\sigma_3}$ with $\sum \sigma_i = n$.

Summary

- For an arithmetically progressed permutation P there is a string T over a unary, binary or ternary alphabet with $\mathsf{SA}_\mathsf{T} = P$.
- We described a class of strings for which the shape of the suffix array (and BWT) is known.

- Outlook
  - Arithmetic properties can be considered for other integer arrays, such as the LCP array, prefix table, border table, ...