Introduction, basics of Complexity, overview of topics

Monday, 1 May 2023 01:05

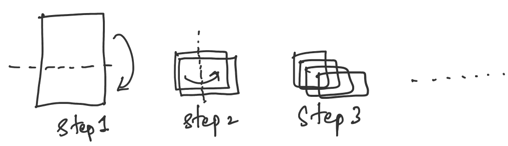
Algebraic Methods in Computational Complexity Theory
— Abhiram Natarajan

"Mathematical Study of Computation"

This course is about how algebraic methods interact with Complexity theory. It is **not** about complexity theory per se.

TED TALK ON COMPLEXITY THEORY
↪ along the lines of saying $1+2+3+\dots = -1/12$

Process A



Process B

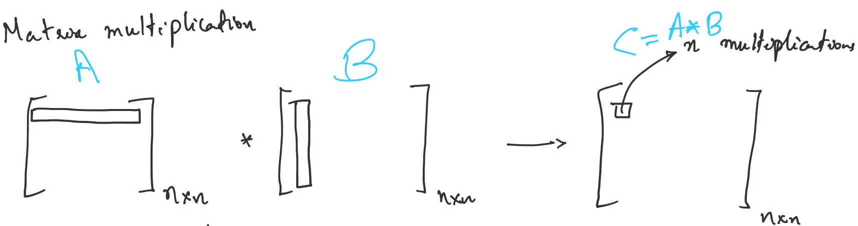
Step 1	Step 2	Step 3	....
10000 sheets of paper	stack 20000 sheets on top of step 1	stack 30000 sheets on top of step 2	

* Process A takes around 41 steps to get to the moon from earth
* Process B " " 17 steps " " " " " " "

* Process A takes ~ 50 steps to get to the sun
* Process B " ~ 387 steps —— " ——

COOL CONTRIBUTIONS OF COMPLEXITY THEORY :-

Matrix multiplication

① 

Can be done using n^3 multiplications

IS THIS UNBEATABLE? NO!

n^3 — trivial method.

Timeline of matrix multiplication exponent

Year	Bound on omega	Authors
1969	2.8074	Strassen[1]
1978	2.796	Pan[11]
1979	2.780	Bini, Capovani [it], Romani[12]
1981	2.522	Schönhage[13]
1981	2.517	Romani[14]
1981	2.496	Coppersmith, Winograd[15]
1986	2.479	Strassen[16]
1990	2.3755	Coppersmith, Winograd[17]
2010	2.3737	Stothers[18]
2013	2.3729	Williams[19][20]
2014	2.3728639	Le Gall[21]
2020	2.3728596	Alman, Williams[3]
2022	2.37188	Duan, Wu, Zhou[2]

* Context :→ 2 is the exponent for Matrix Addition !

MATRIX MULTIPLICATION $\stackrel{?}{=}$ MATRIX ADDITION

(Deepmind has the current best methods for multiplying very small matrices over $\mathbb{Z}_2$)

2. CT has given us the P vs NP question

expert in Donkey Kong / Super Mario Brothers ⟹ Solve the Riemann Hyp.

"THEORY OF NP-COMPLETENESS"

Classic Nintendo Games are (NP-)Hard

Greg Aloupis*   Erik D. Demaine†   Alan Guo†‡

March 9, 2012

Abstract

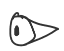
We prove NP-hardness results for five of Nintendo's largest video game franchises: Mario, Donkey Kong, Legend of Zelda, Metroid, and Pokémon. Our results apply to Super Mario Bros. 1, 3, Lost Levels, and Super Mario World; Donkey Kong Country 1–3; all Legend of Zelda games except Zelda II: The Adventure of Link; all Metroid games; and all Pokémon role-playing games. For Mario and Donkey Kong, we show NP-completeness. In addition, we observe that several games in the Zelda series are PSPACE-complete.

3. "Zero Knowledge Proofs."

you can convince that you have solved the Riemann hypothesis without revealing any information about the proof.

A (Blind)

YOU



Goal:- Convince skeptic A that balls are differently coloured

① Ball 1 in left hand & Ball 2 in right ; A shows you.

② A hides balls behind back, and either $\underbrace{\text{swaps or not}}_{\text{unknown to you}}$

③ A shows hands to YOU. YOU answer swap or no swap.

④ REPEAT as many times as needed.

A : - If correct answer provided at all times, then ball differently coloured w.h.p.

* A has no new information about the colours, and which is which.

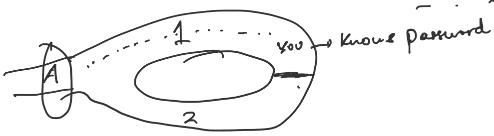
⇓

NO EXTRA INFORMATION

which

⇓

NO EXTRA INFORMATION LEAKAGE



① A lets you get inside Cave
② A asks you to come along side 1 or side 2
③ REPEAT

A : If all calls are completed successfully, then convinced that YOU know pwd.

THIS EXCHANGE CANNOT CONVINCE AN OBSERVER, BECAUSE IT IS TRIVIAL TO ORCHESTRATE

You can convince someone that you have a proof of a mathematical assertion this way!

4. YOU can write down a proof of the R.H. in such a way that anyone can check just ~ 30 bits of the proof to verify it, with very high confidence.

Probability of being falsely convinced
$\leq$ Probability that you are hallucinating

"Probabilistically Checkable Proofs"
"PCP Theorem"

5. P vs NP question

"Is there really a difference between DOING and CHECKING"

≡ Seperating orbit closures, and in turn, by proving things about multiplicities of irreducible representations

→ String theory of computer science

"GEOMETRIC COMPLEXITY THEORY (GCT) PROGRAM"

WE HAVE PROOFS THAT CERTAIN TECHNIQUES WONT WORK

① Relativization
② Natural proofs
③ Algebrization

WORK AROUND THIS CONJECTURE HAS BEEN UNIQUE

OVERVIEW OF TOPICS I AM intend TO COVER

① Matrix Multiplication

$A \times B = C$

① Matrix Multiplication

$$A \times B = C$$

NAIVE METHOD: $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$ takes n multiplications, $(n-1)$ additions per dot product, n^2 dot products $\Rightarrow O(n^3)$ steps

Thm [Klyuyev, Kokovkin-Scherbak '65] Optimal if only allowed to work on rows and columns as a whole.

Thm [Strassen '69] Can do better! $O(n^{2.81})$ operations

Observation: take 2×2 matrices... naive takes 8 products, 4 sums

$$\begin{bmatrix} c_{1,1} \\ c_{1,2} \\ c_{2,1} \\ c_{2,2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} (a_{1,1}+a_{2,2})(b_{1,1}+b_{2,2}) \\ (a_{2,1}+a_{2,2})(b_{1,1}) \\ (a_{1,1})(b_{1,2}-b_{2,2}) \\ (a_{2,2})(-b_{1,1}+b_{2,1}) \\ (a_{1,1}+a_{1,2})(b_{2,2}) \\ (-a_{1,1}+a_{2,1})(b_{1,1}+b_{1,2}) \\ (a_{1,2}-a_{2,2})(b_{2,1}+b_{2,2}) \end{bmatrix}$$

$$\boxed{\text{Takes 7 products, 18 sums}}$$

Observation: Works even if entries of A & B are from a non-comm. ring

entries can be from any non-comm ring

Thus entries can be $n \times n$ matrices

$$\overset{A}{\begin{bmatrix} \square_{n\times n} & \square_{n\times n} \\ \square_{n\times n} & \square_{n\times n} \end{bmatrix}} \times \begin{bmatrix} \square_{n\times n} & \square_{n\times n} \\ \square_{n\times n} & \square_{n\times n} \end{bmatrix}$$

$$M_2(M_n(\mathbb{C})) \cong M_{2n}(\mathbb{C})$$

- Thus you can recurse, and using the above idea, you can do matrix multiplication using $O(n^{\log_2 7})$ operations.
- It is a fact that no. of multiplications govern the complexity of matrix mult

Defn 'ω' - called the exponent of matrix multiplication

$$\omega := \inf\{ h \in \mathbb{R} \mid n \times n \text{ matrices can be multiplied using } O(n^h) \text{ operations}\}$$

$$\underbrace{2 \le \omega}_{\text{easy to show}} \overbrace{< \log_2 7}^{\text{Strassen}} \quad \text{as a first approximation}$$

We now know a much better upper bound on ω, which brings us to

Conjecture $\omega = 2$ b!!!!!!!! , i.e. multiplication of matrix is asymptotically the same as addition!!!!

All improvements begin with the following abstraction. Consider the matrix multiplication map

$$M_{\langle n \rangle} : \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$$

- $M_{\langle n \rangle}$ is a bilinear map, thus it can be thought of as an element of

$$(\mathbb{C}^{n\times n})^* \otimes (\mathbb{C}^{n\times n})^* \otimes (\mathbb{C}^{n\times n})$$

Thm ("geometric" characterization of ω)

$$\omega = \inf \{ \tau \in \mathbb{R} \mid \operatorname{rank} M_{\langle n \rangle} = O(n^\tau) \}$$

→ Thus upper bounds on ω correspond to studying decompositions of the matrix mult. tensor

→ For lower bounds, we must study secant varieties of the segre variety.

② Matrix mult is about solving efficiently. Flip side is — What can't we solve efficiently?

Defn "Efficient Solving" → An algorithm for a problem X is called efficient if the no. of operations of the algorithm is bounded from above by a polynomial in the size of any problem instance of X.

We say X is efficiently solvable if there exists an efficient alg. to solve X.

e.g. Problem Take in a list of 'n' numbers and output any one of them

Algorithm :- ① Input the list → n operations

② Output the first item in the list — 1 operation.

Total = $n+1$ operation
efficient

Problem Given list of n numbers, output the list in descending order

Alg 1 :- ① Generate all permutations of the list $n! \times n$

② check the permutations one by one and output whichever is sorted $n! \times n$

Total $\geq n!$ inefficient

Alg 2 :- ① Find max of $a_1 \dots a_n$, and output $\sim n$

② Delete max from the list and then repeat. $\sim O(1)$

Total $\sim O(n^2)$ operations
efficient

Problem Check if n is prime. length of input $\sim \log n$

Alg check for divisibility with all nos. from 2 to $\sqrt{n}$ — $\sim \sqrt{n}$ operations

Total $\sim \sqrt{n}$ operations
~~efficient~~
$\downarrow$
exponential in $\log n$, thus
INEFFICIENT

Happily, there exists an "efficient" algorithm to test primality.

Agrawal, Manindra; Kayal, Neeraj; Saxena, Nitin (2004). "PRIMES is in P" (PDF). *Annals of Mathematics*. **160** (2): 781–793.

Defn [Polynomial time reducibility]
If we can solve arbitrary instances of problem Y using a polynomial number of steps, plus a polynomial no. of calls to an algorithm that solves X, then we write

$$Y \leq_p X$$

"Y is polynomial time reducible to X" or "X is at least as hard as Y"

Fact If X is solvable efficiently, and $Y \leq_p X$, the Y is efficiently solvable too.

Defn [k-Satisfiability (k-SAT) problem]

Given: $x_1, \ldots, x_n$, $x_i \in \{0,1\}$ variables
$C_1 \ldots C_m$ Clauses $\quad m = O(n^c)$ c constant

$$C_i = \bigvee_{j=1}^{k} t_{i,j}, \text{ where } t_{i,j} \in \{x_1, \ldots, x_n, \overline{x}_1, \ldots, \overline{x}_n\}$$
(negation)

Ques does there exist/find assignment for x_i's $\in \{0,1\}^n$ s.t all Clauses are simultaneously satisfied

e.g. x_1, x_2, x_3

$C_1 = x_1 \vee \overline{x}_2$, $C_2 = \overline{x}_1 \vee \overline{x}_3$, $C_3 = x_2 \vee \overline{x}_3$

$(0,0,0)$ is a satisfying assignment

Trying all assignments takes $\geq 2^n$ time, NOT EFFICIENT

2-SAT is efficiently solvable

2. ^ a b c d e f Krom, Melven R. (1967), "The Decision Problem for a Class of First-Order Formulas in Which all Disjunctions are Binary", *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **13** (1–2): 15–20, doi:10.1002/malq.19670130104.

(≥ 3)-SAT — unknown, conjectured to not be efficiently solvable

Defn "Efficient certification" – A soln. to a problem instance of X can be efficiently certified if the proposed soln can be verified in polynomial time. We'll say X is efficiently certifiable if solutions to arbitrary instances are efficiently certifiable.

e.g. Factoring
→ solving unknown

Unsolved problem in computer science:

? *Can integer factorization be solved in polynomial time on a classical computer?*

(more unsolved problems in computer science)

→ certification efficient!

3-SAT
→ solving unknown
→ certification efficient

Defn P → class of problems efficiently solvable
NP – class of problems efficiently certifiable
$P \subseteq NP$, is converse true?

Conjecture $P \neq NP$

Defn X is an NP complete problem if
(a) $X \in NP$
(b) $\forall Y \in NP \quad Y \leq_p X$

X is the 'hardest' problem in NP.

Thm [Cook-Levin Theorem] ① Circuit-SAT is NP-complete

Corollary Circuit-SAT $\leq_p$ 3-SAT — this gives a recipe for proving NP-completeness
$\Rightarrow$ 3-SAT is NP complete

* If you can give a poly time alg. for 3-SAT (or for any NP-Complete Problem), you've proved $P = NP$

2. ^ a b c d e f Krom, Melven R. (1967), "The Decision Problem for a Class of First-Order Formulas in Which all Disjunctions are Binary", *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **13** (1–2): 15–20, doi:10.1002/malq.19670130104.

* If you show no poly time alg. exists for 3SAT (any NP-Complete Problems), you've proved $P \neq NP$

1000's OF NP-COMPLETE PROBLEMS

P vs NP is out of reach at the moment. We don't even know if 3-SAT requires superlinear time, i.e. $\omega(n)$ is also not known

P vs NP is out of reach at the moment. We don't even know if 3-SAT requires superlinear time, i.e. $\omega(n)$ is also not known.
Also, 3-SAT is conjectured to not even have "slightly" better than 2^n time algorithms.

Exponential time hypothesis

Article Talk — Read Edit View history Tools — 5 languages

From Wikipedia, the free encyclopedia

In computational complexity theory, the **exponential time hypothesis** is an unproven computational hardness assumption that was formulated by Impagliazzo & Paturi (1999). It states that satisfiability of 3-CNF Boolean formulas cannot be solved in subexponential time, i.e., $2^{\epsilon n}$ for all constant $\epsilon > 0$.

Instead we shall work on a "cleaner", "more-algebraic" conjecture that is to be thought of as the algebraic analogue of the P vs NP question

Defn VP – Class of polynomials that are easy to evaluate.
("take polynomial time to evaluate")

VNP – Class of polynomials whose coefficients are easy to evaluate

to define rigorously, we need to define circuit size, but let us look at examples instead

$\det_n \in VP$. (Gaussian elimination $= O(n^3)$ time) — has $n!$ terms

$$\mathrm{Perm}_n = \sum_{\sigma \in S_n} x_{1,\sigma(1)}\, x_{2,\sigma(2)} \cdots x_{n,\sigma(n)} \in VNP$$

Permanent not as natural as det, but does show up.

Question (Algebraic Analogue of P vs NP) [VALIANT '1979]

$$VP \overset{?}{=} VNP$$

$\boxed{\text{Can be investigated over any field}}$

there is evidence that this might be easier over p-adics

For exact relationship b/w VP vs VNP & P vs NP, see

Bürgisser, Peter; Clausen, Michael; Shokrollahi, M. Amin (1997). Algebraic complexity theory. Grundlehren der Mathematischen Wissenschaften. Vol. 315. With the collaboration of Thomas Lickteig. Berlin: Springer-Verlag. ISBN 978-3-540-60582-9. Zbl 1087.68568.

Thm [Valiant '1979]

$$VP \neq VNP \iff \mathrm{Perm}_n \notin VP$$

Notice $\det_n \in VP$ .... $n!$ monomials, but Poly(n) computable

$\Downarrow$

$\det_{\mathrm{poly}(n)} \in VP$ .... Poly(n)! monomials, then still poly(n) computable

Thus if perm_n can be expressed as the det of a matrix of size polynomial in n, then $\mathrm{Perm}_n \in VP \Rightarrow VP = VNP$

Defn [determinantal complexity] for $f \in \mathbb{F}[\bar{x}]$,
$dc(f)$ is min s s.t. there are affine linear forms
$\alpha_{i,j} \in \mathbb{F}[\bar{x}]$ $1 \le i,j \le s$,

such that

$$f = \det \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,s} \\ \vdots & & \vdots \\ \alpha_{s,1} & \cdots & \alpha_{s,s} \end{bmatrix}$$

$$f = \det \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,s} \\ \vdots & & \vdots \\ \alpha_{s,1} & \cdots & \alpha_{s,s} \end{bmatrix}$$

We will show that

$$\frac{n^2}{2} \leq dc(\mathrm{perm}_n) \leq 2^n - 1$$

$\boxed{\text{Conjecture } dc(\mathrm{perm}_n) \text{ grows faster than any polynomial in } n}$

How do we make progress? By rephrasing this in language comfortable to us.

on $VP \stackrel{?}{=} VNP$

FEWNOMIALS

e.g. $f = 7x^{100} - 22x^{32} + 45x^{21} + 9$

Sure we can say $f \leq 100$ real roots, but can we do better? Can we say anything independent of the degree, but depending on the sparsity.

Thm [DESCARTES RULE OF SIGNS (DROS)]
for $f \in \mathbb{R}[x]$ of sparsity t, no. of real roots (counted with multiplicity)

$$\sim 2t$$

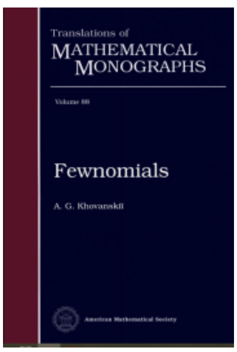
Thm [MULTIVARIATE DROS] (Like a Bezout thm for fewnomials)
Let $f_1, \dots, f_n \in \mathbb{R}[x_1, \dots x_n]$. Also, let the system

$\{ f_i = 0$ have $l+n+1$ distinct monomials.

System has at most

$$2^{\binom{l+n}{2}} (n+1)^{l+n}$$

non-singular solutions in the positive ~~oct~~ octant of $\mathbb{R}^n$.



→ Proved more generally for system of Pfaffians

Conjecture [Real-Tau Conjecture] Consider polynomials of the form, f_{ij} t-sparse $\in \mathbb{R}[x]$

Koiran

$$f = \sum_{i=1}^{m} \prod_{j=1}^{k} f_{ij}$$

no. of real zeros of f poly$(m, t, 2^k)$

Thm Conjecture $\Rightarrow VP_{\mathbb{C}} \neq VNP_{\mathbb{C}}$
[Tavenas]

[Tavenas]

Thm [Briquel-Bürgisser] if $f_{i,j}$ are chosen as follows:-

(1) fix any support of size $\leq t$

(2) choose co-eff to be indep $\mathcal{N}(0,1)$. Then

$\mathbb{E}[\text{zeros of } f] \leq O(m k^2 t)$

(necessary condition for $VP_{\mathbb{C}} \neq VNP_{\mathbb{C}}$)

Geometric Complexity Theory (GCT)

"String theory of computer science" Let $V = (\mathbb{C}^{m^2})^*$.

$GL(V)$ – group of all automorphisms of $(\mathbb{C}^{m^2})^*$ acts on

$Sym^m(V) \rightarrow$ degree m homogeneous polys in m^2 variables by

$\underset{GL(\cdot)}{L} \cdot \underset{Sym(\cdot)}{p(x)} := p(L^T x)$ (b'cos of transpose $L_1 \cdot (L_2 \cdot p) = (L_1 L_2) \cdot p$)

observe $det_m \in Sym^m(V)$.

Let $n \leq m$. Per is a lower degree poly over fewer vars. Define the padded permanent

$Per^*_{m,n} = \underset{\text{auxiliary variable}}{x^{m-n}_{m,m}} \, Per_n \quad \in Sym^m(V)$

Conjecture [Mulmuley-Sohoni] If $m = 2^{n^{O(1)}}$, then for sufficiently large n,

$Per^*_{m,n} \notin \overline{GL(V). det_m}$

Thm $dc(perm_n) \leq m \implies Per^*_{m,n} \in \overline{GL(V). det_m}$

[MS] Orbit closures is required b'coz

$\underset{\text{reducible}}{Per^*_{m,n}} \notin \underset{\text{all irreducible}}{GL((\mathbb{C}^{m^2})^*). det_m}$, also b'coz

So far hardly anything novel has happened. But.

Thm (a) if P, Q are permutation matrices, A, B are diag matrices with $Per(A) Per(B) = 1$, and

$f(X) = f(PXQ) = f(AXB) \;\; \forall P,Q,A,B \implies$ then $f(X)$ is a $\mathbb{C}$-multiple of permanent

(b) if A, B are such that $det(A)\, det(B) = 1$ and $f(X) = f(AXB) \;\; \forall A, B$, then $f(X)$ is a $\mathbb{C}$-multiple of the determinant

Both per & det are characterized by their symmetries. This helps avoid barriers (a) Natural proofs (b) Relativization (c) Algebrization

Per & det satisfy something called "partial stability", thus Luna's Etale Slice theorem from Geom. inv theory can be used to get a handle on their orbit closures

Luna's Etale slice theorem from ...

to get a handle on their orbit closures

↓ This gives

Let $R_{det} := \mathbb{C}[\overline{GL(V).det_m}]$ & $R_{per} := \mathbb{C}[\overline{GL(V).Per^*_{m,n}}]$.

$GL(V)$ acts on both R_{det} & R_{per}

$A \cdot q(p(x)) := q(p(A^T x))$,

$q \in R_{det} | R_{per}$

thus we get two representation of $GL(V)$

ρ_{det} & ρ_{per}

Let $\lambda_{per}(\rho) / \lambda_{det}(\rho)$ denote the multiplicity of the irreducible representation ρ in the isotypic decomp of $\rho_{per} | \rho_{det}$.

Thm Suppose there exists irrep ρ s.t. $\lambda_{per}(\rho) > \lambda_{det}(\rho)$. Then $Per^*_{m,n} \notin \overline{GL(V).det_m}$

(not an iff!)

we have hope to tackle this b'coz there are surprising algorithms to calculate multiplicities of irreps

(∃ ρ s.t $\lambda_{per}(\rho) > 0$ & $\lambda_{det}(\rho) = 0$) – Conjecture has been proved false

---

"NOT ENOUGH YELLOW BOOKS HAVE BEEN WRITTEN"

ULRICH COMPLEXITY

for determinantal complexity of f, we want $f = det(M)$. If we look for $f^r = det(M)$, this falls into the domain of Ulrich Sheaves / Modules (extensively studied)

Defn $uc(f)$ is smallest r s.t there exists M of linear forms with $\det M = f^r$, & there exist N s.t

$$MN = f.I$$

Thm $VP \neq VNP \implies uc(perm_n) \geq 2^{n-2}$

might be easier to prove ($dc(\sum_{i=1} x_i y_i) \leq C+1$, $uc(\sum_{i=1} x_i y_i) = 2^{\lceil c/2 \rceil - 2}$)

---

equivalent way of thinking about UC

Let $f \in k[x_0, \dots, x_n]$ be homogeneous. Repose the standard grading on ($\deg x_i = 1$) $k[x_0, \dots, x_n]$. Let S be the graded ring. Let $R = S/\langle f \rangle$. Let F – be a finitely generated R-module. Defn F is Ulrich module if F has a free resolution of the form

$$0 \to S^n(-1) \xrightarrow{M} S^n \to F \to 0$$

(M - matrix of linear forms)

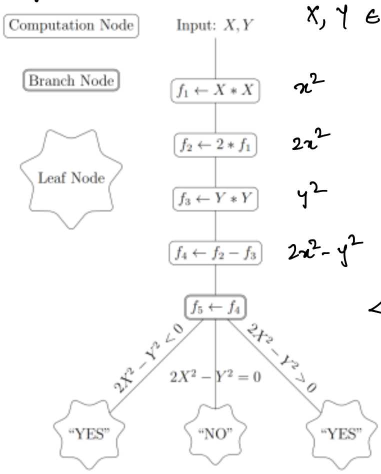
Turns out

$$uc(f) = \inf \{ \operatorname{rank} F \mid F \text{ is an Ulrich module on } R \}$$

MISCELLANEOUS TOPICS

- Algebraic Computation tree - Model of Computation that represents the computational steps that a Turing machine would execute

e.g.



tests if (x,y) belongs to the semialgebraic set $2x^2 - y^2 \neq 0$

($x^3 + 3x^2 + 3x + 1 = (x+1)^3$) . Q: what is the best way to compute ?
$\iff$ what is the least height ?

Thm [Gabrielov - Vorobjov] Consider problem of testing membership in a semialgebraic set $S \subseteq \mathbb{R}^n$.

Singular Betti numbers

height of algebraic comp tree $\geq \dfrac{c_1 b_m(S)}{m+1} - c_2 n$

- Complexity theory of Constructible Sheaves (Basu 2017)
  defines constructible analogs of VP & VNP
- Categorical Complexity (Basu - Isik 2018)
  defines notions of complexity of Categories & functors
  recovers classical notions in complexity theory

"rising sea" approach to complexity theory

The Rising Sea

Contents

The "rising sea" is a metaphor due to Alexander Grothendieck (see the quote below), meaning to illuminate how the development of general abstract theory eventually brings with it effortless solutions to concrete particular problems, much like a hard nut may be cracked not immediately by sheer punctual force, but eventually by gently immersing it into a whole body of water.

- Freedman

ANNALS OF MATHEMATICS

Complexity classes as mathematical axioms

By Michael H. Freedman

Proves $\sim P \neq NP \Rightarrow$ Knots with certain properties exist

Matrix Multiplication

Saturday, 6 May 2023 20:48

Complexity of Matrix Multiplication

$$A_{n\times m} \cdot B_{m\times \ell} = C_{n\times \ell}$$

$$C_{i,j} = \sum_{k=1}^{m} a_{i,k}\, b_{k,j}$$

For $n\times n$ matrices, naive method takes n multiplications & $(n-1)$ additions per dot product, a total of n^2 dot products, thus $O(n^3)$ total operations.

Strassen showed

① You can do 2×2 times 2×2 matrices using 7 products and 18 sums instead of 8 products & 4 sums.

② Since alg. works with entries in any ring, you can recurse and do $n\times n$ times $n\times n$ in $O(n^{\log_2 7})$ operations.
$\overset{\shortparallel}{n^{2.81}}$

note: 7 comes because it is the no. of multiplications in Strassen's algorithm.

It is a fact (check proposition 15.1 in the Burgisser et al. textbook) that the total complexity is governed by the no. of multiplications

Thus we can look at **TENSOR RANK** → a geometric object to count the number of multiplications

Thus $\omega := \inf \{ h \in \mathbb{R} \mid n\times n$ matrices can be multiplied using $O(n^h)$ arithmetic operations$\}$, can be equivalently characterized geometrically.

$M_{\langle k,m,n\rangle} : \mathbb{C}^{k\times m} \times \mathbb{C}^{m\times n} \to \mathbb{C}^{k\times n}$ is a bilinear map, thus

... be thought of as a tensor in

$\langle k,m,n \rangle$

$M_{\langle k,m,n \rangle}$ can be thought of as a tensor in

$$(\mathbb{C}^{k\times m})^* \otimes (\mathbb{C}^{m\times n})^* \otimes (\mathbb{C}^{k\times n})$$

Thm [Strassen '69, geometric characterization of ω]

$$\omega = \inf \{ \tau \in \mathbb{R} \mid \underset{\text{rank}}{R} (M_{\langle n,n,n \rangle}) = O(n^\tau) \}$$

N.B. (1) In principle ω can depend on the characteristic of the field, but I don't know of any fact about matrix multiplication that is different for different fields.

(2) (Conjecture) $\omega = 2$.

(3) ω is defined to be a limit pt. It is known that the limit CANNOT be achieved. An additional $n^{o(1)}$ is required

Plan:-

① Explain Strassen's 2×2 observation as intuitively as possible.

Using upper bounds on

$$R(M_{\langle k,m,n \rangle})$$

② for a specific (k,m,n) triple to get an u.b. of ω.

(recall Strassen explicitly shows $R_{\langle 2,2,2 \rangle} \le 7$) (Pan shows $R_{\langle 70,70,70 \rangle} \le 143{,}240$) → also practical.

③ Border rank + Bini

using upper bounds on

(border rank) $\underline{R}(M_{\langle k,m,n \rangle})$

for a specific (k,m,n) triple to get u.b. of ω

(Bini explicitly shows $\underline{R}(M_{\langle 2,2,3 \rangle}) \le 10$)

④ Schonhage's τ theorem

(4) Schonhage's τ theorem
upper bounds on $R(\oplus M_{\langle * \rangle})$ to get
upper bound on ω.

(5) Briefly mention the path taken by Coppersmith-Winograd and subsequent works

(6) Cohn-Umans group theoretic approach

CU + other authors show how to match the Coppersmith Winograd approach, and also give conjectures that would imply $\omega = 2$.

Strassen's Algorithm. Let's rephrase mat. mult. symmetrically

$$M_{\langle n \rangle} : M_n \times M_n \to M_n \quad \text{(} n \times n \text{ matrices)}$$

Treat it as an element of $M_n^* \otimes M_n^* \otimes M_n^*$, i.e.

abuse of notation $\to$ $$M_{\langle n \rangle} = \sum_{i,j,k \in [n]} E_{i,j}^* \otimes E_{j,k}^* \otimes E_{k,i}^* .$$

Observe given $A \otimes B \otimes C \in M_n \otimes M_n \otimes M_n$

$$\langle M_{\langle n \rangle}, A \otimes B \otimes C \rangle = \text{trace}(ABC)$$

$$(AB)_{i,k} = \langle M_{\langle n \rangle}, A \otimes B \otimes E_{k,i} \rangle$$

Notice $\langle M_{\langle n \rangle}, (Z^{-1} A X) \otimes (X^{-1} B Y) \otimes Y^{-1} C Z \rangle$

$$= \text{trace}(ABC).$$

Fact $M_{\langle n \rangle}$ is the only operator (up to a constant) that has this symmetry.

Defn A set S of n-dimensional vectors is a unitary 2-design if

$$\sum_{v \in S} v = 0 \quad \text{and} \quad \frac{1}{|S|} \sum_{v \in S} |v\rangle\langle v| = \frac{1}{n} I$$

$$\sum_{v\in S} v = 0 \quad \text{and} \quad \frac{1}{|S|}\sum_{v\in S} |v\rangle\langle v| = \frac{1}{n} I$$

Thm Let $S=\{w_1,\dots,w_s\}$ be a unitary 2-design. Then tensor rank of $T_{\langle n\rangle}$ is at most $s(s-1)(s-2)+1$

Lazy proof By defn.

$$① - \frac{s^3}{n^3}\cdot I^{\otimes 3} = \sum_{i,j,k\in[s]} \underbrace{|w_i\rangle\langle w_i|}_{s/n\, I} \otimes \underbrace{|w_j\rangle\langle w_j|}_{s/n\, I} \otimes \underbrace{|w_k\rangle\langle w_k|}_{s/n\, I}$$

$$② - \frac{s^3}{n^3} M_{\langle n\rangle} = \sum_{i,j,k\in[s]} |w_i\rangle\langle w_j| \otimes |w_j\rangle\langle w_k| \otimes |w_k\rangle\langle w_i|$$

	i,j	jk	ki
MM -	1,1	1,1	1,1
+	1,2	2,1	1,1
	1,1	1,2	2,1
	1 2	2,2	2,1
	⋮		

	$i\ i$	$j\ j$	$k\ k$
I	1,1	1,1	1,1
	1,1	1,1	2,2
	1,1	2,2	1,1
	1,1	2,2	2,2
		⋮	

$$② - ① \quad \frac{s^3}{n^3} M_{\langle n\rangle} - \frac{s^3}{n^3} I^{\otimes 3} =$$

$$\sum_{\substack{i,j,k \\ \text{distinct}}} |w_i\rangle\langle w_j - w_i| \otimes |w_j\rangle\langle w_k - w_j| \otimes |w_k\rangle\langle w_i - w_k|$$

$$M_{\langle n\rangle} = I^{\otimes 3} + \frac{n^3}{s^3}\left(\text{sum of } \binom{s}{3} \text{ terms}\right)$$

Thus rank $(M_{\langle n\rangle}) \le \binom{s}{3} + 1$

$\boxtimes$

In $n=2$, the three corners of an equilateral traingle form a 2-design

2-design

$$S = \{ (1,0),\ (-1/2, \sqrt{3}/2),\ (-1/2, -\sqrt{3}/2) \}$$

The outer products are

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 1/4 & -\sqrt{3}/4 \\ -\sqrt{3}/4 & 3/4 \end{pmatrix},\ \begin{pmatrix} 1/4 & \sqrt{3}/4 \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{pmatrix}$$

$|S| = 3$, Thus $R(M_{\langle 2 \rangle}) \leq 7$. $\boxtimes$

Proposition If V is a nontrivial irrep of $\underset{\uparrow \text{ finite group}}{G}$, then for any $v \in V$ with $|v|^2 = 1$, the orbit of v, i.e. $\{gv\}_{g \in G}$ is a unitary 2-design.

Proof Schur's lemma $\boxtimes$

—— Machinery to get u.b. on ω using u.b on any $R(M_{\langle k,m,n \rangle})$ ——

Permutation of tensors:- Suppose $t \in A \otimes B \otimes C$, s.t

$t = \sum_{j=1}^{r} t_j$, where $t_j = a_{j,1} \otimes a_{j,2} \otimes a_{j,3}$. For $\pi \in S_3$

define $\pi(t) = \sum_{j=1}^{r} \pi(t_j)$, where $\pi(t_j) = a_{j,\pi^{-1}(1)} \otimes a_{j,\pi^{-1}(2)} \otimes a_{j,\pi^{-1}(3)}$

Lemma $\pi(t)$ is well-defined, i.e. if there are two decompositions of t, ~~Fa~~

$$t = \sum_{j=1}^{r} a_{j,1} \otimes a_{j,2} \otimes a_{j,3} = \sum_{j=1}^{s} b_{j,1} \otimes b_{j,2} \otimes b_{j,3},$$

then apply π to both should yield the same result. $\boxtimes$

Lemma $R(t) = R(\pi(t))$

Proof $\pi^{-1} \in S_3$, thus $t = \pi^{-1}(\pi(t))$. Result follows from well-definedness. $\boxtimes$

Let $f_1 : A \to A'$, $f_2 : B \to B'$, $f_3 : C \to C'$

well-definedness. $\boxtimes$

Defn Let $t \in A \otimes B \otimes C$. Let $f_1 : A \to A'$, $f_2 : B \to B'$, $f_3 : C \to C'$.

Then $(f_1 \otimes f_2 \otimes f_3).(t) = (f_1 \otimes f_2 \otimes f_3)\left(\sum_{j=1}^{r} a_{j,1} \otimes a_{j,2} \otimes a_{j,3}\right)$

$$= \sum_{j=1}^{r} f_1(a_{j,1}) \otimes f_2(a_{j,2}) \otimes f_3(a_{j,3})$$

The tensor $(f_1 \otimes f_2 \otimes f_3).(t) =: t'$ is called a "restriction" of t.

Denoted $t' \lesssim t$

Lemma $R((A \otimes B \otimes C)\, t) \leq R(t)$ with equality when A, B, C are isomorphisms

Proof: obvious → The transformation could allow us to

$$\left(\begin{array}{l} a \otimes b \otimes c_1 \\ + a \otimes b \otimes c_2 \end{array} = a \otimes b \otimes (c_1 + c_2)\right) \quad \boxtimes$$

Defn [Permutation of slices of tensors] $t \in \mathbb{C}^k \otimes \mathbb{C}^m \otimes \mathbb{C}^n$.

$t = (t_{i,j,k})_{i \in [k], j \in [m], k \in [n]}$. Take $\sigma \in S_k$. Let

$$t' = (t_{\sigma(i),j,k})_{i \in [k], j \in [m], k \in [n]}.$$

Lemma $R(t') = R(t)$

Proof obvious $\boxtimes$

Lemma Take any $\sigma \in S_3$ $R(M_{\sigma(\langle k, m, n\rangle)})$ is unchanged

Proof Recall $M_{\langle k,m,n\rangle} \in \mathbb{C}^{k \times m} \otimes \mathbb{C}^{m \times n} \otimes \mathbb{C}^{k \times n}$.

take $\sigma = (1,2) \in S_3$ for eg. Our goal is to get to $M_{\langle m,k,n\rangle}$. Take $f_1 : \mathbb{C}^{k \times m} \to \mathbb{C}^{m \times k}$, f_2 & f_3 identity.

$$(f_1 \otimes f_2 \otimes f_3)(M_{\langle k,m,n\rangle}) \in \mathbb{C}^{m\times k} \otimes \mathbb{C}^{m\times n} \otimes \mathbb{C}^{k\times n}$$

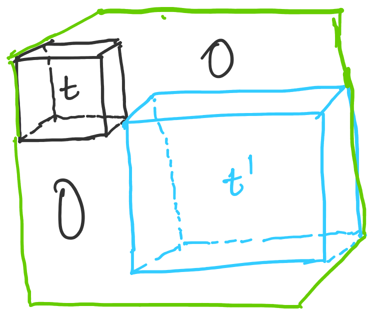
Let $\pi = (2,3) \in S_3$. Subsequently,

$$\pi\Big((f_1 \otimes f_2 \otimes f_3)(M_{\langle k,m,n\rangle})\Big)$$

$$= M_{\langle m,k,n\rangle} = M_{\sim\langle k,m,n\rangle}. \quad \boxtimes$$

Defn [Direct sum of tensors] $t \in \mathbb{F}^k \otimes \mathbb{F}^m \otimes \mathbb{F}^n$, $t' \in \mathbb{F}^{k'} \otimes \mathbb{F}^{m'} \otimes \mathbb{F}^{n'}$

$$t \oplus t' \in \left(\mathbb{F}^{k+k'} \otimes \mathbb{F}^{m+m'} \otimes \mathbb{F}^{n+n'}\right)$$



Lemma $R(t \oplus t') \le R(t) + R(t')$

Proof $t = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$, $t' = \sum_{i=1}^{r'} u_i' \otimes v_i' \otimes w_i'$

define $\hat{u}_i = (u_i, \underbrace{0}_{k' \text{ times}})$ $\quad \hat{u}_i' = (\underbrace{0}_{k \text{ times}}, u_i')$, $\hat{v}_i, \hat{v}_i', \hat{w}_i, \hat{w}_i'$

$$t \oplus t' = \sum_{i=1}^{r} \hat{u}_i \otimes \hat{v}_i \otimes \hat{w}_i + \sum_{i=1}^{r'} \hat{u}_i' \otimes \hat{v}_i' \otimes \hat{w}_i'$$

$\boxtimes$

Defn $t \in \mathbb{F}^k \otimes \mathbb{F}^m \otimes \mathbb{F}^n$, $t' \in \mathbb{F}^{k'} \otimes \mathbb{F}^{m'} \otimes \mathbb{F}^{n'}$

$a = (a_1, \dots, a_k) \in \mathbb{F}^k$, $\quad b = (b_1, \dots, b_{k'}) \in \mathbb{F}^{k'}$

$\mathbb{F}^{kk'} \ni (a \otimes b)_{i,j} = a_i b_j \quad i \in [k], j \in [k']$, extended linearly.

$$t = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i, \quad t' = \sum_{i=1}^{r'} u_i' \otimes v_i' \otimes w_i'$$

$$t \otimes t' = \sum_{i=1}^{r} \sum_{j=1}^{r'} (u_i \otimes u_i') \otimes (v_i \otimes v_i') \otimes (w_i \otimes w_i')$$

Lemma $R(t \otimes t') \le R(t)\, R(t')$

$\otimes M \quad = M$

Lemma $R(t \otimes t') \leq R(t) R(t')$

Notice $M_{\langle k,m,n \rangle} \otimes M_{\langle k',m',n' \rangle} = M_{\langle kk', mm', nn' \rangle}$

Using all the above machinery, we can turn an upper bound on $M_{\langle k,m,n \rangle}$ for any specific k, m, n into an upper bound on ω.

Thm $R(M_{\langle k,m,n \rangle}) \leq r$ then $\omega \leq 3 \cdot \log_{kmn} r$.

Proof $R\left(M_{\langle k,m,n \rangle} \otimes M_{\langle m,n,k \rangle} \otimes M_{\langle n,k,m \rangle}\right) \leq r^3$

$$\Rightarrow R\left(M_{\langle kmn, kmn, kmn \rangle}\right) \leq r^3 = (kmn)^{3 \log_{kmn} r}$$

$\omega \leq 3 \log_{kmn} r$ $\boxtimes$

Strassen $R(M_{\langle 2,2,2 \rangle}) \leq 7 \Rightarrow \omega \leq 3 \log_8 7 \leq 2.8074$

Pan $R(M_{\langle 70,70,70 \rangle}) \leq 143640 \Rightarrow \omega \leq 2.796$

There is an intuitive way of understanding this. This is also practical

BLAS level 3 uses Strassen. Experiments have been run on Pan, but currently not in use.

Timeline of matrix multiplication exponent

Year	Bound on omega	Authors
1969	2.8074	Strassen[1]
1978	2.796	Pan[11]
1979	2.780	Bini, Capovani [it], Romani[12]
1981	2.522	Schönhage[13]
1981	2.517	Romani[14]
1981	2.496	Coppersmith, Winograd[15]
1986	2.479	Strassen[16]
1990	2.3755	Coppersmith, Winograd[17]
2010	2.3737	Stothers[18]
2013	2.3729	Williams[19][20]
2014	2.3728639	Le Gall[21]
2020	2.3728596	Alman, Williams[3]
2022	2.37188	Duan, Wu, Zhou[2]

2020	2.3728596	Alman, Williams[3]
2022	2.37188	Duan, Wu, Zhou[2]

Border rank

General method that was used 40 years ago was to

(a) Do brute force search for small decompositions for $R_{\langle small \rangle}$.

(b) Recurse and hope for the best.

Consider $M_{\langle 2,2,3 \rangle}$

$$\begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix}_{2\times 2} \begin{bmatrix} y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \end{bmatrix}_{2\times 3} = \begin{bmatrix} z_{1,1} & z_{1,2} & z_{1,3} \\ z_{2,1} & z_{2,2} & z_{2,3} \end{bmatrix}_{2\times 3}$$

$M_{reduced\langle 2,2,2 \rangle}$

$$\begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix} *_{red} \begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} = \begin{bmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & \cancel{z_{2,2}} \end{bmatrix}$$

$$M_{reduced\langle 2,2,2 \rangle} \leq 6$$

⊛ Bini wanted to find a rank five expression for $M_{red\langle 2,2,2 \rangle}$

shows $\mathbf{R}(M^{red}_{\langle 2 \rangle}) \leq 6$. Bini et al. attempted to find a rank five expression for $M^{red}_{\langle 2 \rangle}$. They searched for such an expression by computer. Their method was to minimize the norm of $M^{red}_{\langle 2 \rangle}$ minus a rank five tensor that varied (see §4.6 for a description of the method), and their computer kept on producing rank five tensors with the norm of the difference getting smaller and smaller, but with larger and larger coefficients. Bini (personal communication) told me about how he lost sleep trying to understand what was wrong with his computer code. This went on for some time, when finally he realized *there was nothing wrong*

from Landsberg's textbook

$$M_{reduced\langle 2,2,2 \rangle} = \lim_{t \to 0} \frac{1}{t} \Big[(x_{1,2} + t x_{1,1}) \otimes (y_{1,2} + t y_{2,2}) \otimes z_{2,1}$$

$$+ (x_{2,1} + t x_{1,1}) \otimes y_{1,1} \otimes (z_{1,1} + t z_{1,2}) - \big(x_{1,2} \otimes y_{1,2} \otimes (z_{1,1} + z_{2,1} + t z_{2,2})\big)$$

$$- x_{2,1} \otimes \big((y_{1,1} + y_{1,2}) + t y_{2,1}\big) \otimes z_{1,1} + (x_{1,2} + x_{2,1}) \otimes (y_{1,2} + t y_{2,1}) \otimes (z_{1,1} + t z_{2,2}) \Big]$$

$- x_{2,1} \otimes \left((y_{1,1} + y_{1,2}) + t\, y_{2,1} \right) \otimes z_{1,1} + (x_{1,2} + x_{2,1}) \otimes (y_{1,2} + t\, y_{2,1}) \otimes (z_{1,1} + t\, z_{2,2}) \Big]$

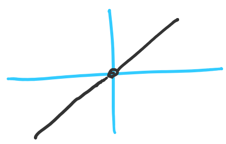
$M_{reduced \langle 2,2,2 \rangle}$ Can be arbitrarily well approximated by a rank 5 tensor

Fact $R(M_{reduced\, \langle 2,2,2 \rangle}) = 6$ $\left(\underline{R}(M_{reduced\, \langle 2,2,2\rangle}) \leq 5 \right)$

Can we use upper bounds on the border rank of $M_{\langle small \rangle}$ and get a upper bound on ω? [Bini] YES! NON-TRIVIALLY SO!

Let us try and define border rank. Why is this happening?

— is the punctured line through the origin, i.e. $\{(x,x) \mid x \in \mathbb{R}\} \setminus \{(0,0)\}$



$y-x$ vanishes on the punctured line, but it has an extra point

THERE IS NO POLYNOMIAL THAT VANISHES EXACTLY ON THE PUNCTURED LINE

"Punctured line is not a variety/algebraic set"

Defn Zariski Closure of the Punctured line is the zero set of $y-x$.

If a set is Zariski closed, then there is a finite set of polynomials whose common zeros vanish exactly on the set.

Zm

Defn Segre Variety (Parameterizes rank 1 tensors)

$$Seg : \mathbb{P}(A_1) \times \cdots \times \mathbb{P}(A_n) \longrightarrow \mathbb{P}(A_1 \otimes \cdots \otimes A_n)$$

$$(a_1, \ldots, a_n) \longmapsto a_1 \otimes \cdots \otimes a_n.$$

$\downarrow \; \downarrow \; \swarrow$ homogeneous co-ordinates

$\dim(A_i)$ homogeneous co-ordinates

$\sigma_1 = \mathrm{Im}(\mathrm{Seg}) \leftarrow$ Segre variety.

Defn Let V be a projective variety. A line L is called a secant line to V if L meets V in two or more points.

If the points coincide, there is a problem. But that pt. lies on V.

Defn Let $X \subseteq \mathbb{P}(V)$ be a projective variety. Define

$$S_r(X)^0 = \{(x_1, \dots, x_r, z) \in X^{\times r} \times \mathbb{P}(V) \mid z \in \mathrm{span}(x_1, \dots x_r)\}$$
$$\subseteq \mathrm{Seg}(X^{\times r} \times \mathbb{P}(V)) \subseteq \mathbb{P}(V^{\otimes r+1}).$$

Let

$$S_r(X) = \overline{S_r(X)^0} \quad \text{(Zariski Closure)}.$$

$S_r(X)$ is called the abstract r^{th} Secant variety of X.

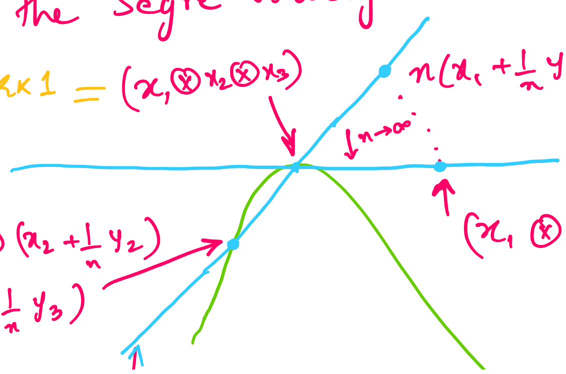
$$\pi^0 : S_r(X)^0 \longrightarrow \mathbb{P}(V) \qquad \text{\& likewise} \qquad \pi : S_r(X) \longrightarrow \mathbb{P}(V)$$
$$(x_1, \dots, x_r, z) \longmapsto z$$

- Image of π^0 is denoted $\sigma_r^0(X)$
- $\sigma_r(X) := \mathrm{Im}(\pi)$ (Called r^{th} Secant variety of X) (note:- $\sigma_1(X) = X$)

When $X = \mathrm{Seg}(\mathbb{P}(A_1) \times \dots \times \mathbb{P}(A_n))$ (Cone in affine space) $\hat{\sigma}_r$

$\sigma_r(X) \rightarrow$ set of tensors of border rank at most r.

Very little is known abt. "actual descriptions" of the secant varieties of the Segre variety.



rk 1 $= (x_1 \otimes x_2 \otimes x_3)$

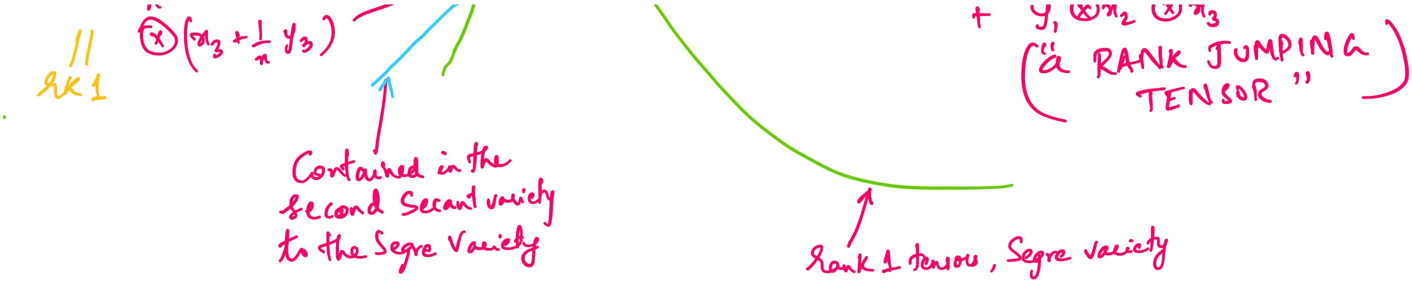
$n(x_1 + \frac{1}{n} y_1) \otimes (x_2 + \frac{1}{n} y_2) \otimes (x_3 + \frac{1}{n} y_3)$ $=$ rk 2

$- n(x_1 \otimes x_2 \otimes x_3)$

$\downarrow n \to \infty$

$n(x_1 + \frac{1}{n} y_1) \otimes (x_2 + \frac{1}{n} y_2) \otimes (x_3 + \frac{1}{n} y_3)$ $=$ rk 1

$(x_1 \otimes x_2 \otimes y_3) + x_1 \otimes y_2 \otimes x_3 + y_1 \otimes x_2 \otimes x_3 = $ rk 3, ("a RANK JUMPING TENSOR")



$(*)$ $M_{\text{reduced } \langle 2,2,2 \rangle}$ is a rank jumping tensor

"a tensor is rank jumping iff it is regular element of the tangential variety of the Segre variety"

Idea:- Interpolate between rank and border rank.

Let ε be an indeterminate.

<u>Defn</u> [h rank, border rank] $t \in \mathbb{F}^k \otimes \mathbb{F}^m \otimes \mathbb{F}^n$

$h \in \mathbb{Z}_{\geq 0}$

(1) $$R_h(t) = \min\left\{ r \;\middle|\; \exists\, u_i \in \mathbb{F}[\varepsilon]^k, v_i \in \mathbb{F}[\varepsilon]^m, w_i \in \mathbb{F}[\varepsilon]^n : \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i = \varepsilon^h t + O(\varepsilon^{h+1}) \right\}$$

$$\Updownarrow$$

$$t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^h} \left(\sum_{i=1}^{r} u_i \otimes v_i \otimes w_i \right)$$

(2) $$\underline{R}(t) = \min_{h \in \mathbb{N}} R_h(t)$$

<u>Observation</u> (1) $R_0(t) = R(t)$

(2) $R_0(t) \geq R_1(t) \geq \dots \dots = \underline{R}(t)$

(3) $R_h(t)$ Can be thought of as degree at most h in ε.

<u>Proof</u> By defn $\boxtimes$

Lemma (1) $\pi \in S_3$: $R_h(\pi t) = R_h(t)$

(2) $R_{\max\{h, h'\}}(t \oplus t') \leq R_h(t) + R_{h'}(t')$

(3) $R_{h+h'}(t \otimes t') \leq R_h(t) R_{h'}(t')$

Proof Same as earlier lemmas, but using defn of $R_h(*)$ instead $\boxtimes$

K_h (*) instead ⊠

Proof Same as earlier lemmas, but using defn of K_h (*) instead ⊠

Lemma* [turn approximate computations into real ones] There is a constant $C_h \le \binom{h+2}{2}$ s.t $\forall\, t$ (if field is infinite $C_h = 1+2h$ works)

$$R(t) \le C_h R_h(t)$$

Proof Let $R_h(t) = r$

$$\sum_{i=1}^{r} u_i \otimes v_i \otimes w_i = \varepsilon^h t + O(\varepsilon^{h+1})$$

triples of Polynomials in ε

t is the coefficient of ε^h in the expression on the LHS

(how do you get ε^h in $u_i \otimes v_i \otimes w_i$?)

$$u_i = \sum_{\alpha=0}^{h} \varepsilon^\alpha u_{i,\alpha}, \quad v_i = \sum_{\beta=0}^{h} \varepsilon^\beta v_{i,\beta}, \quad w_i = \sum_{\gamma=0}^{h} \varepsilon^\gamma w_{i,\gamma}$$

for ε^h, we need $\alpha+\beta+\gamma = h$; $\alpha, \beta, \gamma \in \mathbb{Z}_{\ge 0}$

— no. of such triples $= \binom{h+2}{2}$

Thus each $u_i \otimes v_i \otimes w_i$ has at most $\binom{h+2}{2}$ terms. There are r such $u_i \otimes v_i \otimes w_i$. Thus $R(t) \le \binom{h+2}{2} R_h(t)$ ⊠

Thm $\underline{R}(M_{\langle k,m,n\rangle}) \le r \implies \omega \le 3 \log_{kmn} r$

Proof

$(\underline{R}(T) \le r \iff \exists h\ R_h(T) \le r)$

$$\underline{R}(M_{\langle k,m,n\rangle}) \le r \implies \exists h\ R_h(M_{\langle k,m,n\rangle}) \le r.$$

By tensoring permutations of $M_{\langle k,m,n\rangle}$, we have

By tensoring permutations of $M_{\langle k,m,n\rangle}$, we have

$$R_{3h}\left(M_{\langle kmn,\, kmn,\, kmn\rangle}\right) \le r^3$$

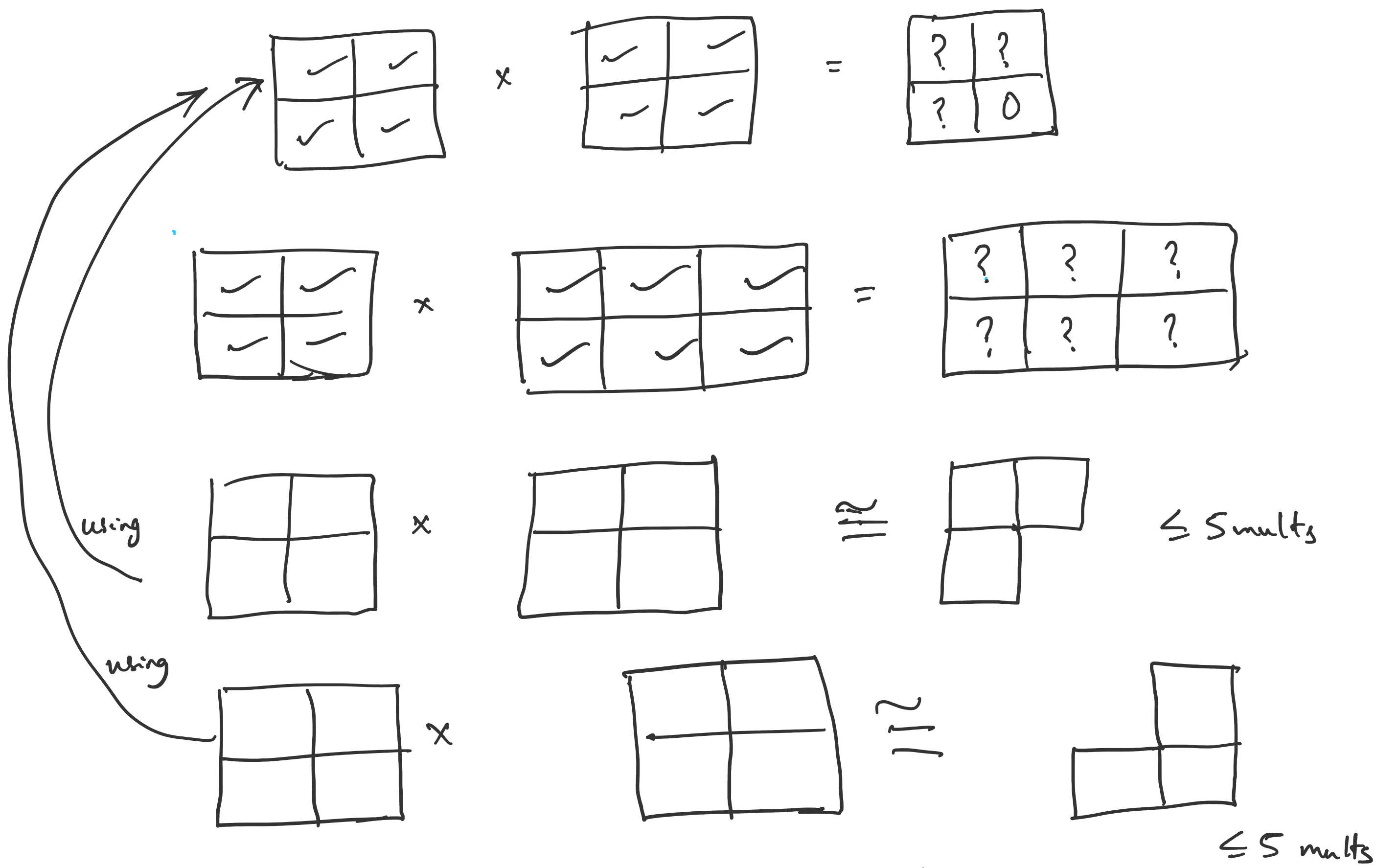
$$\Rightarrow R_{3hs}\left(M_{\langle (kmn)^s,(kmn)^s,(kmn)^s\rangle}\right) \le r^{3s} \quad \forall s$$

$$\Rightarrow R\left(M_{\langle (kmn)^s,(kmn)^s,(kmn)^s\rangle}\right) \le \binom{3hs+2}{2} r^{3s}$$

$$\omega \le \log_{(kmn)^s} r^{3s} + \log_{(kmn)^s}\left(\binom{3hs+2}{2}\right)$$

$$= 3\log_{(kmn)} r + \underbrace{\frac{1}{s}\log(\text{Poly}(s))}_{=0 \text{ as } s\to\infty} \qquad \boxtimes$$

(*) We started by noticing $\underline{R}\left(M_{\text{reduced}\,\langle 2,2,2\rangle}\right) \le 5.$



$$\underline{R}\left(M_{\langle 2,2,3\rangle}\right) \le 10 \Rightarrow \boxed{\omega \le 3\log_{12} 10 \le 2.78}$$

Schonhage's τ theorem

Schonhage's τ theorem

[Strassen] turned u.b. on $R(M_{\langle * \rangle})$ into u.b. on ω

[Bini et al.] turned u.b on $\underline{R}(M_{\langle * \rangle})$ into u.b. on ω

[Schonhage] turn u.b. on $\underline{R}$ (not even a matrix mult. tensor) into u.b. on ω

$\underline{R}$(matrix mult tensor$_1$ $\oplus$ matrix mult. tensor$_2$)

$\Downarrow$ multiple independent matrix multiplications...

$$\begin{bmatrix} \square & 0 \\ 0 & \square \end{bmatrix} \times \begin{bmatrix} \square & 0 \\ 0 & \square \end{bmatrix}$$

Lemma ① $R(M_{\langle k,1,n\rangle} \oplus M_{\langle 1,m,1\rangle}) = k.n+m$

② $\underline{R}(M_{\langle k,1,n\rangle}) = k.n$ and $\underline{R}(M_{\langle 1,m,1\rangle}) = m$

③ $\underline{R}(M_{\langle k,1,n\rangle} \oplus M_{\langle 1,N,1\rangle}) \le k.n+1$ when $N=(n-1)(k-1)$

Defn $t \in \mathbb{F}^k \otimes \mathbb{F}^m \otimes \mathbb{F}^n$, $t' \in \mathbb{F}^{k'} \otimes \mathbb{F}^{m'} \otimes \mathbb{F}^{n'}$

① t is called a restriction of t' if there are homomorphisms $\alpha: \mathbb{F}^{k'} \to \mathbb{F}^k$, $\beta: \mathbb{F}^{m'} \to \mathbb{F}^m$, $\gamma: \mathbb{F}^{n'} \to \mathbb{F}^n$, s.t.

$$t = (\alpha \otimes \beta \otimes \gamma)\, t'$$

$$t \lesssim t'$$

② if α, β, γ are isomorphisms, we say t & t' are isomorphic, we write

$$t \cong t'$$

Defn $\underset{\sim}{I}_{\langle r \rangle} \in \mathbb{F}^r \otimes \mathbb{F}^r \otimes \mathbb{F}^r$ is s.t.

$$(I_{\langle r \rangle})_{i,i,i} = 1 \quad \forall i \in [r] \text{ \& } 0\text{'s elsewhere}$$

Lemma $R(t) \leq r \iff t \lesssim \mathcal{I}_{\langle r \rangle}$

Proof $\Leftarrow$ $R(t) = R(\text{Restriction } \mathcal{I}_{\langle r \rangle}) \leq R(\mathcal{I}_{\langle r \rangle}) = r$ ✓

$\Rightarrow$ $\mathcal{I}_{\langle r \rangle} = \sum_{i=1}^{r} e_i \otimes e_i \otimes e_i$, and since $R(t) \leq r$

$$t = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$$

$\forall i \in [r]$, define

$\alpha : e_i \mapsto u_i$, $\beta : e_i \mapsto v_i$, $\gamma : e_i \mapsto w_i$

$$(\alpha \otimes \beta \otimes \gamma)\, \mathcal{I}_{\langle r \rangle} = \sum_{i=1}^{r} \alpha(e_i) \otimes \beta(e_i) \otimes \gamma(e_i) = t \quad \boxtimes$$

Claim $R\left(M_{\langle k,m,n \rangle}^{\oplus a}\right) \leq g \Rightarrow R\left(M_{\langle k^s,m^s,n^s \rangle}^{\oplus a}\right) \leq \left\lceil \frac{g}{a} \right\rceil^s a \quad \forall s \in \mathbb{N}$

Proof By induction on s. For $s=1$, it is just the hypothesis of the claim.

$$(*)\ R\left(M_{\langle k,m,n \rangle}^{\oplus a}\right) \leq g \Rightarrow M_{\langle k,m,n \rangle}^{\oplus a} \lesssim \mathcal{I}_g$$

$(a \oplus b) \otimes c = a \otimes c + b \otimes c$

$$M_{\langle k^{s+1}, m^{s+1}, n^{s+1} \rangle}^{\oplus a} = M_{\langle k,m,n \rangle}^{\oplus a} \otimes M_{\langle k^s, m^s, n^s \rangle}$$
$$\lesssim \mathcal{I}_g \otimes M_{\langle k^s, m^s, n^s \rangle}$$
$$= M_{\langle k^s, m^s, n^s \rangle}^{\oplus g}$$

Thus

$$R\left(M_{\langle k^{s+1}, m^{s+1}, n^{s+1} \rangle}^{\oplus a}\right) \leq R\left(M_{\langle k^s, m^s, n^s \rangle}^{\oplus g}\right)$$
$$\leq R\left(M_{\langle k^s, m^s, n^s \rangle}^{\oplus a \cdot \lceil g/a \rceil}\right) \leq \left\lceil \frac{g}{a} \right\rceil \left\lceil \frac{g}{a} \right\rceil^s a \quad \boxtimes$$

Lemma $R\left(M_{\langle k,m,n \rangle}^{\oplus a}\right) \leq g \Rightarrow \omega \leq 3 \dfrac{\log \lceil g/a \rceil}{\log (kmn)}$

Proof $R\left(M_{\langle k,m,n \rangle}^{\oplus a}\right) \leq g \Rightarrow R\left(M_{\langle k^s,m^s,n^s \rangle}^{\oplus a}\right)$

$$\leq \left\lceil \frac{g}{a} \right\rceil^s a$$

$\to 0$ as $s \to \infty$

$$\omega \leq 3 \cdot \left(\frac{s \log \left\lceil \frac{g}{a} \right\rceil + \log a}{s \log (kmn)} \right) = 3 \left(\frac{\log \lceil g/a \rceil + \log a/s}{\log (kmn)} \right)$$

Since ω is an infimum, lemma follows $\boxtimes$

Thm [τ-theorem] If $\underline{R}\left(\bigoplus_{i=1}^{p} M_{\langle k_i, m_i, n_i\rangle}\right) \leq r$, and $r > p$, then

$\omega \leq 3\tau$, where τ is s.t. $\sum_{i=1}^{p} (k_i \cdot m_i \cdot n_i)^{\tau} = r$

Proof There is an h such that

$$R_h\left(\bigoplus_{i=1}^{p} M_{\langle k_i, m_i, n_i\rangle}\right) \leq r. \quad \star$$

$$\left(\langle k_1, m_1, n_1\rangle \oplus \langle k_2, m_2, n_2\rangle\right) \otimes \left(\langle k_1, m_1, n_1\rangle \oplus \langle k_2, m_2, n_2\rangle\right)$$
$$= \left(\langle k_1^2, m_1^2, n_1^2\rangle + 2\langle k_1k_2, m_1m_2, n_1n_2\rangle + \langle k_2^2, m_2^2, n_2^2\rangle\right)$$

$$\star \Rightarrow R_{hs}\left(\left(\bigoplus_{i=1}^{p} M_{\langle k_i, m_i, n_i\rangle}\right)^{\otimes s}\right) \leq r^s$$

$$\Rightarrow R_{hs}\left(\bigoplus_{\sigma_1+\dots+\sigma_p=s} \left(M_{\langle k', m', n'\rangle}\right)^{\oplus \frac{s!}{\sigma_1!\dots\sigma_p!}}\right) \leq r^s, \text{ where}$$

$$k' = \prod_{i=1}^{p} k_i^{\sigma_i}, \quad m' = \prod_{i=1}^{p} m_i^{\sigma_i}, \quad n' = \prod_{i=1}^{p} n_i^{\sigma_i}.$$

Convert approximate computation to exact computation

$$\Rightarrow R\left(\bigoplus_{\sigma_1+\dots+\sigma_p=s} \left(M_{\langle k', m', n'\rangle}\right)^{\oplus \frac{s!}{\sigma_1!\dots\sigma_p!}}\right) \leq \binom{hs+2}{2} r^s$$

We know
$$\left(\sum_{i=1}^{p} (k_i \cdot m_i \cdot n_i)^{\tau}\right)^s = r^s$$

$$\Rightarrow \sum_{\sigma_1+\dots+\sigma_p=s} \frac{s!}{\sigma_1!\dots\sigma_p!} (k' \cdot m' \cdot n')^{\tau} = r^s$$

Using lemma

$$\omega \leq 3 \cdot \left(\frac{\tau \cdot \log(k'm'n') + \log\left(\binom{s+p-1}{p-1}\right) + \log\binom{hs+2}{2}}{\log(k'm'n')} \right)$$

$\leq 3\tau$ + term that goes to 0 as $s \to \infty$ $\boxtimes$

Recall $\underline{R}\left(M_{\langle 4,1,3\rangle} \oplus M_{\langle 1,6,1\rangle}\right) \leq 13$

Using the theorem, we get $\omega \leq 2.55$

Coppersmith - Winograd

Defn [easy Cop. W. tensor]

$$T_{q,cw} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0$$

$\mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1}$

This is a nicely symmetric tensor

↳ CW show how to get many little matrix mult. tensors in VERY HIGH powers of $T_{q,CW}$

above expression has $3q$ terms. However, it turns out

$$\underline{R}(T_{q,cw}) = q+2$$

Uses Explicit constructions of dense sets of integers with no 3-term Arithmetic Progression, ADVANCED ARITHMETIC COMBINATORICS + PROBABILISTIC ARGUMENT

Thm

$$\omega \leq \frac{\log\left(\frac{4}{27}\underline{R}(T_{q,cw})^3\right)}{\log(q)}$$

gives $\omega < 2.41$ when $q = 8$

Defn [big Cop. W. tensor]

$$T^+_{q,cw} = T_{q,cw} + a_0 \otimes b_0 \otimes c_{q+1} + a_0 \otimes b_{q+1} \otimes c_0 + a_{q+1} \otimes b_0 \otimes c_0$$

$$+ a_{q+1} \otimes b_0 \otimes c_0$$

$\mathbb{C}^{q+2} \otimes \mathbb{C}^{q+2} \otimes \mathbb{C}^{q+2}$

Thm! $\underline{R}(T^+_{q,cw}) = q+2$ and

$$\omega \leq \log\left(\frac{4}{27}\underline{R}\left(T^{+\,\boxtimes k}_{q,cw}\right)^{3/k}\right)$$

$$\omega \leq \frac{\log\left(\frac{4}{27} R\left(\left(T^{+}_{q,cw}\right)^{\boxtimes k}\right)^{1/k}\right)}{\log q}$$

People have done up to $\left(T^{+}_{q,cw}\right)^{\boxtimes 32}$ $\Rightarrow \omega < 2.372839$

People have shown that taking higher and higher powers won't get you $\omega < 2.30$

Conjecture [Asymptotic Rank Conjecture]

$$\lim_{n\to\infty} R\left(T_{2,cw}^{\otimes n}\right)^{1/n} = 3$$

Theorem ($\Rightarrow \omega = 2$)

Conjecture [No 3 disjoint equivoluminous subsets Conjecture] Let H be an abelian group. Let $m_1, \ldots, m_n \in H$. This satisfies the "no 3 disjoint equivoluminous subsets" property if $\forall$ $S, T, U \subseteq [n]$, S, T, U disjoint, $\sum_{i\in S} m_i \neq \sum_{i\in U} m_i \neq \sum_{i\in T} m_i$. There exists $|H| \leq 2^{o(n)}$.

(Theorem $\Rightarrow$ Asymptotic rank conjecture)

$\Downarrow$

One of the Sunflower Conjectures of Erdos-Rado is false

if we take $H = \mathbb{Z}_2^n$, $m_i := e_i$, this has the N3DES property. But $|H| = 2^n$

Cohn-Umans group theoretic approach

What worked? Powers of direct sums of matrix multiplication tensors, plus recursion.

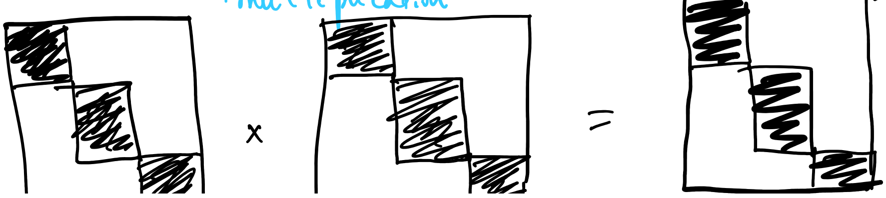
← seems ad-hoc....

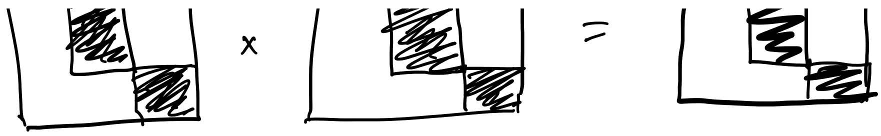
This is not matrix mult, but block-diagonal matrix mult.

Can we do this abstractly?

Idea Embed $M_{\langle n\rangle}$ into semisimple algebra multiplication

Informal defn [Semisimple Alg] Algebra in which multiplication is isomorphic to block-diagonal matrix multiplication





other words :- multiplication in this algebra can be expressed as multiplication of matrices with is block diagonal structure

Hope that the Algebra has a structure such that questions about ω reduce to group-theoretic questions

Defn [Sem. Alg.] "An associative Artinian algebra (over a field) that has a trivial Jacobson radical"
→ Satisfies the descending chain condition on ideals
↳ Ideal of elements that annihilate every simple left-module.

Thm [Wedderburn's Theorem] Any such finite dimensional algebra is isomorphic to a finite product $\prod M_{n_i}(D_i)$, where D_i are Division algebras over the field, and $M_{n_i}(D_i)$ is the algebra of $n_i \times n_i$ matrices over D_i.

(Read "Wedderburn-Artin Ring Theory" in Knapp's Advanced Algebra)

An Example Semi-Simple Algebra

G _ finite group. $\mathbb{C}[G]$ is the group algebra (formal linear combs. $\sum_{g \in G} a_g g$) . ($a_g \in \mathbb{C}$)

$$\left(\sum_{g\in G} a_g g\right) + \left(\sum_{g\in G} b_g g\right) = \sum_{g\in G} (a_g + b_g) g$$

$$\left(\sum_{g\in G} a_g g\right)\left(\sum_{h\in G} b_h h\right) = \sum_{f\in G}\left(\sum_{\substack{g,h\in G\\ gh=f}} a_g b_h\right) f$$

$\mathbb{C}[G]$ is a semi-simple algebra.

★ Notice if $G = C_n$, and g is a generator, then

$$\left(\sum_{i=0}^{n-1} a_i g^i\right) \circledast \left(\sum_{i=0}^{n-1} b_i g^i\right) = \sum_{i=0}^{n-1}\left(\sum_{\substack{j,k\\(j+k\equiv i) \bmod n}} a_j b_k\right) g^i, \text{ i.e.}$$

multiplication in $\mathbb{C}[C_n]$ is a cyclic convolution (think of the elems of $\mathbb{C}[C_n]$ as n-dim co-ordinate vectors) multiplication in $\mathbb{C}[C_n]$ is a cyclic convolution (think of the elems of $\mathbb{C}[C_n]$ as n-dim co-ordinate vectors)

univar. Polynomials

Observe $\left(\sum_{i=0}^{n-1} a_i x^i\right) * \left(\sum_{i=0}^{n-1} b_i x^i\right)$ is very close to multiplication in $\mathbb{C}[C_n]$, except for the wrap-around.

Thus, to avoid wrap around look at a larger group's algebra... If we look at multiplication in $\mathbb{C}[C_m]$, $m \geq 2n$, then polynomial mult. is the same as multiplication in the algebra $\mathbb{C}[C_m]$.

Thm [Fast Fourier Transform Algorithm] There is an invertible linear transformation $D: \mathbb{C}[G] \to \mathbb{C}^{|G|}$ that turns multiplication in $\mathbb{C}[G]$ into pointwise mult. in $\mathbb{C}^{|G|}$. There is a very efficient algorithm to compute the transform.

- So what we do is we embed the polynomials $\sum_i a_i x^i$ & $\sum_i b_i x^i$ into $\mathbb{C}[C_m]$, to get $\sum a_i g^i$, $\sum b_i g^i$, respectively, compute their DFT's, compute the pointwise product of their DFT's, and compute the inverse DFT. Turns out that using $O(m \log m)$ multiplications $(\sim O(n \log n))$, we can compute the product of polynomials.

★ The Cohn-Umans approach is to embed matrix multiplication into group algebra multiplication in an analogous way.

(Vague Plan)

① Mat. Mult. $\xhookrightarrow[\text{embedding}]{}$ $\mathbb{C}[G]$ $\xrightarrow[\text{DFT}]{}$ $\mathbb{C}^{|G|}$

appropriately chosen

② Do mult. in $\mathbb{C}^{|G|}$ and "Come back"...

Defn [Right Quotient] for a subset S of a finite group, define

$$Q(S) = \{ st^{-1} \mid s, t \in S \}$$

→ if S is a subgroup $Q(S) = S$.

Defn Subsets X, Y, Z of G satisfy the "triple product property" if

$\forall\ x \in Q(X),\ y \in Q(Y),\ z \in Q(Z)$,

$$xyz = 1 \longrightarrow x = y = z = 1$$

$\forall\ x \in Q(X),\ y \in Q(Y),\ z \in Q(Z),$

$$xyz = 1 \longrightarrow x = y = z = 1$$

(note that if X, Y, Z are subgroups we can say

$$\underset{\substack{\cap\\X}}{x}\underset{\substack{\cap\\Y}}{y}\underset{\substack{\cap\\Z}}{z} = 1 \iff x = y = z = 1 \;)$$

in some sense, we cannot have x_1^{-1} for a non identity in Y ... the subgroups are disjoint in some sense

HOW TO EMBED?

Suppose S, T, U are subsets of G, and

$$A = (a_{s,t})_{s\in S,\, t\in T}\ , \qquad B = (b_{t,u})_{t\in T,\, u\in U}\ .$$

($|S|\times|T|$ matrix) ($|T|\times|U|$ matrix)

Define $\underset{\substack{\cap\\ \mathbb{C}[G]}}{\bar{A}} = \sum a_{s,t}\, s^{-1}t\ ,\quad \bar{B} = \sum b_{t,u}\, t^{-1}u$. If S, T, U satisfy the 'triple product property'

then we can read off entries of AB from $\bar{A}\bar{B} \in \mathbb{C}[G]$. ☆

☆ $(AB)_{s,u}$ is just the coefficient of $s^{-1}u$ in $\bar{A}.\bar{B}$.

Thm [Wedderburn] $\underset{\text{group alg.}}{\mathbb{C}[G]} \cong \mathbb{C}^{d_1\times d_1} \times \dots \times \mathbb{C}^{d_k\times d_k}$

direct product of square matrices over $\mathbb{C}$

(d_i's are called "character degrees" of G)

Comes from representation theory

(k is the no. of conjugacy classes)

Shows that $|G| = \sum_{i=1}^{k} d_i^2$

"discrete fourier transform".

Thus product of $|S|\times|T|$ times $|T|\times|U|$ matrices reduces to many small matrix multiplications

depends on Character degrees of the group.

Thus product of $|S|\times|T|$ times $|T|\times|U|$ matrices reduces to many small matrix multiplications

depends on Character degrees of the group.

Defn If you can find a G and subsets X, Y, Z satisfying the triple product property, then we say G realizes $M_{\langle |X|,|Y|,|Z| \rangle}$.

e.g. $C_k \times C_m \times C_n$ realizes $M_{\langle k,m,n \rangle}$ via the subgroups $C_k \times \{1\} \times \{1\}$, $\{1\} \times C_m \times \{1\}$ and $\{1\} \times \{1\} \times C_n$.

Thm If G realizes $M_{\langle k,m,n \rangle}$, then $M_{\langle k,m,n \rangle} \lesssim \mathbb{C}[G]$

multiplication here is a bilinear map, so by abuse of notation, $\mathbb{C}[G]$ is the tensor

In particular $R(M_{\langle k,m,n \rangle}) \le R(\mathbb{C}[G])$.

Proof Just read "HOW TO EMBED" $\boxtimes$

Summary :-

1. G realizes $M_{\langle k,m,n \rangle} \Rightarrow M_{\langle k,m,n \rangle}$ is a restriction of $\mathbb{C}[G]$.
2. Wedderburn states $\mathbb{C}[G]$ is iso. to a product of matrix algebras
3. Thus mult. in $\mathbb{C}[G]$ (and in turn mult. of $k\times m$ times $m\times n$ matrices) breaks down into many small matrix multiplications.

Thm For a non-trivial finite group G, define

$$\alpha(G) = \min \left\{ \frac{3 \log |G|}{\log kmn} \;\middle|\; G \text{ realizes } M_{\langle k,m,n \rangle}, \ (\text{one of } k,m,n > 1) \right\}.$$

Called pseudo-exponent

Called pseudo-exponent

Then
(1) $2 < \alpha(G) \leq 3$
(2) If G is abelian, $\alpha(G) = 3$
(3) If the character degrees of G are $d_1 \dots d_t$, then
$$|G|^{w/\alpha(G)} \leq \sum_{i=1}^{t} d_i^{w}.$$

Proof (1a) $\alpha(G) \leq 3$ – trivial. for G, Let $H_1 = H_2 = 1$, $H_3 = G$

$x_1 x_2 x_3 = 1$. Since $x_1 = x_2 = 1$, $x_3 = 1$ ✓

Thus G realizes $M_{\langle |G|, 1, 1\rangle}$. ✓

(1b) $2 < \alpha(G)$ – Suppose G realizes $M_{\langle k,m,n\rangle}$ via S_1, S_2, S_3, i.e where $|Q(S_1)| = k$, $|Q(S_2)| = m$, $|Q(S_3)| = n$. Consider the map.
$$\phi : Q(S_1) \times Q(S_2) \to G,$$
$$(x, y) \mapsto x^{-1} y$$

(a) This is injective ($x_1^{-1} y_1 = x_2^{-1} y_2 \Rightarrow x_2 x_1^{-1} y_1 y_2^{-1} = 1$, by triple prod. prop,
$\Rightarrow x_2 x_1^{-1} = y_1 y_2^{-1} = 1 \Rightarrow x_1 = x_2$ & $y_1 = y_2$)

(b) $\mathrm{Im}(\phi) \cap Q(S_3) = \{1\}$. Suppose not. There exists $z \in Q(S_3)$. $z \neq \{1\}$ s.t
$$\underbrace{x^{-1}}_{\in Q(S_1)} \cdot \underbrace{y}_{\in Q(S_2)} = z^{\in Q(S_3)} \Rightarrow x^{-1} y z = 1 \Rightarrow x^{-1} = y = z = 1$$ by triple prod prop.
Contradiction!

due to (b)

Since (a), we have $|G| \geq km$ (ineq. strict unless $n = 1$)

Symmetrically, $|G| \geq mn$ & $|G| \geq kn$ (ineq. " " ")

$\Rightarrow |G|^3 \geq (kmn)^2$ with ineq. strict unless $n = m = k = 1$.

not. in defn of $\alpha(G)$,

Thus we have $|G| > (kmn)^{2/3}$

$\Rightarrow \alpha(G) > 2$. ✓

(2) $\alpha(\text{abelian } G) = 3$.

the map $\psi : S_1 \times S_2 \times S_3 \to G$
$a, b, c \mapsto abc$ is injective. Because

$$a_1 b_1 c_1 = a_2 b_2 c_2 \underset{\text{b'coz of abelianess}}{\Rightarrow} a_1 a_2^{-1} b_1 b_2^{-1} c_1 c_2^{-1} = 1 \underset{\text{triple prod. prop.}}{\Rightarrow} a_1 = a_2,\ b_1 = b_2,\ c_1 = c_2 .$$

Since ϕ is injective,

$$|G| \geq kmn \Rightarrow \alpha(G) \geq 3$$

(3) Let (k', m', n') be triple responsible for $\alpha(G)$. This means, by defn,

$$\alpha(G) = \frac{3 \log |G|}{\log k'm'n'} \Rightarrow (k'm'n')^{\alpha(G)} = |G|^3.$$

By defn, G realizes $M_{\langle k', m', n' \rangle}$, so

$$M_{\langle k', m', n' \rangle} \lesssim \mathbb{C}[G] \cong \bigoplus_{i=1}^{t} M_{\langle d_i, d_i, d_i \rangle} .$$

Taking l^{th} tensor power

$$M_{\langle (k')^l, (m')^l, (n')^l \rangle} \lesssim \bigoplus_{i=1}^{t} \left(M_{\langle d_i, d_i, d_i \rangle} \right)^{\otimes l}$$

$$= \bigoplus_{i_1, \ldots, i_l = 1}^{t} M_{\langle d_{i_1} \cdots d_{i_l},\ d_{i_1} \cdots d_{i_l},\ d_{i_1} \cdots d_{i_l} \rangle}$$

Taking rank,

$$R\left(M_{\langle (k')^l, (m')^l, (n')^l \rangle}\right) \leq \sum_{i_1, \ldots i_l = 1}^{t} R\left(M_{\langle d_{i_1} \cdots d_{i_l},\ d_{i_1} \cdots d_{i_l}, \ldots \rangle}\right)$$

$$= c . \left(\sum_{i=1}^{t} d_i^{w+\varepsilon} \right)^l$$

$R(M_{\langle n,n,n \rangle}) = O(n^{w+\varepsilon})$, $\varepsilon > 0$

Since $R\left(M_{\langle (k')^l, (m')^l, (n')^l \rangle}\right) \geq (k'm'n')^{lw/3}$, we have after taking l^{th} roots

$$|G|^{w/\alpha} = (k'm'n')^{w/3} \leq \sum_{i=1}^{t} d_i^{w+\varepsilon} .$$ Since $\varepsilon > 0$ was arbitrary, claim follows $\boxtimes$

APPLICATION :-

(*) Let $H = C_n^3$, Let $G = H^2 \rtimes C_2$

(*) Let $H = C_n^3$, Let $G = H^2 \rtimes C_2$

C_2 acts on H^2 by switching the two factors

Let H_1, H_2, H_3 be the three factors of H viewed as subgroups.

$H_1 = C_n \times \{1\} \times \{1\}$, etc...

Define subsets

$$S_i = \{ (a,b) z^j \mid a \in H_i \setminus \{1\},\ b \in H_{(i \% 3 + 1)},\ C_2 = \langle z \rangle,\ j \in \{0,1\} \}.$$

Then G realizes $M_{\langle |S_1|, |S_2|, |S_3| \rangle}$

Setting $n = 17$ gives $\omega \le 2.91$

(*) Using wreath product groups gives $\omega < 2.41$ ← Match Coppersmith-Winograd

$S_n \ltimes A^n$

In general you want $|G| \simeq n^2$,

and subgroups of size n, and

small max (character degrees), so you

have to think about the representation theory

of the group

Sadly, non-trivial results require non-Abelian groups.

Most ideas fail due to large character degrees.

Generalization of all this allows you to use the language of

Coherent Configurations

⟷ "Doing group theory without groups"

The advantage of this language is that you avoid representation theory, group theory ....

Thm if you can embed $M_{\langle n,n,n\rangle}$ into a Commutative Coherent Configuration (an association scheme) of rank $\approx n^2$, then $\omega = 2$.

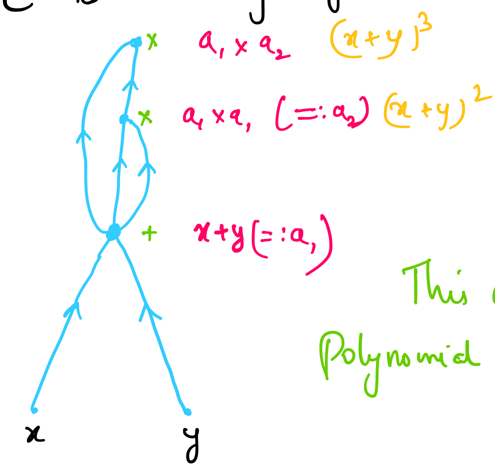
"algebraic Combinatorics"

VP vs VNP, determinantal Complexity, etc.

Wednesday, 24 May 2023 06:25

<u>Defn</u> A arithmetic circuit C is a finite, directed, acyclic graph with vertices of in-degree 0 or 2 and exactly one vertex of out-degree 0.

- The vertices of in-degree 0 are labelled by elems of $\mathbb{C} \cup \{x_1, \dots, x_n\}$, called inputs
- Those of indeg 2 are labelled $+$ or $\times$, called gates.
- If out-degree of a vertex is 0, then it is called an output gate.
- The size of C is the no. of edges



This circuit computes the Polynomial $(x+y)^3$.

This does not look geometrical. But it is a fact that, upto a polynomial factor, (*) the size of the circuit does not change if the inputs are arbitrary linear functions on a vector space.

<u>Defn</u> [VP] Let $d(n), N(n)$ be polynomials, and let $f_n \in \mathbb{C}[x_1, \dots, x_{N(n)}]_{\leq d(n)}$ be a sequence of polynomials. We say $(f_n) \in VP$ if there exists a sequence of circuits C_n of size poly in n computing f_n.

<u>Defn</u> [VNP] A sequence (f_n) is in VNP if there exists a polynomial in n, i.e. $p(n)$ and a sequence $(g_n) \in VP$ s.t.

$$f_n(x) = \sum_{\varepsilon \in \{0,1\}^{p(n)}} g_n(x, \varepsilon)$$

Think of elems of VNP as "projections" of elements of VP, where elems of VP are ... $\{0,1\}^{N(n)} \to \mathbb{C}$. and elems of VNP are are

Think of elems of VNP as "projections" of elements of VP, where elems of VNP are thought of a sequence of maps $g_n : \mathbb{C}^{N(n)} \to \mathbb{C}$, and elems of VNP are are projections of these maps.

※ Projection of varieties can be far more complicated than the original varieties (This is elaborated on in [Basu] "Complexity theory of constructible sheaves")

Prop $\mathrm{Per}_n \in VNP$

Proof Define

$$g_n(x_{1,1}, \dots, x_{n,n}, y_{1,1}, \dots, y_{n,n})$$

$$:= \underbrace{\underbrace{\prod_{\substack{i,j,l,m \in [n] \\ (i=l) \iff j \neq m}} (1 - y_{i,j} y_{l,m})}_{\alpha_n(y)} \underbrace{\left(\prod_{i=1}^n \sum_{j=1}^n y_{ij} \right)}_{\beta_n(y)}}_{\gamma_n(y)} \underbrace{\left(\prod_{i=1}^n \sum_{j=1}^n x_{i,j} y_{i,j} \right)}_{\mu_n(x,y)}$$

① $(g_n) \in VP$ bcoz no. of indeterminates $= 2n^2$, degree of each $g_n = O(n^3)$

② $\gamma_n(e) \neq 0$ iff e is a permutation matrix

- ⊛ iff there is a row or col. that has two or more 1's, then $\alpha_n(e) = 0$
- ⊛ Suppose $\alpha_n(e) \neq 0$. Then $\beta_n \neq 0 \iff$ every row of e contains at least one 1.
- ⊛ Thus $\gamma_n(e) = \alpha_n(e)\beta_n(e) \neq 0$ iff e is a perm. matrix

③ If e is a perm matrix, $\gamma_n(e) = 1$, and $\mu_n(x, e) = \prod_{i=1}^n x_{i,\sigma(i)}$,

where σ is the perm in S_n corresponding to e.

④ $\mathrm{Per}_n = \sum_{e \in \{0,1\}^n} g_n(x, e)$ ⊠

Defn A sequence $(g_m(y_1, \dots y_{M(m)}))$ can be polynomially reduced to $(f_n(x_1, \dots, x_{N(n)}))$ if there is a polynomial $n(m)$ s.t.

$$g_m(y_1, \dots y_{M(m)}) = f_n(\alpha_1(y_1, \dots y_{M(m)}), \alpha_2(y_1, \dots y_{M(m)}), \dots \alpha_{N(n)}(y_1, \dots, y_{M(m)})),$$

where α_i are affine linear functions. Written as

$(f_n) \leq_p (g_m)$

For a Complexity class $\mathcal{C}$, (P_n) is hard for $\mathcal{C}$ if $\forall (f_m) \in \mathcal{C}$,

$$(f_m) \leq_p (P_n).$$

(P_n) is Complete for $\mathcal{C}$ if (P_n) is hard for $\mathcal{C}$ and $(P_n) \in \mathcal{C}$.

Prop $(\det_n) \in VP$. (Gaussian elim works to compute $\det_n$ "efficiently", but that is not a circuit bcoz of division and checking of pivots for non-zeroness)

Proof Let S_n act on $\mathbb{C}[x_1, \dots, x_n]$ naturally, let $\mathbb{C}[x_1, \dots, x_n]^{S_n}$ denote the subspace of $\mathbb{C}[x_1, \dots, x_n]$ that is invariant under this action.

Fact 1 The elementary symmetric funcs,

$$e_r = \sum_{\substack{J \subseteq [n] \\ |J| = r}} x_{j_1} \cdots x_{j_r},$$

are a basis on $\mathbb{C}[x_1, \dots, x_n]^{S_n}$.

Fact 2 The power sum symmetric funcs

$$P_r = x_1^r + \dots + x_n^r$$

are also a basis of $\mathbb{C}[x_1, \dots, x_n]^{S_n}$.

(*) The determinant of a linear map $f: V \to V$ is the product of its eigenvalues: $\det(f) = \lambda_1 \times \lambda_2 \times \dots \times \lambda_v$.

(*) $\operatorname{Trace}(f) = \sum_{i=1}^{v} \lambda_i$; and $\operatorname{trace}(f^k) = \sum_{i=1}^{v} \lambda_i^k$.

The quantities f^k can be computed with small circuits so $\det(f)$ can be computed with small circuits

$\boxtimes$

Thm [Valiant] (Per_n) is VNP-Complete (char $K \neq 2$).

Conjecture [Valiant] $VP \subsetneq VNP$

Conjecture [Valiant] $VP \subsetneq VNP$

Thm [Burgisser] $VP = VNP \implies P/poly = NP/poly$ (assuming GRH)

non-uniform Polynomial time

non-uniform NP

Class of decision problems solvable by a family of poly-size Boolean Circuits. This family can be non-uniform, i.e. there could be a completely different circuit for each input length

— VP vs VNP is "arithmetic circuit complexity"

— P/poly vs NP/poly is "boolean complexity"

— P vs NP is "uniform complexity"

— In any case VP vs VNP has very strong impact on what is called the Polynomial hierarchy

For a graph G with adjacency matrix A, per(A) counts the number of perfect matchings. Note that the presence of a perfect matching can be ascertained in polynomial time, i.e. it is in P. But the enumerator for the same question is hard and, is conjectured to be uncomputable in polynomial time.

Passage to determinantal complexity

Defn [Expression size] $I := \mathbb{C} \cup \{x_1, \ldots, x_n\}$. Every elem of I is an expression.
Also, if ϕ_1 & ϕ_2 are expressions, $\phi_1 \bullet \phi_2$ is an expression, where $\bullet \in \{+, *\}$. Every expression
The size of an expression is the no. of + or * used to build it. For any $f \in \mathbb{C}[x_1, \ldots, x_n]$,
is thus a polynomial in $\mathbb{C}[X_1, \ldots, X_n]$.
define the expression size to be min expression size of all ϕ that exactly compute f, i.e. $f = \mathrm{val}\, \phi$.

Thm [universality of the determinant (& permanent)]
For any $f \in K[\bar{X}]$, there exist matrices M, M' of finite size with
... that are linear forms in $\bar{X}$ such that

uniformly bounded as a poly in d, n.

$\Downarrow$ depending on the expression ... polynomial

For any $f \in K[\bar{X}]$, there exist matrices ... such that entries that are linear forms in $\bar{X}$ such that

$$f = \det(M).$$
$$f = \mathrm{per}(M')$$

If f has expression size u, the $dc(f) \leq 2u+2$, and $pc(f) \leq 2u+2$

$\Downarrow$ depending on the expression size of the polynomial

Same is true with per instead of det (Not the same M necessarily)

Defn The min size of the matrix M in the theorem above is called the determinantal complexity of f, denoted $dc(f)$.

Thm $dc(\mathrm{perm}_n) \leq O(n^c)$, c constant $\Longrightarrow VP = VNP$

Proof $\det_n$ has $n!$ monomials and is $O(n^3)$ computable

Thus $\det_{O(n^c)}$ has $O(n^c)!$ monomials and is $O(n^{3c})$ computable $\boxtimes$

$$\boxed{\text{We shall show that } \quad \frac{n^2}{2} \leq dc(\mathrm{perm}_n) \leq 2^n - 1}$$

$dc(\mathrm{perm}_m) \geq \frac{m^2}{2}$

Assume you know basics of
- Projective space
- Differential geometry
- Multilinear algebra
- Algebraic geometry

Overview

1. Defn of Gauss maps
2. Notion of degeneracy of images under the Gauss map
3. dimension of image of the m permanent $\leq$ dimension of image of the n-determinant, for sufficient n.
4. n-det has Gauss image$_n$ of dimension $2n-2$.
5. Under substitution, Gauss image stays degenerate
6. m-Permanent has Gauss image of dimension m^2-2.
7. from 3, we have $m^2-2 \leq 2n-2 \Rightarrow n \geq m^2/2$

GAUSS MAPS

Maps a point on surface in 3-space to its unit normal vector on the unit sphere

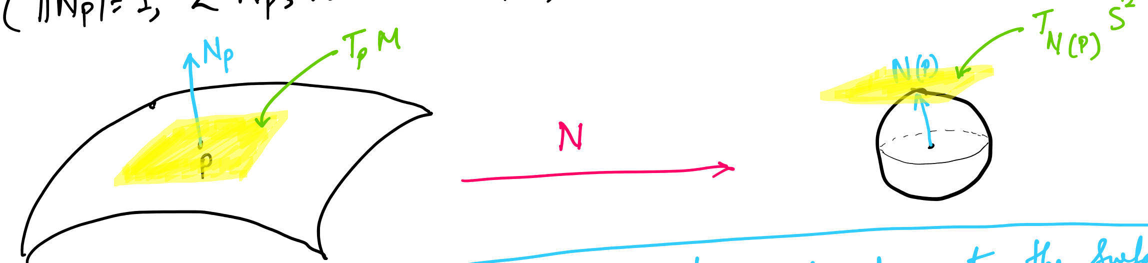
Diff. geometric motivation for this defn → Way the vector varies at and around the point gives you information about the surface (e.g. curvature)

Defn [Gauss Map for surfaces in $\mathbb{R}^3$] $M \subseteq \mathbb{R}^3$ is an oriented surface. N is the Gauss map of M

$$N : M \longrightarrow S^2 \text{ (unit sphere in } \mathbb{R}^3)$$

(Continuous map) $\quad p \longmapsto N_p$ (Oriented unit normal vector at point p)

$(\|N_p\| = 1,\ \langle N_p, \vec{v} \rangle = 0\ \forall v \in T_pM)$



Just a translation of the unit normal at p to the surface of a unit sphere

Both T_pM and $T_{N(p)}S^2$ are the same vector subspace of $\mathbb{R}^3$,

So the derivative of the Gauss map

$$d_pN : T_pM \longrightarrow T_{N(p)}S^2 \ (\cong T_pM)$$

is a linear map

(the determinant of this linear operator gives you information about curvature)

Note: ① Most surfaces have two possible choices for the direction of normal vectors ↗ or ↙

② Since N needs to be continuous (otherwise we can't define derivative), some surfaces such as the Mobius strip don't have a Gauss map

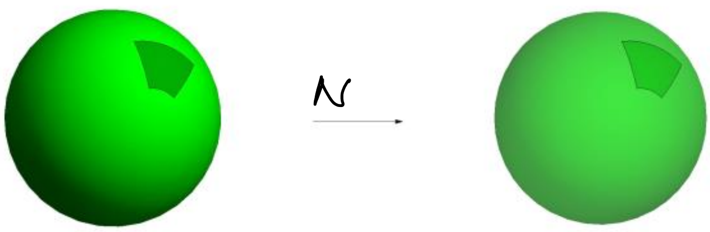
the map N is a differentiable unit normal vector field on an open neighbourhood of p

field on a neighbourhood of p

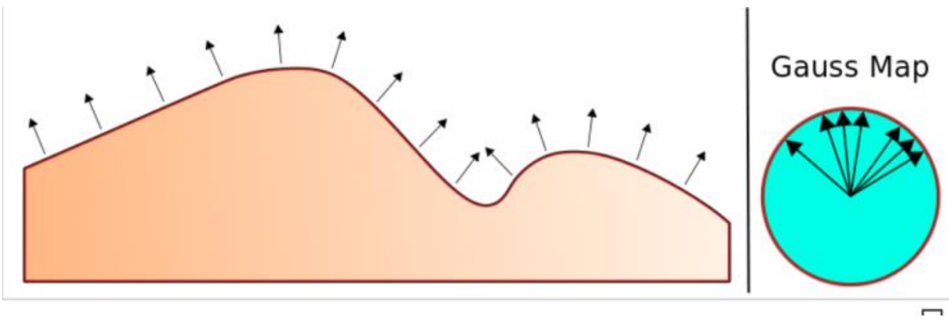
Mobius Band

(*) A surface is called orientable iff Gauss map exists
(*) A choice of a Gauss map for a surface is called an orientation
(*) An oriented surface just means that the orientation is specified

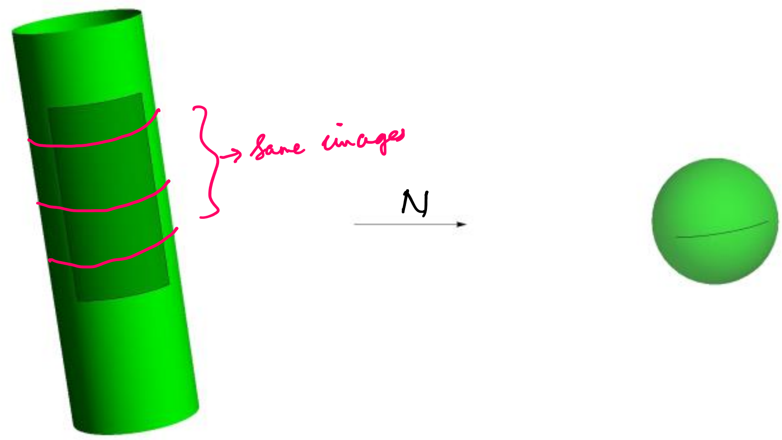
EXAMPLES ① From Landsberg's slides

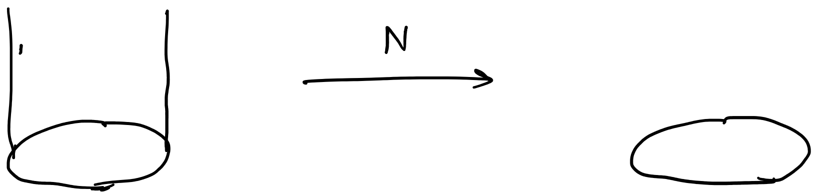


② From Wikipedia



☆ Note that the dimension of the Gauss image can drop





→ through all points on the cylinder, there is a curve along which the tangent plane is constant.

→ Gauss image of a plane in $\mathbb{R}^3$ is just a point on S^2.

Thm [Segre 1910] Let $M \subseteq \mathbb{P}^3$ be a surface with degenerate Gauss image. Then it is one of

(1) A linearly embedded $\mathbb{P}^2$

(2) A cone over a curve C

(3) A tangential variety to a curve C.

* Having a degenerate Gauss image is a pathology

Notations ① V → vector space, $\pi : V \setminus \{0\} \longrightarrow \mathbb{P}V$ is the projection map.
$(y \longmapsto [y])$

② For $X \subseteq \mathbb{P}V$, define

$$\hat{X}^{(\subseteq V)} := \pi^{-1}(x) \cup \{0\}$$

called a variety.

$X \quad := \pi(X) \cup \{0\}$

③ If $\hat{X}$ is a variety, $X \subseteq \mathbb{P}V$ will also be called a variety.

④ For $Z \subseteq V$, $\mathbb{P}Z \subseteq \mathbb{P}V$ will denote its image under π

Varieties in V defined by homogeneous polys. are invariant under scaling, so we just work with their projective versions instead.

Defn Let $X \subseteq \mathbb{P}V$ be an irreducible projective variety.

[Affine tangent space at $[x] \in X$] $\quad \hat{T}_{[x]} X$ is just $\quad T_x \hat{X} = \hat{T}_{[x]} X$

any pt. representing $[x] \in \mathbb{P}V$ in V

[Projective tangent space] $\quad \mathbb{P}(T_x \hat{X})$

Can be done in a completely abstract way using language of schemes and local rings

Defn [Conormal space] For $X \subseteq \mathbb{P}V$ (proj variety), the conormal space at $[x]$ $N^*_{[x]} X \subseteq V^*$ is just the annihilator of $\hat{T}_{[x]} X$, i.e.

$$N^*_{[x]} X = \left(\hat{T}_{[x]} X\right)^{\perp}$$

Defn [Gauss image / Dual variety] $X \subseteq \mathbb{P}V$ is an irred. hypersurface. Define the dual variety of X

$$X^\vee := \overline{\left\{ H \in \mathbb{P}V^* \;\middle|\; \exists [x] \in X_{smooth},\ \hat{T}_{[x]} X \subseteq \hat{H} \right\}}$$

hyperplane in $\mathbb{P}V$ determined by H

$$= \overline{\left\{ H \in \mathbb{P}V^* \;\middle|\; \exists [x] \in X_{smooth},\ H \in \mathbb{P}N^*_{[x]} X \right\}}$$

Point

Union of all conormal lines in $\mathbb{P}V^*$

(*) For our purpose, note that for smooth hypersurfaces other than ... a surface. If $X^\vee$ is not a

(*) For our purpose, note that for smooth hypersurfaces other than hyperplanes, the dual variety is also a hypersurface. If $X^\vee$ is not a hypersurface, we say X has a degenerate dual variety.

* Perm hypersurface does not have a degenerate dual variety
* Det hypersurface has a degenerate dual variety
* We need

$$\begin{tikzcd}
\mathbb{C}^{m^2} \arrow[r] \arrow[dr, "\mathrm{Perm}_m"'] & \mathbb{C}^{n^2} \arrow[r, "\det_n"] & \mathbb{C} \arrow[dl, leftrightarrow, "\text{we want zero sets to coincide}"] \\
& \mathbb{C} &
\end{tikzcd}$$

→ By studying the dimensions of the Gauss images of the hypersurfaces, we get a lower bound.

Since $(\det_n \text{ hypersurface})^\vee$ has low dimension, we need higher 'n' to have any chance of matching the perm hypersurface

Polynomials on the space of polynomials

* Ring of polynomial functions on a vector space V over a field k gives a co-ordinate free analog of a polynomial ring. Denoted $k[V]$.
* $k[V]$ consists of polynomials in t_i, where t_i form a basis of V^*. In other words, $k[V]$ is the comm. k-algebra generated by V^*.
* If k is infinite, $k[V]$ is the symmetric algebra on V^*, i.e. $\mathrm{Sym}(V^*)$
  - $\mathrm{Sym}^q(V^*)$ - v.s. of multilinear symmetric functionals
    $$\lambda : \prod_{i=1}^{q} V \to k$$
  - Any $\lambda \in \mathrm{Sym}^q(V^*)$ gives a homogeneous polynomial func. of degree q.

    $f(v) = \lambda(v, \ldots, v)$. To see that it is a polynomial func, let $\{e_i\}_{i \in [n]}$ be a basis of V & $\{t_i\}$ a basis of V^*.

    $n \quad \lambda(e \quad e)t(v) \quad t(v)$

let $\{e_i\}_{i\in[n]}$ be a basis of V & $\{t_i\}$ a basis ...

$$\lambda(v_1, \ldots, v_q) = \sum_{i_1, \ldots, i_q = 1}^{n} \lambda(e_{i_1}, \ldots, e_{i_q})\, t_{i_1}(v_1) \ldots t_{i_q}(v_q)$$

$$\Downarrow$$

f is a polynomial in the t_i's.

- $\phi : \mathrm{Sym}^q(V^*) \to k[V]_{(q)}$ is an isomorphism

$$\lambda \longmapsto \phi(\lambda)$$
$$\phi(\lambda) : v \mapsto \lambda(v, \ldots, v)$$

I'll fudge V & V^*

Elements of $\mathrm{Sym}^d\, \mathbb{C}^M$ can be thought of as (a) a hom. poly. of degree d on $(\mathbb{C}^M)^*$

(b) a symmetric tensor

(c) a homogeneous differential operator of order d on the space of Polynomials, i.e. $\mathrm{Sym}((\mathbb{C}^M)^*)$

Idea :→ $k[x_1, \ldots, x_n]$ is a vector space.

e.g. $\frac{\partial}{\partial x_i}$ maps an elem of $k[x_1, \ldots, x_n]_{(q)} \cong \mathrm{Sym}^q(V^*)$

to an elem of $k[x_1, \ldots, x_n]_{(q-1)} \cong \mathrm{Sym}^q(V^*)$

(Let $k \geq t$) $\mathrm{Sym}^k(V^*) \otimes \mathrm{Sym}^t(V) \longrightarrow \mathrm{Sym}^{k-t}(V^*)$

[Contraction map]

$\mathrm{Sym}^t(V)$ ↓ homogeneous linear differential operators of order t.

★ In other words, we can say $\mathrm{Sym}^d(\mathbb{C}^M)$ is v.s. of order d differential operators on the space of polynomials $\mathrm{Sym}((\mathbb{C}^M)^*)$.

Let $P \in \mathrm{Sym}^n\, \mathbb{C}^N$. $\mathbb{C}^{N^*}$ can be considered as the space of first order homogeneous differential operators on $\mathrm{Sym}^n\, \mathbb{C}^N$. Define

$$P_{1,n-1} : (\mathbb{C}^N)^* \longrightarrow \mathrm{Sym}^{n-1}(\mathbb{C}^N)$$
$$\frac{\partial}{\partial x_j} \longmapsto \frac{\partial P}{\partial x_j}$$

(... $\subset V \otimes \mathrm{Sym}^{n-1} V = \mathrm{Hom}(V^*, \mathrm{Sym}^{n-1} V)$.)

$\partial x_j \qquad \partial x_j$

(Can be co-ordinate free: $\mathrm{Sym}^n V \subseteq V \otimes \mathrm{Sym}^{n-1} V = \mathrm{Hom}(V^*, \mathrm{Sym}^{n-1} V)$.
For $P \in \mathrm{Sym}^n V$, $P_{1,n-1} \in \mathrm{Hom}(V^*, \mathrm{Sym}^{n-1} V)$)

Define

$$P_{2,n-2} : \mathrm{Sym}^2((\mathbb{C}^N)^*) \longrightarrow \mathrm{Sym}^{n-2} \mathbb{C}^N$$

$$\to \quad \frac{\partial^2}{\partial x_i \partial x_j} \longmapsto \frac{\partial^2 P}{\partial x_i \partial x_j}$$

$$\to P_{k,n-k} : \mathrm{Sym}^k((\mathbb{C}^N)^*) \longrightarrow \mathrm{Sym}^{n-k}(\mathbb{C}^N)$$

$$D \longmapsto D(P)$$

Proposition Let $P \in \mathrm{Sym}^d V^*$ be irreducible, and let $[x] \in \mathrm{Zeros}(P)$ be a general pt. Then

$$\dim \mathrm{Zeros}(P)^\vee = \mathrm{rank}\left(P_{d-2,2}(x^{d-2})\right) - 2$$

$\to P_{d-2,2}(x^{d-2}) \in \mathrm{Sym}^2 V^*$, a symmetric matrix. In co-ordinates, $P_{d-2,2}$ is a symmetric matrix whose entries are polynomials of degree $d-2$ in the co-ordinates of x, called the Hessian.

Prop Let $Q \in \mathrm{Sym}^m \mathbb{C}^M$, and $\tilde{A} : \mathbb{C}^M \to \mathbb{C}^N$ be such that

$Q(y) = P(\tilde{A}(y)) \;\forall y \in (\mathbb{C}^M)^*$, then

$$\mathrm{rank}\left(Q_{m-2,2}(y)\right) \leq \mathrm{rank}\left(P_{m-2,2}(\tilde{A}(y))\right)$$

We will show

(1) $\mathrm{rank}\left(n\text{.perm}_{m-2,2}(x^{m-2})\right)$ for general $[x] \in \mathrm{Zeros}(\mathrm{Perm}_m)$ is full, i.e. there is a pt. where the Hessian of the perm_m has full rank m^2

(2) $\dim \mathrm{Zeros}(\det_n)^\vee = 2n-2$ $\quad (\ldots\ (\ldots n-2))$ for general $[x] \in \mathrm{Zeros}(\det_n)$

(2) $\dim \text{Zeros}(\det_n)^\vee = 2n-2$

$\implies \text{rank}\left(n\text{-}\det_{n-2,2}(x^{n-2})\right)$ for general $[x] \in \text{Zeros}(\det_n)$
$= 2n$

(*) Since by prop $\ddot{X}$, the Gauss map of $\{\det(f(x)) = 0\}$, for $f: \mathbb{C}^{m^2} \to \mathbb{C}^{n^2}$ is as degenerate as the Gauss map of $\{\det(x) = 0\}$, we have $m^2 \leq 2n$
$\implies n \geq m^2/2$

To show that a hypersurface has a non-degenerate Gauss image, it (*) suffices to find a pt. where the Hessian of its defining eqn. has maximal rank

Lemma There exists such a pt.

Proof

Consider
$$y_0 = \begin{pmatrix} 1-m & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$
Easy to check $\text{Perm}_m(y_0) = 0$.

To compute $(\text{Perm}_m)_{m-2,2}(y_0)$, note

$$\frac{\partial^2}{\partial y_{i,j}\,\partial y_{k,l}} \text{Perm}_m \underbrace{\begin{bmatrix} y_{1,1} & \cdots & y_{1,m} \\ \vdots & & \vdots \\ y_{m,1} & \cdots & y_{m,m} \end{bmatrix}}_{Y} = \begin{cases} 0 & \text{if } i=k \text{ or } j=l \\ \text{Perm}_{m-2}\left(\underbrace{Y_{(\hat{i},\hat{j}),(\hat{k},\hat{l})}}_{Y \text{ with rows } i,k \text{ removed \& cols } j,l \text{ removed}}\right) & \text{otherwise} \end{cases}$$

Hessian of the perm at y_0 takes the form

$$\begin{pmatrix} 0 & Q & Q & \cdots & Q \\ Q & 0 & R & \cdots & R \\ & & & & \vdots \end{pmatrix} \in \mathbb{C}^{m^2 \times m^2}$$

$$M = \begin{pmatrix} Q & 0 & R & \cdots & R \\ Q & R & 0 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & R \\ Q & R & \cdots & R & 0 \end{pmatrix}, \text{ where}$$

$$Q = (m-2)\begin{pmatrix} 0 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & & \vdots \\ \vdots & \vdots & & \vdots & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}, \text{ and}$$

$$R = \begin{pmatrix} 0 & m-2 & m-2 & \cdots & m-2 \\ m-2 & 0 & -2 & \cdots & -2 \\ m-2 & -2 & 0 & \cdots & \vdots \\ \vdots & \vdots & & & -2 \\ m-2 & -2 & \cdots & -2 & 0 \end{pmatrix}$$

NTS M is invertible. WLOG assume $Q = Id_m$.

Let $\underset{\text{Ker} M}{V \in} = (\underbrace{V_{1,1} \cdots V_{1,m}}_{\tilde{V}_1}, \underbrace{V_{2,1} \cdots V_{2,m}}_{\tilde{V}_2}, \cdots \underbrace{V_{m,1}, \cdots V_{m,m}}_{\tilde{V}_m})$

We have the eqns.

$$\begin{aligned} \tilde{V}_2 + \cdots + \tilde{V}_m &= 0 \\ \tilde{V}_1 + R\tilde{V}_3 + \cdots + R\tilde{V}_m &= 0 \\ &\vdots \\ \tilde{V}_1 + R\tilde{V}_2 + \cdots + R\tilde{V}_{m-1} &= 0 \end{aligned} \iff \begin{aligned} \tilde{V}_2 + \cdots \tilde{V}_m &= 0 \\ \tilde{V}_1 - R\tilde{V}_2 &= 0 \\ &\vdots \\ \tilde{V}_1 - R\tilde{V}_m &= 0 \end{aligned}$$

$$\Downarrow$$

$$R\tilde{V}_2 + \cdots R\tilde{V}_m = 0$$

$$(m-1)\tilde{V}_1 = 0 \Rightarrow \tilde{V}_1 = 0$$

Therefore all other $\tilde{V}_i = 0 \Rightarrow$ Kernel of M is trivial $\boxtimes$

Gauss image of $perm_n$ has dimension n^2-2.

Zeros(det_n):

$(SL_n(\mathbb{C}) \times SL_n(\mathbb{C}))/\mu_n \rtimes \mathbb{Z}_2$ is the stabilizer of det_n (G_{det_n})

Kernel of the product map $\mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$

: $\det(AXB) = \det(A)\det(B)\det(X)$, and
$\det(X^T) = \det(X)$

$\star$ Any pt. on Zeros(det_n) is in the G_{det_n} orbit of

(1)
$$P_r := \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \text{ for } r \leq n-1$$

(2) also the hypersurface is singular at the pts. $G_{det_n} \cdot P_r$ for all $r < n-1$.

Recall $\sigma_1 = Seg(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n)$
$:= \mathbb{P}\{T \in A_1 \otimes \cdots \otimes A_n \mid R(T) = 1\} \subseteq \mathbb{P}(A_1 \otimes \cdots \otimes A_n)$

Prop Recall $\hat{\sigma}_r^0(Seg(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}))$ is the space of $n \times n$ matrices of rank r.

(1) $\hat{T}_M \hat{\sigma}_r^0(Seg(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})) = \{X \in Mat_{n\times n} \mid X\,Ker(M) \subseteq Im(M)\}$

(2) $N_M^* \hat{\sigma}_r^0(Seg(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})) = Ker\, M \otimes (Image\ M)^{\perp} = Ker\, M \otimes Ker\, M^T$

Lemma $\dim Zeros(det_n)^\vee = 2n-2$.

Proof All smooth pts on Zeros(det_n) are in the G_{det_n}-orbit of

$$P_{n-1} = \begin{pmatrix} I_{n-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Also, $N_{P_{n-1}}^* Zeros(det_n) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ * \end{pmatrix} \otimes (0\ 0\ \cdots\ *)$

$(Ker\, P_{n-1} \otimes Ker\, P_{n-1}^T)$

$(0\ \cdots\cdots\ 0\ |$

$(\text{Ker } \Psi_{n-1} \otimes \text{Ker } \Psi_{n-1})$

$$= \begin{pmatrix} 0 & \cdots & 0 \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots \; 0 & * \end{pmatrix} \leftarrow \text{rank 1 matrix}$$

Any smooth point can be moved to P_{n-1} by a change of basis, thus tangent hyperplanes to $Z(\det_n)$ are parameterized by rank one matrices $\hat{\sigma}_1^0 (\text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}))$ which has dimension $2n-2$

$\downarrow$

(multiplying a col. vector by a row vector) in hom. coordinates

$\boxtimes$

Thm [Mignon-Ressayre]

rank of Hessian of $\det_n$ at smooth pts is $2n$.

Recall rank of Hessian of perm_n at some pt. is m^2

$$\Rightarrow m^2 \leq 2n \Rightarrow dc(\text{perm}_m) \geq \frac{m^2}{2}$$

Grenet's u.b. on $dc(\text{perm}_m)$

Proof 1 [Combinatorial] ① Construct a directed graph whose vertices are subset of $\{1, \ldots, m\}$, and identify ϕ and $\{1, \ldots, m\}$. Thus $2^m - 1$ vertices in total.

② Place a directed edge from S to T with weight $x_{i,j}$ if

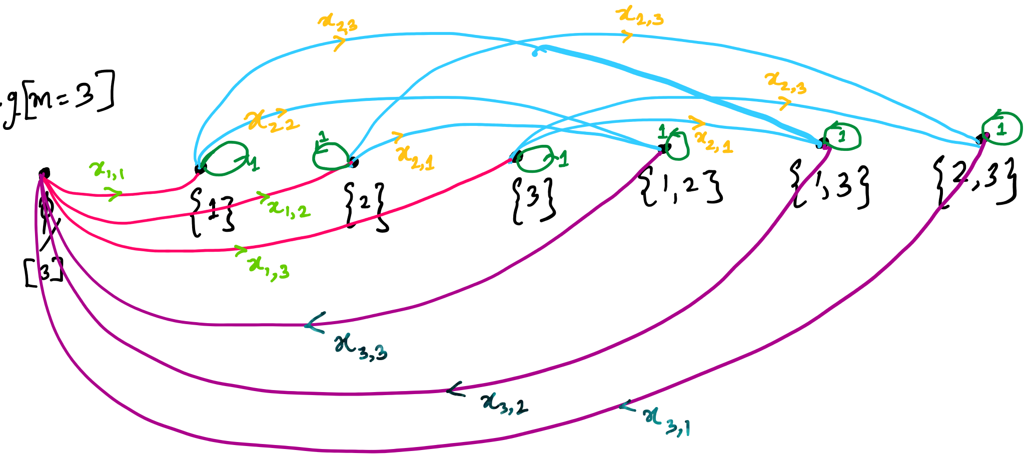
$$|S| = i-1, \; j \notin S, \; T = S \cup \{j\}.$$

(a) Thus node ϕ will have outgoing edges weight $x_{1,j}$ to all nodes $\{j\}$ $j \in [m]$

(b) Since ϕ & $[m]$ are identified, ϕ also has incoming edges of weight $x_{m,j}$ from the nodes $[m] \setminus \{j\}$.

③ All node except ϕ get a self loop of weight 1.

e.g [m=3]



$$G = \begin{bmatrix} 0 & x_{1,1} & x_{1,2} & x_{1,3} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x_{2,2} & x_{2,3} & 0 \\ 0 & 0 & 1 & 0 & x_{2,1} & 0 & x_{2,3} \\ 0 & 0 & 0 & 1 & 0 & x_{2,1} & x_{2,3} \\ x_{3,3} & 0 & 0 & 0 & 1 & 0 & 0 \\ x_{3,2} & 0 & 0 & 0 & 0 & 1 & 0 \\ x_{3,1} & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Observe Vertex cycle covers are in one-to-one correspondence with permutations in S_m.

↳ Because the graph is almost acyclic. Any non-self loop has to pass through ϕ. So any vertex cycle cover has to have exactly 1 non-self loop and $2^m - 1 - m$ self loops.

(*) Also the non-self loop cycles simulate process of adding the numbers $1, \dots, m$ in some particular order to the empty set.

(*) This gives you a permutation

Permanent of adjacency matrix of graph G $= \text{perm}(X)$.

The cycle covers all have the same sign, so $\text{Per}(G) = \pm \det(G)$. $\boxtimes$

Defn R- Comm. ring, E is a free R-module of rank r. Given an R-linear ... the Koszul complex associated to s is.

Defn R - Comm. ring, E is a free R-module of rank n. Given map $s : E \to R$, the Koszul complex associated to s is.

$$K_\bullet(s) : 0 \longrightarrow \bigwedge^{n} E \xrightarrow{d_n} \bigwedge^{n-1} E \longrightarrow \cdots \cdots \bigwedge^{1} E \xrightarrow{d_1} R \longrightarrow 0$$

where $d_k(e_1 \wedge \cdots \wedge e_k) = \sum_{i=1}^{k} (-1)^{i+1} s(e_i)\, e_1 \wedge \cdots \wedge \hat{e}_i \wedge \cdots \wedge e_k$.

Facts (1) above is a chain complex, i.e. $d_k \circ d_{k+1} = 0$.

(2) If $s : E = R^n$ (free mod. generated by a regular sequence of elements $x_1, \ldots, x_n$)
$\downarrow$
R

$s = [x_1 \ldots x_n]$, then

$K_\cdot(s)$ is a free resolution of $R/\langle x_1, \ldots, x_n \rangle$

(3) $d_1 = s$

Proof 2 [Algebraic] Let ϕ_i denote the matrix of d_i of the Koszul complex $x_{i,1}, \ldots, x_{i,m}$. Let $\tilde{\phi}_i$ be the matrix from ϕ_i by replacing minus signs by plus signs.

Let ψ be the direct sum of $\tilde{\phi}_i$, and define

$$M_m = \psi + 1.\, J_n$$

$n \times n$ nilpotent matrix + Jordan matrix with 1's on the sub diagonal

Verify that the determinant is the prem of $x_{i,j}$ upto a $\pm$ sign (the block diagonal structure helps ☺)

$\boxtimes$

```
i1 : R = QQ[x_11,x_12,x_13,x_21,x_22,x_23,x_31,x_32,x_33]

o1 = R

o1 : PolynomialRing

i2 : C = koszul matrix{{x_11,x_12,x_13}}

      1      3      3      1
o2 = R  <-- R  <-- R  <-- R

     0      1      2      3

o2 : ChainComplex

i3 : C.dd

          1                       3
o3 = 0 : R  <-------------------- R  : 1
              | x_11 x_12 x_13 |

          3                            3
     1 : R  <------------------------- R  : 2
              {1} | -x_12 -x_13 0     |
              {1} | x_11  0     -x_13 |
              {1} | 0     x_11  x_12  |

          3                 1
     2 : R  <-------------- R  : 3
              {2} | x_13  |
              {2} | -x_12 |
              {2} | x_11  |
```

← Compare to adj. matrix of G.

Stabilizer of the permanent:

$$G_{perm} = \Big(\big(T(SL_m(\mathbb{C})) \times T(SL_m(\mathbb{C}))\big) \rtimes (S_m \times S_m)\Big)/\mu_m \rtimes \mathbb{Z}_2$$

Maximal torus, i.e. diagonal matrices

(*) $\tilde{A}_{grenet} : \mathbb{C}^{m^2} \to \mathbb{C}^{n^2}$ for $n = 2^m - 1$ (Grenet's embedding) is $(T(SL_m(\mathbb{C}))$ - equivariant$)$

Satisfies equivariance :-
There exist an injective homomorphism $\psi : T(SL_m(\mathbb{C})) \to G_{det_n}$ s.t

$$\tilde{A}_{Grenet}(gY) = \psi(g)(\tilde{A}_{Grenet}(Y))$$

If you impose the restriction that your expression is $T(SL_n(\mathbb{C}))$ - equivariant, Grenet's expression is the best!

$edc(perm_m) = 2^m - 1$, obvs $edc(det_m) = m$.

✱ Thus we have exponential separation b/w perm & det in restricted model

If we can show an equivariant determinantal expression for $perm_m$ of size $dc(perm_m)^c$, then $VP_{\mathbb{C}} \neq VNP_{\mathbb{C}}$

Restricted models of computation are studied because
① Interesting
② Have implications to unrestricted, and are easier to study.

Defn Depth of a circuit is the no. of edges in the longest path from an input to its output.

Defn fan in no. of edges coming into a gate

An example (Waring Rank & $\Sigma\Lambda\Sigma$-Circuits)

Defn [Waring/Symmetric Rank] $P \in \mathbb{C}[\bar{x}]_{(d)}$ smallest r s.t we can write
$$P = l_1^d + \dots + l_r^d \qquad l_i \to \text{linear forms}$$

Defn $[\Sigma \Lambda^\delta \Sigma$ Circuit$]$ Consists of three layers: first addn gates, second powering gates $l \mapsto l^\delta$, third single addn gate.

Prop $P \in \mathrm{Sym}^d \mathbb{C}^n$, Waring rank$(P) = r$
① $\implies P$ admits a $\Sigma\Lambda^d\Sigma$ circuit of size $r(n+2)$

② $dc(P) \le d\,\mathrm{Waring}(P) + 1$.

algebro-geometric way of studying Waring rank

Waring rank can be studied by looking at Secant varieties of the Veronese variety.

Shallow Circuits that can be used for VP ≠ VNP

We shall consider depth 3, 4 & 5 Circuits
($\Sigma\Pi\Sigma$, $\Sigma\Pi\Sigma\Pi$, $\Sigma\Lambda^\alpha\Sigma\Lambda^\beta\Sigma$)

Defn A circuit is called homogeneous if for each + gate, Defn A circuit is called homogeneous if for each … inputs have the same degree

Thm $N = N(d)$ is a polynomial in d. Let (P_d) $\in Sym^d \mathbb{C}^N$ be a sequence of polys. that can be computed by a circuit of poly size $s = s(d)$. Let $\alpha(d) = 2^{O(\sqrt{d}\log ds \log N)}$. Then P_d is computable

(a) by a homogeneous $\Sigma\Pi\Sigma\Pi$ circuit of size $\alpha(d)$.

(b) by a $\Sigma\Pi\Sigma$ circuit of size $\alpha(d)$

(c) by a homogeneous $\Sigma\wedge^{O(\sqrt{d})}\Sigma\wedge^{O(\sqrt{d})}\Sigma$ circuit of size $\alpha(d)$

Cor. if $(perm_m)$ is not computable by any of the above circuits of size $2^{\omega(\sqrt{m}\log^{3/2} m)}$, then $VP \neq VNP$

The point is that all of the above results can be stated geometrically.

e.g. Let's state the result abt. $\Sigma\Pi\Sigma$ circuits geometrically.

$P \in Sym^d \mathbb{C}^N$ has a circuit of size s.

Prop ① Let $d = N^{O(1)}$, then $[\ell^{n-d} P]$ belongs to the r^{th} secant variety of the Chow variety of degree n in $\mathbb{C}^{N+1}$ with $rn \sim s^{\sqrt{d}}$

② If $[\ell^{n-m} perm_m] \notin r^{th}$ secant variety of the degree n Chow variety in $\mathbb{C}^{m^2+1}$, then $VP \neq VNP$

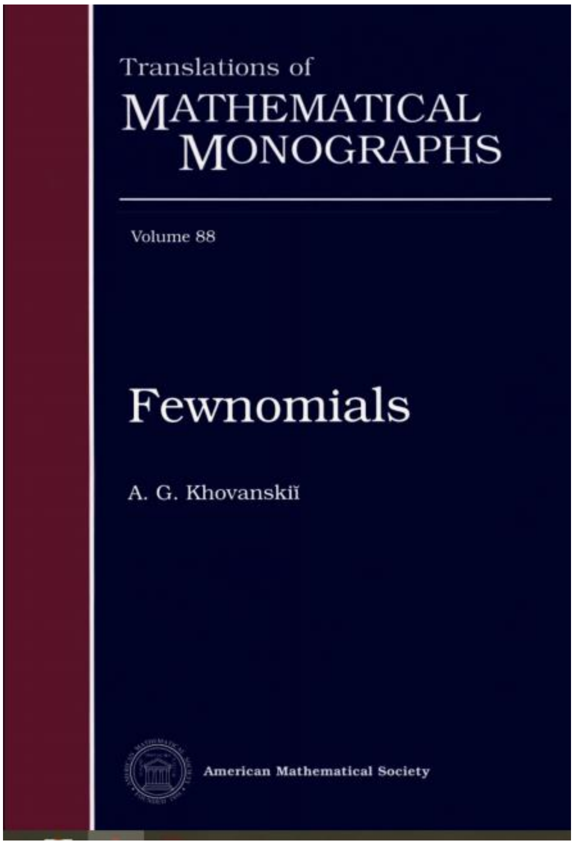
Thm [Gupta et al. "Method of Shifted Partial Derivatives"]
Any $\Sigma\Pi^{O\sqrt{m}}\Sigma\Pi^{O\sqrt{m}}$ circuit that computes $perm_m$ must have top fanin at least $2^{\Omega(\sqrt{m})}$

→ Came very close to $VP \neq VNP$
→ But Efremenko et al. show the method cannot be used to show $VP \neq VNP$

Fewnomials & Real-Tau conjecture

Thm [Descartes' Rule of Signs] A univariate polynomial of any degree, but with at most 't' monomials has $\sim 2t+1$ roots (counted with multiplicities)

⊛ The theory of fewnomials is has been studied very extensively



Conjecture [Real-Tau Conjecture] no. of zeros of

$$\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x), \text{ where } f_{i,j} \text{ are } t\text{-sparse is}$$

$\text{Poly}(k, t, 2^m)$.

Thm Real-Tau $\Rightarrow VP_{\mathbb{C}} \neq VNP_{\mathbb{C}}$

Mathematical problems for the next century

Steve Smale

The Mathematical Intelligencer 20, 7–15 (1998) | Cite this article

4th	Shub–Smale tau-conjecture on the integer zeros of a polynomial of one variable[6][7]	Unresolved.

Thm [Briquel-Bürgisser] If $f_{i,j}$ are chosen as follows :-

① Fix support of size t

② Let coeffs be independent $\mathcal{N}(0,1)$

Then $\mathbb{E}[\text{real zeros}] = O(k m^2 t)$!

Thm [Koiran et al.].

$$\sum_{i=1}^{k} \prod_{j=1}^{m} f_j^{\alpha_{i,j}} \text{ has } O\left(t^{O(k^2 m)}\right)$$ roots, thus restricted

Class of polynomially sized depth-4 Circuits cannot compute the permanent.

e.g. $fg+1 \rightarrow$ Descartes gives $\sim t^2$ bound. open question to show $\sim t$.
f,g t-sparse

MAIN TECHNICAL TOOL - WRONSKIAN

Given $f_1 \ldots f_k$, define

$$W(f_1 \ldots f_k) = \det\left[\left(f_j^{(i-1)}\right)_{i,j \in [k]}\right]$$

Prop If $f_1 \ldots f_k$ are analytic functions, then

$\{f_i\}$ are linearly dependent $\iff W(f_1, \ldots, f_k) = 0$

Thm [Voorhoeve & Van der Poorten] Let $f_1, \ldots, f_k$ be real analytic funcs over an interval I. Then

$$N(f_1 + \cdots + f_k) \le k-1 + \sum_{j=1}^{k-2} N(W(f_1 \ldots f_j)) + \sum_{j=1}^{k} N(W(f_1 \ldots f_j))$$

zeros with multiplicity

(*) We need to count zeros without multiplicity because we don't want $\alpha_{i,j}$ in the bound.

(*) We need to count ... Want $\alpha_{i,j}$ in the bound.

Thm [Koiran et al.]. Same bound holds on no. of zeros without multiplicity if f_i's are linearly independent on I.

Main thm follows by applying the u.b. on the sum along with studying what happens when you have $W(f_1^\alpha, \ldots, f_k^\alpha)$ vs $W(f_1, \ldots, f_k)$

ULRICH COMPLEXITY

(Introduced by Blaser - Eisenbud - Schreyer)

How do you resolve VP vs VNP (so far):

1. Prove that you can or cannot compute the permanent in polynomial time
2. Either a super polynomial lower bound on dc (perm) or a polynomial u.b. on dc (perm)
3. Some specific bounds on sizes of bounded-depth circuits
4. suitable Upper bounds on no. of roots of specific types of fewnomials

Ulrich Complexity: - Suggests how to prove VP = VNP
- Can be thought of in a very abstract & co-ordinate free way.

Defn The Ulrich complexity of a homogeneous polynomial $f \in K[x_0 \ldots x_n]$ of degree d is the smallest r s.t there exists a square matrix M of linear forms with

$$\det(M) = f^r, \text{ and such that}$$

there is a matrix N with $MN = f\, I_{dr}$

The first idea here is that if you can compute f^r easily, then you can also compute f easily by just taking the r^{th} root.

Thus we can study the size of the matrix M that gives you $\det(M) = f^r$. [... the entries are linear forms], so

Thus we can study the size of the matrix M that gives you ...
However, $\deg f \times r = $ size(M) [because the entries are linear forms], so we might as well study $\frac{\text{size}(M)}{d} = r$.

So far, nothing too different has happened. However, by adding the restriction that there is a matrix N with $MN = f I_{dr}$ brings this into the language of Ulrich modules & Ulrich sheaves

$uc(\det) = 1$ (N is the matrix of cofactors)

Conjecture $uc(f) \geq 2^{\lceil \text{codim sing} f / 2 \rceil - 2}$ for all f.

where sing f, the singular locus, is where all partial derivatives of f vanish.

this is a special case of the BGS conjecture

* It is well known that the codim sing $\det_n = 4$ (by the generalized Principal ideal thm, see "Comm. Alg. with a View toward Alg Geom Chap10 by Eisenbud)

So the above predicts $uc(\det_n) \geq 1$, which is correct ☺

The codimension of the singular locus of the permanent is not known. Conjectured to be $2n$. This gives

Conjecture $uc(\text{perm}_m) \geq 2^{m-2}$

Trivially true for $m=2$, also true for $m=3$.

Thm

Polynomial upper bound on $uc(\text{perm}) \implies VP = VNP$

Stated as subexponential in the original paper, but author confirmed to me in personal communication that it should be polynomial

Defn [Ulrich bundles] Let $X \subseteq \mathbb{P}^n$ be a smooth variety of degree d.

A rank r vector bundle E on X is Ulrich if any of the foll. equiv conds are satisfied:-

① The cohomology $H^i(X, E(-p))$ vanishes for $1 \leq p \leq \dim(X)$

(large number of vanishing cohomology groups)

[... Macaulay (aCM) bundles]

(large number of vanishing cohomology groups)

[Read about arithmetically Cohen-Macaulay (aCM) bundles]

" Ulrich, in some sense mean they have largest permitted number of global sections "

② If $\pi : X \to \mathbb{P}^{\dim(X)}$ is a finite linear projection, the vector bundle $\pi_* E$ is trivial

Defn [Co-ord free Ulrich Complexity] Let $f \in k[x_0 \dots x_n]$ homogeneous define an integral hypersurface $X \subseteq \mathbb{P}^n$. The Ulrich complexity $uc(f)$ is the min. rank of an Ulrich bundle on X. (Agrees with earlier defn of $uc(f)$)

Defn [Linear Mat fac] for $f \in k[x_0 \dots x_n]$ be a homogeneous poly of degree $d \geq 2$. f has a matrix factorization of size 'm' if $\exists$ $\alpha_1 \dots \alpha_d$ in $M_m(k)$ (entries are linear forms)

$$f I_m = \alpha_1 \cdots \alpha_d$$

Defn [Chow rank, Waring rank] $f \in k[\bar{x}]$, homogeneous & degree d. Waring rank of f, denoted $w(f)$ is the min. number s.t there exists

$$f = \sum_{i=1}^{w(f)} l_i^d$$

Chow rank, denote $ch(f)$ – min. number s.t. there exist

$$f = \sum_{i=1}^{ch(f)} l_{i,1} \cdots l_{i,d}$$

(All l_* are linear forms)

Thm ① f has a linear mat factorization of size $d^{w(f)-1}$, and of size $d^{ch(f)-1}$

② f has l.m.f of size $m \Rightarrow$ hypersurface supports Ulrich bundle of rank $\frac{m}{d} \Rightarrow uc(f) \leq \frac{m}{d}$

③ X supports an Ulrich bundle of rank $d^{w(f)-2}$ & $d^{ch(f)-2}$

(Ulrich complexity can be studied by studying secant varieties (c.f. Prev. lecture)

Ulrich complexity can be studied by studying secant ...
of Veronese and/or Chow varieties (c.f. prev. lecture)

Geometric Complexity Theory (GCT)

Friday, 9 June 2023 16:12

"... String theory of computer science ..." — Aaronson

Introduced by Mulmuley - Sohoni :

GCT publications:

Overviews of GCT

- The GCT program toward the P vs. NP problem, CACM, vol. 55, issue 6, June 2012, pp. 98-107.
- On P vs. NP, and Geometric Complexity Theory, JACM, vol. 58, issue 2, April 2011.
- FOCS 2010 Tutorial based on this overview.

GCT Papers

- Lower Bounds in a Parallel Model without bit operations, SIAM J. Comput., 28, (1999), pp. 1460-1509.
- Geometric complexity theory I: An approach to the P vs. NP and related problems (with M. Sohoni), SIAM J. Comput., vol 31, no. 2, pp. 496-526, (2001).
- Geometric complexity theory II: Towards explicit obstructions for embeddings among class varieties (with M. Sohoni) , SIAM J. Comput., Vol. 38, Issue 3, June 2008.
- Geometric complexity theory, P vs. NP and explicit obstructions (with M. Sohoni), in "Advances in Algebra and Geometry", Edited by C. Musili, the proceedings of the International Conference on Algebra and Geometry, Hyderabad, 2001.
- Geometric complexity theory III: on deciding nonvanishing of a Littlewood-Richardson coefficient (with H. Narayanan and M. Sohoni), Journal of Algebraic Combinatorics, pages 1-8, November, 2011.
- Geometric complexity theory IV: nonstandard quantum group for the Kronecker problem (with J. Blasiak and M. Sohoni), to appear in Memoirs of American Mathematical Society. Preprint available as arXiv:cs/0703110[cs.CC], June 2013.
- Geometric Complexity Theory V: Efficient algorithms for Noether normalization, to appear in the Journal of the AMS.
- Explicit Proofs and The Flip, Technical Report, Computer Science Department, The University of Chicago, September 2010.
- Geometric Complexity Theory VI: the flip via positivity, Technical Report, computer science department, The University of Chicago, January 2011.
- Geometric Complexity Theory VII: Nonstandard quantum group for the plethysm problem, Technical Report TR-2007-14, computer science department, The University of Chicago, September, 2007.
- Geometric Complexity Theory VIII: On canonical bases for the nonstandard quantum groups, Technical Report TR-2007-15, computer science department, The University of Chicago, September, 2007.

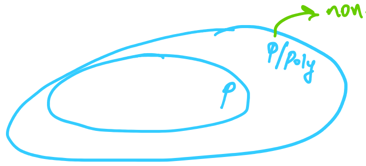
Lecture notes on GCT

- On P vs. NP, Geometric Complexity Theory, and the Riemann Hypothesis, Technical Report, Computer Science department, The University of Chicago, August, 2009. cs.ArXiv preprint cs.CC/0908.1936

This overview is based on a series of three lectures. Video lectures in this series are available here.

- Geometric Complexity Theory: Introduction (with M. Sohoni), Technical Report TR-2007-16, computer science department, The University of Chicago, September, 2007. Lecture notes for an introductory graduate course on geometric complexity theory in the computer science department, the university of Chicago.
- On P vs. NP, Geometric Complexity Theory, and The Flip I: a high-level view, Technical Report TR-2007-13, computer science department, The University of Chicago, September, 2007.

☆ Most expositions describe GCT as an approach towards VP vs VNP, but GCT-like approach is feasible for even P vs NP.



non-uniform version of P

$P \subsetneq P/poly$

Thus if $NP \not\subseteq P/Poly \Rightarrow P \neq NP$

Example of P/Poly algorithm (Miller-Rabin primality test)

1. ^ a b Miller, Gary L. (1976), "Riemann's Hypothesis and Tests for Primality", *Journal of Computer and System Sciences*, **13** (3): 300–317, doi:10.1145/800116.803773, S2CID 10690396
2. ^ a b Rabin, Michael O. (1980), "Probabilistic algorithm for testing primality", *Journal of Number Theory*, **12** (1): 128–138, doi:10.1016/0022-314X(80)90084-0

Testing against small sets of bases [edit]

When the number n to be tested is small, trying all $a < 2(\ln n)^2$ is not necessary, as much smaller sets of potential witnesses are known to suffice. example, Pomerance, Selfridge, Wagstaff[4] and Jaeschke[11] have verified that

- if n < 2,047, it is enough to test a = 2;
- if n < 1,373,653, it is enough to test a = 2 and 3;
- if n < 9,080,191, it is enough to test a = 31 and 73;
- if n < 25,326,001, it is enough to test a = 2, 3, and 5;
- if n < 3,215,031,751, it is enough to test a = 2, 3, 5, and 7;
- if n < 4,759,123,141, it is enough to test a = 2, 7, and 61;
- if n < 1,122,004,669,633, it is enough to test a = 2, 13, 23, and 1662803;
- if n < 2,152,302,898,747, it is enough to test a = 2, 3, 5, 7, and 11;
- if n < 3,474,749,660,383, it is enough to test a = 2, 3, 5, 7, 11, and 13;
- if n < 341,550,071,728,321, it is enough to test a = 2, 3, 5, 7, 11, 13, and 17.

Using the work of Feitsma and Galway enumerating all base 2 pseudoprimes in 2010, this was extended (see OEIS: A014233), with the first result shown using different methods in Jiang and Deng:[12]

- if n < 3,825,123,056,546,413,051, it is enough to test a = 2, 3, 5, 7, 11, 13, 17, 19, and 23.
- if n < 18,446,744,073,709,551,616 = 2^{64}, it is enough to test a = 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, and 37.

Sorenson and Webster[13] verify the above and calculate precise results for these larger than 64-bit results:

- if n < 318,665,857,834,031,151,167,461, it is enough to test a = 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, and 37.

(efficiently computing the list 'a' for any n is not possible)

High-level overview of GCT:- Consider P vs NP. Construct, for each n, algebraic varieties $X_{P,n}$ & $X_{NP,n}$ s.t.

$$P \supseteq NP \iff X_{NP,n} \subseteq X_{P,n^k} \text{ for } \forall n \geq n_0 \text{ \& some } k.$$

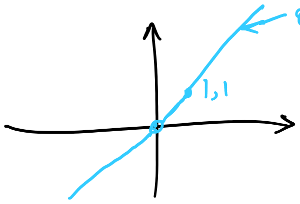
Make sure $X_{NP,n}$ & X_{P,n^k} are symmetric under the action of GL_n so that we can use tools from representation theory.

Caveat:→ So far, representation theoretic reasons are sufficient but not necessary

Orbit Closures Consider the action of $\mathbb{R}^\times$ on $\mathbb{R}^2$

$$a \cdot (x,y) = (ax, ay).$$

Then $\mathbb{R}^\times \cdot (1,1)$ is



orbit of (1,1) under $\mathbb{R}^\times$ action

NOT CLOSED IN EUCLIDEAN & ZARISKI TOPOLOGY

In GCT, orbit closures come up a LOT. It is not ... difficulty is to understand

In GCT, orbit closures come up a LOT. It is not exaggeration to say that main difficulty is to understand ORBIT CLOSURES vis-a-vis ORBITS

CLASSICAL COMPLEXITY	GCT
A problem / Funct to be computed	Pt. on an algebraic variety
$f \sim g$	pt_f and pt_g lie in the same orbit
$\Downarrow$	$\Downarrow$
Reduction b/w f & g	action of group element
Reduction b/w arbitrary funcs	action of limits of group elems.
$\Downarrow$	$\Downarrow$
$f \le g$	f lies in $\overline{G.g}$ ← orbit closure

$V = (\mathbb{C}^{m^2})^*$. End (V) acts on $Sym^m(V)$ ← degree m hom. polynomials in m^2 vars.

$$\underset{\in}{L} . p(x) = P(L^T x)$$

End (V) is not a group.

[Padded n-permanent] $z^{m-n}\,\mathrm{perm}_n \in Sym^m(V)$. Also, obviously $\det_m \in Sym^m(V)$

Prop $dc(\mathrm{perm}_n) \le m = n^{O(1)} \iff \mathrm{End}\,V.\det_m \ni z^{m-n}\,\mathrm{perm}_n$

↑ any variable other than $(x_{ij})_{i,j\in[n]}$

Also,

$$z^{m-n}\,\mathrm{perm}_n \in \mathrm{End}(V).\det_m \iff \mathrm{End}(V).\,z^{m-n}\,\mathrm{perm}_n$$

$GL(V)$ ← group of invertible linear transformations $\subseteq \mathrm{End}(V)$, and is dense in End (V), i.e. $\overline{GL(V)} = \mathrm{End}(V)$,

$\left(\forall A \in \mathrm{End}(V) \setminus GL(V),\ \exists \text{ sequence } (A_i)_{i\in\mathbb{N}} \text{ s.t.} \right)$

dense in $End(V)$, i.e. $\overline{GL(V)} = End(V)$,

$\left(\forall A \in End(V) \setminus GL(V),\ \exists \text{ sequence } \underset{GL(V)}{(A_i)}_{i \in \mathbb{N}} \text{ s.t. } \lim_{i \to \infty} A_i = A \right)$

Thus $\underset{\text{group orbit}}{GL(V)}.det_m$ is dense in $End(V).det_m$

$$DET_m := \overline{GL(V).det_m} = \overline{End(V).det_m}$$

$$\left(\text{Infact } End(V).det_m \subsetneq DET_m\right)$$

$$PER_m^n = \overline{GL(V).\, z^{m-n} perm_n}$$

Conjecture [Strengthening of Valiant's conjecture] when $m = n^{O(1)}$, then

$PER_m^n \notin DET_m$ (equiv. to $z^{m-n} perm_n \notin DET_m$)

orbit closure is the smallest set that is closed (Euc. & Zar)

Padded perm is not the limit of a sequence of pts in the $GL(V)$ orbit of det_m.

"Cannot be approximated"

The above conjecture implies Valiant's conjecture, but not equivalent because it involves orbit closures (End-orbit $\subsetneq$ orbit closure)

Why do we move to orbit closures? Because, by virtue of being closed, they are algebraic varieties.

This came up when we studied matrix multiplication

Also, the padded perm cannot be in $GL(V).det_m$ since it is reducible

(LATER) We can use any two functions complete for two complexity classes and ask if there is inclusion b/w their orbit closures.

This is what allows us to use this approach for $NP \not\subseteq P/poly$, and thus showing $P \neq NP$.

WHERE IS THIS GOING:-

Q. Why should we hope that the language of orbit closures is more promising?

↓ (that's why)

CHARACTERISATION BY SYMMETRIES

+ Partial Stability & Stability

⇓ By Luna's étale slice thm

You can look at multiplicities of irreps in the isotypic decomposition of the representations obtained by considering actions of $GL(V)$ on the co-ordinate rings of the determinant and the padded permanent

There are surprising algorithms to calculate Littlewood-Richardson coeffs. There is a simple linear programming based algo. that tests positivity of LR coeff. MS suggest that this is the best way forward.

Let us work through some examples for motivation.

APPROACH TO USE ALGEBRAIC GEOM. FOR COMPLEXITY CLASS SEPARATION

Idea to show $x \notin X$ ① Find a polynomial P that vanishes on all of X
② Show that $P(x) \neq 0$

'P' is called a seperating polynomial

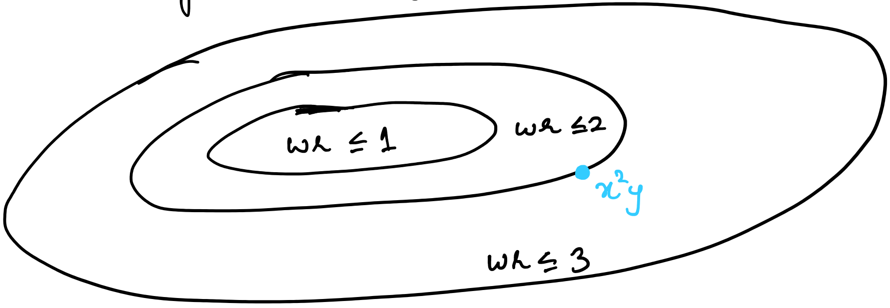
This strategy could be used for any complexity class separation

Consider the polynomial x^2y:

$$x^2y = \frac{1}{6}\left[(x+y)^3 + (y-x)^3 - 2y^3\right] \Rightarrow \mathrm{wr}(x^2y) \le 3.$$

$$x^2y = \frac{1}{6}\left[(x+y)^3 + (y-x)^3 - 2y^3\right] \implies wr(x^2y) \le 3.$$

Infact $wr(x^2y) = 3$!



Check that $\frac{1}{3\varepsilon}\left((x+\varepsilon y)^3 - x^3\right) = x^2y + \varepsilon xy^2 + \frac{\varepsilon^2}{3}y^2$

$$\downarrow \varepsilon \to 0$$

$$x^2y$$

$\Downarrow$

x^2y is the limit of a sequence of polynomials of $wr = 2$

Thus because of continuity, any poly that vanishes on $wr \le 2$ must also vanish at x^2y. Thus we define border Waring rank

smallest n s.t. there is an approximation by a sequence of polys of $wr \le n$, denoted $\underline{wr}$

Similarly, we define border determinantal complexity, border formula size, etc.

Border complexity measures are convenient b'cos $S := \{p \in \mathbb{k}[\bar{x}] \mid \underline{wr}(p) \le n\}$ is closed (in Euclidean & Zariski topology), thus finding a seperating polynomial is plausible! ☺ Also, GL acts of such sets.

Aside It is natural to use the language of orbit closures for sets of type S.

here's why

$$x^2y = \lim_{\varepsilon \to 0} \frac{1}{3\varepsilon}\left((x+\varepsilon y)^3 - x^3\right).$$ Let $s_\varepsilon = \left(\frac{1}{3\varepsilon}\right)^{1/3}$ & $\omega^3 = -1$, then

$$x^2y = \lim_{\varepsilon \to 0}\left[(s_\varepsilon x + \varepsilon s_\varepsilon y)^3 + (\omega s_\varepsilon x)^3\right]$$

$(s_\varepsilon x + \varepsilon s_\varepsilon y)^3 + (\omega s_\varepsilon x)^3$

... the polynomial

$(s_\varepsilon x + \varepsilon s_\varepsilon y)^3 + (w s_\varepsilon x)^3$

Can be thought of evaluating the polynomial x^3+y^3 at the pt. $(x \;\; y)\begin{pmatrix} s_\varepsilon & w s_\varepsilon \\ \varepsilon s_\varepsilon & 0 \end{pmatrix}$

This is denoted
$$\begin{pmatrix} s_\varepsilon & w s_\varepsilon \\ \varepsilon s_\varepsilon & 0 \end{pmatrix} \circ (x^3+y^3)$$

Generalizing, we can say

$M_2(\mathbb{C}) \circ (x^3+y^3)$ exactly gives you the set of all polynomials of $wr \le 2$ "Monoid orbit"

$x^2 y \notin M_2(\mathbb{C}) \circ (x^3+y^3)$, but $x^2 y \in \overline{M_2(\mathbb{C}) \circ (x^3+y^3)}$

(*) We don't like $M_2(\mathbb{C})$, but we do like $GL_2(\mathbb{C})$, which is dense in $M_2(\mathbb{C})$.

Thus $x^2 y \in \overline{GL_2(\mathbb{C}) \circ (x^3+y^3)}$

TAKEAWAY:- Language of orbit closures, which we like for mathematical reasons, comes up when we try to attack complexity class separations using algebraic geometry

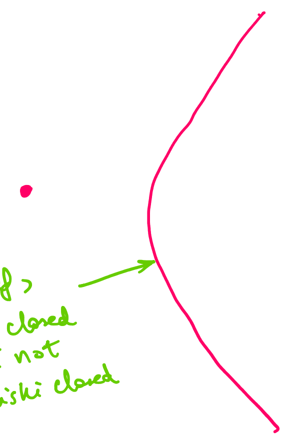
Proof that such sets are algebraic varieties

Zariski closure is more coarse than Euclidean topology, so $\overline{Y}^Z \supseteq \overline{Y} \supseteq Y$ (Zariski closure)

$W_n = \{ p \in k[\bar{x}] \mid wr(p) \le n \}$

$\overline{Y}^Z \neq \overline{Y}$ in general. e.g. over $\mathbb{R}$

Consider zeros over $\mathbb{R}$ of $y^2 = x(x+1)^2$



By itself, this is closed but not Zariski closed

Chevalley's Structure theorem tells orbit closures are varieties $\boxtimes$

Tiny example: Consider $\mathrm{Sym}^2(\mathbb{C}^2) \longrightarrow$ 3 dimensional vec. space with basis x^2, xy, y^2

Tiny example: Consider $\mathrm{Sym}^2(\mathbb{C}^2) \to$ 3 dimensional vec. space ...
x^2, xy, y^2

Let $X_1 = \{h \in \mathbb{C}[x,y] \mid \exists \alpha, \beta \in \mathbb{C}, \text{ s.t } h = (\alpha x + \beta y)^2\}$

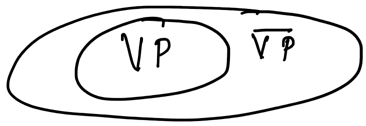
It can be seen that that $X_1 = \{ax^2 + bxy + cy^2 \mid b^2 - 4ac = 0\}$

★ $b^2 - 4ac \in \mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2))$ is a separating polynomial.

Claim $wr(xy) > 1$

Proof $xy = \underset{a}{0}x^2 + \underset{b}{1}.xy + \underset{c}{0}.y^2$, $b^2 - 4ac \neq 0$ ⊠

For all complexity measures we care abt, we know examples where
- border complexity is different from normal complexity
- We can define $\overline{VP}$
- Strengthening of Valiant's conjecture: $\overline{VP} \neq VNP$
- We don't know if $\overline{VP} \subseteq VNP$



we don't know if
$VP \overset{?}{=} \overline{VP}$

GCT proposes that we study orbit closures via representation theory
(remember, in our cases, GL_n acts)

Recap. (1) We have the vector space of polynomials. Has GL_n action
(2) We have a Zariski closed subset X inside (orbit closure)
(3) Suitable functions on X will help ascertain membership in X.
(4) GL_n action carries over to functions on X, so it is a representation
(5) Use multiplicities .....

Let us introduce aspects (4) & (5) via examples again

Consider $\mathrm{Sym}^2(\mathbb{C}^2)$. Consider the action of S_2: $e \neq \rho \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$, $\rho \in S_2$

Then $\mathrm{Sym}^2(\mathbb{C}^2)$ is a 3-dimensional representation of S_2.

Thus $Sym^2(\mathbb{C}^2)$ is a 3-dimensional representation of S_2.

$\left(\rho : S_2 \longrightarrow GL\left(Sym^2(\mathbb{C}^2)\right) \text{ is a group homomorphism}\right)$

$Sym^2(\mathbb{C}^2)$ has a basis x^2, y^2, xy. Another basis is x^2+y^2, x^2-y^2, xy

observe $\underbrace{\rho(xy) = xy, \rho(x^2+y^2) = x^2+y^2}_{\text{invariants under } S_2}$, $\rho(\underbrace{x^2-y^2}_{\text{skew-invariant under } S_2}) = y^2-x^2$

$$Sym^2(\mathbb{C}^2) \cong \mathrm{span}(xy, x^2+y^2) \oplus \mathrm{span}(x^2-y^2)$$

under action of S_2, this vector subspace is closed
other words this vector subspace is fixed by S_2.
It is a <u>subrepresentation</u>

multiplicity of invariants of $Sym^2(\mathbb{C}^2) := \dim \mathrm{span}(xy, x^2+y^2) = 2$

multiplicity of skew-invariants of $Sym^2(\mathbb{C}^2) := \dim \mathrm{span}(x^2-y^2) = 1$

Another example

Consider $Sym^2(Sym^2(\mathbb{C}^2))$: $Sym^2(\mathbb{C}^2)$ has basis x^2, y^2, xy. An elem

– of $Sym^2\mathbb{C}^2$ corresponds to choices of coeffs for x^2, y^2, xy, say a, b, c respectively

– Thus elem of $Sym^2(Sym^2(\mathbb{C}^2))$ can be thought of as degree 2 polys in a, b, c, thus $\{a^2, ab, ac, b^2, bc, c^2\}$ is a basis of $Sym^2(Sym^2(\mathbb{C}^2))$

– Take the S_2 action on $Sym^2(\mathbb{C}^2)$, i.e. $x \longleftrightarrow y$. Thus, the corresponding coeffs must be transformed correspondingly, i.e.

$$a \longleftrightarrow c \qquad b \longleftrightarrow b.$$

Thus $e \neq \rho \in S_2$ acts on $Sym^2(Sym^2(\mathbb{C}^2))$ as follows

$$a^2 \overset{\rho}{\longleftrightarrow} c^2, \quad ab \overset{\rho}{\longleftrightarrow} bc, \; ac \overset{\rho}{\longleftrightarrow} ac, \; b^2 \overset{\rho}{\longleftrightarrow} b^2.$$

[illegible] a different basis, we have

— Taking a different basis, we have

$$Sym^2(Sym^2(\mathbb{C}^2)) \cong \underbrace{\langle ac, b^2, a^2+c^2, ab+bc \rangle}_{\text{invariants under } S_2} \oplus \underbrace{\langle a^2-c^2, ab-bc \rangle}_{\text{skew-invariants under } S_2}$$

here $\text{mult}_{inv} = 4$, $\text{mult}_{skew-inv} = 2$

* note $b^2-4ac \in$ invariants of $Sym^2(Sym^2(\mathbb{C}^2))$ under S_2 action

Instead of just S_2, we could have analyzed this for GL_2 itself

a computation shows that

$$\underset{\substack{\cap \\ GL_2}}{g} \circ \langle b^2-4ac \rangle = (\det(g))^2 (b^2-4ac)$$

. Thus, we have

$$Sym^2(Sym^2(\mathbb{C}^2)) \cong \underbrace{\text{span}(b^2-4ac)}_{A} \oplus \underbrace{\text{span}(a^2, ab, b^2+4ac, bc, c^2)}_{B}.$$

Furthermore, neither A nor B have non-trivial subrepresentations

A & B are called irreducible representations

GL_n is a reductive group ... any finite dim representation of GL_n has such a decomposition into irreducibles

Q: Great! Why did we go through all this?

A: We want to do complexity class separation. For this study the representation obtained by the "Vanishing ideal"

Let $X \subseteq Sym^2(\mathbb{C}^2)$ be the set of polys of Waring rank ≤ 1.

Defn [Vanishing ideal] $I(X) := \{ f \in \mathbb{C}[Sym^2(\mathbb{C}^2)] \mid Z(f) \supseteq X \}$

— thus if $f \in X$, then f vanishes on all elems of $I(X)$

(if X is Zariski closed, then above is iff)

(Recall the language of orbit closures makes an appearance here)

$I \cap \dots \; I(X) \cap Sym^d(Sym^2(\mathbb{C}^2))$

Define $I(X)_d = I(X) \cap \mathrm{Sym}^d(\mathrm{Sym}^2(\mathbb{C}^2))$

$I(X)_d$ is a subrepresentation because X is an orbit closure

In this case $I(X)_2 = \mathrm{span}(b^2 - 4ae)$. Juxtapose with

$$\mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2)) = \underbrace{\mathrm{span}(b^2-4ae)}_{\text{irred}} \oplus \underbrace{\mathrm{span}(a^2, ab, b^2+4ae, be, c^2)}_{\text{irred}}$$

★ This illustrates that the separating polynomial can specifically be taken from an irreducible representation (irrep) of the vanishing ideal

What is special abt. b^2-4ae? It is highest weight vector (hwv)

Thm + Defn [highest weight vector] Every irrep of GL_n contains (up to scale) exactly one hwv. Defn of hwv: $f \in$ irrep of GL_n s.t.:

① $\mathrm{diag}(\alpha_1, \dots \alpha_n)\, f = \alpha_1^{\lambda_1} \dots \alpha_n^{\lambda_n} f$

② $\begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} f = f$

$(\lambda_1, \dots, \lambda_n)$ is called the weight of f.

★ Two irreps of GL_n are called isomorphic if their hwvs coincide.

Since $g(b^2-4ae) = (\det(g))^2 (b^2-4ae)$, b^2-4ae is a hwv of wt. $(2,2)$

Also, a^2 is hwv of the irrep span $(a^2, ab, b^2+4ae, bc, c^2)$ of weight $(4,0)$

★ You can use hwv for a border complexity lower bound in restricted cases

Thm + defn The no. of isomorphic copies of irreducibles in a decomposition is independent of the decomposition. This no. is called the multiplicity.

$$\mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2)) = \underbrace{\mathrm{span}(b^2-4ac)}_{\text{irrep of type}} \oplus \underbrace{\mathrm{span}(a^2, ab, b^2+4ae, bc, c^2)}_{\text{irrep of type } (4,0)}$$

$\mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2)) = \underbrace{\mathrm{span}(b^2-4ac)}_{\text{irrep of type } (2,2)} \oplus \underbrace{\mathrm{span}(\quad)}_{\text{irrep of type } (4,0)}$

$$\left.\begin{aligned} &\mathrm{mult}_{(2,2)}(\mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2))) = 1 \\ &\mathrm{mult}_{(4,0)}(\mathrm{Sym}^2(\mathrm{Sym}^2(\mathbb{C}^2))) = 1 \\ &\mathrm{mult}_{(3,1)}(\quad '' \quad) = 0 \end{aligned}\right\} \text{Called plethysm coefficients}$$

multiplicity obstructions can circumvent having to compute hwvs

☆ In GCT, we take the co-ordinate rings of the orbit closures.

$$\mathbb{C}[X] = \frac{\mathbb{C}[\mathrm{Sym}^2(\mathbb{C}^2)]}{I(X)}$$

$$\mathrm{mult}_\lambda(\mathbb{C}[X]_d) = \underbrace{\mathrm{mult}_\lambda(\mathbb{C}[\mathrm{Sym}^2(\mathbb{C}^2)]_d)}_{\text{Plethysm coefts}} - \mathrm{mult}_\lambda(I(X)_d)$$

The co-ordinate rings are interesting :- (the co-ord. ring of the orbit closure has a convenient localization that can be studied using the algebraic Peter-Weyl thm)

The above was a mix of examples, heuristics, etc. to show you what GCT aims to do. Let's get more concrete

1. Understand representations in co-ordinate rings
2. Partial stability (generalizes notion of stability in Geom. inv. theory)
3. PS leads us to the symmetries of perm & det and the fact that they are characterized by their symmetries
4. Natural proofs barrier & actual P vs NP in GCT

Recall: looked at the set of polys of wt ≤ 1. We wanted a separating ... with

Prev. we looked at the set of polys of $wx \le 1$. We wanted a separating polynomial, and we saw that obtaining one had something to do with representation theory. Let's do reverse – start with representations in the co-ord. ring and see where that takes us.

e.g. b^2-4ac is invariant under SL_2 action on $Sym^2(\mathbb{C}^2)$

$b^2-4ac = 0 \iff ax^2+bxy+c^2$ is a perfect square $\equiv wx = 1$

(*) $\text{span}\{b^2-4ac\} \subseteq Sym^2(Sym^2(\mathbb{C}^2))$ is trivial rep of SL_2

We want to look at other reps in these co-ord rings generally ...

Claim G rep V in the co-ordinate ring $\longleftrightarrow$ G. invariant property (Zariski closed)

in our e.g. b^2-4ac is invariant, so the set of pts where it vanishes in G-inv.
This claim says if a set itself is G-inv (here the set is V), then the set of pts where all these pts. vanish is going to be G.invariant

---

Let us look at a more complicated example than b^2-4ac.

e.g. $GL_n \times GL_n$ acting on M_n

$(A,B) \cdot X = AXB^T$ (left action)

Fact:- Orbits are completely determined by rank. This is an example where we know a lot abt. going in...

$X = g.Y \iff \text{rank } X = \text{rank } Y$

(*) Invariant ring is $\mathbb{C}[\det]$ (one invariant)

defn: for a G-rep V, $N_G(V) \subseteq V$ pts. s.t. orbit closure contains zero

(*) The null cone is the set of singular matrices (det vanishes on them)

(*) So this invariant by itself does not separate orbits

(*) But we have an invariant to separate orbits, i.e rank

(*) Rank per se is not a polynomial invariant

... correspondence, there

⊛ Rank per se is not a polynomial invariant

⊛ It is however a G-invariant, and by our correspondence, there should be some G-invariant subspaces in our co-ordinate ring that this corresponds to

⊛ What are those? MINORS...

$\{X \mid \text{rank } X < r\}$ is closed and is the $Z(r \times r \text{ minors of } X)$

⊛ Thus the span of $r \times r$ minors is G-representation

Abstractly, if our original space was $V \otimes W$ acted on by $GL(V) \times GL(W)$

the G-rep would be written as

$$\Lambda^r V \otimes \Lambda^r W$$

Our co-ord ring here is $\mathbb{C}[X_{ij}]$

⋆ $\Lambda^r V \otimes \Lambda^r W$ sits inside $\mathbb{C}[X_{ij}]$ and we can use vanishing or non vanishing to EXACTLY DISTINGUISH ORBITS

Thus:

X, Y in the same orbit $\iff$ set of r s.t $\Lambda^r V \otimes \Lambda^r W$ vanish on X is the same as that for Y

— Pts in the null cone are of the form

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix}$$

— permanent_m sits inside a space with more than m^2 vars ... i.e P can depends only on a few variables

Vaguely, these facts are analogous
⋮
This leads to the notion of Partial stability

⊛ "non-full rank are in the null cone, and this is undesirable"

⊛ In GCT ... similar thing happens ... the smaller permanent sits

(*) "non-full rank are in the null cone, and this is undesirable

(*) In GCT, a similar thing happens ... the smaller permanent sits inside a larger space, and it will also be in the null cone

Defn. [Stable] x stable $\iff G.x$ is closed

$\xRightarrow{\text{Matsushima}}$ $\mathrm{Stab}_G\, x$ is a reductive subgroup.

we like b/c we can talk abt their rep theory, multiplicities, etc.

Defn [Partial Stability] $[x] \in \mathbb{P}V$ is partially stable if $\exists$ maximal parabolic subgroup P of G (for GL_n it is diagonal matrices like $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$) s.t $P \supseteq \mathrm{Stab}_G(x)$ and

① $\mathrm{Stab}_G(x) \supseteq R \begin{pmatrix} 0 & * \\ & 0 \end{pmatrix}$ [Called unipotent radical of P]

② $[x]$ is stable w.r.t $K \subseteq \begin{pmatrix} * & 0 \\ & * \end{pmatrix} = L \leftarrow$ Called "Levi Subgroup" of G

where K is reductive & $\mathrm{rank}\, K = \mathrm{rank}\, L - \delta \leftarrow$ think of this as small

& $\mathrm{Stab}_L(x)$ is reductive

③ $L.x \cong L/\mathrm{Stab}_L(x)$ is not too small relative to $G.x$

$$\frac{\dim L.x}{\dim G.x} \geq \Delta$$

Δ & δ are parameters

Remark if $P = G$, $\delta = 0$, $\Delta = 1$, this is equivalent to Stability

The pf. is if we allow "slacks" of δ & Δ, then

(*) The pt. is if we allow "slacks" of δ & Δ, then we can handle padded permanent even though it is in the null cone

(*) Partial stability generalizes the notion of stability, and thus refines information within the null cone.

(*) Partially stable pts with δ small & Δ close to 1 are very nice pts in the null cone

Let us illustrate this notion of partially stable on matrices with rank.

Recall: $GL(V) \otimes GL(W)$ acting on $V \otimes W$

$$(A, B) . X = A X B^T$$

→ here all matrices of non-full rank are in the null cone

→ lower the rank, worse δ, Δ get (partial stability "quantifies" how "deep" your pts are in the null cone)

Q: what is Stabilizer of $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ ← recall all pts of rank r are G-equiv to this

Ans Action looks like this

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} B_{11}^T & B_{12}^T \\ B_{21}^T & B_{22}^T \end{pmatrix}$$

We get $A_{11} B_{11}^T = I_r$, $A_{11} B_{21}^T = A_{21} B_{11}^T = A_{21} B_{21}^T = 0$

Thus A_{11} & B_{11}^T are invertible, thus

$$A_{21} = B_{21} = 0$$

A_{21} $\quad$ A_{21}

Thus stabilizer is

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \begin{pmatrix} A_{11}^{-T} & B_{12} \\ 0 & B_{22} \end{pmatrix}$$

both parabolic! This will be the P for the Partial stability condition

✓ Cond 1 of partial stability is satisfied since A_{11} & B_{12} are arbitrary
(b'coz we are taking transpose)

✓ For Cond 2, take action of $\begin{pmatrix} GL_r & \\ & GL_{n-r} \end{pmatrix}$ on $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ is actually stable under action of

$$\begin{pmatrix} SL_r & 0 \\ 0 & GL_{n-r} \end{pmatrix} \times \begin{pmatrix} SL_r & 0 \\ 0 & GL_{n-r} \end{pmatrix}$$

which is reductive & Rank is one less than

R ↗

$(S = 2)$

$$\begin{pmatrix} GL_r & \\ & GL_{n-r} \end{pmatrix}$$

inverse of other SL_r, so not free

L ↗ (Levi subgroup)

effectively, stabilizer is $\begin{pmatrix} \overset{A_{11}}{SL_r} & \\ & GL_{n-r} \end{pmatrix} \times \begin{pmatrix} \boxed{}^{A_{11}^{-T}} & \\ & GL_{n-r} \end{pmatrix}$

$\dim L / \mathrm{stab}_L(x) = r^2 - 1$ ← decreasing as r gets smaller

✓ Cond 3

Smaller

✓ Cond 3

UPCOMING

* Notion of partial stability leads us to symmetries.
* GCT for P vs NP

Natural Proofs [Razborov-Rudich 1990's]

- People were looking at resolving P vs NP by showing $NP \not\subseteq P/poly$, a harder problem. Because circuits are "nicer" than Turing machines
- Raz-Rud. realized that all circuit l.b. techniques have some similarities. They called such proofs "natural"
- There cannot be a natural proof that $NP \not\subseteq P/poly$

Defn A natural property of Boolean functions is a subset $C_n \subseteq \text{funcs}(\{0,1\}^n \to \{0,1\})$ that is

1. large
2. Constructive (deciding if a func belongs to C_n can be done efficiently)
3. useful against P/poly
   (if you have a sequence of funcs $f_1, f_2, \ldots$ where $f_n \in C_n$, then $(f_i) \notin P/poly$)

Similar to separating polynomials

So a proof of $NP \subseteq P/Poly$ necessarily HAS to violate one of these properties.

- Meaningless to expect to violate (3)
- There is evidence that you can't really violate (2)

Intuition :- A lower bound that is general, i.e., one that would work for ... be used for $NP \not\subseteq P/poly$.

Intuition :- A lower bound that is general, i.e., one that would work for "randomly" chosen functions cannot be used for NP $\not\subseteq$ P/poly.
In other words one lower bound technique should crucially depend upon the function for which we are trying to prove a lower bound.

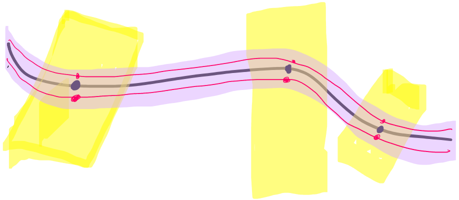
Stability & Partial Stability

Thm [Luna's étale slice thm]. Let G act on V & let $x \in V$ be stable.
Also $H := \mathrm{Stab}_G(x)$.

A neighbourhood of the orbit Gx is (almost) isomorphic to $G \times_H N$

direct product, but "h" can move $(g, hn) \sim (hg, n)$ freely

normal vector space to the orbit



This means that if there is a pt y in the infinitesimal nbhd of Gx, then $\exists\, y' \in G.y$ s.t $\mathrm{Stab}_G\, y \subseteq \mathrm{Stab}_G\, x$

So if you want to separate orbit closures, it is sufficient to look at stabilizers

Since x is stable, orbit is closed

Pun [Abhiram]: To understand a theorem abt. groups, see it in action.

Claim $\mathrm{perm}_n \notin \overline{GL_n \circ \det_n}$. Also perm_n is "far" from the orbit of $\det_n$

FIRST PART :-

① $\mathrm{stab}_G(\mathrm{perm}_n) \sim (\mathbb{C}^\times)^{2n-1} \rtimes (S_n \times S_n) \rtimes \mathbb{Z}_2$ ← transpose

scale columns, rows, but the product of all scaling factors must be 1

permute rows and/or columns

② $\mathrm{Stab}_G(\det_n) \sim S(GL_n \times GL_n) \rtimes \mathbb{Z}_2$ ← transpose

(A, B) s.t $\det(A) \times \det(B) = 1$

③ Suppose $\exists\, g$ s.t $\mathrm{perm}_n = g \circ \det_n$.

③ Suppose $\exists g$ s.t $\quad perm_n = g \cdot det_n$.

$stab_G(perm_n) = g\, stab_G(det_n)\, g^{-1} \Rightarrow$ stabs are conjugate

$\Rightarrow$ stabilizers need to be isomorphic

$\dim (\mathbb{C}^\times)^{2n-1} = 2n-1 \quad , \quad \dim S(GL_n \times GL_n) = 2n^2-1$

So stabs are not even isomorphic. Contradiction

SECOND PART:

By contradiction, Luna's étale slice theorem entails that the $stab(det_n)$ is a conjugate to a subgroup of $stab(perm)$.

Impossible bcz of dimensions. $\boxtimes$

e.g. $x = z\, perm_m$

$y = det_n$

(If the padded perm was stable, the above would tell us that some conjugate of the stabilizer of the determinant is contained in the stabilizer of the permanent)

→ Stability implies higher cohomology vanishes.

Partial stability gives you a weaker LES thm.

– If x is partially stable, Gx is now a fiber bundle over a Grassmanian with affine fibers

$G/P \qquad\qquad L/L'$

"The orbit $G.x$ looks like this:- for every d-dim subspace (elems of the Grassmanian), you can pick a vector in L/L' and you consequently get a point in the orbit.

Also, the association from d-dim subspaces to vectors is 'continuous'.

→ Partial stability implies cohomology of $G.x$ is essentially the same as that of the Grassmanian

Partially stable implies orbit has a nice structure

... $G.x$ (need not be Zariski closed)

Partially stable ...

Ring of regular functions on the orbit $G.x$ (need not be Zariski closed)
↳ locally are ratios of polynomials

$\mathbb{C}[G]^{Stab_G(x)}$ ← all stuff inside the co-ord ring that is invariant under $Stab_G(x)$
↳ Co-ord ring of the group

again stabilizer shows up

Intuition Stability/Partial Stability ⟹ Stabilizer is important when looking at orbits

⊛ Most polynomials have trivial stabilizer
→ So we can be optimistic that GCT avoids the natural proofs barrier

⊛ Even more so, the permanent & determinant have the property of being "characterized by symmetries" which is a stronger property.

Defn [Characterization by symmetries] G acts on V. $v \in V$ is char. by its symm. if

$$\forall v' \text{ s.t. } Stab_G(v') \geq Stab_G(v) \implies v' = \lambda v$$

⊛ Here it means that if you know $Stab_G(v)$, you don't need to know anything else.

Thm Both determinant & permanent are characterized by their symmetries :-

(a) $\det(X) = \det(AXB)$ when $\det(A) = 1/\det(B)$.

(b) $\mathrm{per}(X) = \mathrm{per}(PXQ)$ when P, Q are permutation matrices
$\mathrm{Per}(X) = \mathrm{per}(AXB)$ when A, B are diagonal with $\mathrm{per}(A) = 1/\mathrm{per}(B)$

This is a very strong property:
① Gives us a way to get around the natural proofs barrier
② Stability, Partial Stability, Luna's ---, etc. all make it clear that there is some sort of duality between orbits and stabilizers.

Here there is a very strong correspondence, so there is justification to belive that the perm & det have "special" orbits justification to belive that The [

orbits

③ Caveat:- All heuristic justification, no theorem . . . .

A sufficient condition for separating orbit closures

For a group G and two reps V & W, let $\mathrm{Hom}_G(V, W)$ denote

Defn $\phi : V \to W$ s.t. $\forall g \in G$

$$\begin{tikzcd} V \arrow[r, "g"] \arrow[d, "\phi"'] & V \arrow[d, "\phi"] \\ W \arrow[r, "g"] & W \end{tikzcd}$$

Commutes.

Let $Z \subseteq \mathbb{C}^n$ be an algebraic set such that $I(Z)$ is a homogeneous ideal.

Thus $\mathbb{C}[Z] = K[x_1 \ldots x_n] / I(Z)$ has a grading via

$$\mathbb{C}[Z]_\delta = K[x_1 \ldots x_n]_\delta \Big/ I(Z)_\delta$$

GCT:- $A := \overline{GL(V) \circ z^{m-n}\,\mathrm{perm}_n}$, $A' := \overline{GL(V) \circ \det_m}$. If it is true that for some m $A \subseteq A'$, then $I(A') \supseteq I(A)$; and consequently, $I(A')_\delta \supseteq I(A)_\delta$.

This gives us a canonical $GL(V)$-equivariant surjection $\mathbb{C}[A']_\delta \twoheadrightarrow \mathbb{C}[A]_\delta$.

$\mathbb{C}[A]_\delta$ & $\mathbb{C}[A']_\delta$ are obviously $GL(V)$-representations :-

$g \in GL(V)$ acts on $p(q(x))$

$$g \circ p(q(x)) = p\left(q\left(g^{T} x\right)\right)$$

Thus we have a $GL(V)$-equivariant surjection of representations here

Defn For an irrep ρ of $GL(V)$, and another G-representation W, denote $\mathrm{mult}_\rho(W) = \dim \mathrm{Hom}_{GL(V)}(\rho, W)$, where in turn, $\dim \mathrm{Hom}_{GL(V)}(\rho, W)$

Can be calculated as follows

$\ldots \oplus U_i$ be a decomposition into irreducibles.

Can be calculated as follows

Let $W = U_1 \oplus \dots \oplus U_t$ be a decomposition into irreducibles.

Then $\dim \operatorname{Hom}_{GL(V)}(\rho, W) = \left| \{ i \mid U_i \cong \rho \} \right|$

By Schur's lemma, we have in the above case that for all irreps ρ of $GL(V)$, we would have $\operatorname{mult}_\rho(\mathbb{C}[A]) \leq \operatorname{mult}_\rho(\mathbb{C}[A'])$.

Of course this would be if the orbit closure of the padded permanent was indeed contained in the orbit closure of the determinant, so

Thm Suppose there is an irrep ρ of $GL(V)$ s.t. $\operatorname{mult}_\rho(\mathbb{C}[A]) > \operatorname{mult}_\rho(\mathbb{C}[A'])$, then $y^{m-n}\operatorname{perm}_n \notin \overline{GL(V).\det_n}$

☆ Why should I look at this language of multiplicities? We gave evidence the orbits of det & perm are special, so alg. geom of the varieties is determined by the representation theory. Thus look at multiplicities

☆ Furthermore, there are algorithms to calculate multiplicities, specifically the Littlewood-Richardson coefficients

☆ So MS suggest to explicitly work on algorithms to calculate multiplicities to constructively find an irrep that gives you an obstruction, i.e. an explicit poly that vanishes on $\overline{GL(V).\det_m}$, but not on $y^{m-n}\operatorname{perm}_n$.

They call this "the flip", i.e. getting lower bounds via upper bounds.

☆ MS give an algorithm to test the positivity of Littlewood-Richardson coefficients. Alg. is "easier" and uses linear programming. So they hoped:

Conjecture [Initially by MS] For any poly $p(m)$ $\exists$ infinitely many m and $n \geq p(m)$ s.t. $\operatorname{mult}_\rho(\mathbb{C}[A]) > \operatorname{mult}_\lambda(\mathbb{C}[A']) = 0$

Conjecture was falsified by BIP.

Proof path:- ① By contradiction, assume there exists an irrep with ... and >0 in the perm case.

Proof path:- ① By contradiction, assume the exists an irrep with mult = 0 in the determinant case and > 0 in the perm case. FOR A FIXED m & n

② They do induction to find "obstructing" irreps for higher and higher n, eventually violating Granet's upper bound on dc($perm_m$), i.e. $2^m - 1$. This gives the contradiction required. ☒

Toward P vs NP in GCT

→ In principle, you can ask these sort of orbit closure separation questions for any complexity separations. The alg. geom approach should work for <u>anything</u>.

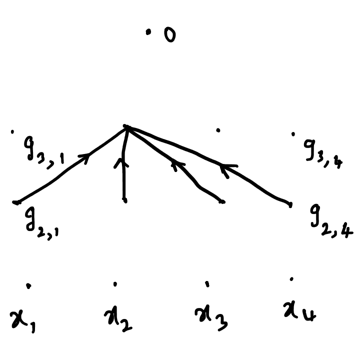
→ But your best hope is to choose "nice" functions.

→ These algebraic techniques transfer nicely to finite fields, although rep. theory over finite field is nasty.

✱ They propose candidate functions for P/poly & NP

For P/poly. 'H' which is P complete. 'E' ∈ NP, but not known to be complete

H: – Layered circuit 'n' inputs, 'n' levels, 'n' gates on all levels except last.



$$h_{\ell,i} = \sum_{j,k\in[n]} \alpha_{\ell,i,j,k}\, h_{\ell-1,j}\, h_{\ell-1,k}$$

weights

Polynomials

H_n → function computed at o/p vertex

<u>Thm</u> H_n is P/poly-Complete over any finite field

Proof Set α's to "anything" ☒

Thm ... ~ ... poly

Proof Set α's to anything $\boxtimes$

Since we are over any finite field, it is really a boolean complexity class! This is whilst still being algebraic!

E (for NP)

Let X_0 & X_1 be two $n \times n$ matrices over the finite field $\mathbb{F}_q$.

For any $s = (s_1, \ldots s_n) \in \{0,1\}^n$, define X_s to be the matrix whose i^{th} column is the i^{th} column of X_{s_i}. Define $E(X) = \prod_{s \in \{0,1\}^n} X_s$.

Thm Over $\mathbb{F}_q$, $E(X)^{q-1}$ is $\{0, 1\}$-valued. and in NP.

Thm $E(X)$ & $H(X)$ are characterized by their symmetries.

E: highly; H: almost

If someone shows $E(X)$ cannot be computed by H after setting any weights, then $P/poly \not\supseteq NP$.

Conjecture No function that is a projection of $E(X)$ is in the orbit closure of $H(Y)$

But alg. geom/rep. theory is harder over finite fields

But at least we can make the statements

Miscellaneous Topics

Sunday, 18 June 2023 01:33

Hardness of approximation and the Unique Games Conjecture

Let $\mathcal{I}$ be an optimization problem & let I be an instance of $\mathcal{I}$ of size N. Let OPT(I) be the value of the optimal soln and let A be an alg. that is an "approximation" algorithm. A returns A(I).

⊛ if $\mathcal{I}$ is a maximization problem, then

$$A(I) \leq OPT(I).$$

Suppose

$$\forall I, \quad A(I) \geq \alpha \; OPT(I) \quad (\alpha < 1),$$

then we say A achieves an approx factor of α

⊛ If $\mathcal{I}$ is a min prob, and

$$A(I) \leq \beta \; OPT(I) \quad (\beta > 1)$$

Unique Games Conjecture:-

$$\left.\begin{aligned} x_1 &\equiv 2.x_2 \pmod 7 \\ x_2 &\equiv 4.x_5 \pmod 7 \\ &\vdots \\ x_1 &= 2.x_7 \pmod 7 \end{aligned}\right\} \text{System of linear eqns over a finite field}$$

Conjecture Even if a large fraction of a system of linear equations over a finite field is satisfiable, it is NP-hard to even certify that at least a small fraction of them is satisfiable.

This is equivalent to a flavour of the following problem

This is equivalent to a flavour of the following problem

[1-homology localization] Given a simplicial complex / combinatorial CW complex X, and a 1 cycle $a \in Z_1(X, A)$, determine the sparsest homologous representative of a, i.e. a 1-cycle a' that is homologous (a & a' differ by the boundary of a 2-cycle) to a with min possible support.

If UGC / 1-hom.loc is true, then

Problem	Poly.-time approx.	UG hardness
Max 2SAT	$0.940\ldots$[5]	$0.9439\cdots+\varepsilon$[7]
Max cut	$0.878\ldots\ldots$[8]	$0.878\cdots+\varepsilon$[7]
Min vertex cover	2	$2-\varepsilon$[10]
Feedback arc set	$O(\log n \log\log n)$[11]	All constants[13]
Max acyclic subgraph	$\frac{1}{2}+\Omega(1/\sqrt{\Delta})$[14]	$\frac{1}{2}+\varepsilon$[13]
Betweenness	$\frac{1}{3}$	$\frac{1}{3}+\varepsilon$[17]