

数学の対象としての 文法

多重文脈自由文法に関する最近の成果と未解決問題

金沢 誠

主な関心は言語を題材とした応用数学

人間の言語の科学的研究にどこまで数学を取り込めるか

人間の言語に動機づけられた数学的課題によって数学を発展させる

数学の対象

$$\frac{3}{7}$$

$$\frac{31}{19}$$

$$\sqrt{2}$$

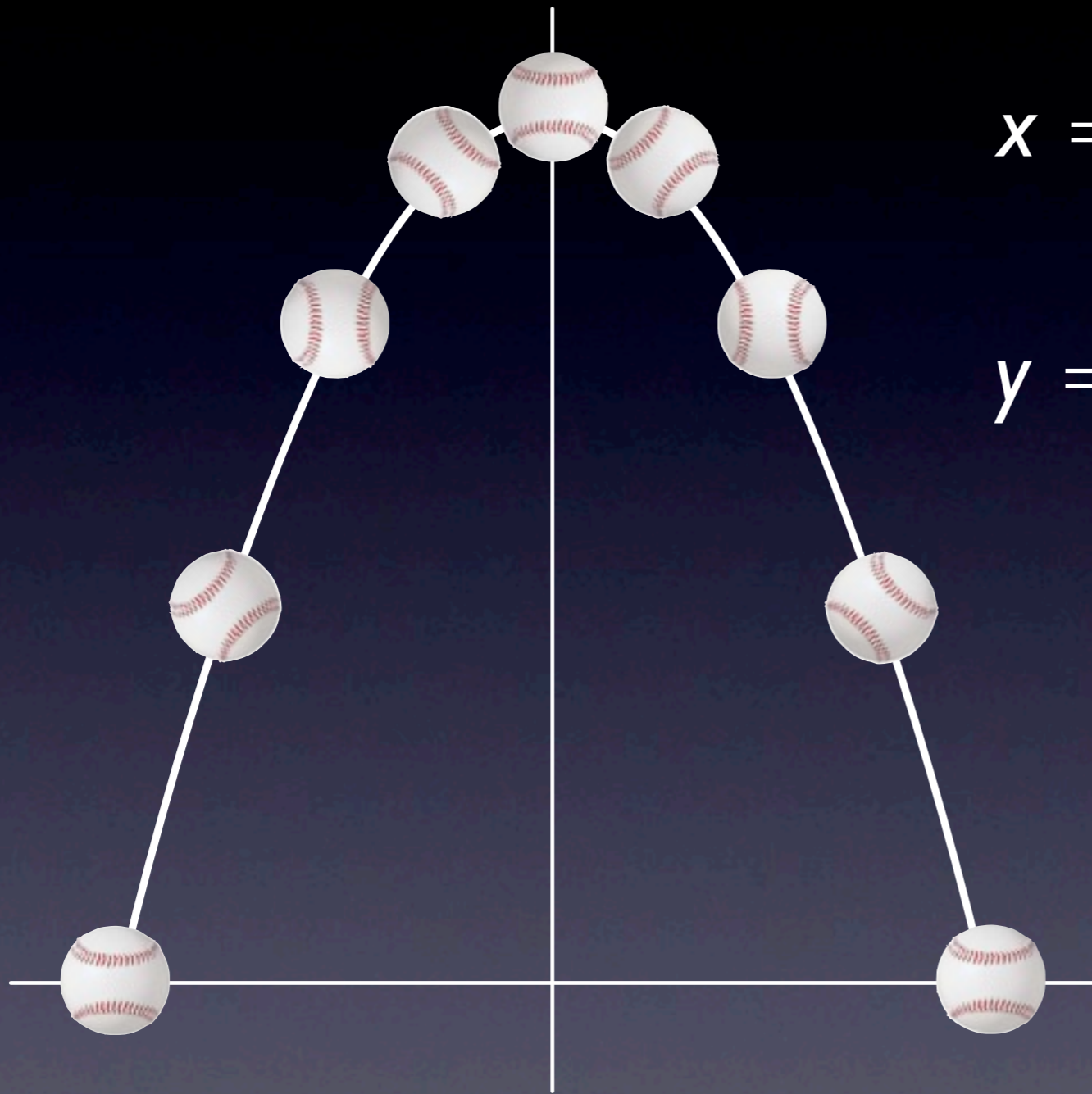
$$3 - \sqrt{5}$$

実数

$$\pi$$

$$e$$

$$\sqrt{2}^{\sqrt{2}}$$



$$x = \frac{1}{2}t - 2$$

$$y = -\frac{1}{4}t^2 + 2t$$

関数

モデル

実数から実数への関数

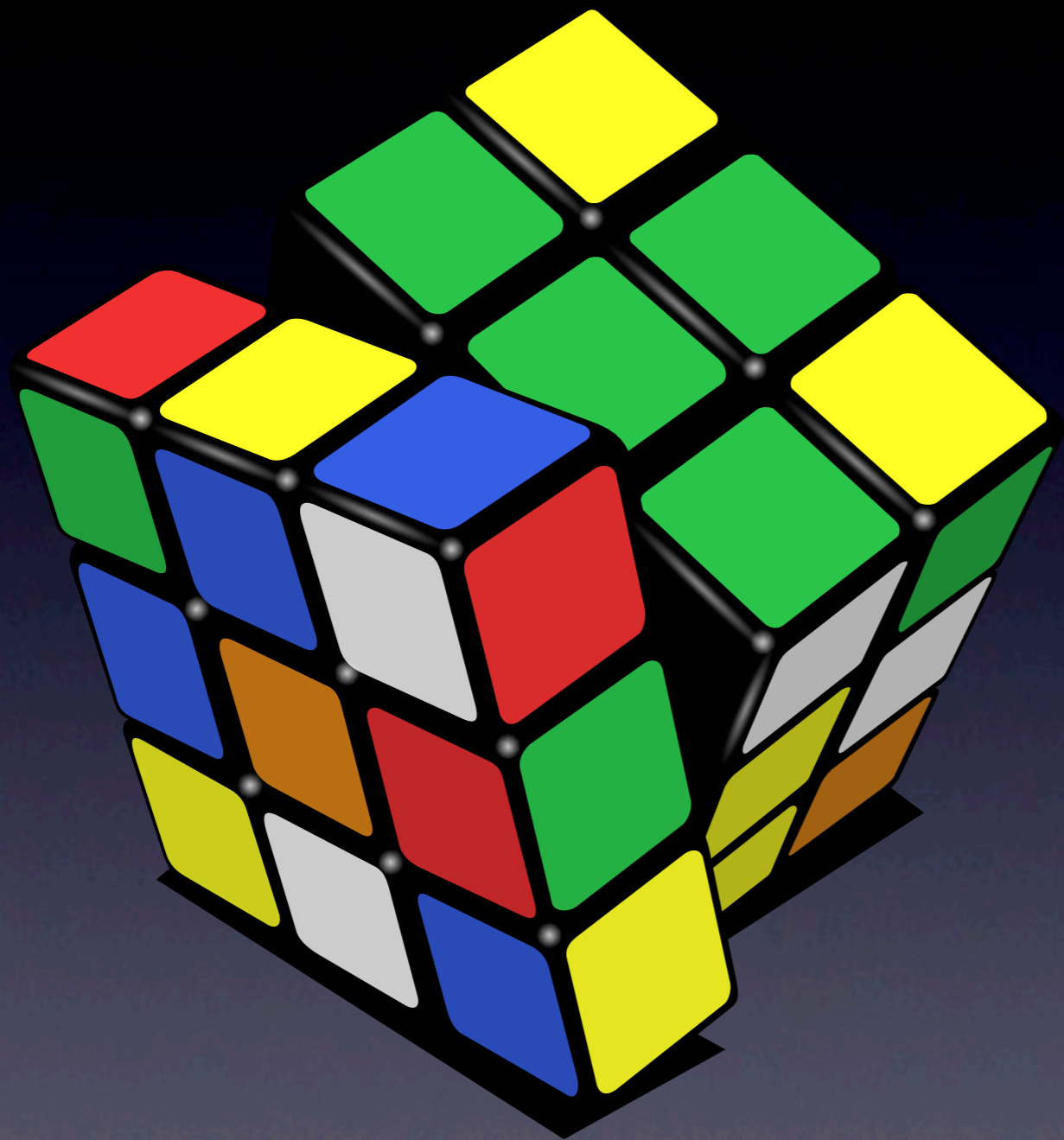
数学の外の事象をモデル化する



位相幾何学

より抽象的な数学の対象

結び目



群論

抽象代数

ルービックキューブに潜んだ群の構造

形式言語理論

「文法」を数学的対象として研究する分野は形式言語理論

a	aaaa	baaab	abbabba
b	abba	babab	abbbbba
aa	baab	bbabb	baaaaab
bb	bbbb	bbbbbb	baabaab
aaa	aaaaa	aaaaaaaa	bababab
aba	aabaa	aaabaaa	babbbbab
bab	ababa	aababaa	bbaaaabb
bbb	abbba	aabbbbaa	bbababb
		abaaaba	bbbabbb
		abababa	bbbbbbb

...

「回文」

形式言語

形式言語とは、数学の対象としてはかなり具体的
文字列の集合
これは回文の集合

a	aaaa	baaab	abbabba	
b	abba	babab	abbbbba	
aa	baab	bbabb	baaaaab	
bb	bbbb	bbbbbb	baabaab	
aaa	aaaaa	aaaaaaa	bababab	
aba	aabaa	aaabaaa	babbbab	...
bab	ababa	aababaa	bbaaabb	
bbb	abbba	aabbbaa	bbababb	
		abaaaba	bbbabbb	
		abababa	bbbbbbb	

$$\subset \{a, b\}^*$$

*a, b*からなる
文字列全体の
集合

形式言語

一般に、文字集合が与えられたときそれに属する文字を並べてできる文字列の集合を形式言語と言う

人間の言語

Gelukkig nieuwjaar!

Bonne année!

שנה טובה!



形式言語を研究する理由はいろいろある

人間の言語がもともとの動機のひとつだがそれだけではない

ここでは、形式言語を人間の言語のある側面のモデルと考える

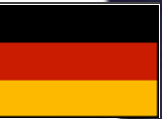
語順の制約

チャールズがメアリーにピーターがジョンに泳ぎを教えるのを手伝わせること

that Charles lets Mary help Peter teach John to swim



daß der Karl die Maria dem Peter den Hans schwimmen lehren helfen läßt



dat Karel Marie Piet Jan laat helpen leren zwemmen



dass de Karl d'Maria em Peter de Hans laat hálfe lárne schwüme



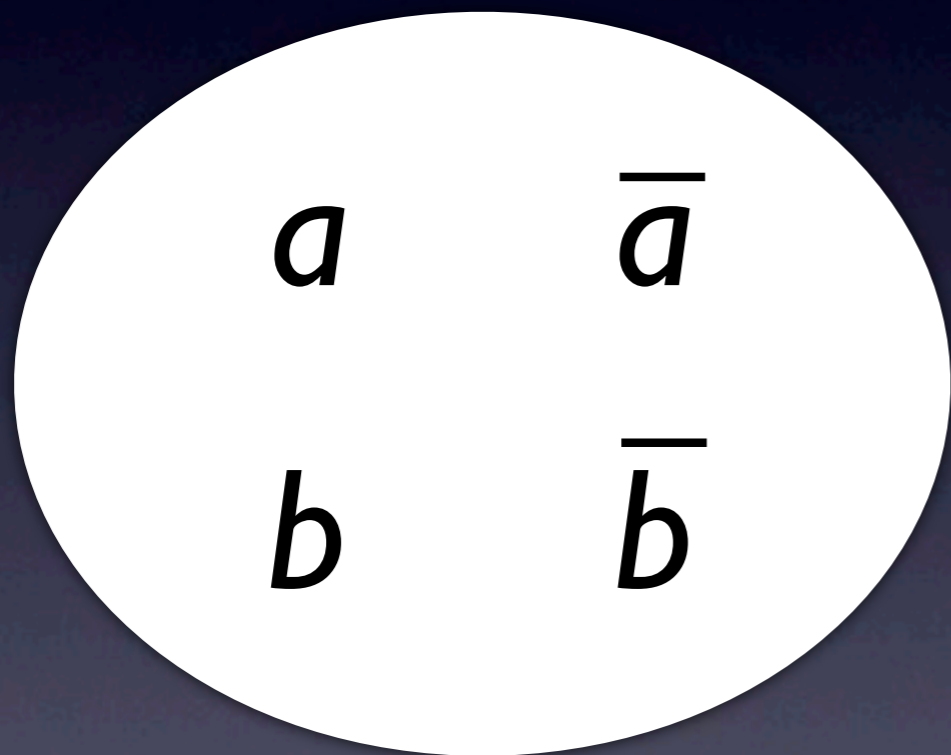
語順の制約を形式言語でモデル化する

文法的依存関係にある単語（句）の対の出現の仕方に関する制約

ひとつの他動詞／目的語の位置関係ではなく、複数の他動詞／目的語の対がおたがいに関連している場合の位置関係

使役の意味を持つ動詞を使った構文

文字対集合



[]

{ }

かっこ

他動詞／目的語のように文法的依存関係にある語句の対を表すために「文字対」を考える
おたがいを要求する点でかっこの対のようなものだがかっこのような決まった配置を仮定しない

原点回帰文字列

a と \bar{a} を同数含む
 b と \bar{b} を同数含む

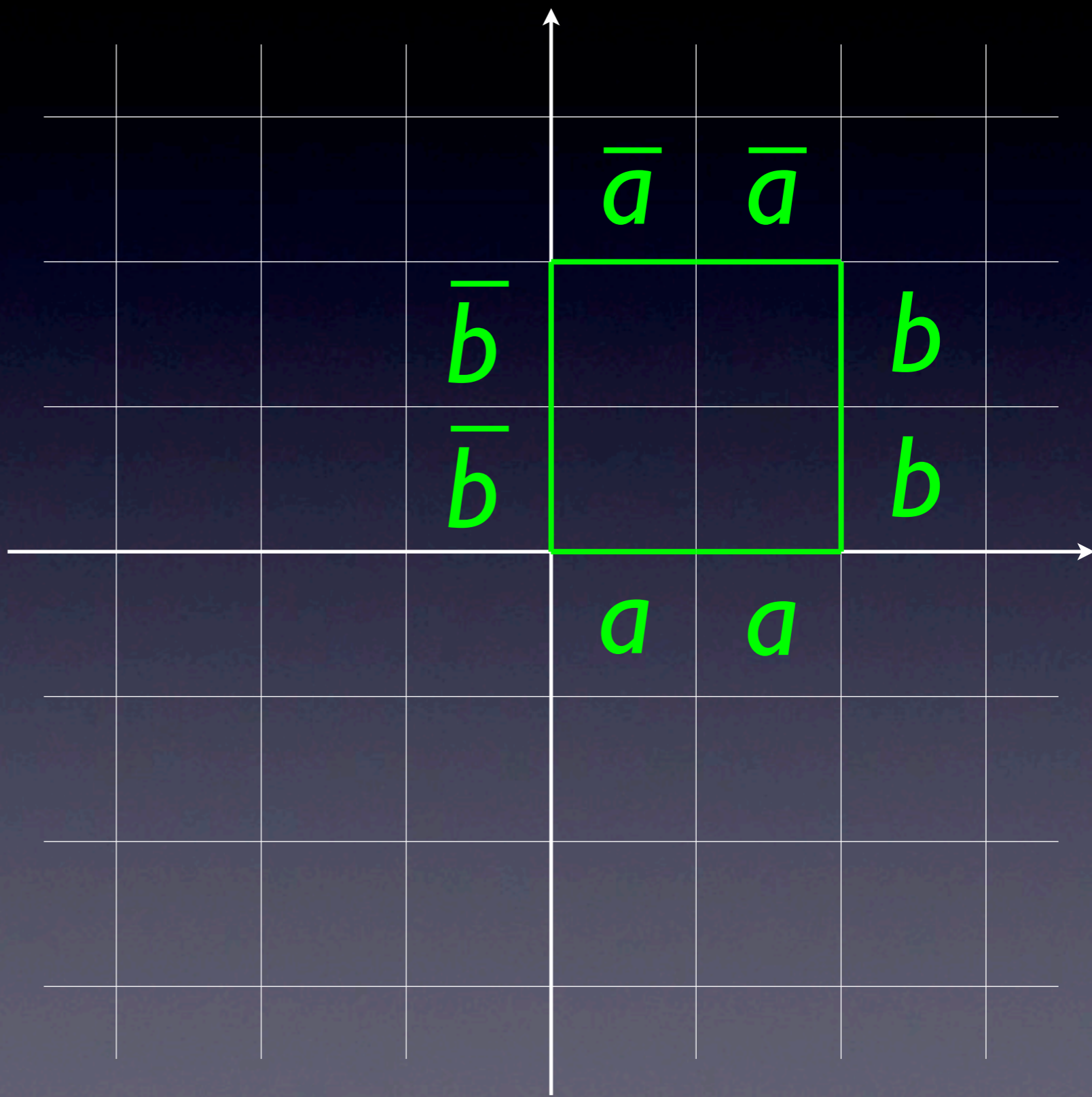
$abb\bar{b}\bar{a}$

$a\bar{b}\bar{a}b\bar{b}\bar{a}\bar{a}$

$\bar{b}\bar{a}ab\bar{b}\bar{a}\bar{a}\bar{a}$

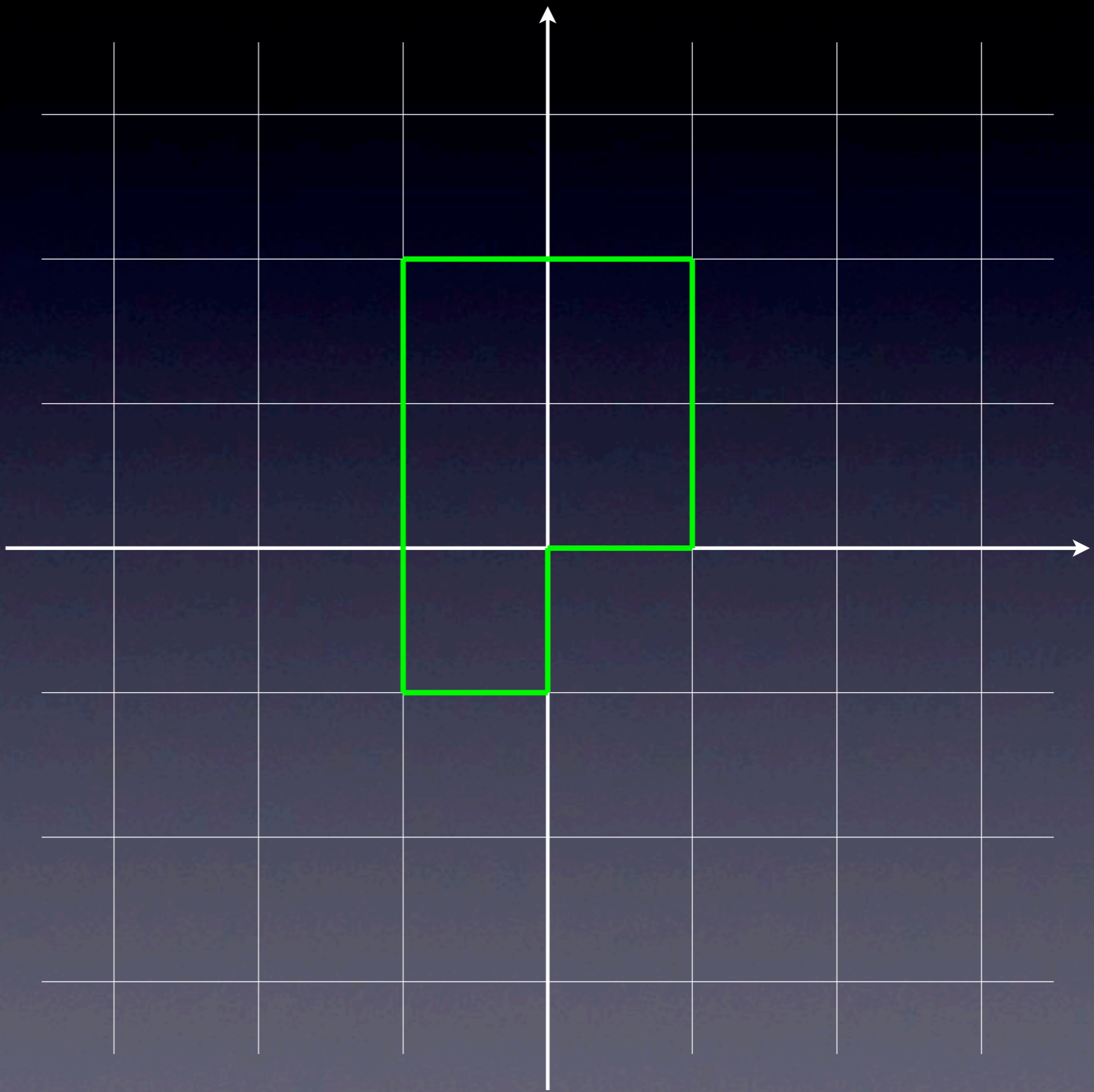
文字対集合に属する文字を並べてできる文字列のうち、ある条件を満たすものに着目する

$aabb\overline{aabb}$

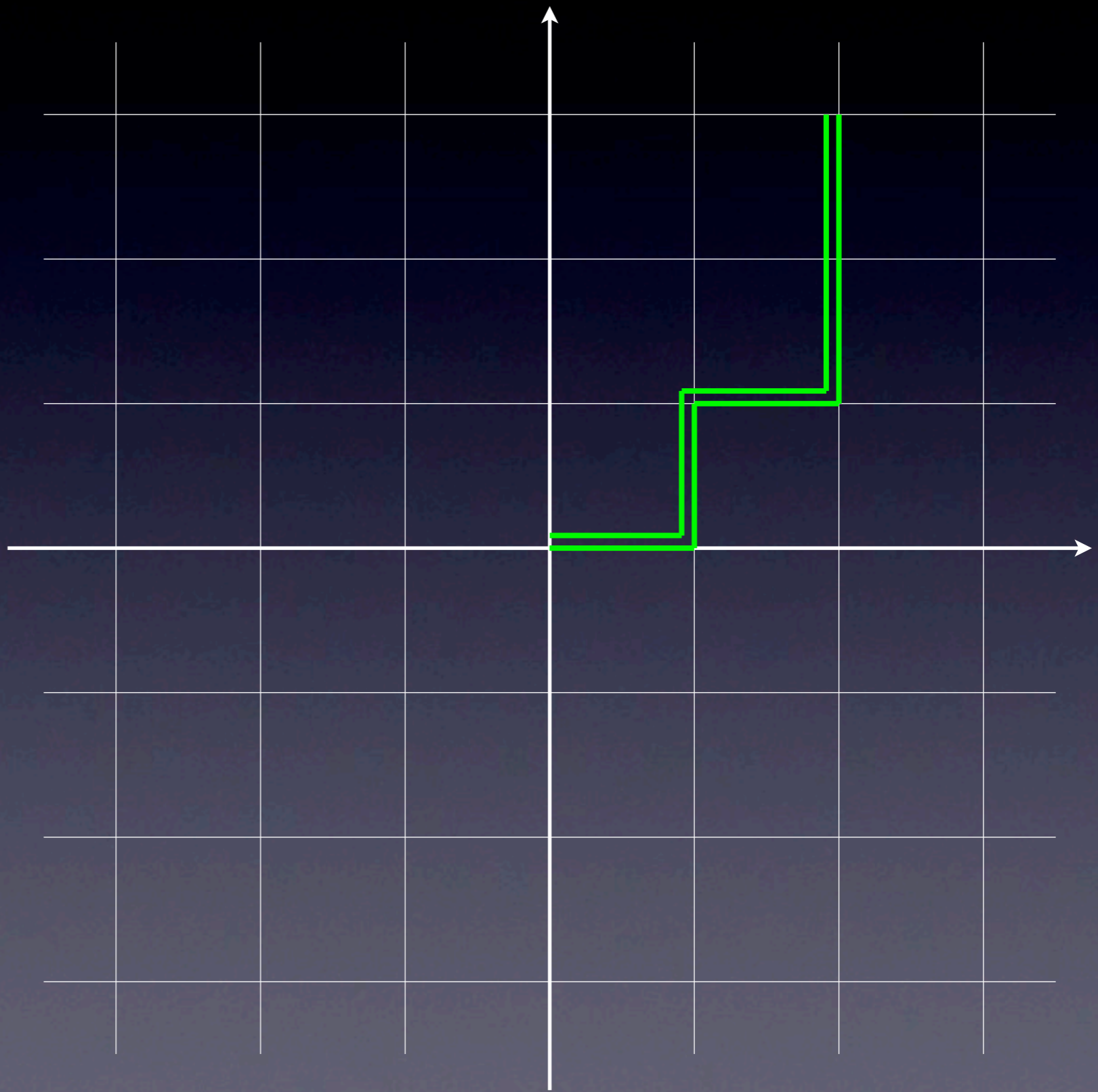
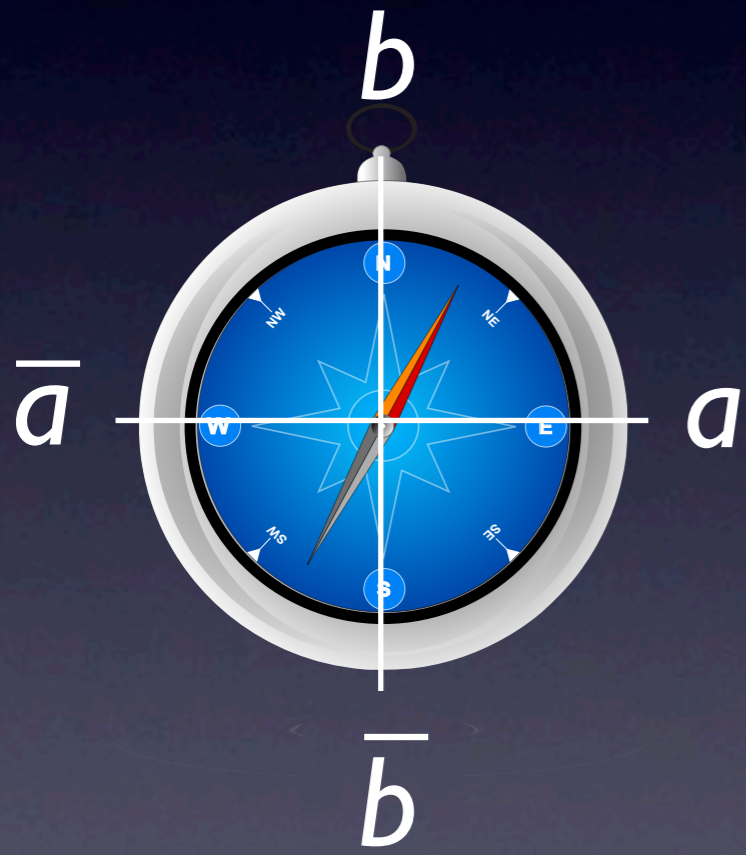


原点回帰文字列という呼び方は次のような幾何学的 (?) 解釈による
文字列は格子の上を動く軌跡で表される

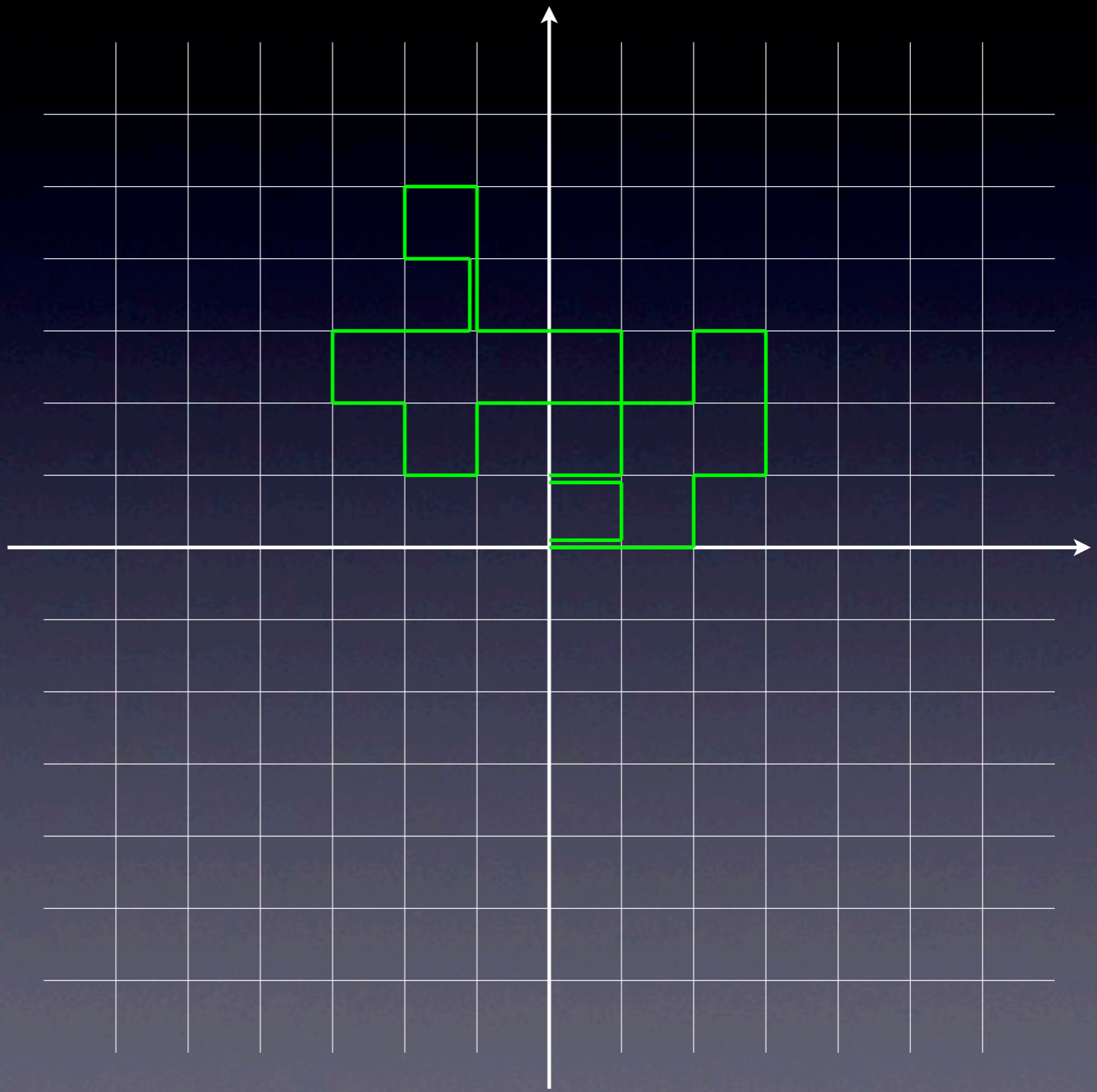
$ab\overline{a}abb\overline{b}ab$



$ababbb\bar{a}\bar{b}$



$aababb\bar{a}\bar{b}a\bar{a}\bar{a}\bar{b}\bar{a}\bar{b}\bar{a}\bar{b}a\bar{a}\bar{b}\bar{a}\bar{b}\bar{b}a\bar{a}\bar{b}\bar{b}\bar{a}\bar{a}\bar{b}\bar{a}$



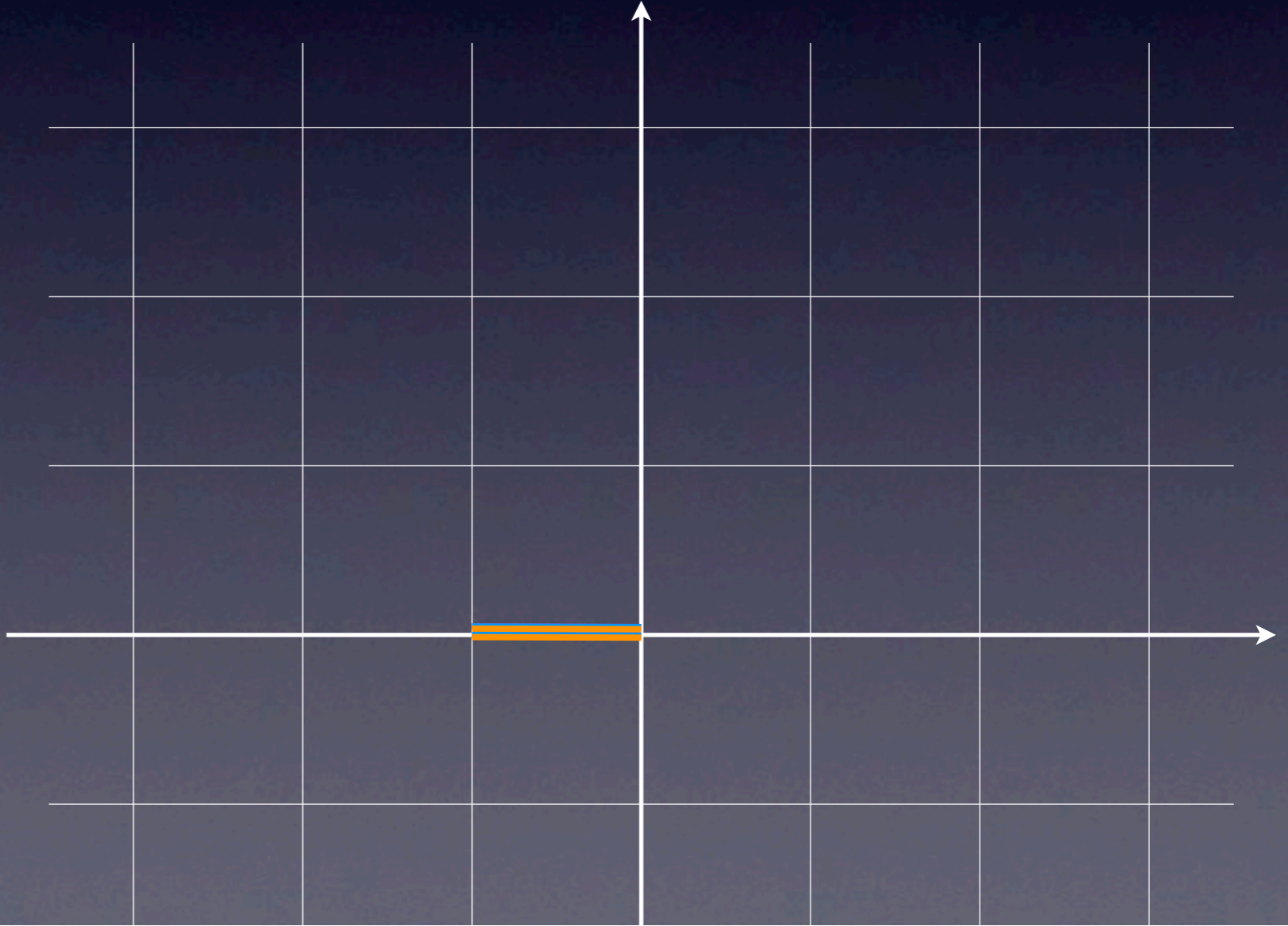
文字列よりも軌跡の方が特徴が見えやすい

V O V O V O

that Charles lets Mary help Peter teach John to swim



\bar{a} a \bar{a} a \bar{a} a



英語の使役構文の語順は単純

対格 与格 対格
～を ～に ～を

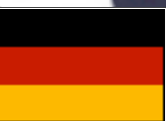


daß der Karl **die Maria** **dem Peter** **den Hans** schwimmen **lehren** **helfen** läßt



a *b* *a* *a* *b* *a*

daß der Karl **die Maria** **den Hans** **dem Peter** schwimmen **helfen** **lehren** läßt



a *a* *b* *b* *a* *a*

ドイツ語では、英語と違って目的語の名詞に格の区別がある
他動詞も対格を要求する動詞と与格を要求する動詞に分かれる
2つの文字対を使って表す

daß der Karl die Maria dem Peter den Hans schwimmen lehren helfen läßt



a

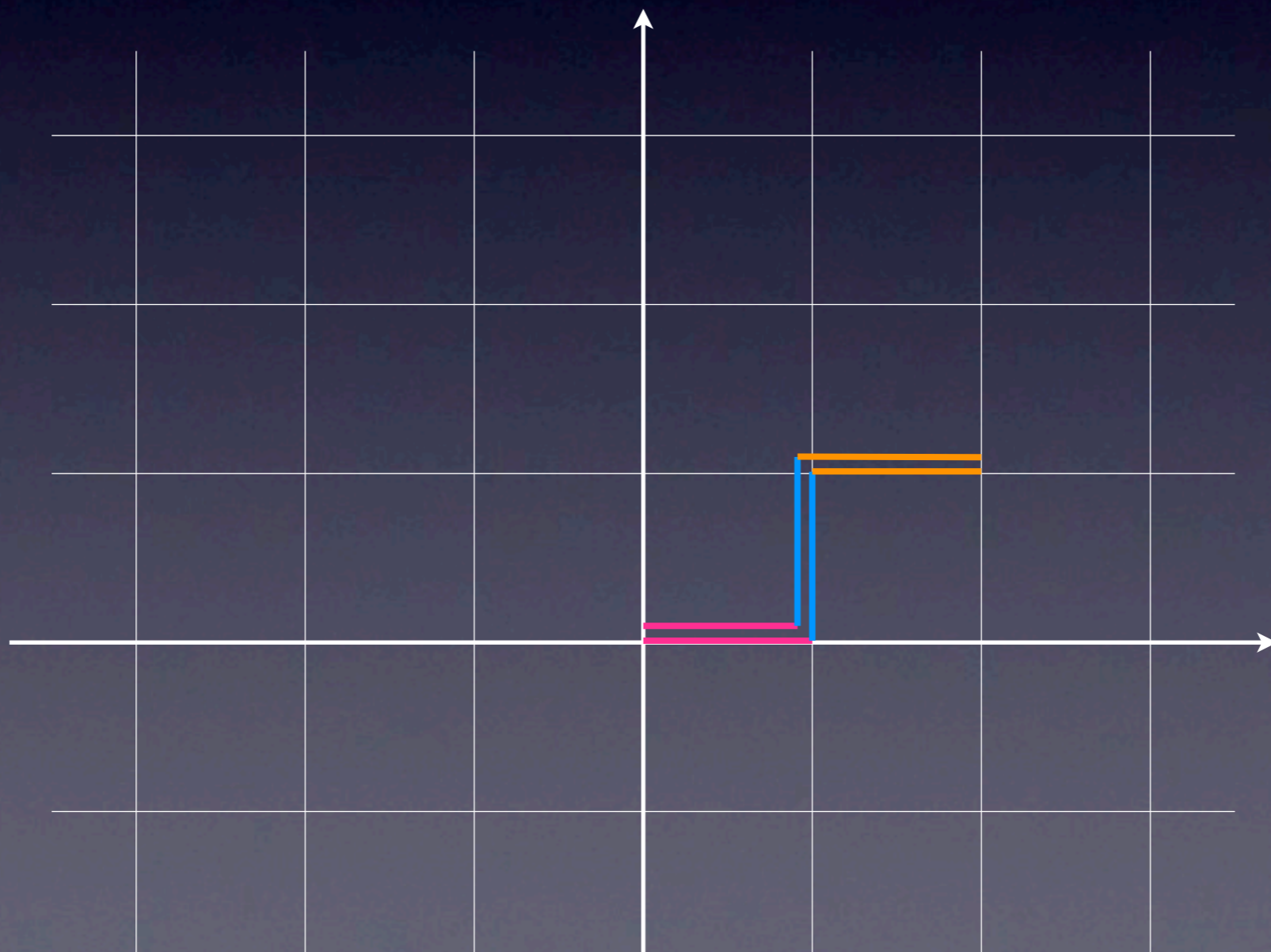
b

a

\bar{a}

\bar{b}

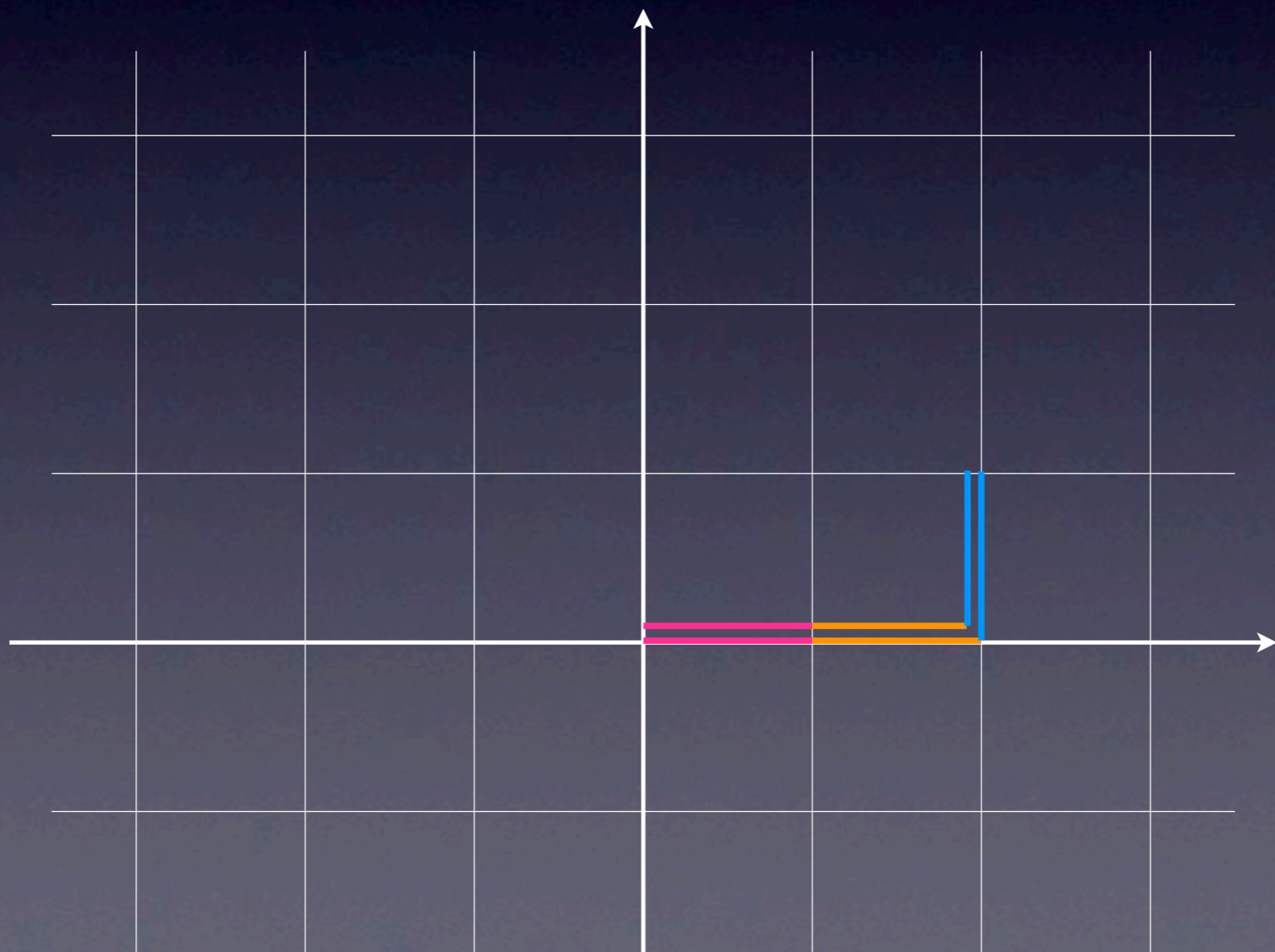
\bar{a}



daß der Karl die Maria den Hans dem Peter schwimmen helfen lehren läßt



a a b \bar{b} \bar{a} \bar{a}



対格 与格 対格
～を ～に ～を

○

○

○

∨

∨

∨

dass de Karl d'Maria em Peter de Hans laat hälfe lärne schwüme



a *b* *a* \bar{a} \bar{b} \bar{a}

dass de Karl d'Maria de Hans em Peter laat lärne hälfe schwüme



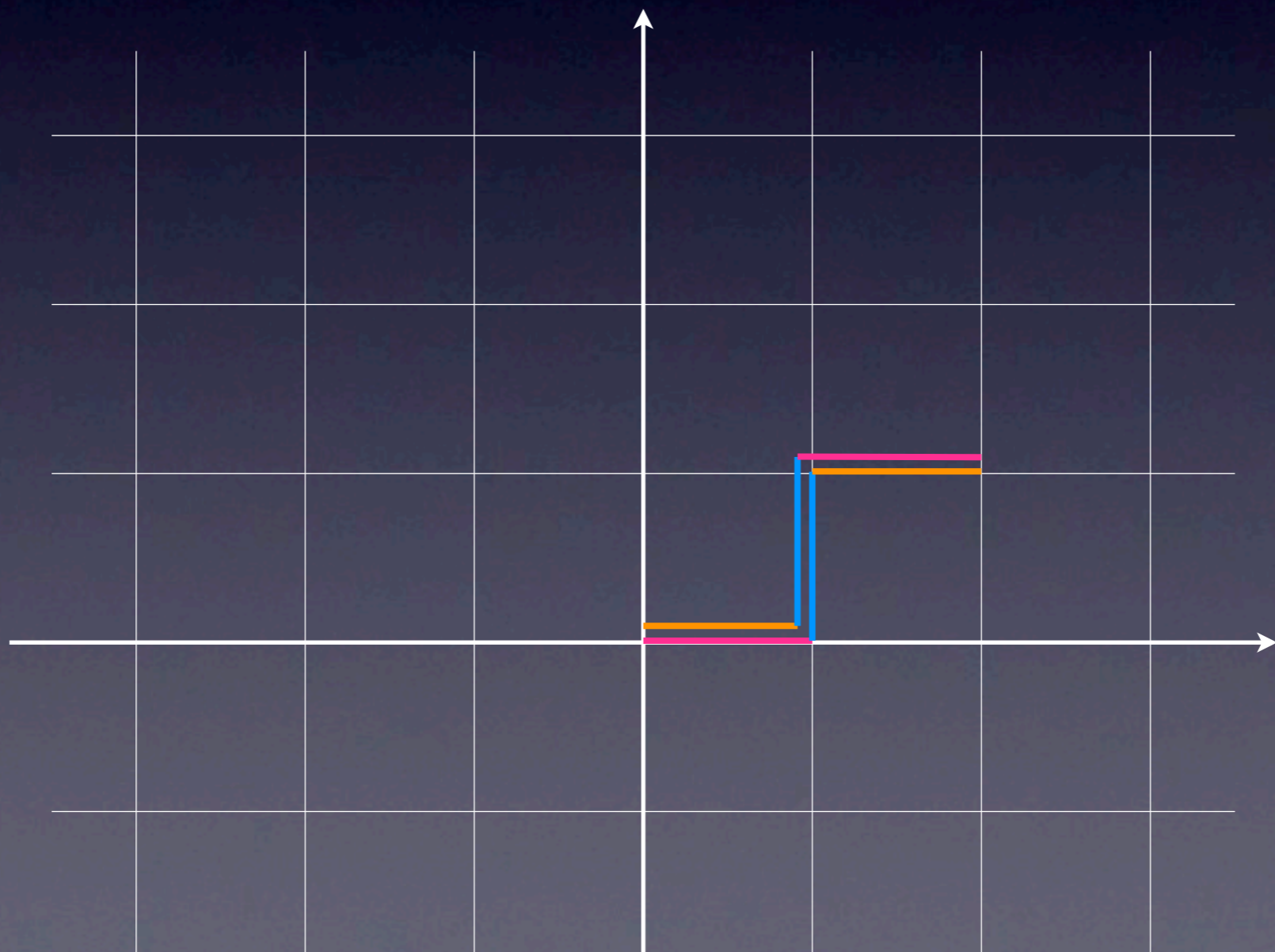
a *a* *b* \bar{a} \bar{a} \bar{b}

ドイツ語チューリッヒ方言も格の区別を持つ

dass de Karl d'Maria em Peter de Hans laat halfe larne schwume



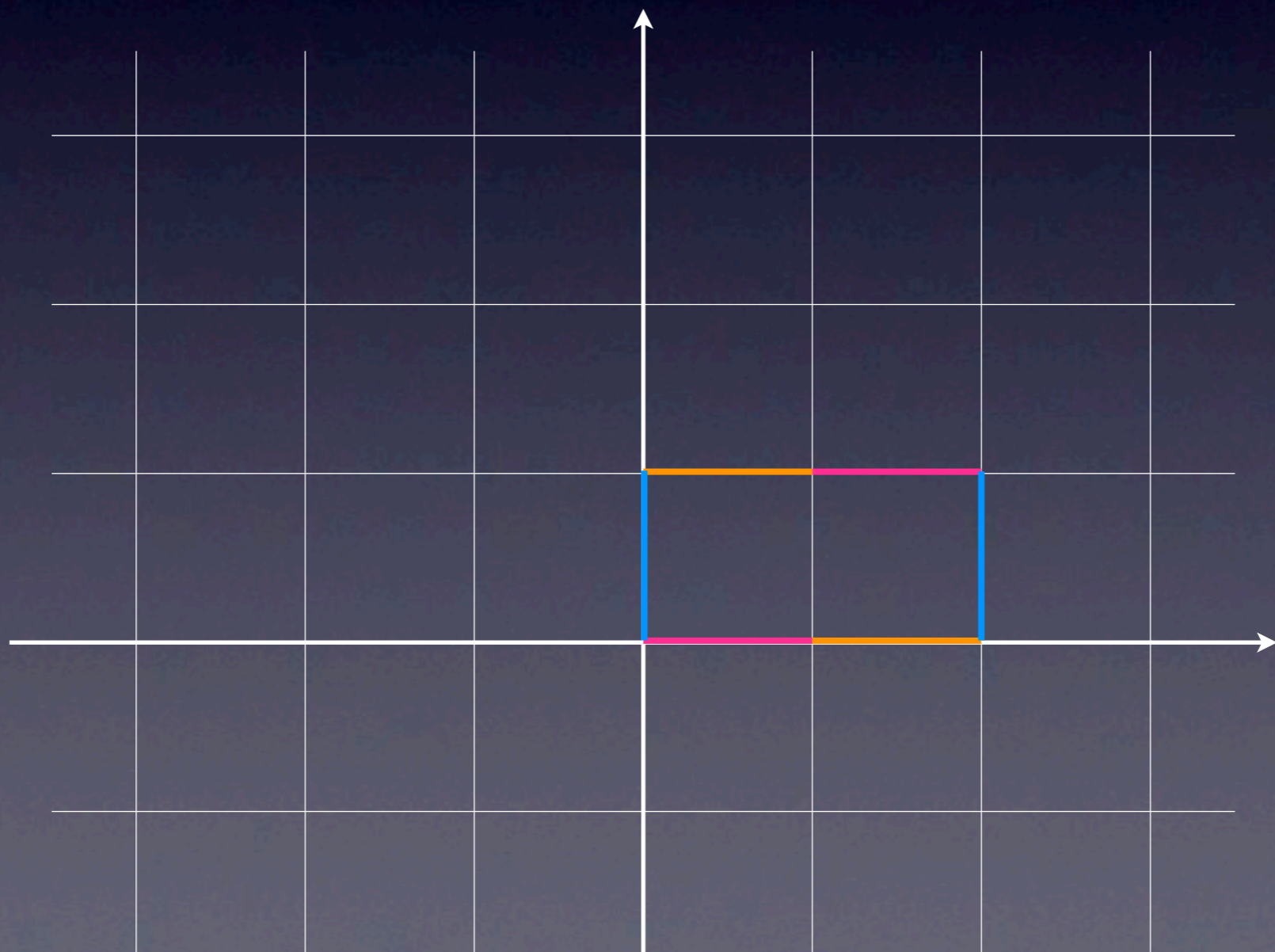
a b a \bar{a} \bar{b} \bar{a}

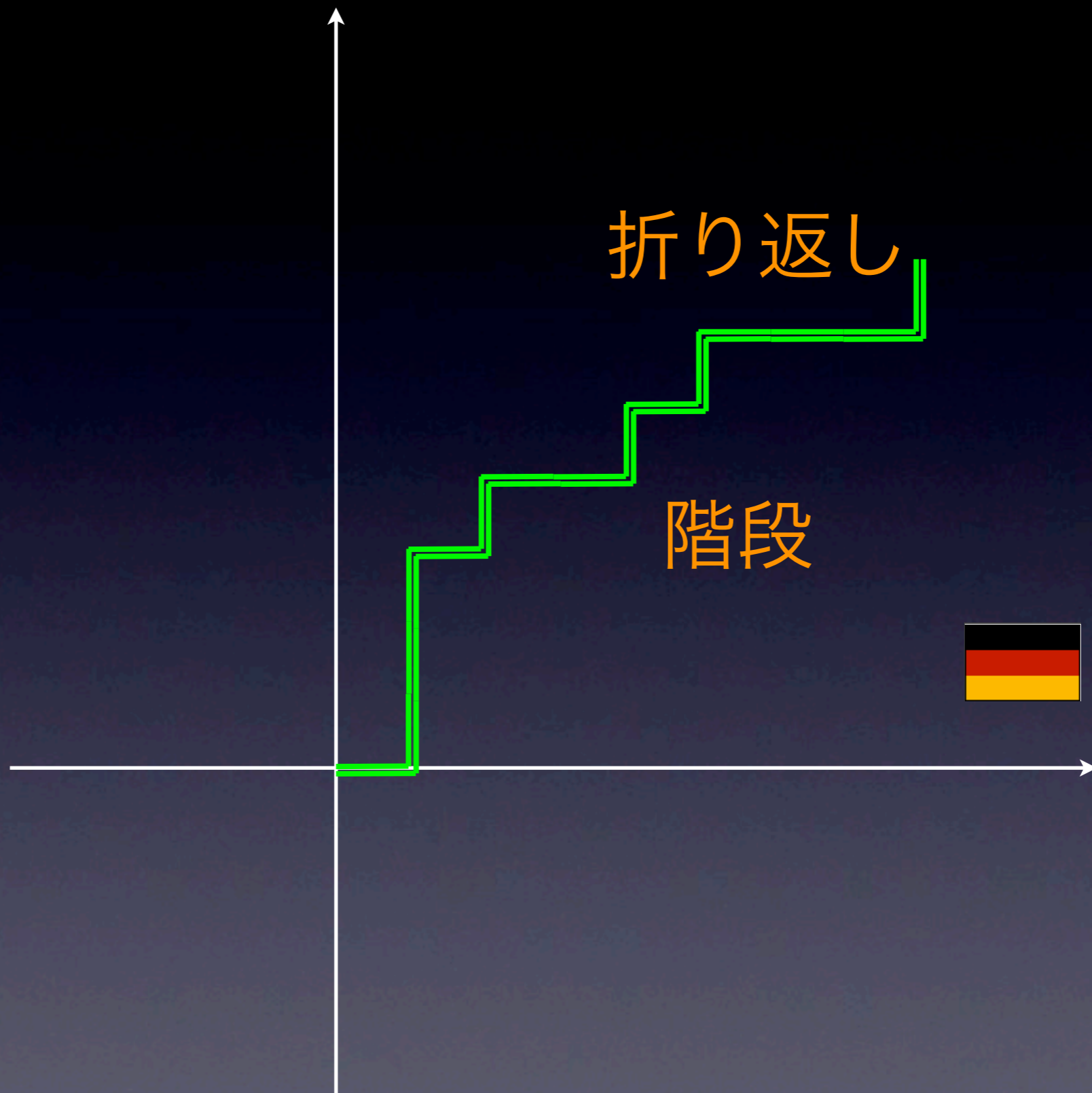


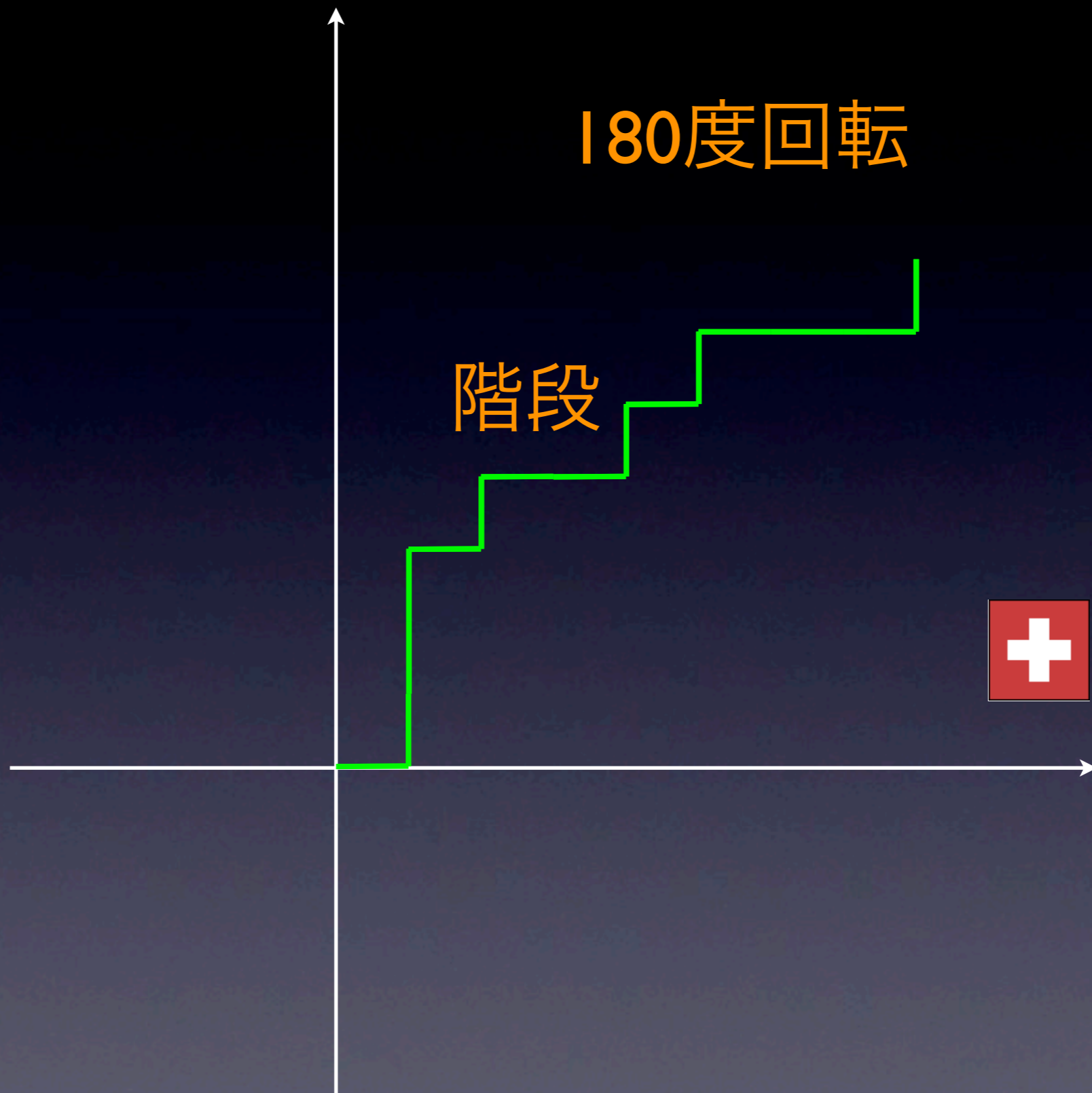
dass de Karl d'Maria de Hans em Peter laat lärne hälfe schwüme

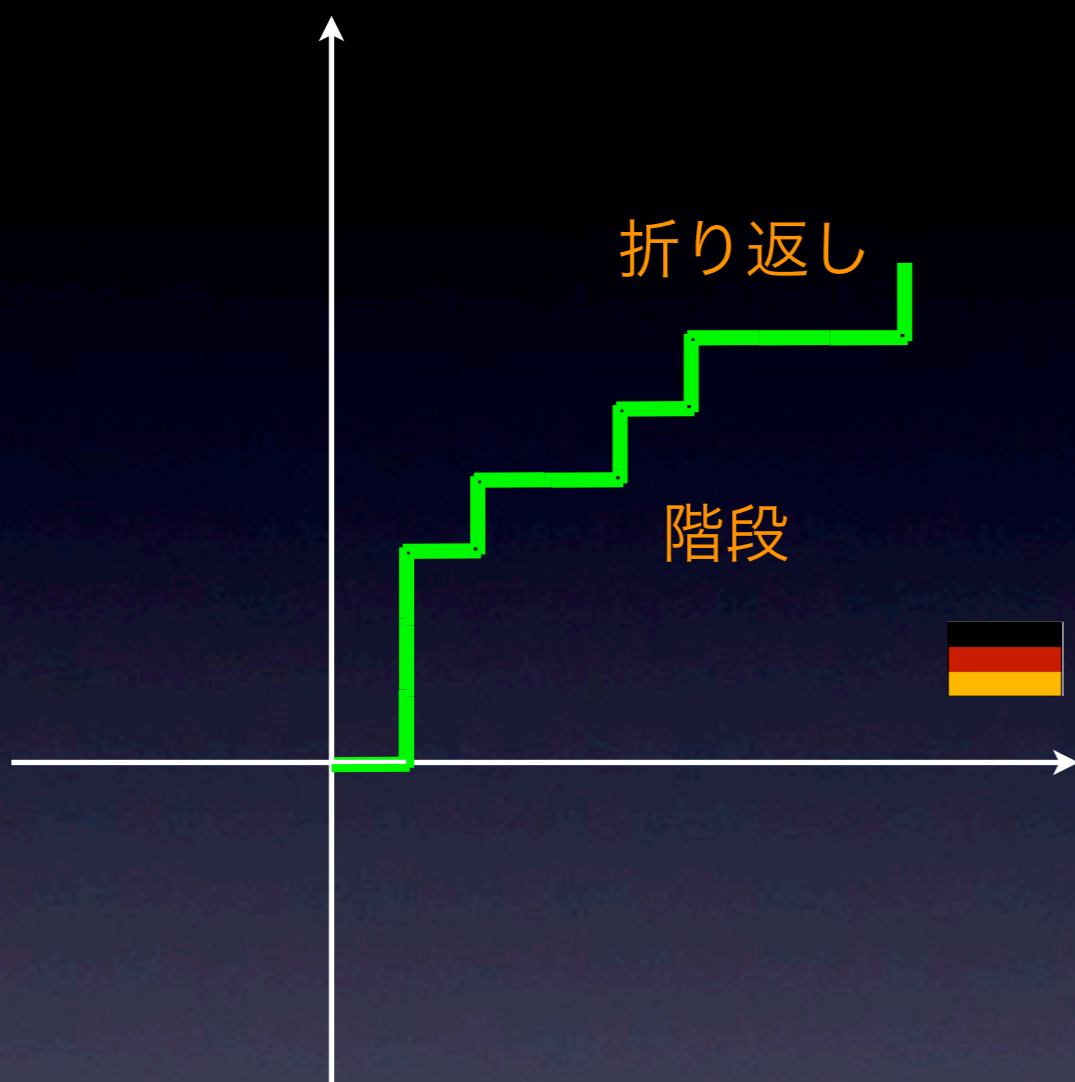


a a b \bar{a} \bar{a} \bar{b}

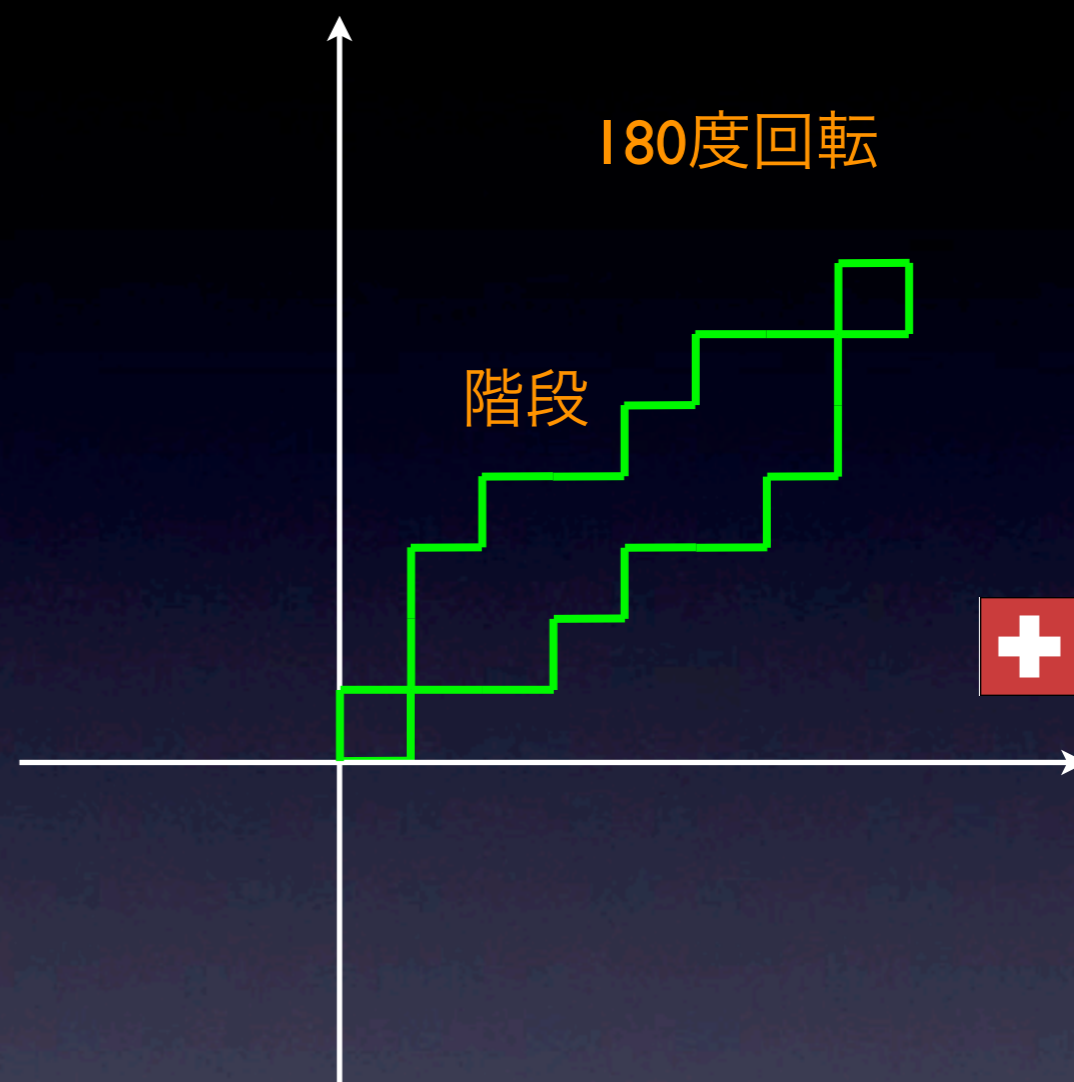









入れ子依存
 nested dependencies




交差依存
 crossed dependencies



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$
$b\bar{b}$	$abb\bar{a}$	$aabb\bar{a}\bar{a}$
	$ba\bar{a}\bar{b}$	$aba\bar{a}\bar{b}\bar{a}$
	$bbbb$	$abbbb\bar{a}$
		⋮
		$baaa\bar{a}\bar{b}$
		$babb\bar{a}\bar{b}$
		$bba\bar{a}\bar{b}\bar{b}$
		$bbbbbb$

入れ子依存

ありふれている



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$
$b\bar{b}$	$ab\bar{a}\bar{b}$	$aab\bar{a}\bar{a}\bar{b}$
	$bab\bar{a}$	$aba\bar{a}\bar{b}\bar{a}$
	$bbbb$	$abb\bar{a}\bar{b}\bar{b}$
		⋮
		$baab\bar{a}\bar{a}$
		$babb\bar{a}\bar{b}$
		$bbab\bar{a}\bar{b}\bar{a}$
		$bbbbbb$

交差依存

ごくまれ

入れ子の形は人間の言語や人工言語にいたるところに現れる
 交差は、人工言語にはない 人間の言語でもまれ
 2つの集合を区別するものは何か (複雑さではない)

$$\frac{3}{7}$$

$$\frac{31}{19}$$

有理数

$$x^2 - 6x + 4 = 0$$

$$\sqrt{2}$$

$$3 - \sqrt{5}$$

無理数


代数的数

$$\pi$$


$$e$$

$$\sqrt{2}^{\sqrt{2}}$$

超越数



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$abb\bar{a}$	$aabb\bar{a}\bar{a}$	
	$ba\bar{a}\bar{b}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bbbb$	$abbbb\bar{a}$	\vdots
		$baaa\bar{a}\bar{b}$	
		$babb\bar{a}\bar{b}$	
		$bba\bar{a}\bar{b}\bar{b}$	
		$bbbbbb$	



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$ab\bar{a}\bar{b}$	$aab\bar{a}\bar{a}\bar{b}$	
	$bab\bar{a}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bbbb$	$abb\bar{a}\bar{b}\bar{b}$	\vdots
		$baab\bar{a}\bar{a}$	
		$babb\bar{a}\bar{b}$	
		$bbabb\bar{a}$	
		$bbbbbb$	

代数的言語
= 文脈自由言語

$$S = aS\bar{a} \cup bS\bar{b} \cup \{a\bar{a}, b\bar{b}\}$$

形式言語にも「代数的」の概念がある

方程式によって定義可能

演算は文字列集合に対する演算（和／接続）



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$abb\bar{a}$	$aabb\bar{a}\bar{a}$	
	$ba\bar{a}\bar{b}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bb\bar{b}\bar{b}$	$abb\bar{b}\bar{b}\bar{a}$	\vdots
		$baaa\bar{a}\bar{a}\bar{b}$	
		$babb\bar{a}\bar{b}$	
		$bba\bar{a}\bar{b}\bar{b}$	
		$bbbb\bar{b}\bar{b}$	



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$ab\bar{a}\bar{b}$	$aaba\bar{a}\bar{b}$	
	$bab\bar{a}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bb\bar{b}\bar{b}$	$abb\bar{a}\bar{b}\bar{b}$	\vdots
		$baab\bar{a}\bar{a}$	
		$babb\bar{a}\bar{b}$	
		$bbabb\bar{a}$	
		$bbbb\bar{b}\bar{b}$	

$$S(ax\bar{a}) \leftarrow S(x)$$

$$S(bx\bar{b}) \leftarrow S(x)$$

$$S(a\bar{a}) \leftarrow \text{文脈自由言語}$$

$$S(b\bar{b}) \leftarrow$$

方程式のかわりに帰納的定義の形でも書ける
文脈自由文法と言う



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$abb\bar{a}$	$aabb\bar{a}\bar{a}$	
	$ba\bar{a}\bar{b}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bb\bar{b}\bar{b}$	$abb\bar{b}\bar{b}\bar{a}$	\vdots
		$baaa\bar{a}\bar{a}\bar{b}$	
		$babb\bar{a}\bar{b}$	
		$bba\bar{a}\bar{b}\bar{b}$	
		$bbbb\bar{b}\bar{b}$	

$$S(ax\bar{a}) \leftarrow S(x)$$

$$S(bx\bar{b}) \leftarrow S(x)$$

$$S(a\bar{a}) \leftarrow \text{文脈自由言語}$$

$$S(b\bar{b}) \leftarrow$$



$a\bar{a}$	$aa\bar{a}\bar{a}$	$aaa\bar{a}\bar{a}\bar{a}$	
$b\bar{b}$	$ab\bar{a}\bar{b}$	$aaba\bar{a}\bar{b}$	
	$bab\bar{a}$	$aba\bar{a}\bar{b}\bar{a}$	
	$bb\bar{b}\bar{b}$	$abb\bar{a}\bar{b}\bar{b}$	\vdots
		$baab\bar{a}\bar{a}$	
		$babb\bar{a}\bar{b}$	
		$bbabb\bar{a}$	
		$bbbb\bar{b}\bar{b}$	

$$S(x_1x_2) \leftarrow A(x_1, x_2)$$

$$A(ax_1, \bar{a}x_2) \leftarrow A(x_1, x_2)$$

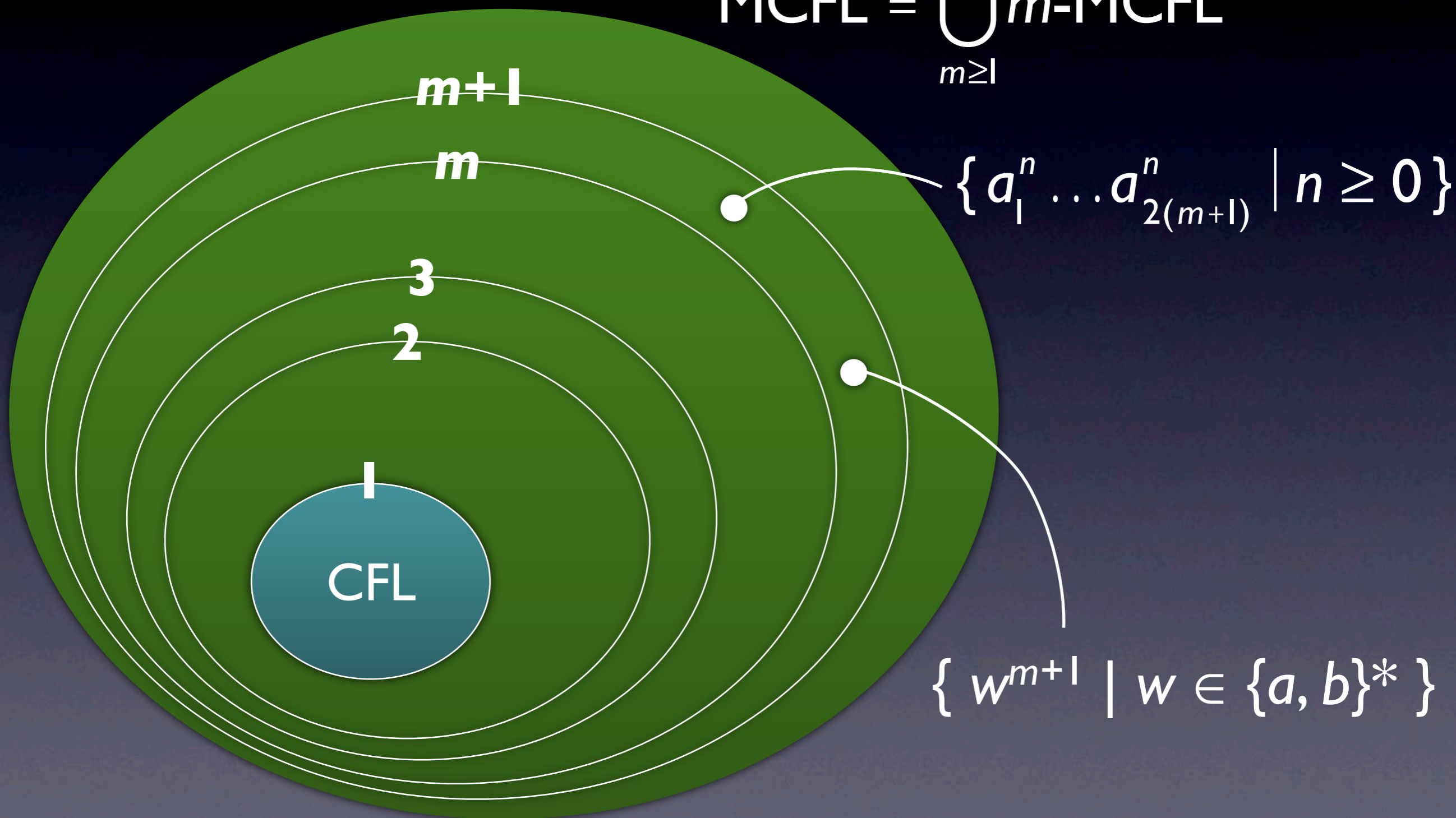
$$A(bx_1, \bar{b}x_2) \leftarrow A(x_1, x_2)$$

$$A(a, \bar{a}) \leftarrow \quad A(b, \bar{b}) \leftarrow$$

2-多重文脈自由言語

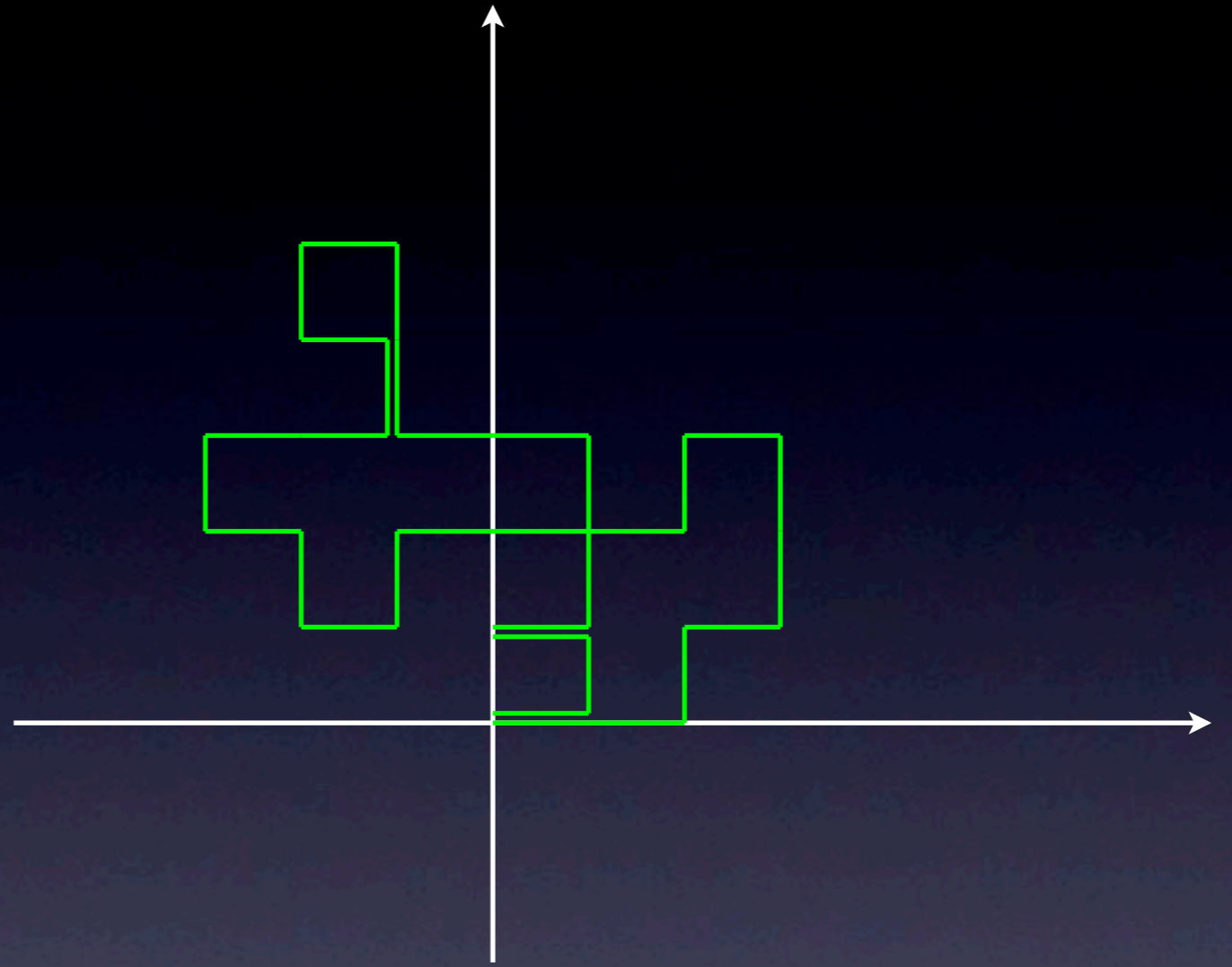
多重文脈自由言語の階層

$$\text{MCFL} = \bigcup_{m \geq 1} m\text{-MCFL}$$





Aravind K. Joshi



すべての原点回帰文字列の集合は
多重文脈自由言語か？

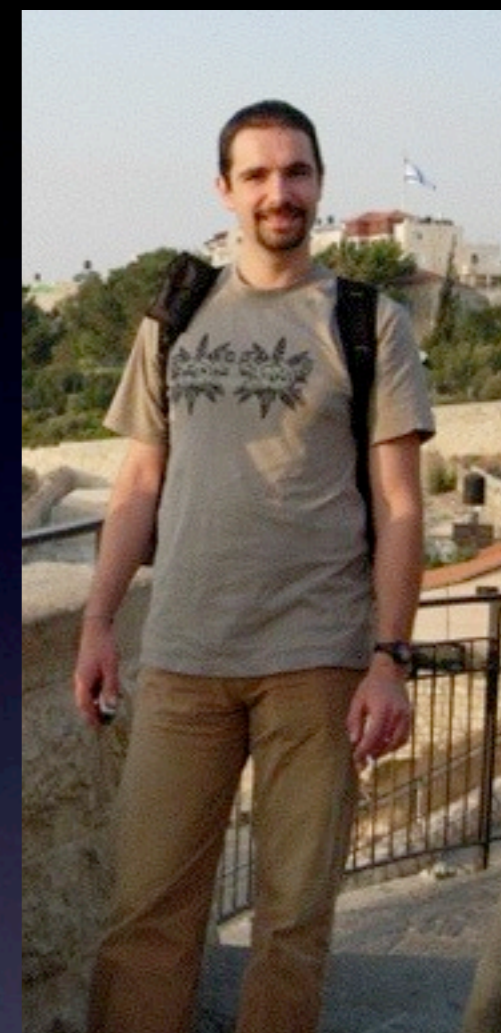
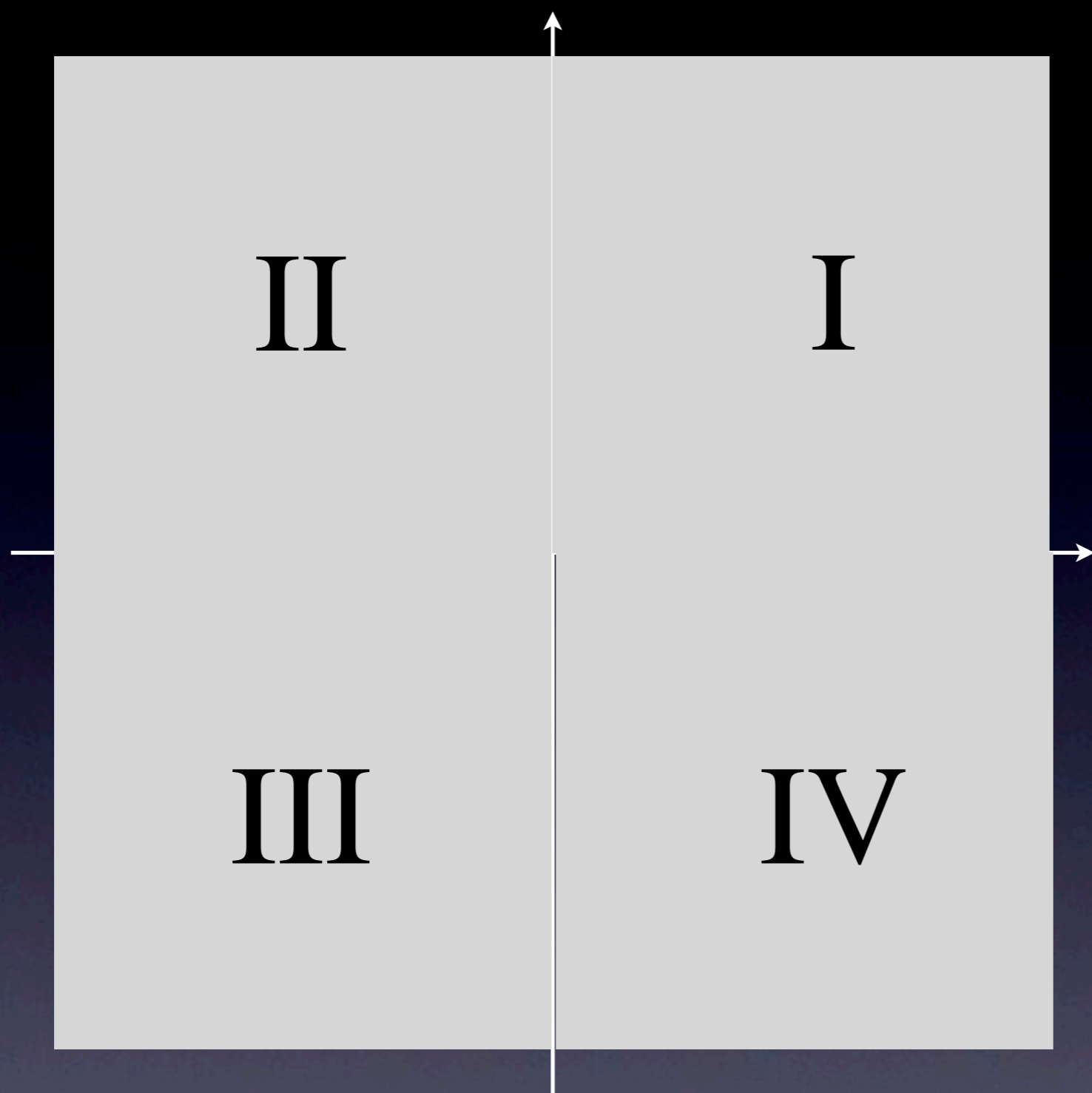
80年代からの予想 (否定的)



I

第I象限言語は
2-多重文脈自由言語でない

2009年に証明できた (?)



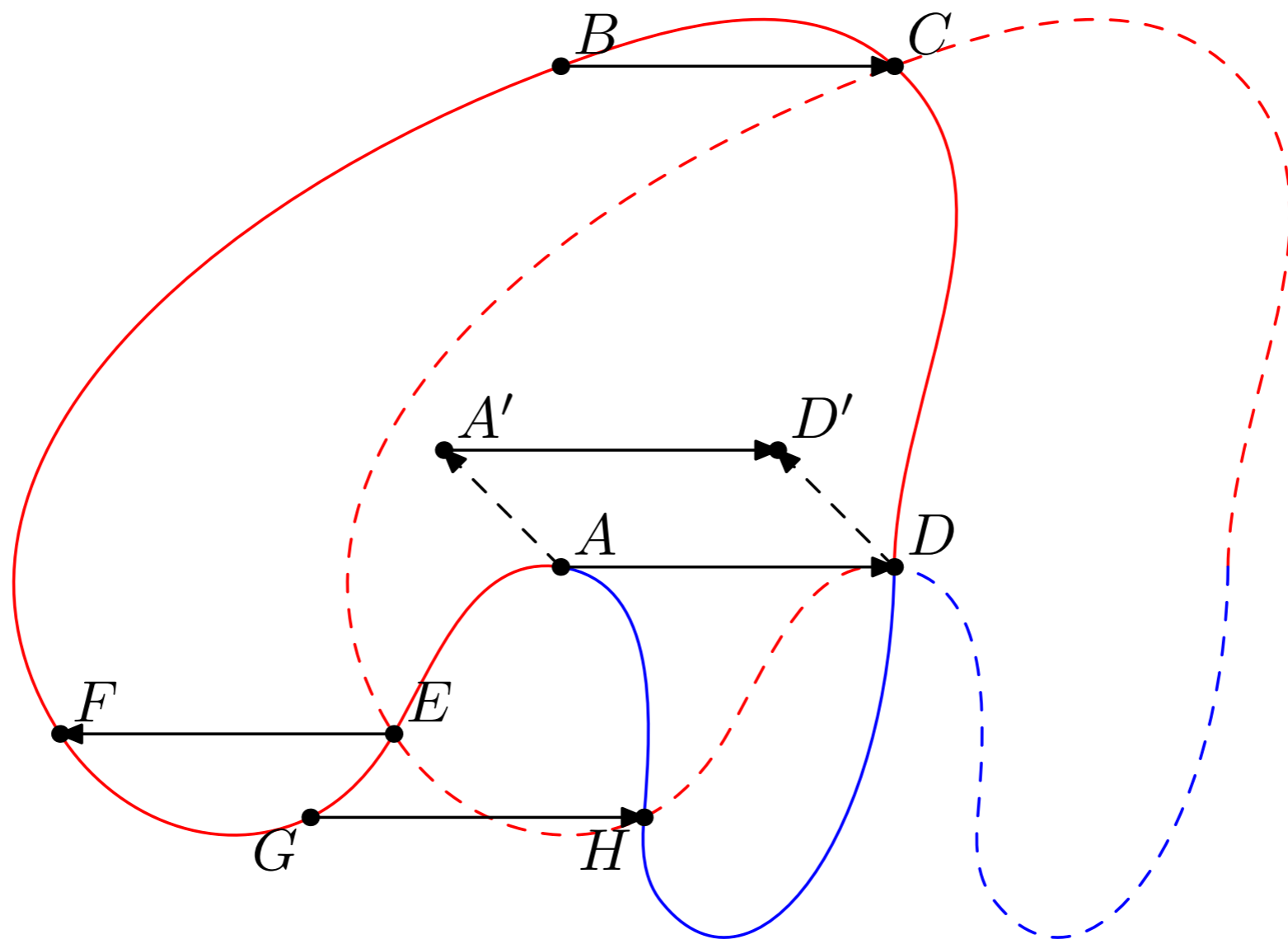
Sylvain Salvati

すべての原点回帰文字列の集合は
2-多重文脈自由言語！

今年証明が完成した

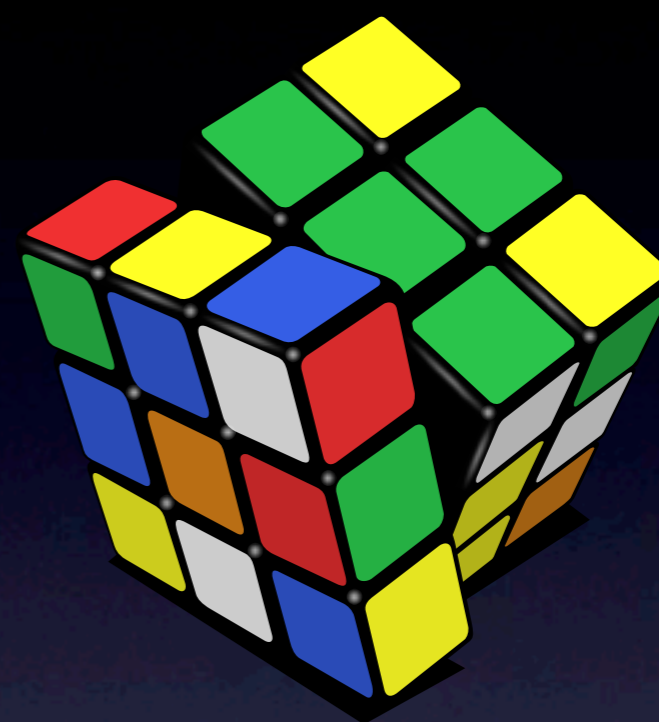
A theorem on Jordan curves

Theorem: If A and D are two points on a Jordan curve J such that there are two points A' and D' inside J such that $\overrightarrow{AD} = \overrightarrow{A'D'}$, then there are two points B and C pairwise distinct from A and D such that $A, B, C,$ and D appear in that order on one of the arcs going from A to D and $\overrightarrow{AD} = \overrightarrow{BC}$.





位相幾何学



群論

代数的位相幾何、群論が関連する

未解決問題

- 3次元の場合の原点回帰文字列の集合は多重文脈自由言語か？

その他未解決の問題がたくさんある

多重文脈自由言語がどんなものかまだよくわかっていない

人間の言語の記述にふさわしい文法形式は？

