Problems for this century

Michael Shub, City University of New York

FoCM 2014 Meeting, Montevideo, December 11, 2014

Smale's Problems for the next century, 1998

- Problem 3: Does **P**=**NP**?
- "Hilbert's Nullstellensatz": Does a system of m equations in n complex (or real) unknowns have a solution?

Hilbert's Nullstellensatz is **NP**-Complete (over any field). So **P**=**NP** if and only Hilbert's Nullstellensatz is in **P**. The model of computations is a BSS-machine (see Blum,L.,F.Cucker,M.Shub,S.Smale Complexity and Real Computation) Branching is on $=$ or $\neq$ for unordered fields as $\mathbb{C}$ and on $\geq$ or $<$ for ordered fields as $\mathbb{R}$ Complexity theory measures the cost of finding a solution for a problem instance in terms of the input size. The class of problems **P** are those problems for which there is an algorithm which solves the problem in polynomial cost. The input size is the dimension and the cost the number of arithmetic operations and comparisons.

NP-Complete and **NP**-Hard Problems

We need a big list of **NP**-Complete or Hard problems. Here are a few trivial ones to get started. I will restrict myself to $\mathbb{C}$ for the moment.

- Hilbert's Nullstellensatz for n quadratic equations in n complex unknowns is **NP**-Complete.
- Let $f : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ be a polynomial mapping. Does f have a fixed point? or a point of period k for some fixed integer k? Both are **NP**-Complete problems.
- Is H_0 of an algebraic set 0. This is **NP**-Hard. So computing homology groups of algebraic sets should be difficult.

These are trivial. Here is one more interesting.

Homogeneous Hilbert's Nullstellengsatz

Problem: "Homogeneous Hilbert's Nullstellensatz" (HHN) Does a system of m homogeneous equations in n complex unknowns have a non-zero solution?

Question: Is HHN **NP**-Complete over $\mathbb{C}$?

Some references

- Basu,S., A Complexity Theory of Constructible Functions and Sheaves, Found.Comp. Math. OnLine
- Basu,S. and T.Zell, Polynomial Hierarchy, Betti Numbers and a Real Analogue of Toda's Theorem, Found.Comp.Math. 10 (2010),429-454
- Basu,S., A Complex Analogue of Toda's Theorem, Found.Comp.Math. 12 (2012) 327-362.
- Cucker,F., A Theory of Complexity, Condition and Roundoff (Arxiv)
- Heintz,J., B. Kuijpers and A.R.Paredes,Software Engineering and Complexity in Effective Algebraic Geometry, Journal of Complexity, 29 (2013), 92-138.
- Mulmuley,K., The GCT Program Towards the **P** vs **NP** Problem, Communications of the ACM 55 (2012), 98-107.

In the next section on Smale's Problem 4. There will be more connections with the **P**,**NP** problem.

Problem 4: Integer zeros of a polynomial of one variable.

A straight line program to compute a polynomial $f \epsilon \mathbb{Z}[t]$ of one variable with integer coefficients is the sequence of elements $u_0, u_1, u_2, ...u_k \in \mathbb{Z}[t]$ such that $u_0 = 1, u_1 = t, u_l = u_j * u_k$ for all $l \geq 2$ where $j, k < l$, $u_k = f$ and $*$ is a ring operation in $\mathbb{Z}[t]$ i.e. $+, -, x$. Let $\tau(f)$ be the minimum k for all straight line programs to compute f.

For $f \in \mathbb{Z}[t]$ let $N(f)$ be the number of distinct integer zeros of f.

Tau Conjecture. There is a constant $c > 0$ such that $N(f) \leq \tau(f)^c$ for any $f \in \mathbb{Z}[t]$.

The Tau Conjecture II

If the Tau conjecture is true then $\mathbf{P} \neq \mathbf{NP}$ over $\mathbb{C}$ (Shub-Smale) and the Permanent is hard to compute (Koiran-Buergisser).

Buergisser,P, On Defining Integers and Proving Arithmetic Circuit Lower Bounds, comput.complex. 18(2009),81-103.
Other versions of the Tau conjecture appear in
Koiran,P.,N.Portier, S. Tavenas, S. Thomasse, A τ-Conjecture for Newton Polygons, Found. Comp. Math. Online
Koiran,P., Shallow circuits with high powered inputs, in Proc. Second Symposium on Innovations in Computer Science (ICS2011) 2011
with similar results.

Tau Conjecture III

If we allow arbitrary constant in the definiton of τ $u_{-l}, ..., u_0, u_1, u_2, ...u_k \in \mathbb{Z}[t]$ where $u_{-l}, ..., u_0$ are integer constants then we define $L(f)$ as the minimum k of such a computation of f. Clearly $L(f) \leq \tau(f)$. Comparable theorems concerning **P** vs **NP** or the permanent are not known nor conjectured about $L(f)$. $L(f)$ was considered by Strassen and its relation to the number of zeros was raised by him.

Growth of N(f) with L(f)

A potential method to produce exponentially many zeros:
Let $F_i \in \mathbb{Z}[t]$ have degree d_i, $i = 1, ..., n$ then evaluating the composition is $O(\sum d_i)$ while the number of complex roots is $\prod d_i$ If we can make a large fraction of $\prod d_i$ integer roots by judicious selection of F_i we would get exponential growth of zeros with respect to $L(f)$. Are there such judicious selections?

For all $d_i = 2$ can one find n quadratics with 2^n integer zeros of the composite? Yes, $n = 1, 2, 3, 4$ (Richard Bumby, Carlos DiFiore). 5 and bigger?

Finding Hay in the Haystack

Let $f = (f_1, \ldots, f_n)$ be a system of homogeneous complex polynomial equations with unknowns $X_0, \ldots, X_n$ and degrees $d_1, \ldots, d_n$. Denote by $\mathcal{H}_{(d)}$ the vector space of such systems by $\mathbb{P}(\mathcal{H}_{(d)})$ the associated projective space.

Note $N = dim\mathcal{H}_{(d)} = \sum \left(\begin{array}{c} n + d_i \\ n \end{array} \right)$

While the number of solutions is given by the Bezóut number $\mathcal{D} = \prod d_i$.

For all $d_i = 2$, $N \sim n^3$ while $\mathcal{D} = 2^n$. Let

$$\mu(f, \zeta) = \|f\| \left\| \left(Df(\zeta) \mid_{\zeta^\perp} \right)^{-1} Diag \left(d_i^{1/2} \|\zeta\|^{d_i - 1} \right) \right\|$$

and

$$\mu(f) = max_{\zeta | f(\zeta) = 0} \mu(f, \zeta)$$

Finding Hay in the Haystack II

On $\mathbb{P}(\mathcal{H}_{(d)})$ we put the probabilty structure given by the Fubini-Study Riemannian structure defined by the Bombieri-Weyl (L^2) Hermitian structure on $\mathcal{H}_{(d)}$, $\sum < f_i, g_i >= \sum \int f_i \bar{g}_i$ (normalized so that $\|z_0^{d_i}\| = 1$.)

Problem Find an algorithm and a polynomial P which on input $(d_1, \ldots, d_n)$ outputs $f \in \mathcal{H}_{(d)}$ with $\mu(f) \leq P(n, N, \mathcal{D})$

With Probability greater than $1/2$ in $\mathbb{P}(\mathcal{H}_{(d)})$, $\mu(f) \leq 2n^2 N^{1/2} \mathcal{D}^{1/4}$.

Distribution of points on the two sphere

Even for $n = 1$ and $d > 2$, $\mu(f) < d$ with probability $1/2$. But we don't know an algorithm and a polynomial P which outputs f of degree d and $\mu(f) < P(d)$.

We can express the problem in terms of the roots of the polynomial which are points on the Riemann sphere S^2, which is the sphere in $\mathbb{C}x\mathbb{R}$ with center $(0, \frac{1}{2})$ and radius $\frac{1}{2}$.

Distribution of points on the two sphere

Let $\zeta_i = (w_i, s_i) \in S^2$, $i = 1, \ldots, d$, $g(x,y) = \prod (s_i x - w_i y)$ and $\hat{g} : S^2 \to \mathbb{R}$, $\hat{g}(z) = \prod_{i=1,\ldots,d} |z - \zeta_i|$.

In terms of the roots ζ_i $\mu(g, \zeta_i) = \frac{(d(d+1))^{1/2}}{\Pi^{1/2}} \frac{||\hat{g}||_{L_2}}{\prod_{j \neq i} |\zeta_i - \zeta_j|}$.

So our problem becomes to find
$(*)(\zeta_1, \ldots, \zeta_d)$ such that $max_i \frac{||\hat{g}||_{L_2}}{\prod_{j \neq i} |\zeta_i - \zeta_j|} < P(d)$

Smale's 7th problem is to find points satisfying a more classical inequality.

Elliptic Fekete Points

Let $V : (S^2)^d \to \mathbb{R}$
$V(\zeta_1, \ldots, \zeta_d) = \prod_{1 \le i < j \le d} ||\zeta_i - \zeta_j||$
and V_d the max value of V.
Smale's 7th Problem is:

Find an algorithm and a constant $c > 0$ which on input d outputs $(\zeta_1, \ldots, \zeta_d)$ such that $\frac{V_d}{V(\zeta_1, \ldots, \zeta_d)} < d^c$.

There has been a lot of progress recently on Smale's problem, see references below. We know that $max_i \frac{||\hat{g}||_{L_2}}{\prod_{j \neq i} ||\zeta_i - \zeta_j||} \le \pi^{1/2} \frac{V_d}{V(\zeta_1, \ldots, \zeta_d)}$ so a solution to Smale's problem is a solution to $(*)$ but $(*)$ may be easier.

References Smale's 7th

- Shub,S. and S.Smale, Complexity of Bezout's Theorem III: Condition Number and Packing, Journal of Complexity Vol. 9 (1993), pp. 4-14.
- Beltrán,C. The State of the Art in Smale's 7th Problem, in F.Cucker et al, Foundations of Computational mathematics, Budapest 2011, LMS Lecture Notes 403, Cambridge, 1-15
- Borodachov,S.V.,Hardin,D.P. and E.B. Saff, Low Complexity Methods for Discretizing Manifolds Via Riesz Energy Minimization, Foundations of Computational Math. 14 (2014) 1173-1208.
- Brauchart, J.S., Hardin,D.P. and E.B. Saff, The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere, preprint

More references

- Bétermin,L., Renormalized Energy and Asmptotic Expansion of Optimal Logarithmic Energy on the Sphere, preprint
- Serfaty,S., Ginzburg-Landau Vortices, Coulomb Gases, and Renormalized Energies,Journal of Statistical Physics, 154 (2013), 660-680.
- Erwin Schrodinger International Institute for Mathematical Physics, Programme "Minimal Energy Point Sets, Lattices, and Designs", October-November, 2014

Smale's 17th Problem

Can **a** zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?

Here we will take an approximate zero to mean one for which Newton's method is quadratically converging, so:
$d(N_f^k(z), \zeta) \leq (\frac{1}{2})^{2^k-1} d(z, \zeta)$
We let $\mathcal{H}_{(d)}$ and $\mathbb{P}(\mathcal{H}_{(d)})$ be as above and $\mathbb{P}(\mathbb{C}^{n+1})$ be the projective space of $\mathbb{C}^{n+1}$.
And average means with respect to the probabilty induced on the space of systems, $\mathbb{P}(\mathcal{H}_{(d)})$,by the Fubini-Study Riemannian structure as above.
Recent Progress by Beltrán-Pardo and Buergisser-Cucker.
Homotopy methods play a big role.

Elimination theory

N is our input size. When $d >> n$ symbolic techniques can be used to reduce the problem to solving a univariate polynomial of degree $\mathcal{D}$ in polynomial time. (Elimination theory, Groebner bases, Resultants- Renegar, Grigoriev-Vorobjov, Heintz-Pardo-Roy, Canny,...) Then the univariate polynomial may be solved in polynomial time by many methods (Renegar, Pan, Neff, Manning, Hubbard-Schleicher-Sutherland, Shub-Smale,Kim...) But Caveat!!! Moreover, when $n >> d$ as in quadratic system the Bezout number is exponential in N so these techniques seem to be intrinsically exponential in the general case.

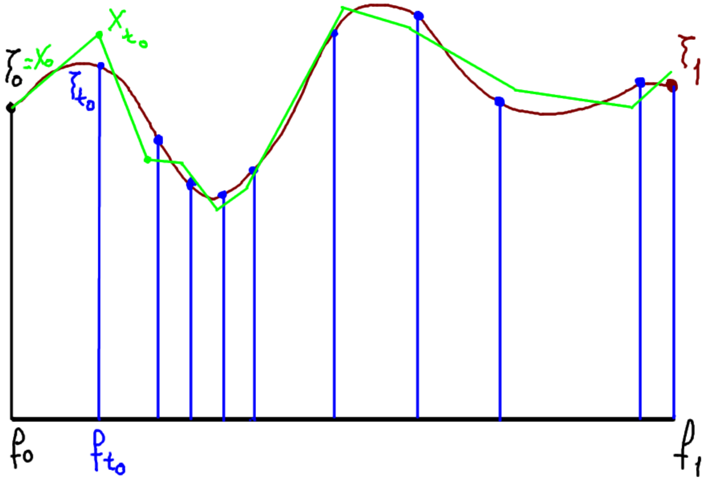
Homotopy method I

- Let f_1 be a system you want to solve, and let f_0 be a system you can solve.
- Construct a path of systems f_t joining f_0 and f_1.
- Choose some solution ζ_0 of f_0. Let $z_0 = \zeta_0$ or a close enough approximation to it.
- Choose a small step size t_0. Apply Newton's projective method

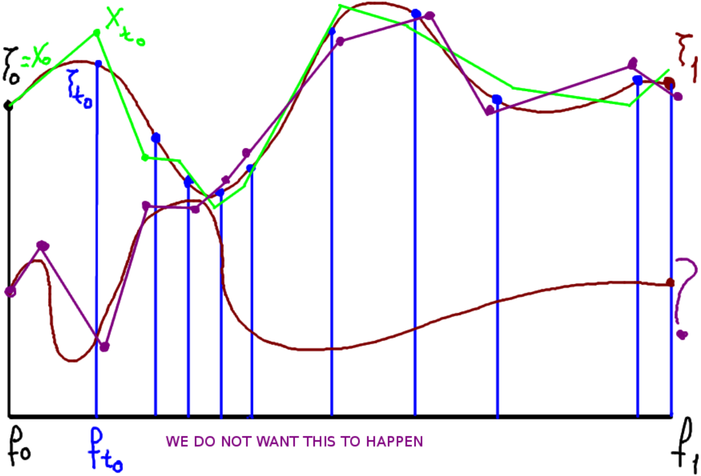
$$z_1 = N_{f_{t_0}}(z_0) = z_0 - (Df_{t_0}\mid_{z_0^\perp})^{-1} f(z_0)$$

- Continue the process until you are close to f_1. Generate $z_2, z_3, \ldots$.
- Output the last value z_j.

Homotopy method II

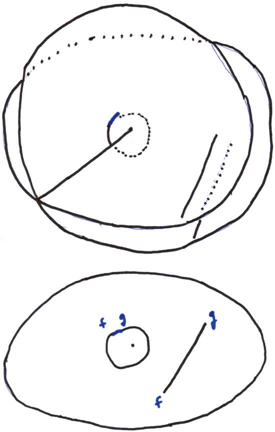


Homotopy method III



WE DO NOT WANT THIS TO HAPPEN

Homotopy method IV



V as Double Fibration

$$\hat{V} \subset \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$$

$$V \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$$

$$= \{(f,\zeta) \mid f(\zeta) = 0\}$$

$$\begin{tikzcd} & V \arrow[dl] \arrow[dr] & & \hat{V} \arrow[dl] \arrow[dr] & \\ \mathbb{P}(\mathcal{H}_{(d)}) & & \mathbb{P}(\mathbb{C}^{n+1}) & & \mathcal{H}_{(d)} \end{tikzcd}$$

Solution variety and condition number

Let

$$V = \{(f,\zeta) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1}) : f(\zeta) = 0\}$$

the solution variety. and let

$$W = \{(f,\zeta) \in V : Df(\zeta) \text{ is of maximal rank } \},$$

and

$$\mu(f,\zeta) = \|f\| \left\| \left(Df(\zeta)\mid_{\zeta^\perp}\right)^{-1} Diag\left(d_i^{1/2}\|\zeta\|^{d_i-1}\right)\right\|$$

be the condition number, defined for $(f,\zeta) \in W$.
An important point is that for the Hermitian structure on the solution variety V and W the unitary group acts by isometries and preseves μ, $(f,\zeta) \to (f \circ U^{-1}, U(\zeta))$.

Condition number and number of homotopy steps

[S.]
The number of Newton homotopy steps necessary to follow a homotopy path $\Gamma_t = (f_t, \zeta_t),\ 0 \leq t \leq 1$ is bounded by

$$Constant\ d^{3/2} \int_0^1 \mu(f_t, \zeta_t) \|(\dot{f}_t, \dot{\zeta}_t)\|\ dt,$$

that is the length of the path Γ_t in the condition metric.

Choice of (f_0, ζ_0)

Now we take the simplest paths possible. Let (f_0, ζ_0) be a known pair of system-solution. For any system f_1, define the path

$$f_t = (1-t)f_0 + tf_1.$$

Then, define the complexity measure:

$$A(f_0, \zeta_0) = \mathrm{E}_{f \ \ system}\left[\int_0^1 \mu(f_t, \zeta_t)\|(\dot{f}_t, \dot{\zeta}_t)\| \, dt\right].$$

We say that (f_0, ζ_0) is a **good starting pair** for the homotopy if $A(f_0, \zeta_0)$ is "small".

Choice of (f_0, ζ_0)

[Beltrán & Pardo]
A randomly chosen initial pair is indeed a good starting point.
That is,

$$\mathrm{E}_{g \ \ a \ \ system}\left[\frac{1}{\mathcal{D}}\sum_{\zeta : g(\zeta)=0} A(g,\zeta)\right] \leq 16\pi nN,$$

where N is the number of monomials of a generic system and $\mathcal{D} = d_1 \cdots d_n$ is the number of solutions of a generic system.

Moreover, the variance is also small
[Beltrán & S.] the variance of the number of steps is at most $O(d^3 n^2 N^2 \ln(\prod(d_i)))$.

The Theorem of Buergisser-Cucker

Let $\epsilon > 0$.

- There is a deterministic starting point for the homotopy algorithm with the following property. Let $D = max(d_i)$. If $D \leq n^{\frac{1}{1+\epsilon}}$ then the average cost of the algorithm is polynomial in the input size N.
- If $D \geq n^{1+\epsilon}$, the algorithm is polynomial cost (here it is based on Renegar's u-resultant based algorithm).
- The average cost is always $\leq N^{O(log(logN))}$.

Three ways to choose the initial pair (f_0, ζ_0):

1) Choose (f_0, ζ_0) at random, which guarantees average number of Newton steps $O(nN)$.
2) Use the "most simple" ie best conditioned (system,root) pair:

$$g = \begin{cases} d_1^{\frac{1}{2}} X_0^{d_1 - 1} X_1 = 0, \\ \cdots \\ d_n^{\frac{1}{2}} X_0^{d_n - 1} X_n = 0, \end{cases} \qquad e_0 = (1, 0, \dots, 0)$$

Conjectured by [S. & Smale] to satisfy $A(g, e_0) \leq$ "Small".
3)

$$h = \begin{cases} X_0^{d_1} - X_1^{d_1} = 0, \\ \cdots \\ X_0^{d_n} - X_n^{d_n} = 0, \end{cases} \qquad e_0 = (1, 1, \dots, 1)$$

Experiments (Beltrán and Leykin, 2012) suggest 2) is best.

Smooth version of μ

Consider the smooth counterpart of the condition number μ:

$$\mu_F(f,\zeta) = \|f\| \left\| \left(Df(\zeta) \mid_{\zeta^\perp} \right)^{-1} Diag \left(d_i^{1/2} \|\zeta\|^{d_i - 1} \right) \right\|_F ,$$

so that we take the Frobenius norm instead of the operator norm. Note that μ_F is a smooth function defined on W.

Smooth version of μ

[Beltrán,S.]
μ_F is a non-degenerate equivariant Morse function with a unique orbit of non-degenerate minima. This orbit is the orbit of the pair (g, e_0) under the action of the unitary group
$(U, (f, \zeta)) \mapsto (f \circ U^*, U\zeta)$.

Smooth version of μ

[Beltrán,S.]
μ_F is a non-degenerate equivariant Morse function with a unique orbit of non-degenerate minima. This orbit is the orbit of the pair (g, e_0) under the action of the unitary group
$(U,(f,\zeta)) \mapsto (f \circ U^*, U\zeta)$.
Optimistic Conjecture!

$$A(f_0, \zeta_0) = \mathrm{E}_{f \quad system} \left[\int_0^1 \mu(f_t, \zeta_t) \|(\dot{f}_t, \dot{\zeta}_t)\| \, dt \right].$$

is a non-degenerate equivariant Morse function with a unique orbit of non-degenerate minima. This orbit is the orbit of the pair (g, e_0) under the action of the unitary group $(U,(f,\zeta)) \mapsto (f \circ U^*, U\zeta)$.

Back to the condition metric.

Recall our theorem:
The number of Newton homotopy steps necessary to follow a homotopy path $\Gamma_t = (f_t, \zeta_t),\ 0 \leq t \leq 1$ is bounded by

$$Constant\ d^{3/2} \int_0^1 \mu(f_t, \zeta_t) \|(\dot{f}_t, \dot{\zeta}_t)\|\ dt,$$

that is the length of the path Γ_t in the condition metric.

Back to the condition metric.

Recall our theorem:
The number of Newton homotopy steps necessary to follow a homotopy path $\Gamma_t = (f_t, \zeta_t),\ 0 \leq t \leq 1$ is bounded by

$$Constant\ d^{3/2} \int_0^1 \mu(f_t, \zeta_t) \|(\dot{f}_t, \dot{\zeta}_t)\|\ dt,$$

that is the length of the path Γ_t in the condition metric. Understanding geodesics in the condition metric give us some idea of "good" homotopies (not necessarily straight lines!) and also (at least as far as this estimate is concerned) lower bounds for how well homotopy methods may work!

[Beltrán & S.] The distance in the condition metric from the (g, e_0) to any system (f, ζ) is bounded by $O(nd^{3/2} \log \mu(f, \zeta))$.
The average number of steps following geodesics for the condition number, is at most

$$O(nd^{3/2} \log(N)).$$

Thus, much faster average than the linear homotopy $O(nN)$.

Convexity aspects of μ

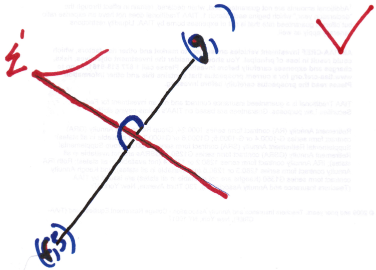
What are the geodesics like? μ is comparable to the distance in V to the degenerate (system,root) pairs. Is the condition number maximized at the endpoints? (Quasi-convexity) or even:
Consider W with the condition metric. Let γ be a geodesic. Is the function

$$t \mapsto \log \mu(\gamma(t))$$

convex? We shall say "μ is a self-convex function in W".

[Beltrán &Dedieu &Malajovich &S.]

Complexity Implications



Convexity aspects of μ

- Let $\mathbb{GL}_{m,n}$ be linear space of m by n matrices with the condition metric (here, the condition number of a matrix A is $\|A^\dagger\|$). Then, the answer to the question above is Yes: $\|A^\dagger\|$ is self-convex in $\mathbb{GL}_{m,n}$.
- The same is true for the condition number $\kappa(A) = \|A\|_F \|A^\dagger\|$ in the projective set of matrices $\mathbb{P}(\mathbb{GL}_{m,n})$.
- The same is true in the solution variety for the linear case, i.e. $W = \{(A, \zeta) \in \mathbb{P}(\mathbb{GL}_{n,n+1}) \times \mathbb{P}(\mathbb{C}^{n+1})\}$.
- Is it true for the non-linear case?

The Eigenpair Problem

Let $\mathbb{M}_{n,n}$ be the *nxn* complex matrices, with Hermitian structure $< A, B >= trace(B^* A)$. The eigenpair problem is: On input $A \in \mathbb{M}_{n,n}$ output approximations to one or all eigenvalue, eigenvector pairs (λ, v) where $\lambda \in \mathbb{C}$ and $v \in \mathbb{P}(\mathbb{C}^n)$. Actually, because of the bilinear nature of the problem it is convenient to be redundant and on input A to output $((A, \lambda), v)$ so we may consider $A \in \mathbb{P}(\mathbb{M}_{n,n})$ and $((A, \lambda), v) \in \mathbb{P}(\mathbb{M}_{n,n} x \mathbb{C}) x \mathbb{P}(\mathbb{C}^n)$.

The Bilinear Eigenpair Problem

$$((A,\lambda), v)$$

$$
\begin{tikzcd}
& V \subset P(M_{n,n} \times \mathbb{C}) \times P(\mathbb{C}^n) \arrow[dl] \arrow[dr] & \\
(A,\lambda) \in P(M_{n,n} \times \mathbb{C}) \arrow[d] & & P(\mathbb{C}^n) \ni (v) \\
A \in P(M_{n,n}) & &
\end{tikzcd}
$$

$$V = ((A,\lambda), v) \mid (A - \lambda I)v = 0).$$

Polynomial Time for the Eigenpair Problem

Theorem:(Armentano, Beltrán, Buergisser, Cucker,S.) Homotopy algorithms provide stable, average polynomial time randomized and deterministic algorithms to find one or all approximate eigenvalues for *nxn* complex matrices.

For more on the eigenpair problem,attend Diego's talks and Felipe's talk!

References

- Beltrán, C. and Shub, M. The Complexity and Geometry of Numerically Solving Polynomial Systems, Contemporary Mathematics Vol 604 (2013), pp 71-104.
- Buergisser, P. and Cucker, F. Condition: The Geometry of Numerical Algorithms, Springer, 2013
- Armentano, D., Beltrán, C., Bürgisser, Cucker, F. and Shub, M. A stable, polynomial time algorithm for the eigenpair problem, in preparation

Thank you for your attention