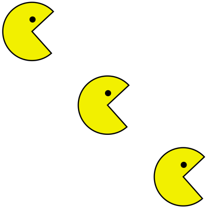
Deformations, torus actions and complexity of matrix multiplication

Joachim Jelisiejew

MIM UW Colloquium

University of Warsaw
MIM
Faculty
of Mathematics, Informatics
and Mechanics

Connection with classical game theory (Pac-man)



complexity
deformations
torus actions

Please don't eat me when I skim over details!

How many multiplications do we need to multiply two $n \times n$ matrices? $*$

$A \cdot B = C = [c_{ij}]$ ← n^2 entries

Usual algorithm: n^3 multiplications

$n \cdot n^2 = n^3$

Strassen '69: for $n = 2$

Complexity of matrix multiplication

How many multiplications do we need to multiply two $n \times n$ matrices?

Usual algorithm: n^3 multiplications

$2^3 = 8$

Strassen '69: for $n = 2$ need $\leq \underline{7}$ multiplications

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_1 + c_4 - c_5 + c_7 & c_3 + c_5 \\ c_2 + c_4 & c_1 - c_2 + c_3 + c_6 \end{bmatrix}$$

for $c_1 = (a_{11} + a_{22})(b_{11} + b_{22})$, $c_2 = (a_{21} + a_{22})b_{11}$,
$c_3 = a_{11}(b_{12} - b_{22})$, $c_4 = a_{22}(b_{21} - b_{11})$, $c_5 = (a_{11} + a_{12})b_{22}$,
$c_6 = (a_{21} - a_{11})(b_{11} + b_{12})$, $c_7 = (a_{12} - a_{22})(b_{21} + b_{22})$.

Complexity of matrix multiplication

How many multiplications do we need to multiply two $n \times n$ matrices?

Usual algorithm: n^3 multiplications

Strassen '69 & CW'70: for $n = 2$ need 7 multiplications

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_1 + c_4 - c_5 + c_7 & c_3 + c_5 \\ c_2 + c_4 & c_1 - c_2 + c_3 + c_6 \end{bmatrix}$$

for $c_1 = (a_{11} + a_{22})(b_{11} + b_{22})$, $c_2 = (a_{21} + a_{22})b_{11}$,
$c_3 = a_{11}(b_{12} - b_{22})$, $c_4 = a_{22}(b_{21} - b_{11})$, $c_5 = (a_{11} + a_{12})b_{22}$,
$c_6 = (a_{21} - a_{11})(b_{11} + b_{12})$, $c_7 = (a_{12} - a_{22})(b_{21} + b_{22})$.

Exact number of multiplications for $n = 3$ still open, in range $\{19, 20, 21, 22, 23\}$ (Bläser, Laderman).

Strassen's estimate

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} = ?$$

Strassen's estimate

$$\left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \cdot \left[\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{array}\right] = ?$$

$$\left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \cdot \left[\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{array}\right] =$$

$$\left[\begin{array}{c|c} A_{12,12} & A_{12,34} \\ \hline A_{34,12} & A_{34,34} \end{array}\right] \cdot \left[\begin{array}{c|c} B_{12,12} & B_{12,34} \\ \hline B_{34,12} & B_{34,34} \end{array}\right] = \text{use Strassen's trick twice!}$$

2×2 mats

7 multp for block mats

$A_{12,12} \cdot B_{12,12}$ each

require 7 multp $\boxed{49}$

Strassen's estimate

$$\left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \cdot \left[\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{array}\right] =$$

$$\left[\begin{array}{c|c} A_{12,12} & A_{12,34} \\ \hline A_{34,12} & A_{34,34} \end{array}\right] \cdot \left[\begin{array}{c|c} B_{12,12} & B_{12,34} \\ \hline B_{34,12} & B_{34,34} \end{array}\right] = \text{use Strassen's trick twice!}$$

for $n = 4$ need $\leq 7^2$ multiplications
for $n = 8$ need $\leq 7^3$ multiplications
for general n need $\approx n^{\log_2 7} < n^{2.81}$: there is an algorithm using $O(n^{2.81})$ multiplications called Strassen's algorithm

Complexity: what is known

'69 $\omega \leq 2.81$

Who: Bini, Schönhage, Coppersmith-Winograd, . . . , Alman-V.Williams.
What: proved existence of algorithm $O(n^{2.3729})$.

'89 $\omega \leq 2.376$

$\omega = 2$ conjecture

For every $\varepsilon > 0$ there is an algorithm in $O(n^{2+\varepsilon})$.
$\omega = \inf\{\tau \mid \text{algorithm in } O(\underset{n}{\cancel{2}}^{\tau}) \text{ exists}\}$.

Who: Landsberg-Michałek, . . .
What: needs at least $2n^2 - \mathit{errorTerm}$ multiplications.

Tensors

Tensor in $\mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^c = a \times b$ matrix with entries from $\mathbb{C}^c$.
Example: matrix multiplication tensor M_n lives in $\mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2}$.

$$\begin{array}{cc} & \begin{array}{cccc} b_{11} & b_{12} & b_{21} & b_{22} \end{array} \\ \begin{array}{c} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{array} & \begin{bmatrix} E_{11} & E_{12} & 0 & 0 \\ 0 & 0 & E_{11} & E_{12} \\ E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & E_{21} & E_{22} \end{bmatrix} \end{array}$$

$n = 2$ matrix multiplication as a tensor

A *rank one tensor* is a nonzero matrix as above with all entries proportional such that treating then as numbers one gets a usual rank one matrix.

no of multiplications = no of rank one tensors summing to M_n.

Laser method, Coppersmith-Winograd

Let T some tensor (e.g. matrix multiplication M_n).

1. *Rank* of T = minimal no of rank one tensors summing to T. Denote $\mathbf{R}(T)$.
2. *Border rank* of T = minimal r such that T is a limit of rank r tensors. Denoted $\underline{\mathbf{R}}(T)$.

Proposition (Bini)

We have $n^\omega \sim \mathbf{R}(M_n)$ and even $n^\omega \sim \underline{\mathbf{R}}(M_n)$.

Laser method, Coppersmith-Winograd

Let T some tensor (e.g. matrix multiplication M_n).

1. *Rank* of T = minimal no of rank one tensors summing to T. Denote $\mathbf{R}(T)$.
2. *Border rank* of T = minimal r such that T is a limit of rank r tensors. Denoted $\underline{\mathbf{R}}(T)$.

$\mathbb{C}^{abc}$

Proposition (Bini)

We have $n^\omega \sim \mathbf{R}(M_n)$ and even $n^\omega \sim \underline{\mathbf{R}}(M_n)$.

Laser method. Gives best known estimates on ω:

1. find a tensor $T \in \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ with $\underline{\mathbf{R}}(T) = d$. Then $\underline{\mathbf{R}}(T^{\otimes k}) \leq d^k$ for any k.
2. degenerate $T^{\otimes k}$ to M_n with $n = n(k)$ large. Get $\underline{\mathbf{R}}(M_n) \leq d^k$. So $\omega \leq \log_n(d^k)$.

Tensors from algebras

Let A be a d-dimensional vector space with a multiplication $\underline{A \times A \to A}$. Fix its basis $a_1, \ldots, a_d$. *Multiplication tensor* μ_A of A has $\underline{a_i \cdot a_j}$ in the (i, j) entry.

Example

For $A_{\text{gen}} = \underbrace{\mathbb{C} \times \mathbb{C} \times \ldots \times \mathbb{C}}_{d}$ with standard basis get

$$\mu_A = \begin{bmatrix} e_1 & 0 & 0 & \ldots & 0 \\ 0 & e_2 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & e_d \end{bmatrix}$$

$e_i := (0 .. 0, \underset{i}{1}, 0 .. 0)$

This is a rank d tensor!

Tensors from algebras

Let A be a d-dimensional vector space with a multiplication $A \times A \to A$. Fix its basis $a_1, \ldots, a_d$. *Multiplication tensor* μ_A of A has $a_i \cdot a_j$ in the (i, j) entry.

Example

For $A_{\mathrm{gen}} = \mathbb{C} \times \mathbb{C} \times \ldots \times \mathbb{C}$ with standard basis get

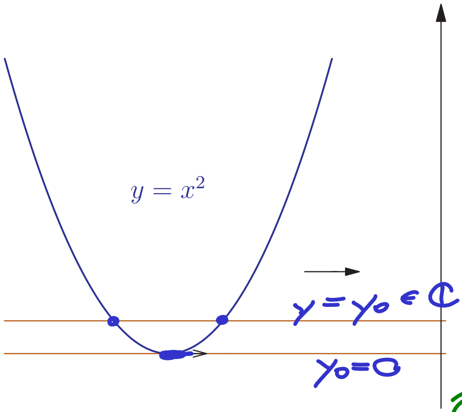
$$\mu_A = \begin{bmatrix} e_1 & 0 & 0 & \ldots & 0 \\ 0 & e_2 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & e_d \end{bmatrix}$$

This is a rank d tensor!

Idea: smoothability

If we have a degeneration of algebras $A_{\mathrm{gen}} \rightsquigarrow A$ then also $\mu_{A_{\mathrm{gen}}} \rightsquigarrow \mu_A$, hence multiplication tensor of A has border rank $\leq d$. Such A are called *smoothable*.

Example of degeneration, $d = 2$



For $y_0 \in \mathbb{C}$ consider

$$\mathbb{C}[x]/(y_0 - x^2).$$

In basis $1, x$ its multiplication tensor is

$$\begin{array}{cc} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} 1 & x \\ x & y_0 \end{bmatrix} \end{array}$$

For $y_0 \neq 0$ get $\mathbb{C} \times \mathbb{C}$,
for $y_0 = 0$ get $\mathbb{C}[x]/x^2$.

THIS IS A DEGENERATION OF $\mathbb{C} \times \mathbb{C} \rightsquigarrow \mathbb{C}[x]/x^2$

Expert slide one

Current best algorithm starts from a special algebra

$$A_{CW} = \frac{\mathbb{C}[x_1, \dots, x_{d-2}]}{(x_i x_j \mid i \neq j) + (x_i^2 - x_j^2 \mid i \neq j) + (x_1^3)}$$

CW 1-89

whose multiplication tensor is *big Coppersmith-Winograd* tensor:

$$\mathrm{CW}_{d-2} := \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_{d-2} & Q \\ x_1 & Q & 0 & \dots & 0 & 0 \\ x_2 & 0 & Q & \dots & 0 & 0 \\ & & & \dots & & \\ x_{d-2} & 0 & 0 & \dots & Q & 0 \\ Q & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Proposition (Hoyois-J-Nardin-Yakerson'21)

The multiplication tensor of every nice (=Gorenstein, smoothable) algebra gives bounds on ω which are better or equal than the Coppersmith-Winograd tensor.

Further in this talk algebras are commutative, associative and with identity. How to parametrize those?

Idea one: multiplication tensor.

An algebra A with a basis $a_1, \dots a_d$ is uniquely determined by $[\lambda_{ij}^k]_{1 \le i,j,k \le d}$ where

d^3 parameters

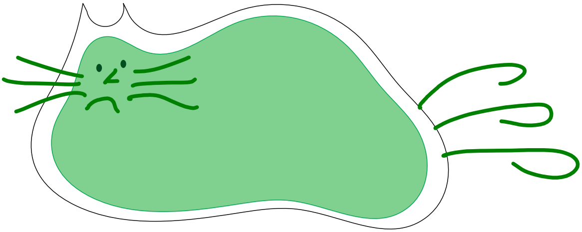
$$a_i \cdot a_j = \sum_{k=1}^{d} \lambda_{ij}^k a_k.$$

1. a_1 is the identity (1_A) iff $\lambda_{1j}^k = \delta_{jk}$, $\lambda_{i1}^k = \delta_{ik}$, with δ being Dirac delta,
2. A is commutative iff $\lambda_{ij}^k = \lambda_{ji}^k$,
3. A is associative iff $\sum_m \lambda_{ij}^m \lambda_{mk}^\ell = \sum_m \lambda_{im}^\ell \lambda_{jk}^m$.
4. A is Gorenstein iff *exists* a functional $f : A \to \mathbb{C}$ such that $(a_1, a_2) \mapsto f(a_1 a_2)$ is a perfect pairing.

Quadratic eqs
$(a_i \cdot a_j) \cdot a_k = a_i \cdot (a_j a_k)$

Deformations and degenerations of algebras

The set of rank d algebras with a basis is cut out by polynomial equations in $\mathbb{C}^{d^3}$. Has a natural topology! Even scheme structure.

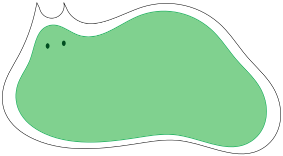


Problem: this topology is mighty complicated.

Theorem (J'20)

"Murphy's Law": every possible singularity (up to retraction) appears in this space for some d.

Expert slide two



Theorem (CEVV'08)

Every algebra is smoothable for $d \leq 7$.

Theorem (Casnati-J-Notari'13)

Every Gorenstein algebra is smoothable for $d \leq 13$.

Theorem (Szachniewicz, J.Marcinkiewicz & mFundacja prizes'21)

Already for $d = 13$ this space is nonreduced and at special points exhibits fractal-like behaviour (I am vague, read this preprint!).

Deformations and degenerations of algebras

How to parametrize algebras?
Recall:

$$A_{CW} = \frac{\mathbb{C}[x_1, \ldots, x_{d-2}]}{(x_i x_j,\ i \neq j) + (x_i^2 - x_j^2,\ i \neq j) + (x_1^3)}$$

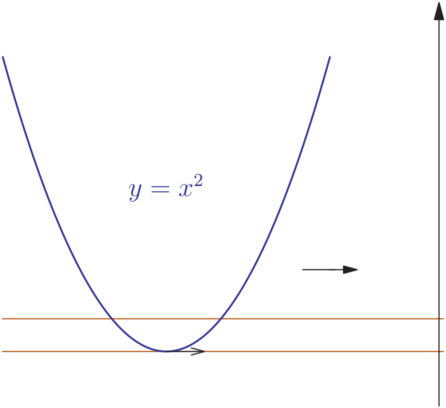
Idea two (Grothendieck): as quotients of a polynomial ring:

$$\{ I \lhd \mathbb{C}[x_1, \ldots, x_n] \mid A = \mathbb{C}[x_1, \ldots, x_n]/I,\ \dim_{\mathbb{C}} A = d \}$$

↑ the cat

with n, d fixed.
This goes under the fancy name: *Hilbert scheme* (gives +5 respect, -2 readability to paper).

The two ideas are equivalent: the spaces they give have same components and singularities for $n \geq d - 1$.

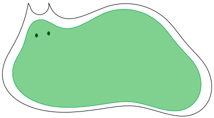


We have

$$\mathbb{C}[x, y]/(y - x^2),$$

where y is the parameter.
For $y = y_0 \neq 0$ get $\mathbb{C} \times \mathbb{C}$,
for $y = 0$ instead $\mathbb{C}[x]/x^2$.

This is a degeneration of $\mathbb{C} \times \mathbb{C}$ to $\mathbb{C}[x]/(x^2)$. This is a map from $\mathbb{C}$ to the space of algebras.



Torus actions

multi ↑

Action of the algebraic group $\mathbb{C}^*$ on a space X.

Assume X is complete

for every $x \in X$ the limit $\lim_{t\to\infty} t \cdot x$ exists

the function $x \mapsto \lim_{t\to\infty} t \cdot x$ may be NOT continuous.

think: $\mathbb{R}^*$-action on a manifold

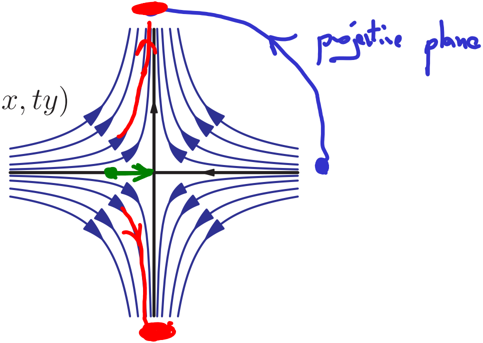
think: compact

for every $x \in X$ the limit $\lim_{t\to\infty} t \cdot x$ exists

the function $x \mapsto \lim_{t\to\infty} t \cdot x$ may be NOT continuous.

$t \to \infty$

$$t \cdot (x, y) = (t^{-1}x, ty)$$



The limit of any point is $\mathbb{C}^*$-fixed. Let $F_1, \ldots, F_r$ be the subdivision of fixed points into connected components and let

$$X_i = \left\{ x \in X \mid \lim_{t \to \infty} t \cdot x \in F_i \right\}.$$

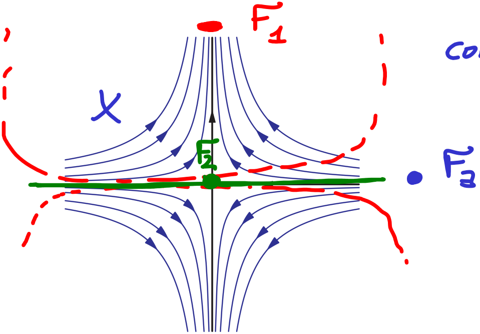
$X_i \xrightarrow{\lim} F_i$ pt

constant.

$X_1 = \mathbb{RP}^2 \setminus \mathbb{RP}^1 = \mathbb{R}^2$

$X_2 = \mathbb{R}^1$

$X_3 = \mathbb{R}^0$



Białynicki-Birula decomposition

Theorem (ASzBB'73)

If X is smooth then the limit function $X_i \to F_i$ is continuous, regular and its fibers are isomorphic to $\mathbb{C}^{n_i}$ (+local trivialization).

Theorem (Drinfeld'13)

The limit function $X_i \to F_i$ is continuous, regular and its fibers are isomorphic to cones in $\mathbb{C}^{n_i}$. (No smoothness assumption!)

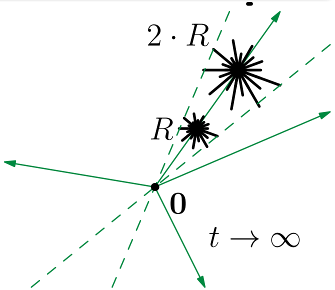
J-Sienkiewicz'19-21: some generalizations to groups other than $\mathbb{C}^*$.

Fix the $\mathbb{C}^*$-action known from linear algebra: $\mathbb{C}^* \times \mathbb{C}^n \to \mathbb{C}^n$ given by $t \cdot (v_1, \ldots, v_n) = (tv_1, \ldots, tv_n)$.

This gives a $\mathbb{C}^*$-action on the space of quotients $\mathbb{C}[x_1, \ldots, x_n]/I$.

1. x_i is v_i^*, so $t \cdot x_i = t^{-1} v_i$
2. $\mathbb{C}^*$ acts on $S = \mathbb{C}[x_1, \ldots, x_n]$
3. $\mathbb{C}^*$ acts on the space of its quotient rings: $t \cdot [S/I] = [S/(t \cdot I)]$

"S/I are functions on zero set of I"



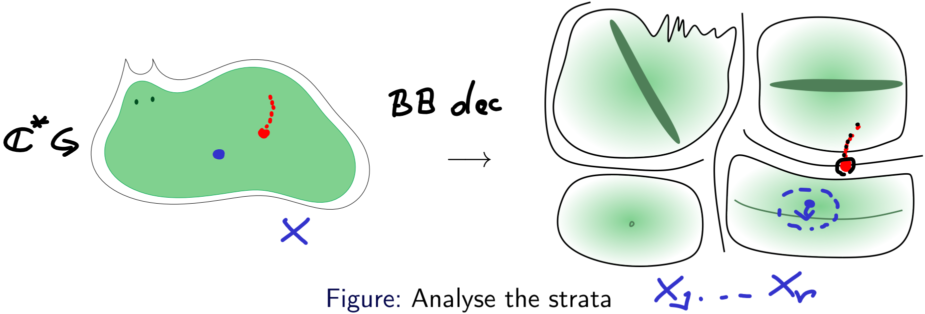


Figure: Analyse the strata

1. Each stratum converges to its fixed points: up to retraction can reduce to fixed points.
2. To prove smoothability of given point it is profitable to deform so as to escape the current stratum. **Effective** for smoothability:

   Previous approach: $d \leq 16$ some cases

   Now: $d \leq 100$ usually doable

Thanks for your attention!