

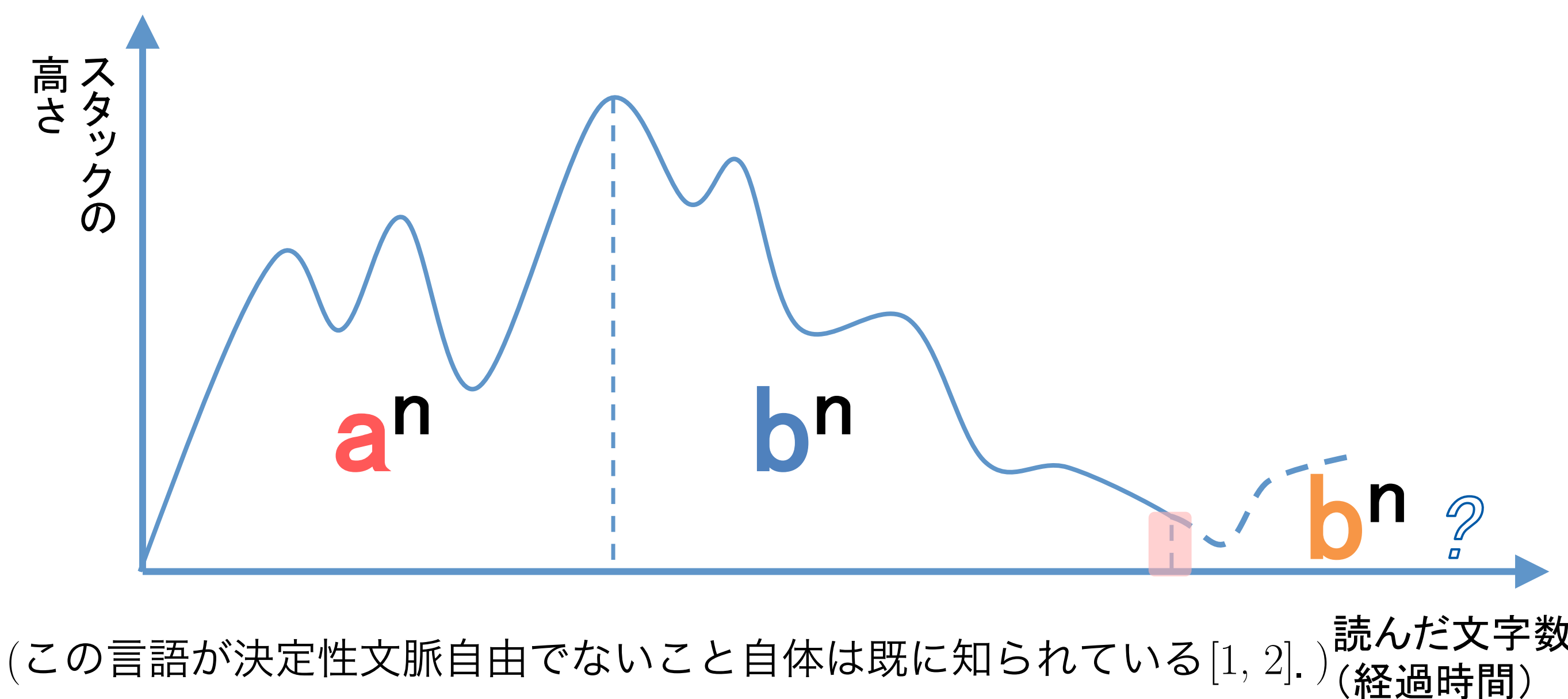
スタックの高さの特徴付けによる 言語が決定性文脈自由でないことの証明

上里 友弥(発表者)
南出 靖彦

言語が決定性プッシュダウン機械で
受理できない原因を一部解明！

$\{a^n b^n, a^n b^n b^n \mid n \geq 1\}$ が決定性文脈自由言語でないことが、
次のように示せる。

- $a^n b^n$ を読み終わった段階でのスタックは **殆ど空** である。
- n の情報がスタックに残っていないので b^n を検査できない。

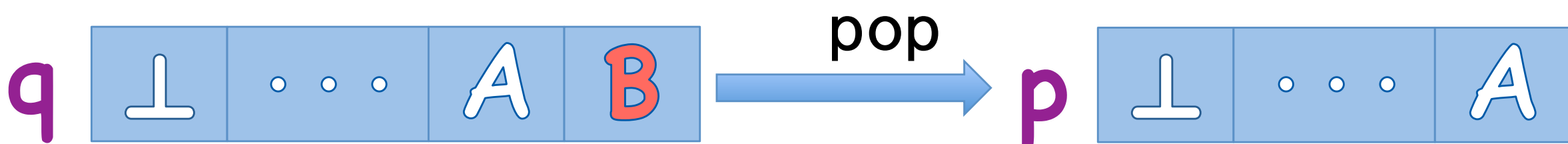


決定性プッシュダウンオートマトン(DPDA)の復習

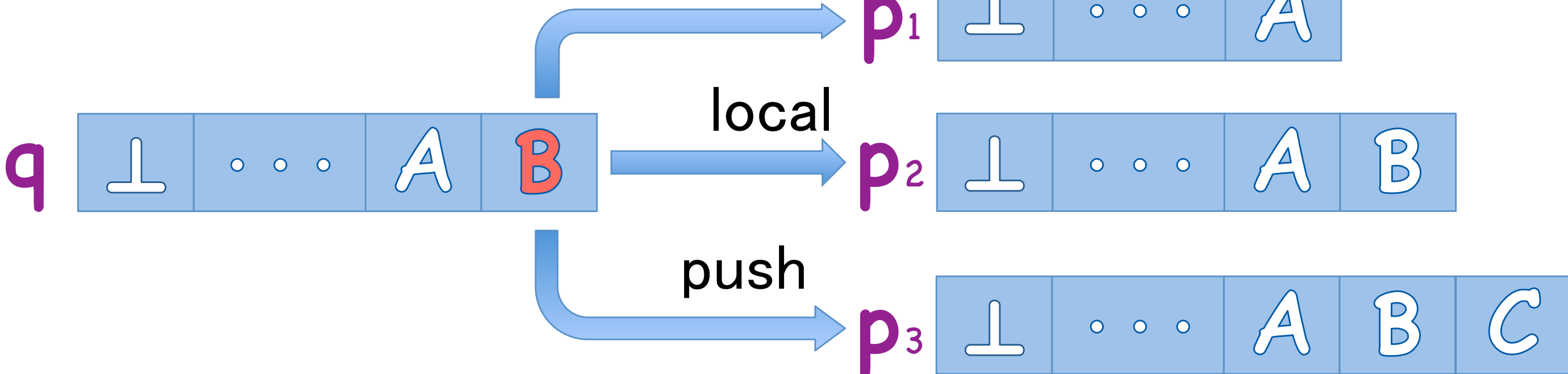
DPDA[3] = プッシュダウンオートマトン + 決定性
= 有限オートマトン + 状態 + 決定性

状態とスタックトップの記号が、遷移の種類 (ε or Σ) を決定する。
$$\delta: Q \times \Gamma \rightarrow (\{\varepsilon\} \rightarrow Q \times [\text{POP}]) \uplus (\Sigma \rightarrow Q \times [\text{POP}, \text{LOCAL}, \text{PUSH}(\Gamma)])$$

ε (入力を消費しない)遷移



Σ (入力を消費する)遷移



文字列 w を受理 $\Leftrightarrow w$ を読んだあとに受理状態 $F(\subseteq Q)$ に入る。

DCFLの例

$\{a^n b^n \mid n \geq 1\}, \{a^n b^m d c^m, a^n b^m e c^n \mid n, m \geq 1\}$ 。

DCFLではない例

$\{a^n b^n c^n \mid n \geq 1\}, \{w \mid w \in \{a, b\}^*, w \text{ は回文}\}$ 。

入力テープの1方向性とスタックの揮発性に基づく直感

$a^m b^n$ について $m = n$ を検査するには、

- a^m (相当の情報) を push し、
- pop しつつ b^n とのマッチングを行うと思われる。

結果として、受理時点のスタックは殆ど空になっているはずである。

実際に以下が成立する：

定理

x, y を空でない文字列とする。任意の自然数 n について

$x^n, x^n y, x^n y^2, \dots$ は受理されないが、 $x^n y^n$ は受理される
(マッチングを行っている、個数を比較している)

ならば、 $x^n y^n$ 受理時点のスタックの高さが定数 $\mathcal{H}_{x,y}$ 以下になる。

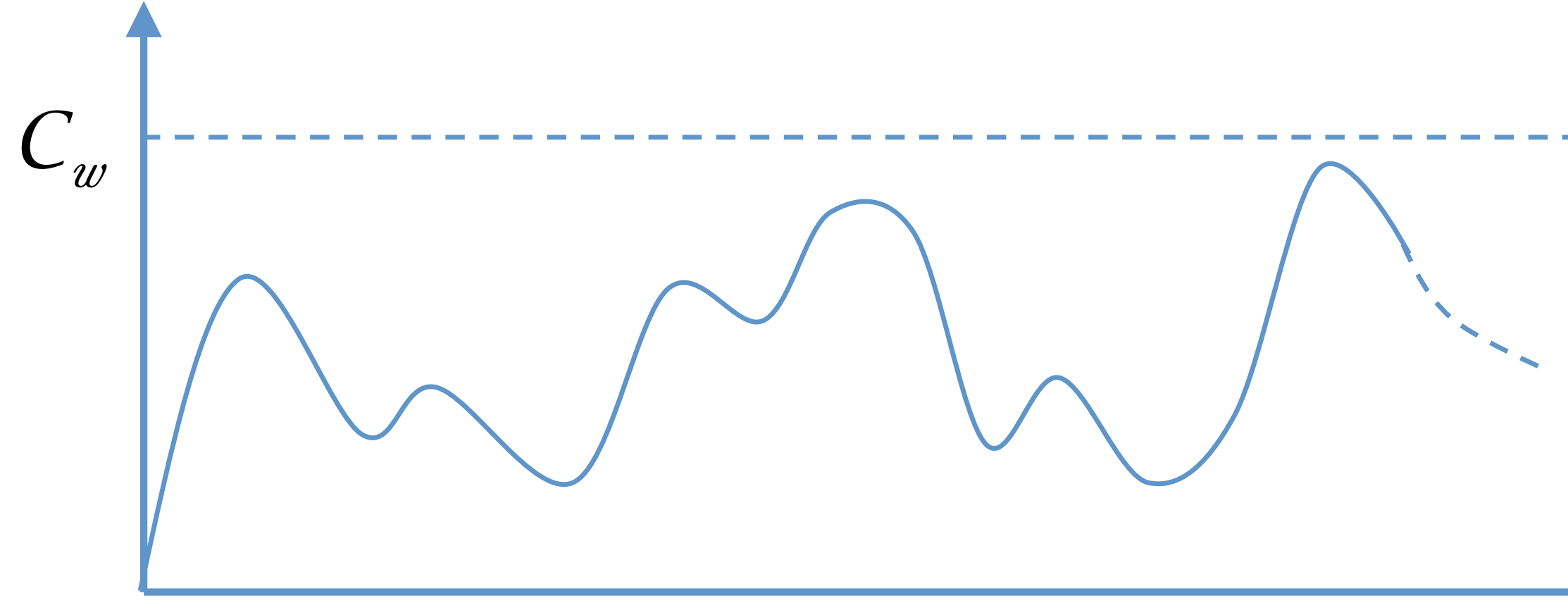
系

$\{a^n b^n, a^n b^n b^n \mid n \geq 1\}$ を受理する DPDA が存在したとすると、
 $a^n b^n$ 受理時点のスタックの高さが定数 $\mathcal{H}_{a,b}$ 以下となる。

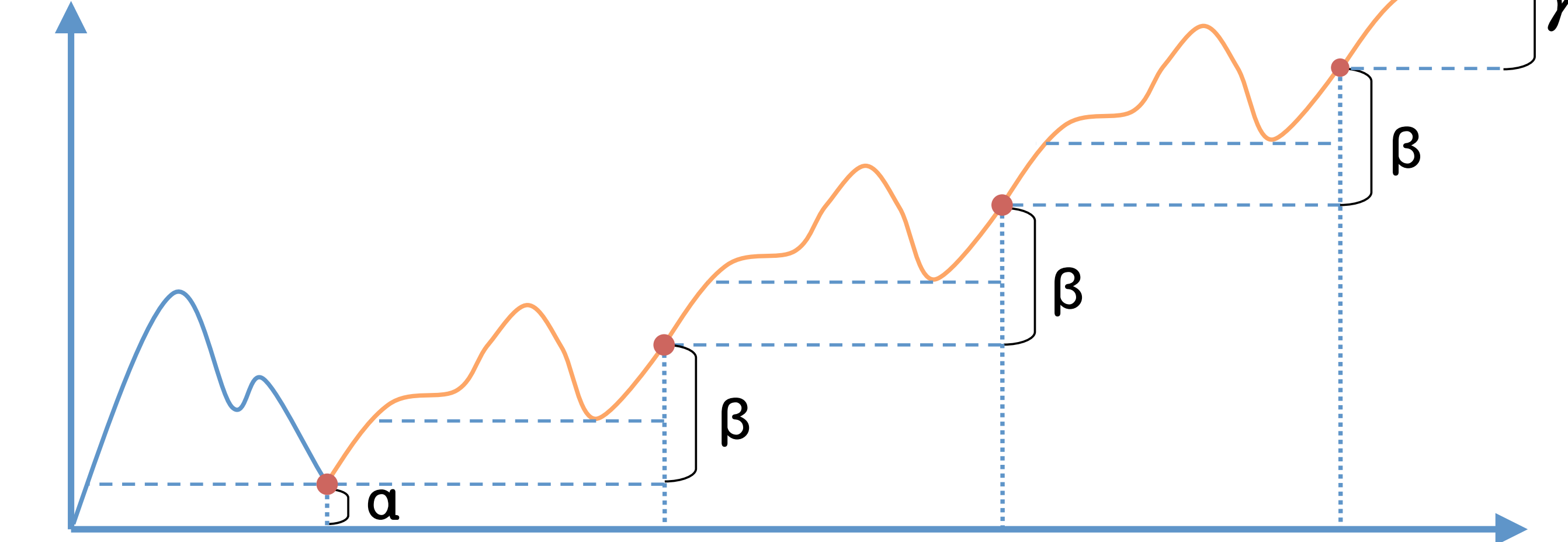
単純な入力の生成するスタックの周期性

入力が文字列 w の繰り返し w^+ であるときには、
以下のどちらかの、分かりやすい計算の構造があらわれる：

- スタックの高さは常に C_w 以下である；



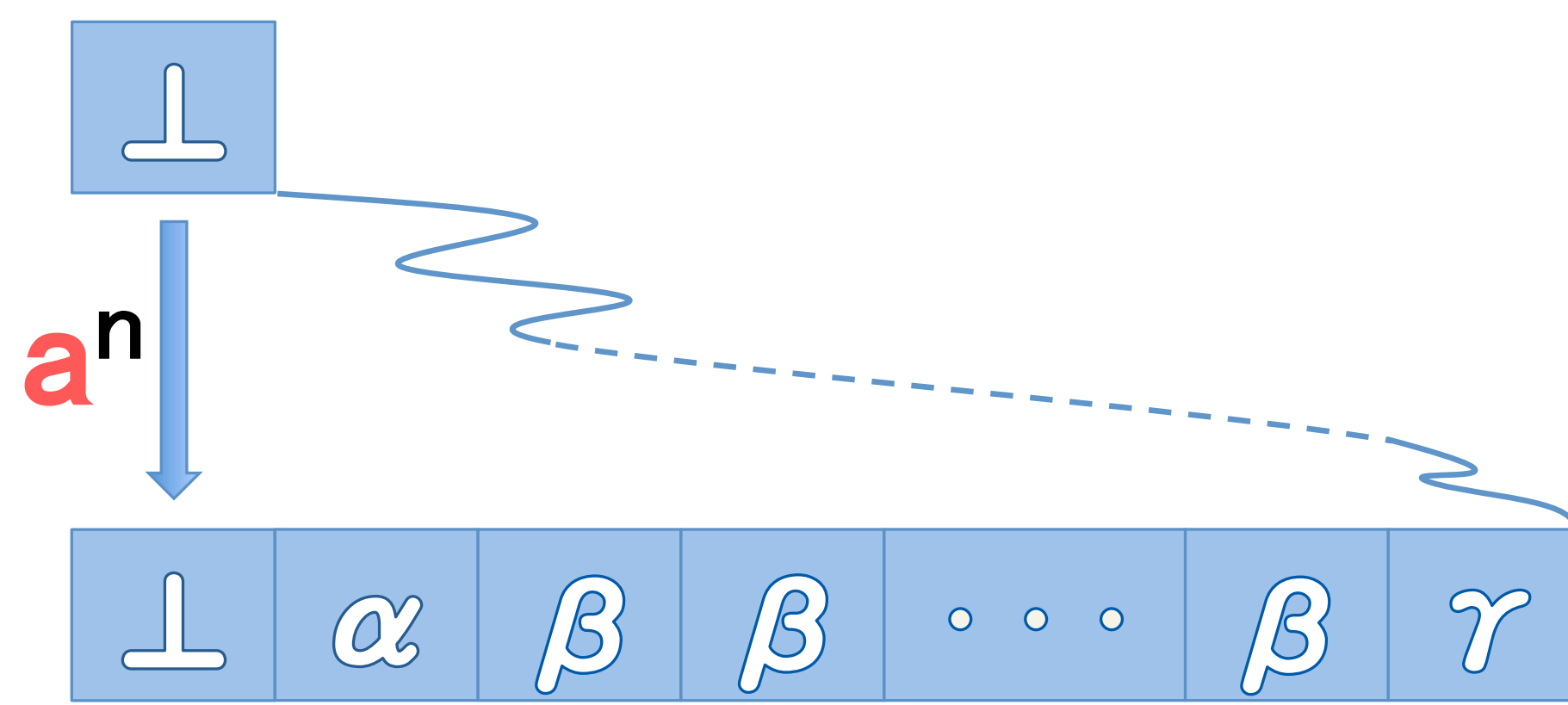
- または、スタックの内容に周期性が生じ $\alpha\beta^*\gamma$ となる。



(昔から知られている性質であることがわかったので、再発見をしたに過ぎない[1].)

a^n はスタック全体で表現される

直前の結果から以下のことが分かる：



b^n とのマッチングのためには α 部まで pop すると思われるが、
これは実際に以下の形で証明される：

定理(下図も参照)

x, y を空でない文字列とする。任意の自然数 n について

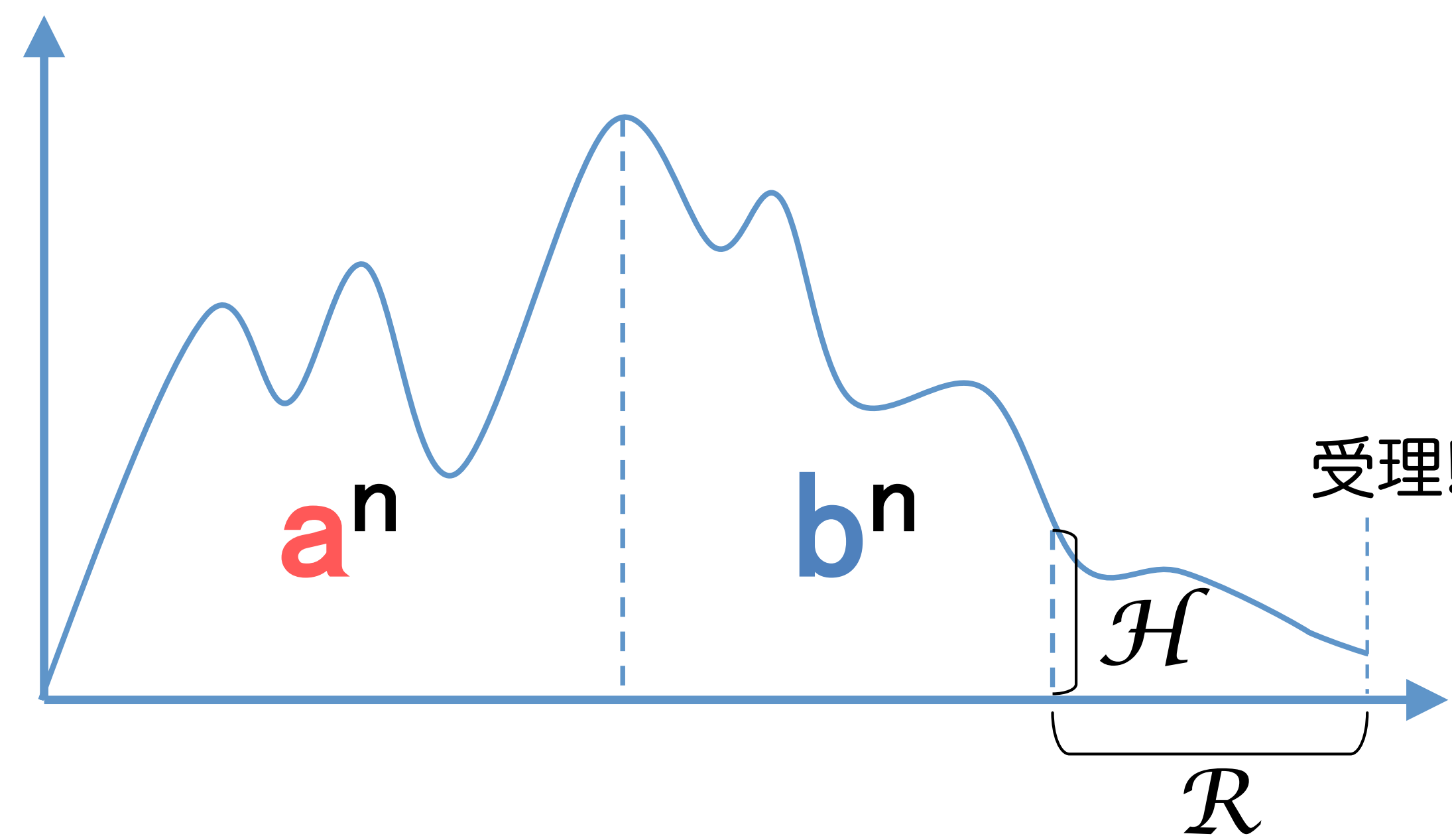
$x^n, x^n y, x^n y^2, \dots$ は受理されないが、 $x^n y^n$ は受理される

ならば、 y^n を読んでいる間にスタックの高さが定数 $\mathcal{H}_{x,y}$ 以下になる。

(スタック全体で a^n の表現を丁度1つ分持っていることが本質で、スタックの内容を補助テープとして持つようなチューリング機械によるマッチングを考えても、同様の結果が示される。)

マッチングするとスタックは殆ど空になる

スタックの高さが $\mathcal{H}$ 以下となるタイミングも分かる。すなわち、
受理状態に入る $\mathcal{R}$ 文字前には高さが $\mathcal{H}$ 以下になることが示せる：



したがって、受理時点のスタックの高さは高々 $\mathcal{H} + \mathcal{R}$ 程度になる。
巨大な n について $\mathcal{H} + \mathcal{R}$ は無視でき、 $a^n b^n$ を受理した時点で：

- スタックが殆ど空になっていることが分かった。
- すなわち、スタックには n の情報が残っていないことが分かった。

Bibliography

- [1] Seymour Ginsburg and Sheila Greibach. Deterministic context free languages. *Information and Control*, 9(6):620 – 648, 1966.
- [2] M. A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley Longman Publishing Co., Inc., 1978.
- [3] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley, 1979.