Geometric Complexity Theory

Milind Sohoni[1]

An approach to complexity theory
via
Geometric Invariant Theory.

Research Institute for Mathematical Sciences
Kyoto University

[1] joint work with Ketan Mulmuley

Talk Outline

- Mainly Valiant
- Mainly stability and obstructions
- Mainly Representations
- Largely hard

The satisfiability problem

- Boolean variables $x_1, \ldots, x_n$
- Term $t_1 = (\neg x_1 \vee x_3 \vee x_7)$, and so on upto t_m.
- Formula $t_1 \wedge t_2 \wedge \ldots \wedge t_m$

Question: Decide if there is a satisfying assignment to the formula.

No known algorithm which works in time polynomial in n and m.

- The problem belongs to an equivalence class called **NP-complete** problems.
- The question of **P v. NP** asks:
  - Either produce an efficient algorithm.
  - Or prove none exists.
- **This has been an outstanding question for the last 50 years.**

Decision vs. Counting

Equivalence: Solve One $\Leftrightarrow$ Solve All
Unsolvable One $\Leftrightarrow$ Unsolvable All

Many relatives of P v. NP. We look at the *counting version*.

- Boolean variables $x_1, \ldots, x_n$
- Term $t_1 = (\neg x_1 \vee x_3 \vee x_7)$, and so on upto t_m.
- Formula $t_1 \wedge t_2 \wedge \ldots \wedge t_m$

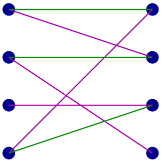
Question: Decide if there is a satisfying assignment to the formula.

Harder Question: Count the number of satisfying assignments.
Thus we have the decision problem and its counting version.

Matchings

Question: Given a bipartite graph on n, n vertices, check if the graph has a complete matching.

This problem has a known polynomial time algorithm.



Harder Question: Count the number of complete matchings.

- There is no known polynomial time algorithm to compute this number.
- Even worse, there is no proof of its non-existence.

Thus, there are decision problems whose counting versions are hard.

The permanent

If X is an $n \times n$ matrix, then the permanent function is:

$$perm_n(X) = \sum_{\sigma} \prod_{i} x_{i,\sigma(i)}$$

The relationship with the matching problem is obvious. When X is 0-1 matrix representing the bipartite graph, then $perm(X)$ counts the number of matchings.

- There is no known polynomial time algorithm to compute the permanent, and worse, no proof of its non-existence.
- The function $perm_n$ is $\#P$-complete. In other words, it is the hardest counting problem whose decision version is easy to solve.

Our Thesis

- Non-existence of algorithm $\implies$ existence of a mathematical structure (obstructions)
- These happen to arise in the GIT and Representation Theory of Orbits.

Example

- Hilbert Nullstellensatz : Either polynomials $f_1, \ldots, f_n$ have a common zero, or there are $g_1, \ldots, g_n$ such that

$$f_1 g_1 + \ldots + f_n g_n = 1$$

- Thus $g_1, \ldots, g_n$ obstruct $f_1, \ldots, f_n$ from having a common zero.

Computation Model-Formula Size

Let $p(X_1, \ldots, X_n)$ be a polynomial.
A formula is a particular way of writing it using $*$ and $+$.

```
formula = formula*formula | formula+formula
```

- Thus the same function may have different ways of writing it.
- The number of operations required may be different.

Example:

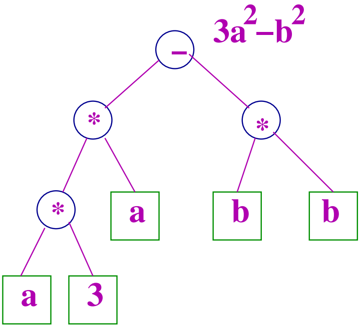
- $a^3 - b^3 = (a-b)(a^2 + a * b + b^2)$.
- Van-der-Monde $(\lambda_1, \ldots, \lambda_n) = \prod_{i \neq j}(\lambda_i - \lambda_j)$.

Formula size: the number of $*$ and $+$ operations.

- LHS1 is 5, RHS1 is 7, RHS2 is n^2.

Formula size

- A formula gives a formula tree.
- This tree yields an algorithm which takes time proportional to *formula size*.



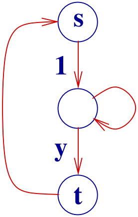
Does $perm_n$ have a formula of size polynomially bounded in n?
(This also implies a polynomial time algorithm) No Answer

Valiant's construction: converts the tree into a determinant.

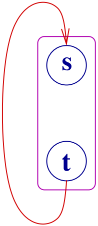
Valiant's Construction

If $p(Y_1, \ldots, Y_k)$ has a formula of size $m/2$ then,

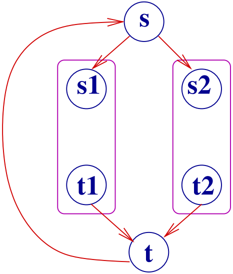
- There is an inductively constructed graph G_p with atmost m nodes, with edge-labels as (i) constants, or (ii) variable Y_i.
- The determinant $det(A_p)$ of the adjacency matrix of G_p equals p.



A simple formula.



The general case.



Addition

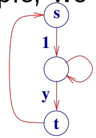
The Matrix

In other words:

$$p(Y_1, \ldots, Y_k) = det_m(A)$$

where $A_{ij}(Y)$ is a degree-1 expression on Y.

For our example, we have:



$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & y \\ 1 & 0 & 0 \end{bmatrix} \quad det(A) = y$$

- Note that in Valiant's construction $A_{ij} = Y_r$ or $A_{ij} = c$.

$$\text{formula size} = m/2 \implies p(Y) = det_m(A)$$

The homogenization

Lets homogenize the above construction:

- Add an extra variable Y_0.
- Let $p^m(Y_0, \ldots, Y_k)$ be the degree-m homogenization of p.
- Homogenize the A_{ij} using Y_0 to A'_{ij}.

We then have: $p^m(Y_0, \ldots, Y_k) = det_m(A')$

For our small example:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & y \\ 1 & 0 & 0 \end{bmatrix} \quad A' = \begin{bmatrix} 0 & y_0 & 0 \\ 0 & y_0 & y \\ y_0 & 0 & 0 \end{bmatrix} \quad det(A') = y_0^2 y$$

Valiant-conclusion

If a form $p(Y)$ has a formula of size $m/2$ then

- There is an $m \times m$-matrix A with linear entries

$$det(A) = p(Y)$$

- There is an $m \times m$-matrix A' with homogeneous linear entries

$$det(A') = p^m(Y)$$

  where p^m is the m-homogenization of p.

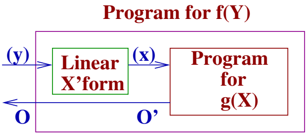
The $\preceq_{hom}$

Let $X = \{X_1, \ldots, X_r\}$.

For two form $f, g \in Sym^d(X)$, we say that $f \preceq_{hom} g$, if $f(X) = g(B \cdot X))$ where B is a fixed $r \times r$-matrix.

Note that:

- B may even be singular.
- $\preceq_{hom}$ is transitive.

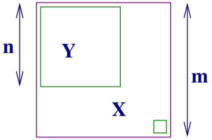


If there is an efficient algorithm to compute g then we have such for f as well.

The insertion

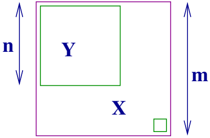
Suppose that $perm_n(Y)$ has a formula of size $m/2$. How is one to interpret Valiant's construction?

- Let Y be $n \times n$.
- Build a large $m \times m$-matrix X.
- Identify Y as its submatrix.



The "inserted" permanent

For $m > n$, we construct a new function $perm_n^m \in Sym^m(X)$.

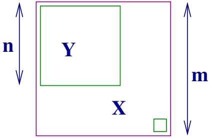


- Let Y be the principal $n \times n$-matrix of X.
- $perm_n^m = x_{mm}^{m-n} perm_n(Y)$

Thus $perm_n$ has been inserted into $Sym^m(X)$, of which $det_m(X)$ is a special element.

The "inserted" permanent

For $m > n$, we construct a new function $perm_n^m \in Sym^m(X)$.



- Let Y be the principal $n \times n$-matrix of X.
- $perm_n^m = x_{mm}^{m-n} perm_n(Y)$

Thus $perm_n$ has been inserted into $Sym^m(X)$, of which $det_m(X)$ is a special element.

- formula of size $m/2$ implies $perm_n = det_m(A)$
- Use x_{mm} as the homogenizing variable

Conclusion
$perm_n^m = det_m(A')$

$perm_n^m \preceq_{hom} det_m$

Group Action and $\preceq_{hom}$

Let $V = Sym^m(X)$. The group $GL(X)$ acts on V as follows. For $T \in GL(X)$ and $g \in V$

$$g_T(X) = g(T^{-1}X)$$

Two notions:

- The orbit: $O(g) = \{g_T | T \in SL(X)\}$.
- The projective orbit closure $\Delta(g) = \overline{cone(O(g))}$.

If $f \preceq_{hom} g$ then $f = g(B \cdot X)$, whence

- If B is full rank then f is in the $GL(X)$-orbit of g.
- If not, then B is approximated by elements of $GL(X)$.

Thus, in either case,

$$f \preceq_{hom} g \implies f \in \Delta(g)$$

The Δ

- Thus, we see that if $perm_n$ has a formula of size $m/2$ then $perm_n^m \in \Delta(det_m)$.

- On the other hand, $perm_n^m \in \Delta(det_m)$ implies that for every $\epsilon > 0$, there is a $T \in GL(X)$ such that $\|(det_m)_T - perm_n^m\| < \epsilon$. This yields a poly-time approximation algorithm for the permanent

Thus, we have an almost faithful algebraization of the formula size construction.

The Obstruction and its existence

To show that $perm_5$ has no formula of size $20/2$, it suffices to show:

$$perm_5^{20} \notin \Delta(det_{20})$$

In other words:

- V is a $GL(X)$-module.
- f and g are special points.
- What is the witness to $f \notin \Delta(g)$?

It is clear that such witnesses or **obstructions** exist in the coordinate ring $k[V]$.
Real Question: How do I find this family and prove that it is indeed so.

What is the structure of such obstructions?

The Obstruction

So let $g, f \in V = Sym^d(X)$. How do we show that $f \notin \Delta(g)$.

- Exhibit a homogeneous polynomial $\mu \in Sym^r(V^*)$ which vanishes on $\Delta(g)$ but not on f.

This μ is then the required **obstruction**. We would need to show that:

- $\mu(f) \neq 0$.
- $\mu(g_T) = 0$ for all $T \in SL(X)$.

Check μ on every point of $Orbit(g)$

False start: Use the $SL(X)$-invariant elements of $Sym^r(V^*)$ for constructing such a μ.

Invariants

- V is a space with a group G acting on V.
- $Orbit(v) = \{g.v | G \in G\}$.
- Invariant is a function $\mu : V \to \mathbb{C}$ which is constant on orbits.

Existence and constructions of invariants has been an enduring interest for over 150 years.

Example:

- V is the space of all $m \times m$-matrices.
- $G = GL_m$ and $g.v = gvg^{-1}$.
- Invariants are the coefficients of the characteristic polynomial.

Invariants and orbit separation

To show that $f \notin \Delta(g)$

Exhibit a homogeneous *invariant* μ which vanishes on g but not on f. This μ would then be the desired obstruction.

- Easy to check if a form is an invariant.
- Easy to construct using age-old recipes.
- Easy to check that $\mu(g) = 0$ and $\mu(f) \neq 0$.

$$\mu(g) = 0 \implies \mu(g_T) = 0 \implies \mu(\Delta(g)) = 0$$

Important Fact

If g and f are stable and $f \notin \Delta(g)$, then there is a homogeneous invariant μ such that $\mu(g) \neq \mu(f)$.

Stability

- g is stable iff $SL(X)$-$Orbit(g)$ is Zariski-closed in V.
- Most polynomials are stable.
- It is difficult to show that a particular form is stable.

Hilbert : Classification of unstable points.

- For matrices under conjugation, precisely the diagonalizable matrices are stable.

$perm_m$ and det_n are stable.

Proof:

- Kempf's criteria.
- Based on the stabilizers of the determinant and permanent.

Rich Stabilizers

The stabilizer of the determinant:

- The form $det_m(X)$:
  - $X \to AXB$
  - $X \to X^T$
- $det_m \in Sym^m(X)$ determined by its stabilizer.

The stabilizer of the permanent:

- The form $perm_m(X)$:
  - $X \to PXQ$
  - $X \to D_1 X D_2$
  - $X \to X^T$
- $perm_m \in Sym^m(X)$ determined by its stabilizer.

Tempting to conclude that the homogeneous obstructing invariant μ now exists.

The Main Problem

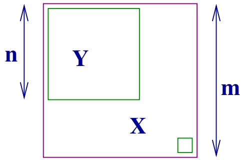
Recall we wish to show

$$perm_n^m \notin \Delta(det_m)$$

where
$perm_n^m = x_{mm}^{m-n} perm(Y)$.

$perm_n^m$ is unstable, in fact in the null-cone, for very trivial reasons.

- Added an extra degree equalizing variable.
- Treated as a polynomial in a larger redundant set of variables.



Two Questions

- Thus every invariant μ will vanish on $perm_n^m$.
- There is no invariant μ such that $\mu(det_m) = 0$ and $\mu(perm_n^m) \neq 0$.

Homogeneous invariants will never serve as obstructions. They dont even enter the null-cone

Two Questions:

- Is there any other system of functions which vanish on $\Delta(det_m)$?
- Can anything be retrieved from the superficial instability of $perm_n^m$?

Part II

- Is there any other system of functions which vanish on $\Delta(det_m)$?
  Yes. The admissibility argument.

- Can anything be retrieved from the superficial instability of $perm_n^m$?
  Yes. Partial or parabolic stability.

Two key focal points:

- **Representations as obstructions**
- **Stabilizers**

Question 1

Is there any other system of functions which vanish on $\Delta(det_m)$ *and* enter the null-cone?

- We use the stabilizer $H \subseteq SL(X)$ of det_m.
- For a representation V_λ of $SL(X)$, we say that V_λ is H-admissible iff $V_\lambda^*|_H$ contains the trivial representation 1_H.

For g stable:

Fact: $k[Orbit(g)] \cong k[G/H] \cong \sum_{\lambda\ H\text{-admissible}} n_\lambda V_\lambda$

Thankfully: $k[\Delta(g)] \cong \sum_{\lambda\ H\text{-admissible}} m_\lambda V_\lambda$

Thus a fairly restricted class of G-modules will appear in $k[\Delta(g)]$. We use this to generate some elements of the ideal for $\Delta(g)$.

G and H

Consider next the G-equivariant surjection:

$$\phi : k[V] \longrightarrow k[\Delta(g)]$$

We see that (i) ϕ is a graded surjection, and (ii) if $V_\mu \subseteq k[V]^d$ is not H-admissible, then $V_\mu \in ker(\phi)$.

Let Σ_H be the ideal generated by such V_μ within $k[V]$.
Clearly Σ_H vanishes on $\Delta(g)$.

How good is Σ_H?

The Local Picture

G-separability: We say that $H \subseteq G$ is G-separable, if for every non-trivial H-module W_α such that:

- W_α appears in some restriction $V_\lambda|_H$.

then there exists a H-non-admissible V_μ such that $V_\mu|_H$ contains W_α.

Theorem: Let g and H be as above, with (i) g stable, (ii) g only vector in V with stabilizer H, and (iii) H is G-separable. Then for an open subset U of V, $U \cap \Delta(g)$ matches $(k[V]/\Sigma_H)_U$.

Applying this ...

The conditions: (i) stability of g, (ii) $V^H = < g >$ and (iii) G-separability of H.

- det_m and $perm_n$ satisfy conditions (i) and (ii) above.
- For $n = 2$, stabilizer of det_2 is indeed SL_4-separable.
- For $V = \bigwedge^d$ and g the highest weight vector, $\Delta(g)$ is the grassmanian. For this Σ_H generates the ideal.
- For $g = det_m$, the data Σ_H does indeed *enter* the null-cone.

Still open:

- Look at $H = SL_n \times SL_n$ sitting inside $G = SL_{n^2}$. Is H G-separable?
- Does Σ_H determine $\Delta(det_m)$?

To conclude on Question 1

- Stabilizer yields a rich set Σ_H of relations vanishing on $\Delta(det_m)$.
- Given G-separability, Σ_H does determine $\Delta(det_m)$ locally.

Now suppose that $perm_n^m \in \Delta(det_m)$ then:

- Look at the surjection $k[\Delta(det_m)] \rightarrow k[\Delta(perm_n^m)]$.

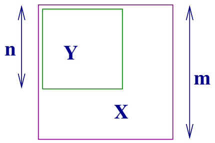
$V_\mu \subseteq k[\Delta(perm_n^m)]$ and V_μ non-H-admissible, then V_μ is the required obstruction.

If $k[\Delta(perm_n^m)]$ is understood then this sets up the representation-theoretic obstruction.

Question 2-Partial Stability

Can anything be retrieved from the superficial instability of $perm_n^m$?

- Let's consider the simpler function $f = perm(Y) \in Sym^n(X)$, i.e., with useful variables Y and useless $X - Y$.
- Let parabolic $P \subseteq GL(X)$ fix Y.
- $P = LU$, with U the unipotent radical.



We see that:

- f is fixed by U.
- f is L-stable.

The form f

Recall $f = perm(Y) \in V = Sym^n(X)$ and P fixing Y.
We see that f is partially stable with $R = L = GL(Y) \times GL(X - Y)$.

With $W = Sym^n(Y)$, we have the P-equivariant diagrams:

$$\begin{array}{ccc} W & \xrightarrow{\iota} & V \\ \uparrow & & \uparrow \\ \Delta_W(f) & \xrightarrow{\iota} & \Delta_V(f) \end{array} \qquad \begin{array}{ccc} k[W]^d & \xleftarrow{\iota^*} & k[V]^d \\ \downarrow & & \downarrow \\ k[\Delta_W(f)]^d & \xleftarrow{\iota^*} & k[\Delta_V(f)]^d \end{array}$$

where $\Delta_W(f)$ is the projective closure of the $GL(Y)$-orbit of f, and $\Delta_V(f)$ is that of the $GL(X)$-orbit of f.

The Theorem

Lifting

- The $GL(X)$-module $V_\mu(X)$ occurs in $k[\Delta_V(f)]^{d*}$ iff $(V_\mu(X))^U$ is non-zero. Thus the $GL(Y)$-module $V_\mu(Y)$ must exist.
- Next, the multiplicity of $V_\mu(X)$ in $k[\Delta_V(f)]^{d*}$ equals that of $V_\mu(Y)$ in $k[\Delta_W(f)]^{d*}]$.

Now recall that $f = perm_m(Y)$, and let $K = stabilizer(f) \subseteq GL(Y)$.

But f is $GL(Y)$-stable, and

- the $GL(Y)$-modules which appear in $k[\Delta_W(f)]^d$ must be K-admissible.

The Grassmanian

Consider $V = V_{1^k}(\mathbb{C}^m) = \bigwedge^k(\mathbb{C}^m)$ and the highest weight vector v.

- v is stable for the $GL_k \times GL_{m-k}$ action.
- $\Delta_V(v)$ is just the grassmanian.
- v is partially stable with the obvious P.
- $W = \mathbb{C}^k \subseteq \mathbb{C}^m$ and $\Delta_W(v)$ is the line through v.
- whence

$$k[\Delta_W(v)] = \sum_d \mathbb{C}$$

- The above theorem subsumes the Borel-Weil theorem:

$$k[\Delta_V(v)] \cong \sum_d V_{d^k}(\mathbb{C}^m)$$

The general partially stable case

Recall: Let V be a G-module. Vector $v \in V$ is called partially stable if there is a parabolic $P = LU$ and a *regular* $R \subseteq L$ such that:

- v is fixed by U, and
- v is R-stable.

In the general case, there is a *regular* subgroup $R \subseteq L$, whence the theory goes through

$$\Delta_W(v) \to \Delta_Y(v) \to \Delta_V(v)$$

- The first injection goes through a Pieri branching rule.
- The second injection follows the lifting theorem.

In Summary

In other words, the theory of partially-stable $\Delta_V(f)$ lifts from that of the stable case $\Delta_W(f)$.

The Obstruction

Let $H \subseteq GL(X)$ stabilize det_m and $K \subseteq GL(Y)$ stabilize $perm_m^n$.

The representation-theoretic obstruction $V_\mu^*(X)$ for $perm_n^m \in \Delta(det_m)$

- V_μ is such that $V_\mu(X)^U$ is non-zero.
- $V_\mu(Y, y)|_R$ has a K-fixed point.
- $V_\mu(X)|_H$ does not have a H-fixed point.

Philosophically-Two Parts

- Identifying structures where obstructions are to be found.
- Actually finding one and convincing others.

Two different types of problems:

- Geometric
  - Is the ideal of $\Delta(g)$ determined by representation theoretic data.
  - Does Σ_H generate the ideal of $\Delta(g)$?
  - Is the stabilizer H of g, G-separable?
    - Larsen-Pink: do multiplicities determine subgroups?
- Representation Theoretic
  - Is this G-module H-admissible!

The subgroup restriction problem

- Given a G-module V, does $V|_H$ contain 1_H?
- Given an H-module W, does $V|_H$ contain W?

The Kronecker Product Consider $H = SL_r \times SL_s \to SL_{rs} = G$, when does $V_\mu(G)$ contain an H-invariant?

This, we know, is a very very hard problem. But this is what arises (with $r = s = m$) when we consider det_m.

Another problem-Strassen

Links invariant theory to computational issues.

- Consider the 2×2 matrix multiplication $AB = C$. To compute C, we seem to need the 8 bilinear forms $a_{ij} b_{jk}$.
- Can we do it in any fewer?

A bilinear form on A, B is rank 1 if its matrix is of rank 1. Let S denote the collection of all rank 1 forms.

- Let $S^k = S + S + \ldots + S$ (k times). These are the so called secant varieties.
- Strassen showed that S^7 contains all the above 8 bi-linear forms.

Consequence

There is an $n^{2.7}$-time algorithm to do matrix multiplication.

Specific to Permanent-Determinant

Negative Results

- von zur Gathen: $m > c \cdot n$
  - Used the singular loci of *det* and *perm*.
  - Combinatorial arguments.
- Raz: $m > p(n)$, but multilinear case.
- Ressayre-Mignon: $m > c \cdot n^2$
  - Used the curvature tensor.

For a point $p \in M$, hyper-surface $\kappa : TP_m \to TP_m$.

- For any point of det_m, $rank(\kappa(det_m)) \leq m$.
- For one point of $perm_n$, $rank(\kappa(perm_n)) = n^2$.
- A section argument.

Any more geometry?

Is there any more geometry which will help?

- The Hilbert-Mumford-Kempf flags: limits for affine closures.
  - Extendable to projective closures?
  - Something there, but convexity of the optmization problem breaks down.

- The Luna-Vust theory: local models for stable points.
  - Extendable for partially stable points?
  - A finite limited local model exists, but no stabilizer condition seems to pop out.

In Conclusion

- Complexity Theory questions and projective orbit closures.
  - stable and partially stable points.
  - obstructions
- obstruction existence
  - Representations as obstructions
  - Distinctive stabilizers
  - local definability of $Orbit(g)$
- partial stability
  - lifting theorems
- subgroup restriction problem
  - tests for non-zero-ness of group-theoretic data

Thank you.