SIAM AAG 15 and ICIAM 2015

James H. Davenport

3–7 August 2015
10–14 August 2015

Contents

Part I

SIAM Applications of Algebraic Geometry 2015 (not fully spell-checked etc.)

Chapter 1

3 August 2015

Opening remarks: 350 registered, biggest conference we have hosted here.

1.1 The Euclidean Distance of an Algebraic Variety: Ottaviani

Theorem 1 (Spectral) *Given a real $n \times n$ symmetric matrix A there exists a diagonal D and $U \in O(n\mathbf{R})$ such that $A = U^{-1}DU$.*

The i-th column of U is an *eigenvector* with *eigenvalue* λ_i.

Real matrices are a vector space with scalar product AB^T.

Theorem 2 (Spectral) *Decomposition of operators form (for physicists).*

The variety of rank one matrices is the cone over the Segre variety $\mathbf{P}(\mathbf{R}^m) \times \mathbf{P}(\mathbf{R}^m)$. The Euclidean scalar product extends to the scalar product between rank one matrices. Then by linearity this extends to any rank.

Theorem 3 *Let u_i be normalised eigenvectors of A. Then the critical points of the distance function from A are $\lambda_i u_i \otimes u_i$*

An analogue extends to unsymmetric $m \times n$ matrices A. The critical points are singular pairs $x \otimes y$ of vectors of A and in particular they are al real.

The number of critical points of the distance from general $p \in iA^n$ to X is called the *Euclidean Distance Degree* of X ($ED(X)$).

The variety of rank one matrices is much better-behaved than general varieties.

$ED(X) = 1 \Leftrightarrow X$ is a linear space. Spheres have ED=2 and ED=2 implies X is a sphere (if smooth), or a few other quadric cones.

Plane curves may have ED 2,3 or 4. 2 are circles or pairs of lines. See Apollonius.

Call the *ED discriminant* the locus of points u such that at least two critical points of the distance from u coincide. The ED discriminant of plane curves is a classical object, the *evolute*.

When an ellipse degenerates to a circle, the evolute disappears to a point.

Theorem 4 (Catanese-Trifogli) *Let X be a variety with projective closure $\overline{X} subset \mathbf{P}^n =]A^n \cup H_\infty$. let $\dim X = m$. The ED degree of a general translate $g \cdot \overline{X} \subset \mathbf{P}^n$ is*

$$\sum_{i=0}^{m} \ldots \textit{Chern classes}$$

But Chern classes cannot distinguish the circle from the ellipse, so we need $\overline{X}$ to be transversal to the isotropic quadric.

If the *desingularization map* is linear, the the Catanese-Trifogli formula can be applied with the Chern classes of the desingularization. In general we need to replace Chern classes by Mather classes (tricky, no software).

Now replace the matrices with tensors.

Theorem 5 *The critical pints of the distance from a tensor t to X correspond to tensors $(x_1 \otimes \cdots x_d)$ such that $t(x_1, \ldots, \hat{x}_i, \ldots, x_d) = \lambda x_i$.*

For (2,2,2) we have 6 critical pints but for (3,3,3) we have 37, which is more than the dimension of the ambient space.

Symmetric tensors are polynomials.

1.2 The Optimal Littlewood–Richardson Homotopy: Sottile

This is all about Numerical Homotopy Continuation.

- Bézout Homotopy: optimal[1] in the generic case
- Polyhedral homotopy. Optimal for Sparse systems with the BKK bound
- equation-by-equation with regeneration, as in `Bertini`

But enumerative geometry problems aren't usually square, and are well below BKK Bound.

For me classical 19th century work by specialisation is just homotopy in reverse.

Consider Schubert problems. The set of linear spaces having position α with respect to a flag of subspaces F is a Schubert variety $X_\alpha F$. We are interested in counting points in

$$X_{\alpha^1} F^1 \cap X_{\alpha^2} F^2 \cap \cdots X_{\alpha^n} F^n.$$

[1] In the sense of never following a redundant or dead-end path.

Ravi Vakil's interpretation of Littlewood–Richardson. Transforms the intersection of two Schubert varieties into a union of them. This is done via "checker-board games" explaining how the flags interact. The aim is to end up with a diagonal checker-board, which [JHD thinks] means we can read off the solutions. He compares this with bubblesort. Simple example, then one where there are two checker-board patterns.

There are three kinds of Homotopy.

- Geometrically constant (coordinate change)
- Simple Homotopy (subspace rotates with flag)
- Subtle Homotopy (read paper!)

FS/RV/Jan Verschelde met several times over three years to write this down in terms of linear algebra. See [SVV10].

1.3 Sparse Gröbner Bases: the Unmixed Case: Spaenlehauer

See ISSAC 2014. $f_1, \ldots, f_m \in K[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$. What exact solutions in $(K \setminus \{0\})^n$. . By unmixed, we mean that 1 is a monomial and is a vertex of the Newton polynomial of the set of monomials. [KipnisShamir1999,Joux2013] in cryptology. Kushnirenko's theorem states that the number of solutions is much smaller than Bézout.

General strategy is a grevlex basis $O(m \begin{pmatrix} n+d \\ d \end{pmatrix}^{\omega})$ then FGLM $O(nDEG^3)$.

1.3.1 Semigroup algebras

An *affine semigroup* S is a finitely generated additive subsemigroup of $\mathbf{Z}^n$ containing $0 \in \mathbf{Z}^n$ and no non-zero invertible elements.

Then we get semigroup algebras (toric rings) Usual concept of admissible ordering. $\langle \mathrm{LM}(G) \rangle = \mathrm{LM}(I)$ iff G is a *Sparse Gröbner Basis* of I.

Let M be a finite "generating set" of S (not necessarily minimal). $\mathrm{rank}_M(X^s)$ is the smallest integer k such that X^s is the product of k elements of M. The rank of a polynomial is the maximal rank of its monomials.

Mimic F_5 by replacing "degree' by "rank" The same theorems apply, and we can read off a sparse Gröbner basis from the row-echelon form of the Macaulay matrix of *rank d*.

Shows huge (10^5) speedups in some examples over classical F_5, less (80) in other examples. Computing over a finite field.

Understand w.r.t. toric homogenisation: $M^{(h)} = \{(s, 1) : s \in M\}$. If $K[S_M^{(n)}]$ is Cohen–Macaulay, we have a theorem of Hochster ....

Then the complexity of the FGLM-equivalent can have a similar bound. There is an extension of Fröberg's conjecture in the Cohen–Macaulay case.

In positive dimension this is not easy, as the "lcm" of two monomials is a non-principal ideal. Might be related to [Stu95].

1.4 Algorithms for the Computation of Chern–Schwartz–MacPherson Classes and the Euler Characteristic: Helmer

Consider subschemes of certain smooth complete toric varieties X_Σ. Work over k an algebraically closed field of characteristic 0. We will find χ (Euler) via c_{SM}. This has useful functional properties.

When V is a subscheme of 'P^n the class $c_{SM}(V)$ can bethought of as a more refined version of the Euler characteristic since it contains the Euler characteristics of ....

Need the Segre class. LetX_Σ be an n-dimensional smooth complete toric variety defined by a fan Σ. Let R be the graded coordinate ring (Cox ring) of X_Σ with irrelevant ideal B and assume that all Cartier divisors associated to generating rays are nef (needed for counting purposes). Also one technical assumption. We work in the Chow ring of X_Σ.

Write the Chow ring as $A^*(X_\Sigma) \equiv \mathbf{Z}[x_1, \ldots, v_m]/(I + J)$ where I is the Stabley–Reisner ideal and J is the ideal generated by all the linear relations of the rays $\Sigma(1)$.

Aluffi had an algorithm by blowups. Eklund/Jost/Peterson had a probabilistic algorithm by saturation. We had one by counting points. Moe/Qviller have one, with same nef-restriction.

Let I be an ideal in R homogeneous with respect to the grading on R. Choose generators f_i such that $[V(f_i)] = \alpha \in A^1(X_\Sigma)$ for all i.

Cut with (the right) general hyperplanes until we get something zero-dimensional.

Claims stunningly better times for Segre classes (all in Macaulay 2). Can do degree 12 in $\mathbf{P}^{16}$ (over a finite field of course). His Bertini timings are much worse, probably an issue with his implementation, he said[2].

How to get from Segre to CSM? There's a formula (Aluffi) for hypersurfaces, then use inclusion/exclusion. But this needs exponentially many computations.

1.5 Some Current Directions in Coding Theory: Walker

Introducer: speaker is famous for book on Algebraic Codes.

Encoder is an injective map, the channel transmits a garbled version of this codeword, which the decoder has to

[2]But a questioner made the same comment.

Theorem 6 (Shannon) *Every channel has a capacity c such that for all RC and every $\epsilon > 0$ there is a code of rate R such that the probability of error using this code is $< \epsilon$.*

Hence "Shannon's Challenge" — find this.

Definition 1 *A linear code of length n, dimension k and minimum distance d over $\mathbf{F} - q$ is a k-dimensional subspace C of $\mathbf{F}_q^n$ such that any two distinct elements of C differ in at least d positions.*

$$\frac{k}{n} \leq 1 - \frac{d}{n} + \frac{1}{n}.$$

1.5.1 Reed–Solomon Codes

For $k < n < q$ the Reed–Solomon Code of length n and dimension k over 'F_q is $C - \{(f(a_1), \ldots, f(a_n) | f \in L_k\}$.

1.5.2 Goppa

Suppose a curve $X/\mathbf{F}_q$ of genus g, $\mathcal{P}$ a set of points, G a divisor with $supp(G) \cap \mathcal{P} = \emptyset$ Then '$frackn`ge1 - \frac{d}{n} + \frac{1}{n} - \frac{g}{n}$.

The rank distance between two codes has distance $\frac{1}{2}$ of the distance between the matrices.

1.5.3 New developments

- Quantum codes. If $C_1 \subset C_2$ are linear codes of length n over $\mathbf{F} - q$ then there is a quantum code such as ...
- Locally recoverable codes, as in node failure in a cloud. Replication is an expensive answer, so Facebook [TB14] uses Reed–Solomon (10/14 code: 10 data bits spread across 14 raw bits) with 40% overhead rather than 200% for replication. We want codes in which every symbol is a function of a small number of others.[3] [There was a debate over why 200% rather than 100%: Speaker is quoting original paper. JHD subsequently: 100% would be simple replication, but that's not ECC at all. See also `https://storagemojo.com/2013/06/21/facebooks-advanced-erasure-codes/`.]
- MIMO interference, as in neural nets.

Q Algebraic geometry codes?

[3] JHD's memory of the original talk is this. In 2015 a Facebook data centre is 40PB, made up of 1TB discs, so 40,000 of them. The Mean Time Between Failure of a disc is five year, say 2000 days. So 20 discs fail a day. each disc failure means recovering the 13 copies, so we are shipping 260TB across the backbone — as of course the replicas are in different racks. Call it 11 TB/hour, or 3GB/sec. That's 24Gb/sec, which takes up a large chunk of 40Gb Ethernet. The speaker noted that disc sizes were growing faster than Ethernet bandwidth, so the problem was getting worse.

A I doubt it will be used for channel encoding.

Introducer But see PQC.

Q–JHD (afterwards) Aren't locally recoverable codes really saying that the decoding matrix should be (uniformly) sparse.

A Essentially, yes. That's why LDPC (Low Density Parity Check) codes work.

1.6 Advances in Software in Numerical Algebraic Geometry: Brake

Aim is to show developments since 2013. Defines this as "the use of numerical tools to study and use zero-sets of polynomials". Want a bridge, rather than a wedge, with the symbolic tools[4].

1.6.1 solvers

Bertini Interfaces with Macaulay2, Singular, MatLab. C under redevelopment into C++. Uses MPI parallelism. Does Numerical Irreducible Decomposition. Also doing bindings for Python scripts. Bertini 2 will be GPL3.

Hom4PS-3 Has automatic multiple precision, MPI Parallelism GPU acceleration. Author: "Tropical geometry has inspired new ...".

NAG4M2 Runs inside Macaulay2. Again Numerical Irreducible Decomposition

PHCPack Sage, Maple etc. Also GPU acceleration.

Polynomial System Solver This has mixed volume computations and sparse condition numbers.

Liddell etc. are looking at "how do you know you've got all the *real* solutions. Note also that lots of people are looking at GPU [VY15]. Quoted [GXD+14] as his favourite weird application.

1.7 Critical Points via Monodromy and Local Methods: Martin del Campo

The critical point is where the Jacobian has full rank, and rank$[\nabla\Phi_u(x)^T, \cdots,]$'$lem$. These conditions are additive so can assume X is irreducible.

Note this re;ates to opening talk. The *monodromy Group* is the subgroup of S_d generated by permutations arising from lifts of loops.

[4] JHD: see note 2.

1. Find one critical point: we use witness sets.
2. Find random loops
3. Trace test

When should I stop? S_{r-1}‘$subseteq S_r$‘$subseteq S$: when are we at S?

Theorem 7 ([Sommeseetal2003]) *The trace of X with respect to ‘$calL_t$ is affine linear on t. Moreover the coordinate-wise sum of any proper subset of ‘$calL_t$‘$cupX$ is not linear on t.*

Once we have found all the critical points over ‘C^n for u we can find the critical points for u' by parameter homotopies.

1.8 A lifted square formulation for certifiable Schubert calculus: Hein

Schubert calculus is the study of linear spaces incident to fixed linear spaces.

Example 1 (Four lines) *Given four lines in $\mathcal{P}^3$, what are the lines that meet all. Generically, two.*

$|w| = \text{codim}_{Gr(a,n(}\, X_w F$. Then a Schubert problem is a list of Schubert conditions (w^i) such that $\sum |w^i| = \dim(Gr(a,n))$. In Plücker coordinates these are heavily over-determined: problems for numerical methods.

Our [HHS12] primal-dual formalisation is square but adds a ton of new variables.

So we lift the problem from the Grassmanian to a more general flag variety.

Example 2 *$Gr(3,8)$: $w = (3,5,8)$, then x must meet three conditions:* $\dim x \cap F_3 \geq 1$, $\dim x \cap F_5 \geq 2$ *and* $\dim x \cap F_8 \geq 3$ *(trivial).*

Theorem 8 (H–Sottile) *This lifted formulation defines each $X_w Fas$ a complete intersection.*

Dimension+—Equation—=—Variables—

Example 3 *$X_{(5,9,10)}F \subset GR(3,10)$. Deyerminental formulation has 45 degree 3 polynomials in 21 variables (10 linearly independent polynomials). Primal–dual uses 21 bilinear polynomials in 39 variables. Lifted uses ....*

Chapter 2

4 August 2015

2.1 p-adic Integration and Number Theory: Kim

Example 4 *Polynomial equations in two variables.* $f(x,y) = 0 : f \in \mathbf{Z}[x,y]$. *This is still an inaccessible problem. Call the solutions* $X_f(\mathbf{Q})$.

genus 0 *Can parametrise all solutions: method of sweeping lines.* $x^2 + y^2 = 1$ *is* $\left(\frac{m^2-1}{m^2+1}, \frac{2}{m^2+1}\right)$.

genus 1 *Birch–Swinnerton-Dyer conjectures. There is*[1] *a finite set* S *of solutions such that all others can be generated from* S *by chord/tangent (elliptic curve arithmetic). The "algorithm" for finding* S *terminates if (the relevant port of) the conjecture is true.*

genus≥ 2 *The solutions are finite (Faltings). But there is no known algorithm for finding them. This is often known as "Effective Mordell Conjecture". We want* $C(f)$ *such that all solutions have numerators/denominators bounded by* $C(f)$. *A precise enough version of the ABC Conjecture would yield such.*

- *I am trying a non-Archimedean approach to this.*

So I want to descibe $X(\mathbf{Q}) \subset X(\mathbf{Q}_p)$, via a non-Archimedean approach using p-adic analytic equations. Sometimes write $\mathbf{R} = \mathbf{Q}_\infty$ and write Q_v for a general completion.

Example 5 $\sum_{n=1}^{\infty} n!$ *converges in any* $\mathbf{Z}_p$. *Open question: is it algebraic?*

$\mathbf{Q}_p = \mathbf{Z}_p[1/p]$. If $p \not| n$, then $n \in \mathbf{Z}_p^\times$. $\mathbf{Z}_p$ has the topologyof a Cantor set: a fact that is probably under-utilised.

$X/\mathbf{F}_p$ is a smooth projective variety. What is $|X(\mathbf{F}_p)|$. Of course, we can use brute force here. So assume X is liftwable, ie. assume there is a

[1] Mordell–Weil Theorem.

smaooth projective scheme $\mathcal{X}/\mathbf{Z}_p$ such that $\mathcal{X} \pmod p \equiv X$. Then $H^i_{cr}(X) := H^i(X_{\mathbf{Q}_p}, \Omega^*_{X_{\mathbf{Q}_p}})$ is the crysttaline cohooology. Then we have a Lefschetz trace formula:

$$|X(\mathbf{F}_{p^n})| = \sum_i (-1)^i Tr[\phi \ldots]^n.$$

Profound, but very hard to use. In fact, can use any Grothendieck cohomology theory, in particular étale cohomology theory. But this is notoriously hard to compute. Kedlaya noticed that crysttaline cohomology is much easier because of the relation to differential forms.

Example 6 *X hyperelliptic with affine model $y^2 = f(x)$ with f having odd degree d. $H^{0\ p}_{cr}(X) = \mathbf{F}_p$. The basis for the de Rham cohomology is $\{x^k dx/y | O \leq k \leq d-2\}$. Given any closed differential form, we can write it in terms of the basis (Kedlaya has an algorithm for this). Then the action of ϕ in $H^1_{cr}(X)$ is represented by the matrix (c_{ij}).*

Note that $\sqrt{-1}$ exists in $\mathbf{Q}_5, \mathbf{Q}_{13}$ etc., even though $\notin \mathbf{Q}$.

Hasse–Minkowski Theorem, underpinned by Class Field Theory. There are local reciprocity maps:

$$Rec_v : \mathbf{Q}_v^\times \leftharpoonup Gal(\overline{\mathbf{Q}}_p/\mathbf{Q}_p).$$

Having done the preliminaries, let us look at p-adic line integrals.

$$\log_p(x) = \int_1^z dt/t : \mathbf{Q}_p^* \to \mathbf{Q}_p.$$

For $z \in 1 + p\mathbf{Z}_p$ gives us a power series which converges. Then $\log_p(u) = \frac{1}{p-1} \log_p u^{p-1}$ Then define $\log_p(p) = 0$ and we get a group homomorphism: $\log_p : \mathbf{Q}_p^* \to \mathbf{Q}_p$.

$$\log_p(z) = \cdots$$

as a consequence of local class firld theory.

Consider the connection $\nabla \ldots$. The thing that actually matters is the parallel transport operator $T_b^z =$ matrix. Sometimes known as Coleman integration.

Example 7 *$X = \mathbf{P}^1 \setminus \{0, 1, \infty\}$. Then $\ell_k(z) = \sum_{n=1}^\infty \frac{z^n}{n^l}$ is only defined for $|z|_p < 1$, but we can do analytic continuation. This relates to the p-adic dilogarithm via usual equations (different proofs!).*

There is currently a strategy in place for describing rational points $X(\mathbf{Q}) \subset X(\mathbf{Q}_p)$. This is "Arithmetic Chern-Simons theory".

2.2 Fast Scalar Multiplication in Pairing Groups: Ionica

Let E be a pairing-freindly elliptic curve over $\mathbf{F}_q$ wher $r||E(\mathbf{F}_q)|$ and $r|q^k - 1$, where k os the embedding degree. Then $\mathbf{G} \subseteq E(\mathbf{F}_q)$ and $\mathbf{G}_2 \subseteq E(\mathbf{F}_{q^k})$.

$$e : \mathbf{G}_1 \times \mathbf{G}_2 \to \mathbf{G}_T$$

is thefundamental operation. We need both this and multiplication in $\mathbf{G}_1\mathbf{G}_2$. We wuld like all this to be efficient!

Multi-sclar multiplication: $[s]P + [t]Q$. Write S and t in binary with bits $s_i t_j$. Precompute $T = P + Q$. For each bit we add *one* of P, Q and$T$.

GLV: Assumethere is an efficient endomorphism $\phi : E \to E : \phi(P) = [\lambda_\phi]P$. Then this makes multiplication faster.

Example 8 *E_α $y^2 = x^3 + \alpha x$. Assume $q \equiv 1 \pmod 4$, let $i \in \mathbf{F}_q$ with $i^2 \equiv 1$. $\phi : (x, y) \to (x, iy)$ is an automorphism.*

Then the GLS construction (2009). There is also 4-GLV if we have two endomorphisms ϕ and ψ with different eigenvalues. Store 16 points, but sve $\frac{3}{4}$ of the doublins (and $\frac{17}{32}$ of the additions.

[LongaSica2009] require the eigenvalues to have good lattice reduction. Consider GLS curves defned over $\mathbf{F}_{q^2}$ with CM by a small $D \to 2$ endomorphisms $\psi^2 + [1] = 0$, $\phi^2 + [D] = 0$ for points over $\mathbf{F}_{q^2}$.

Various friendly curves, e.g. KSS $k = 18$ uses 2-GLV in $\mathbf{G}_1$ and 6-GLV in $\mathbf{G}_2$. No known one with 4-GLV in $\mathbf{G}_1$.

$\mathcal{C}_1/\mathbf{F}_q : y^2 = x^5 + ax^3 + bx$: Satoh curves. $J_{\mathcal{C}_1}$ is the Weil restriction of $E_c/\mathbf{F}_{q^2} : y^2 = x^3 + 27(3c - 10)x + 108(14 - 9c)$ where $c = a/\sqrt{b}$. $D = 2D'$ Gives $I_2 : F_c/\mathbf{F}_{q^2} \to E_{-c}/\mathbf{F}_{q^2}$. Vélu's forumlae give this 2-isogeny explicitly. We also have $I_{D'}$, again with Vélu. Then '$phi_2^2 \pm 2 = 0$ and $\phi_{D'}^2 \pm D' = 0$ for points defined over $\mathbf{F}_{q^2}$. Table of Multiplication/Squaring counts showing, roughly, 20% improvement.

Recipe for constructing such curves.

1. Pick discriminant
2. Search for m, n
3. take Hilbert polynomial and get roots in $\mathbf{F}_{q^2}$
4. ...

But these are not pairing-friendly. This requirement imposes additional constraints. Current work is applying these.

2.3 Pairings and Arithmetic: Schwabe

[Naehrigetal2013a] and other work. However, pairing-based software doesn't always depend on these for critical timngs. Pairing computation is often not the bottleneck (any more). There is a widely used benchmark: `http://bench.cr.yp.to`. Supports benchmarking of many primitives, but not pairing or other group arithmetic. Filling this gap is our aim.

Need a C-API: bench does and everyone serious about performance wants this. In theory, everyone writes multiplicative, but we write additively.

Note that a major attack is via timing, so we are always interested in constant-time software. [Osvietalk206] stole Linux `dmcrypt` AES key. [Brumley-Tuveri2011] took a few minutes to teal OpenSSL ECDSA key *over the network*.

`bgroup_g1e`, `bgroup_g2e`, `bgroup_g3e` and `bgroup_scalar`. Parameters are byte arrays, but `pack` and `unpack` are explicit. Scalar and multi-sclar multiplication, but again constant time. Hashing versions into $\mathbf{G}_1$ and $\mathbf{G}_2$. Note that the public and private versions are allowed to be different, since a datum should never be both!

Also scalar arithmetic. This is not normally the bottleneck so only has constant-time versions. Also no non-constant-time version of pairings, for same reason.

Reference version is 254-bit BN curve ([Aranhaetal2011]) Reference C implementation and AMD-64 version. Points are in Jacobian coordinates. Doubling is 2M+5S. This works for infinity as well. To make addition constant time, we compute $P + Q$ *and* $2P$, then choose which (makes it expensive[2]!). Tried using [Hı10], but these actually have other special cases, so are not truly constant-time.

Bos–Coster algorithm.

Various figures: `g1e_scalarmult`: 347024 cycles. Pairing 2.6M cycles, but pairing product ($n = 2$) is only 3.8M, not 2×2.6M. Note that this shows that simple "pairing count" is naïve.

Examples of how to use: they all fit on one slide. `http://cryptojedi.org/papers/#panda`.

2.4 Applications of Numerical Algebraic Geometry: Hauenstein

One question: how do we know that we have all the solutions?. Many engineers wonder why?

See [BlekhermanHauensteinOttenRanestadSturmfels2012]. We had the correct calculations for a Hilbert SoS case. This was numerical, and correct, whereas hand symbolic was wrong. [SommeseVerscheldeWampler2002]: For a union of irreducible, the centripd of interstecion points moves linearly. This is easy tosee for hypersurfaces. Showed example with 87K (but an exact number) of points,

[HauensteunRodriguez] "fix" the trace test by taking linear slices in the Segre embedding, so the condition here is linear,

Alt's problem had 8652 solutions, but neded to add 23706 extra solutions to get the verified trace test: 32358.

My favourite homotopy is local descent. Note that there is also modular local solutions followed by reconstruction to verify. See [Campo-Rodriguez2015a].

Example 9 $\sigma_8(\mathbf{C}^3 \times \mathbf{C}^6 \times \mathbf{C}^6)$ *[ChiantiniMellaOttaviani2014] proved at least six*

[2] There's an internal-only version which only adds, for cases known to have $P \neq Q$.

solutions. Have verified it's actually 6, or more precisely $528 \cdot 8! = 6 \cdot 8! + 522 \cdot 8!$*, where the last element is terms at infinity.*

[...]Littlewood's problem of sevn touching cylinders.
Bertini2 workshop May 23-25 at Notre Dame.

Q Suppose there is no solution.

A Gröbner then is often faster: run both/all and see which is first.

Q How do we argue for Bertini rather than Newton?

A Give examples where the convergence is so bad that a random choice will fail: easy in high dimension. [GriewankOsborne]

2.5 Theta ranks for Matroids: Sanyal

[Change of title and subject given other talk.]

The minimal degree necessary for h_i such that any linear $l(p) \geq 0$ for all $p \in V$ a finite configuration. $Lev(V)$ depends only on facet-definiing linear functions. $Lev(V) \leq k \Rightarrow Th(V) \leq k-1$ but the converse is not true. However, if $Th(V) = 1$, $Lev(V) = 2$.

Matroid $M = (E, B)$ has a finite ground set E, and B with $B - 1, B_2 \in B$

$$\forall e \in B_2 \setminus B_2 \exists f \in B_2 \setminus B_1 : (B_1 \setminus \{e\}) \cup \{f\} \in B.$$

Theorem 9 *The following three are equivalent.*

1. $Th(M) = 1$
2. *M has no minor isomophic to four cases*
3. *M is constructed from uniform matroids by taking sums or 2-sums.*

2.6 Exact Algorithm for Polynomial Optimisation: Safey El Din

$F = (f_1, \ldots, f_p) \subset \mathbf{Q}[x_1, \ldots, x_n]$, variety V, , $G \in \mathbf{Q}[\ldots]$. All degrees $\leq D$.

$$G^* = \inf_{x \in V \cap \mathbf{R}^n} G(x)$$

In the worst case, G^* is an algebraic number of degree D^n.

Example 10 *[BaiZimmermann20112] suze optimisation of sextic polynmialsin the number field sieve. Degree 12 162 monomials. Coefficents of bit size* ≈ 254*. Output: 14 local minimizers* very *close and of large magintude, 200 digits needed to distinguish these points.*

Quantifier elimination by CAD is doubly exponential and in practice limited to $n \approx 4$ There is $D^{O(n)}$ [BPR06]. The key tool is polar varieties. From now on, we assume that $G = x_1$ (this is wlog, since we just add another variable/equation). Let $\pi_i : (x_1, \ldots, v_n) \to (x_1, \ldots, x_i)$. Polar variety W_{i-1} assiciation to $(f_1, \ldots, f)p)$ and π_i. $\bigwedge f_i = 0 \wedge$ truncated Jacobian $=0$. $W_1 \subset W_2 \subset \cdots W_i$ with $d = \dim V$ (under regularity assumptions).

W_1 is the critical locus of the restriction of $x \to X_1(x)$ to V. Let $\mathcal{C}'_i = W_i \cap V(X_1 \ldots, X_{i-1}$. Most of the time this has dimension1. The let '$\mathcal{C}' = \bigcup_{i=2}^{d} \mathcal{C}'_i$, and most of the time this has dimension 1. It contains W_1.

Let $\mathcal{C} = \overline{\bigcup_{i=2}^{d} (\mathcal{C}_i - W_1)}$. Up to a generic linear change of coordinates in $X_2 \ldots, X_n)$, $\mathcal{C}$ has dimension at most 1.

[SafeyElDinSchost2004] has a topological invariance property.

Canuse symbolic homotopy or geometric resolution. This is quadratic in intrinsic geometrical degree bounds and linear in cost of evaluation.

2.7 Optimality Conditions using Newton diards and sums of squares: Sekiguchi

Theorem 10 ([Nie2013]) *If a Hessian condition and some constraint qualifications hold at each global minimiser of (POP) the* $f - f_{\min} = \sigma_0 + \sum \sigma_i^2 g_i$.

Let $\Delta(f) = \bigcup \{\alpha + \mathbf{R}^n_+ | \alpha \in supp(f)\}$.

Theorem 11 ([Vasiliev1977]) *If lots of conditions, then* f *has an isolate zero at 0.*

I am interested in analogous results. $\mathbf{R}[x]_{\frac{1}{2}\gamma} = \{f : \deg(f) \leq \frac{1}{2}\gamma\}$.

Theorem 12 *Sufficient conditions for* $f \in \sum \mathbf{R}[[x]]^2$ *are*

1. *Every vertex of* Γ *is even*
2. *...*
3. *...*

Theorem 13 *Let* f_{2m} *be the lowest homogeneous part of* f. *If this is a sum of squares (bounded by* m*) then* f *is a SOS in* $\sum \mathbf{R}[[x]]^2$.

The problem is terms with odd degrees separately but even total degree. There seems to be a Newton polygon-based technique for massaging these terms.

The key seems to be "binarily regular Newton polyhedron" conditions. Future work includes asking what sort of Newton diagrams have this condition.

2.8 Gap Vectors of Real projective varietes: Juhnke-Kubitzke

Recall "p is non-negative" "sum of squares". How are these related.

Theorem 14 (Hilbert 1888) *Nonnegative p is a sum of squares if*

1. *p is bivariate (univariate non-homogeneous)*
2. *p is quadratic*
3. *p is of degree 4 in 3 variables*

In all other cases, there exist non-negative polynomails that are not sums of squares.

Let X be a real projective variety, $I(X)$ its ideal, $R = \mathbf{R}[x_i]/I(X)$, $P_X, \Sigma_X = \ldots$ When is $P = \Sigma$?

Theorem 15 *Iff $X(\mathbf{C})$ is a variety of minimal degree.*

So we now ask when are the faces of P_X and Σ_X equal? For $\Gamma \subset X$, let $P(\Gamma)$ be the set of forms of P_X that vanish on Γ. Let $\Sigma_X(\Gamma)$ be the forms of Σ_X that vanish on Γ.

Theorem 16 *Let Γ be a finite set of points. Let Y be the prohection of X away from Γ.*

Definition 2 Γ *is* independent *if*

1. *$\langle\Gamma\rangle \cap X = \Gamma$*
2. *The points in Γ are projectively independent*
3. *$\langle\Gamma\rangle$ and X intersect transversally.*

Then we want to cosnider the dimensions of these $P(\Gamma)$. Le $g_\ell(X) = \dim P(\Gamma) - \dim \Sigma(\Gamma)$ for $1 \leq$
ell $\leq$??.

Suppose X has codimension c. Let $\epsilon(X) = \begin{pmatrix} c+1 \\ 2 \end{pmatrix} - \dim I(X)_2$. Then $g_c(X) = \epsilon(X)$ and $g_{c-1}(X) = 0$ if X is a variety of minimal degree, else $\epsilon(X) - 1$. Also the g_i are weakly increasing with i.

Let g be the vector of g_ℓ. $g = 0$ iff X is avariety of minimal dgeree. g has only one non-zero component iff $\epsilon(X) = 1$: then $g = (0, \ldots, 1)$.

Chapter 3

5 August 2015

3.1 Algebraic Codes and Invariance: Sudan

This is an ex-coding theorist's prespective.

3.1.1 Codes and Algebraic Codes

Linear codes over $\mathbf{F}_q$. Encodingfunction $E : \mathbf{F}_q^K \to \mathbf{F}_q^n$. The associated code C is the image of E.

Reed–Solomon: regard message as a polynomial, and evaluate at n points. Reed–Muller: multivariate generalisations of Reed–Solomon. Algebraic–Geometric Codes: the domain is the set of rational points of an irreducible curve.

3.1.2 Combinatorics of Algebraic Codes

Rate $R(C) = k/n$. *Distance* $= \delta(C) := \min_{x \neq y} d(x, y)$.

Pigeonhole principle implies $R(C) + \delta(c) \leq 1 + \frac{1}{n}$. Note that Reed–Solomon can be made to hit this bound exactly.

For Reed–Solomon and Reed–Muller, distance equates to scarcity of roots. In higher dimension there is a lot of underpinning algebar/geometry Stichtenoth etc.

It is true, non-trivial, that there are infinitely many algebraic-geometric codes with $R(C) + \delta(C) \geq 1 - \frac{1}{\sqrt{q}-1}$.

3.1.3 Algorithmics of Algebraic Codes

Want efficient encoding: matrix-vector product, ehcih is generally efficient. Testing ("is it a code word") is also easy. Decoding (with correction): given $r \in \mathbf{F}_q^n$ find m with $\delta(E(m), r)$ minimal. This is not obviously easy. There are codes for which decoding *is* NP-complete. Let $U * v$ denote coordinate-wise product. For linear codes $A, B \leq \mathbf{F}_q^n$, define $A * B = span\{a * b | a \in A, b \in B\}$. For every known algebraic code C of distance δ there is acode E of codimension

$\approx \frac{'}{d}elta2n$ sich that $E * C$ is a code of distance $\frac{\delta}{2}$. For Reed–Solomon this is algorithmic. Call $(E, E * C)$ an error-locating pair.

Given $r \in \mathbf{F}_q^n$.

1. Find $e \in E$ $f \in E * C$ such that $e * r = f$
2. Find $\hat{x} \in C$ such that $e * \hat{x} = f$ — againa linear system

There is a solution to step 1 if $\dim E > |errors|$.

A list decoding abstraction is an increasing basis sequence, such as x^i for Reed–Solomon.

3.1.4 Locality of (some) Algebraic Codes

This will become more challenging. Want to perform tasks in $o(n)$ time. Notethat we expect to corrects $\frac{\delta}{2}$ errors.

Does correcting a linear fractio of errors require scanning the whole code? does testing? Deterministically: yes, but probabilistially, not necessarily.

The codes I am talking about are less used in practice (yet) than the locally-repairable codes of Section 1.5. Note that Reed–Solomon must require reading at least k elements. Reed–Muller is better. $RM[m, r, q]]\{(\langle f(\alpha)\rangle_{\alpha \in \mathbf{F}_q^m} | f \in \mathbf{F}_q[x_1, \ldots, x_m] \deg(f) \leq r\}$. Restrictions of low-degree polynomials to lines yields low-degree (univariate) polynomials. I have $\frac{n}{1+m}$-locality.

Locality implies small (local) cobstraints. Do these lead to local decoding: No!. Reed–Muller has a lot of transitivity so we need to consider $Aut(C)$.

If a code has l-local constraints and 2-transitivity the the code is l-locally-decodable from $O(\frac{1}{l})$-fraction errors.

Suppose my constraint is $f(a) = f(b) + f(c) + f(d)$. To find $f(x)$, we find a random π with $\pi(a) = x$. Then $f(x) = f(\pi(b)) + f(\pi(c)) + f(\pi(d))$.

Recent progress /cpciteYekhannEfrmenko2006 3-locally decodable cods of sub-exponentail ength. Not great, but best we can do.

[Koppartyetal2013] $n^{o(1)}$-locally decodable codes with $R + \delta \to 1$.

3.1.5 Aside: Symmetric Ingredients

Message is a bivariate polynomial. Ecode f by evaluations of (f, f_x, f_y) and the rate goes to 1 as we take more and more derivatives.

3.1.6 Conclusions

Q Many cnstructions require large q? How about $q = 2$

A Restrict to the small input (as in BCH), or use concatenation of codes, i.e. code $\mathbf{F}_2^\alpha$ into F_q first.

Q Finite fields only for the degree testing??

A Yes.

3.2 Root isolation: Yap

"near optimal" is a code word in this area. Root isolation means finding ϵ-approximations, i.e. an ϵ-disc containing exactly one root. Distinguish

Global find all the roots

Local those in a given region.

3.2.1 selective history

Classical: Descartes etc.. Benchmark problem: isolate all roots of an integer polynomial F. $\tilde{O}(d^2l)$ [Schoenhage(unpubl),Pan]. Based on circle method (global) but now have good bisection methods (local).

Three bisection methods.

1. Sturm. Non-adaptive; limited to polynomials.
2. Descartes; Collins–Akritas. This is the method of choice for Computer Algebra [SM15]
3. Evaluation [BKY09, SY12] Does complex roots [SY11]

Note that representing analytic funtions is a problem. Also whereas we have sqfr for polynmials we have no analogy. We need a replacement for the C_1 (monotonicity) -predicate. Our tool here is Pellet's theorem. We also need a replacment for sign evaluation.

Given $F : \mathbf{C} \to \mathbf{C}$, have $F : \mathbf{C} \to \mathbf{C}$.. Should be *conservative*: $F(B) \subseteq F(B)$, and convergent.

For an analytic function, we want box functions fro all its derivatives.

We will change the problem into root clustering. Descartes etc. rquire only simple zeros (in the box under consideration). To be unconditional, we allow a k-cluster to be k simple roots, a k-multiple root, or inbetween.

A disc D is isolating for F if $3D \setminus D$ contains no roots. Then clusters are either disjoint or containing. n roots have at most $2n-1$ clusters.

Root clustering problem: given a box $B_0 \subset \mathbf{C}$ and find an ϵ-isolating system of cliusters for B_0.

3.2.2 Pellet Predicates

Given $k \geq 0$, reals $r, K \geq 1$

$$C_k(m, r; K) : |F_k(m)|r^K > K \sum_{i \neq k} \dots$$

Lemma 1 (Pellet 1881) *If $C_k(m, r; k)$ holds then teh disc $D_m(r)$ contains ...*

Theorem 17 (Darboux) *$F : D_0 \to \mathbf{C}$ be anaytic in D_0. ...*

Box version of Pellet.

$$C_k(m, r; K) : \ldots$$

Define firstC(B, n) to return the smallest $k \in 0 \ldots N$ such that $D(2k \cdot B)$ is isolating. Then this gives a "split if necessary" algorithm. The problem is that uses exact evaluation: C.

By "soft evaluation" we return $A < B$, $A > B$ or $\frac{1}{2}A < B < 2A$. If C succeeds then C succeeds (but with different parameters)

Use Abbott's improvement of quadratic to give Newton–Bisection processes. If the multiplicity is k we use a k-step Newton.

We can't analyse the analytic algorithm in general but in the case of polynomials it is near-optimal.

3.3 Continuous Amortization: Intrinsic Complexity for subduvsion-bsed ALgorithms: Burr

Prototypical: subdivision for real roots. Many multivatiate analogies.

1. How many subdivisions
2. bit complexity

Challenge is that the algorithm is adaptive. The tree varies in depth and max-depth is not the appropriate measure. Width is often used. Condition numbes are also used.

Definition 3 *$f : \mathbf{R} \to \mathbf{R}^{\geq 0}$ is a*local size bound *if*

$$\forall x \in \mathbf{R} F(x) \leq \min_{J \ni x; C(J) false} w(J)$$

Doesn't depend on onput interval. Measuses local worst-case complexity, We can count the number of sibdivisions by integating the local size function.

Theorem 18 (BurrKramerYap) *$B(I, f) + 1 = \int \frac{2dx}{F}$.*

It matches or improves all known techniques.

Example 11 (Sturm) *$S_{Sturm}(J) = true$ iff J has 0 or 1 roots. So $w(J) \geq dist_{2I}(x, \mathbf{R} - roots(f))$, distance to second-closest root.*

Get an expression in terms of $\ln(\alpha_{i+1} - \alpha_i)$ etc. DMM bound used for these gives equivalent of best results in literature.

Let (X, μ) be a measure space. S a cllection of finite measure subsets of X. Input: $I \in S$, and a stopping criterion. Local Size bound

$$\forall x in \mathbf{R} : F(x) \leq -min_{J \in S; J \ni x; \ldots} \mu(J)$$

$$\sum_{J \in P} g(\mu(J)) \leq \max \left\{ g(\mu(i)), \int_I \frac{g(K \cdot \ldots)}{\ldots} \right\}$$

Again, can match/improve literature's bounds.

3.3.1 Developments

Unequal-sized subdivisions. See [BC11]. Homotopy continuation subdivides the time division. From their paper we know the (lower bound on) size of interval that contains a point and he can regard this as a bound, and reconstruct their research.

Continuous amortisation lead to intrinsic and geometric complexity bounds. Our complexity results are in terms of the actual geometry of the roots.

[PlantingaVegter2004] subdivision-based curve approximation algorithm.

3.4 Davenport

3.5

Drew attention to Alicia Dickenstein conference August 2016.

The previous talk had a lot of resultants! Introduced by Sylvester.

Example 12 *Square $n \times n$ system $\deg(f_i) = d_i$. Consider $\mathrm{Res}_{1,d_1,\ldots,d_n}(1 - tx_i, f_1, \ldots, f_n)$, the roots in x_i.*

What happens in the sparse world?

Consider $n+1$ polynomials in n variables where $A_i \subset \mathbf{Z}^n$ being the support of f_i. What would we mean by sparse resultant? GKZ, Sturmfels. Given us irreducibilty, homogeneities, extremal coefficients etc. but we lose some geometric information.

Self/Galligo/Sombra in AJM2014. Want

$$\mathrm{Res}_{\{0a\}.A_1,\ldots,A_n}(t - x^a, f_1 \ldots, f_n)$$

for all $a \in \mathbf{Z}^n$.

We define the resultant as the definitin equation of the direct image $\pi_* W$.

Example 13 $A_0 =_1 = A_2 = \{(0,0),(0,2),(2,0)\}$.

Now the degree is *always* the mixed volume. We have a toric variety The classic resultant is connected withtheChwo form of thsi variety. We use Rémond's resultant of cycles as a building block. LNM1752 2001.

Define $\mathrm{Res}_X(F_0, \ldots, F_n)$ such that if $|X| \cap V(F_0)$ cuts properly, te

$$\mathrm{Res}_X = \mathrm{Res}_{X \cdot Z(F_0)}$$

The Poisson formula works always with no strange exponents.

D'AndreaJeronimoSombra have generalised work of Sturmfels, which simplifies a lot.

For systems with parameters, the degrees on the parameters are well-controlled.

Chapter 4

6 August 2015

4.1 Algebraic Vision: Reka Thomas

problems in Computer Vision form the point of view of algebraic geometry or algebraic methods in optimisation. Today I will present joint work with Google. See book: Multiple View Geometry in Computer Vision. Hartley,R. & Zisserman,A., C.U.P., 2000.

Definition 4 *A* camera *is a (central projection) map from* $\mathbf{R}^3$ *to* $\mathbf{R}^2$.

Example 14 $(x, y, z) \to (\frac{fx}{z}, \frac{fy}{z})$. *But move to* $\mathbf{P}^3 \to \mathbf{P}^2$*: have* $(x, y, z, 1) \to (fx, fy, z)$.

Now linear, so a camera is essentially a 3×4 matrix P of rank 3. The centre of the matrix is the right kernel. Write $P = [A|b]$. If A is non-singular we say that the camera is *finite*. Use $\sim$ to mean "equal in projection space".

So $P = K[R|t]$ where K is upper triangular with positive diagonal. R is the rotation matrix. A camera is *calibrated* if we know K.

4.1.1 Fundamental Questions [HZ00]

1. **Resectioning** Given $X_i \leftrightarrow x - i$ find P.
2. **Triangulation** Given P_i and x_i find X that gets mapped to each x_i
3. **Reconstruction** $x_i^{(j)}$ find P_j and X_i.

In practice all data are noisy, se we need MLEs. Depends on the noise model.

4.1.2 Two View Geometry

Suppose $x \sim P_1 X$, $y \sim P_2 X$. There is a 3×3 matrix F of rank 2 such that $y^T F x = 0$: the epipolar equation. 7 degrees of freedom. If the cameras are calirbated, ther is an essential matrix E: $y^T E x = 0$ where E has five degrees of freedom.

Theorem 19 (Projective Reconstruction) *Suppose $y_i^T F x_i = 0$. Then we can resocnstruct 2 cameras and the world point X_i up to a projective transformation unless the X_i lie on a livne joining the camera centres.*

Suppose we are given m correspondences $(x_i, y_i) \in \mathbf{R}^2 \times \mathbf{R}^2$: does there exist a reconstruction.

Let a be the vectorisation of the matrix A.

$$\mathcal{R}_2 := \{a \in \mathbf{P}^8_{\mathbf{C}} : \text{rank}(A) \leq 2 : dimension7degree3.$$

$$\mathcal{R}_1 := \{a \in \mathbf{P}^8_{\mathbf{C}} : \text{rank}(A) \leq 1$$

$\mathcal{F} = \mathcal{R}_2 \setminus \mathcal{R}_1) \cap \mathbf{P}^8_{\mathbf{R}}$ = fundamental matrices.

Let Z be the vectorisation of $[x_i \otimes y_i]$. Then the question sis whether $\ker_R(Z) \cap \mathcal{F} \neq \emptyset$

$\text{rank}(Z) = 9$ No solution.

$\text{rank}(Z) = 8$ Compute $a = \ker_R(Z)$. $\exists F \Leftrightarrow \text{rank}(A) = 2$

$\text{rank}(Z) \leq 7$ The intersection is non-empty but it might be entirely in $\mathcal{R}_1$.

$\text{rank}(Z) \leq 5$ Almost always have one: details.

$\text{rank}(Z) = 6$ Might or might not exist.

$\text{rank}(Z) = 7$ Might or might not exist.

Now what happens if we have calibrated cameras? $\sigma_1 = \sigma_2$; $\sigma_3 = 0$. This is a variety $\mathcal{E}$ [Demazure1988]. This is cut out by $2EE^TE - \text{trace}(EE^T)E = 0$, $\det(E) = 0$. The complex $\mathcal{E}$ is a secant variety. $\mathcal{E}_c$ (Zariski closure) is irreducible: dimension 5 and degree 10. So we ask if $\mathcal{E} \cap \ker_R(Z) \neq \emptyset$.

Theorem 20 $\text{rank}(Z) \leq 4 \rightarrow \exists E$.

For rank 8 we just compute $\{A\} = \ker$ and check the Demazure equations. 5,6,7 are the hard cases.

$$\sum a_{i,j}^2 = 4 \sum M^2_{i.j.k.l} : \text{M } 2 \times 2 \text{ minors}$$

There's also a rotational formulation.

The real question is hard: what about the complex form? Rank 5 always exists. For higher ranks we need the Chow form. Suppose $V \subset \mathbf{P}^n_{\mathbf{C}}$ is irreduicble if dimension d. Let $L \subset \mathbf{P}^n_{\mathbf{C}}$ be a linear space of simension $n - d - 1$. Usually $V \cap L = \emptyset$. The Chow form of V, $Ch_V(A)$ is a homogeneous polynomial in A of degree $(d+1)\delta$ such that $V \cap L \neq \emptyset \Leftrightarrow L$ satisfies $Ch_V(A)$.

Shows an example in Macaulay 2: [Nister2004] KTH PhD uses *ad hoc* techniques on Demazure cubics. These days Gröbner trace techniques are used. Kukelova's thesis does this in Android 'phones!

[Sturmfels2014] Hurwitz form of a projective variety. $V \subset \mathbf{P}^n_{\mathbf{C}}$ irreducible of dimension d and degree δ. Intersect with L linear of dimension d. Expect

δ points: when do we get fewer? This defines $\mathcal{H}_V \subset GR(d, \mathbf{P}^n_{\mathbf{C}})$. This gives (Bürgisser) some useful information on the condition number of the reconstruction.

Example 15 (Rome wasn't built in a day) *Project at Washington to reconstruct Rome from all its images in Flickr. First problem is doing the matching.*

Q Real points?

A These are hard questions. Vision considers these problems solved in practice, as they are reconstructing objects that exist!

4.2 Twisted Hessian Curves: Lange

Paper at `http://cr.yp.to/papers.html#hessian`, slides at `https://www.hyperelliptic.org/tanja/vortraege/20150806-squished.pdf`. Note that Google does use elliptic curve signatures. Note that we normally use large finite fields, and Weierstrass forms. Note the problem with addition if $P_1 = \pm P_2$, and special cases of infinity: special cases imply timing risks and bugs.

Hence Edwards[1] curves (example $d = 30$), where all points are equivalent.

$$(x_1, y_1) = (x_2, y_2) = \left(\frac{x_1 y_2 + y_1 x_2}{1 - 30 x_1 x_2 y_1 y_2}, \frac{\dots}{1 - 30 x_1 x_2 y_1 y_2} \right).$$

We say that the addition law is *strongly unified*, i.e. it can be used for doubling. For *complete* [no special caes at all] we need d to be a non-square, but this is a property of d that is relative to the field. Doesn't worry us too much, provided we stick to the appropriate extension fields.

Note [CC86]: much neglected. For some reason they missed Edwards curves.

Weierstrass

Edwards

Jacobi Quartic

Hessian Credited to Sylvester by [CC86]. [JQ01] $2(x_1, y_1 z_1) = (z_1, x_1, y_1) + (y_1, z_1, x_1)$. Claimed to be "unified", but we still need to make sure that we do a rotation for every addition.

We write $H/k : ax^3 + y^3 + z^3 = dxyz$ with $a(27a - d^3) \neq 0$. Use $(0, -1, 1)$ as the neutral element. Addition still fails for doubling. But we have a variant: $x'_3 = z_2^2 x_1 z_1 - y^2 x_2 y^2$ etc. involving a. This is a *complete* addition law if a is not a cube. If a is a cube, they have a strongly unified formulation.

There is a very efficient tripling (no use for constant time encryption, but useful for signature verification etc.). Note that these curves have cofactor 3 in the order and this is the first time this has helped. [BL95, But note there are typos] on addition laws in standard formats.

[1] Today is Harold Edwards' 79th birthday.

Q Analogy for Edwards for hyperelliptic?

A Not known. We have tried, but end up with genus 3.

4.3 Computational algebraic number theory tackles lattice-base cryptolography: Bernstein

Note that the standard sales brochure for lattice-based is powerful.

Problem 1 (Short generator) *Take a degree n number field K. Given the principal ideal $g\emptyset$ find a short g.*

Examples: $\mathbf{Q}$, $\mathbf{Q}[i]$, $\mathbf{Q}[\zeta_n]$, $\mathbf{Q}[\sqrt{2},\sqrt{3}]$. Need $\emptyset$ which might or might not be $\mathbf{Z}[\theta]$, e.g. $\mathbf{Q}[\sqrt{5}]$ has $\mathbf{Z}[\frac{1+\sqrt{5}}{2}]$.

SVP is usually solved by LLL. But for large n, LLL finds short veectors, but not the shortest, and the gap grows exponentially in n. BKZ doesn't actually solve this. [LaarhovendeWeger2015] $\approx 1.23^n$ ([NguyenVidick2008] has $\approx 1.33^n$).

But we can exploit factorisation. Suppose we find α short but $\alpha\emptyset \neq g\emptyset$. Hence produce a lot of short $\alpha_i\emptyset$ and do factor base work in the $\alpha\emptyset/g\emptyset$. Only interested in α with smooth factorisations.

Variation: just igore the $g\emptyset$ and factor $\alpha\emptyset$ into small primes. Does every prime have generators? Also compute $\emptyset^*$ via generators.

[SmartVercauteren????] "exponential time" quote.

There are n ring maps $\psi_i : K \to \mathbf{C}$. Defibe $\mathrm{Log} : K^* \to \mathbf{R}^*$ as $(\log|\phi_i|)$,a nd then $\mathrm{Log}\,\emptyset^*$ is a lattice of rank $r_1 + r_2 = 1$. Use CVP to find elements of $\mathrm{Log}\,\emptyset^*$ close to $\mathrm{Log}\,g\emptyset$. Had a blog post about subfields and relevance. If we know $\mathrm{Log}\,\mathrm{norm}_{k:F}\,g$ for such an F. Then this constrains $\mathrm{Log}\,u$ to a shifted sublattie of $/Log\emptyset^*$: constraints are unit rank of F.

Example 16 ζ_{661} *gives a maximal rank of 8 to be attacked.*

Example 17 $\mathbf{Q}[\sqrt{2},\ldots]$ *degree* 2^{10}*, means the whole problem is trivial. Admittedly no-one's proposed this field but all the old hardness arguments work for it!*

[Campbelletal2014a] shows a textbook attack on cyclotomics. The analysis is bogus, but the algorithm is very fast. Plagiarised by [Crameretal2015], which does the analysis correctly. [Song2015] produced a polynomial-time quantum algorithm.

4.4 Encryption based on card shuffle: Lee

Consider block ciphers and the indistinguishability framework: adversary capable of making adaptive forward and backward queries.

Example 18 *Credit card numbers. Need to transform to/from bit strings, but we'd rather have* `{0-9}`^{16} *transforms. Feistel networks are standard, but are only secure up to* $2^{n/2}$ *queries for a sufficient number of rounds. Credit cards are too small!*

Consider card shuffle: a Markov process whose mixing time is the number of rounds. We want it to be oblivious: trace one card ignoring the rest. Claims the Thorp [Crypto2009] this is secure up to $2^n/n$ queries for $O(n^2)$ rounds.

"swap or not" shuffle [Crypto2012] Chose a round key $K \neq 0$ for $\{1,1\}^3$. the cards at $x, x \oplus K$ are matched, and swapped or not as $\max(x, x \oplus K)$. Can view this as choosing permutations.

New construction "partition and mix": For each elements, choose $D-1$ distinct elements at random, and arrange elements in block. Need an "ϵ-alomst D-uniform" partition, and thsi reduces the number of rounds by $\log \frac{D}{1+\epsilon}$.

But finding such partitions is not trivial.

Definition 5 *A family of permutations on N elements in* perfect D-wise independent *if it acts uniformly on tuples of D elements.*

However, there are no non-trivial subgroups of S_n for $n \geq 25$ which are 4-wise independent.

Alternative technique via Hamming codes. Can extend to $[2^s-1, 2^s-s-1, 3]$-Hamming, which is 2^{D-n}-almost 2^s-uniform. Claims 60 rouns rather than 450 for the SN-shuffle.

Conclusion: claims that this is useful for format-preserving encryption.

4.5 A class of constacyclic codes over $\mathbf{F}_{p^r} + u\mathbf{F}_{p^r} + v\mathbf{F}_{p^r} + uv\mathbf{F}_{p^r}$: Bandi

Classically we consider codes over finite fields. [Hammendsetal1994] initiated study over rings. Various authors have studied the title ring. Apparently several good codes have been produced over rings, better than over fields. Note that $u^2 - v^2 = 0$ and $uv = vu$.

Let $\tau : (c_0, \ldots, c_{n-1}) \mapsto (\delta c_{n-1}, c_0, \ldots, c_{n-2})$. Costacyclic means invariant under τ for a suitable δ.

Let $R_n = \frac{R[x]}{\langle x^n - (1+\lambda u)\rangle}$. This is a local ring (but not a chain) with maximal ideal $\langle \ldots$.

Can count all these constacyclic codes, and can produce the duals.

Q Why is this ring interesting?

A So far unexplored.

Q What about efficient decoders?

A Not solved, apparently.

Q Why $1 + \lambda u$.

A It's a generic linear after scaling.

4.6 Challenges in the Development of Open Source Computer Algebra Systems: Decker

In charge of the Singular project since 2009. I learned to use computer algebra via the original Macaulay. Numerical methods (e.g. Section 9.3) are important, but not the only answer.

Report on OS software from the DfG priority project. Unlike 20 years ago, methods from cmputer algebra are now firmly established in the toolbox of the pure mathematician. A decisive feature of the current developments is that more and more abstract mathematical concepts are being made constructive. Algebraic geometry by itself is not sufficient.

In this project, Malle has considerably strengthened Cohen–Lenstra, and I now call then Cohen–Lenstra–Malle. Demonstrates the Zbl citations for Singular by MSc categories. These days over half the citations are outside Algebraic Geometry/Commutative Algebra.

Want to intertwine Singlaur, GAP and Polymake. I believe that Antic will become a major tool in number theory.

4.6.1 First Challenge: Faster Algorithms

First need to convert curve into primary decomosition, hence factoring. Need lots of algrotiths as there is no universal one.

Example 19 *Phylogenetic modelling. Engelmann in Singular solved the problem: 26 CPU days and a GB with 416812 elements. Used a Hilbert-driven GB computation.*

Example 20 (Gröbner Bases over Number Fields) *Use modular reconstruction, but choose primes such that the minimal polynomial factorises: many more smaller problems. Note this is a highly parallel problem.*

1. Coarse-grained. Comparatively easy.
2. Fine-grained. This needs thread-safe, but optimal, memory management. A major project.

GAP Largely done: see HPC-GAP. Developer nowmoved to Kaiserslautern.

* Racing multiple algorithms. Needs Coarse-grained, and `ParallelWaitFirst`. There is also `ParallelWaitAll`.

Note Villamayor's constructive version of Hironaka. One major problem is choosing the "right" order of blowups.

Example 21 (De Rham) *Use the Weyl lagebra to compute the de Rham cohomology of complements of affive varieties.*

But this needs the BGG correspondence [Bernstein–Gelfand–Gelfand]. Let V be a vector space of dimension $n+1$ with dual space W. $S = Sym_K(W$ and $E = \bigwedge V$. We grade S and E by letting elements of W have degree 1 and V have degreee -1.

Shows a session. Starts with GAP's `SmallGroup(1000,93)`. Then calls Singular from Gap, to get Tate resolutions.

4.6.2 Third Challenge: Making More of the Abstract Concentps Constructuve

Aim: Fourier–Mukai transforms and their generalisations.

4.6.3 Integration of Systems

One example already. Also shows interatcion with Polymake. Computes a GIT-fun from Polymake and Sigular.

FLINT Important basic operations.

ANTIC fast number feild arithmetic.

HECKE an implementation of algebraic number theory in Julia.

Also tools in tropical geometry essentially a piece-wise linear version.

FAN computes tropical varieties, and trivial valuations.

ATINT Tropical intersection theory.

Shows an example involving Chow Rings and its `TopChernClass`.

Q How does this relate to SAGE?

A We want to conect systems together directly, rather than via the SAGE kernel.

Q Representation Theory?

A GAP's Chevy system does something, but also Cohen's Lie.

Q Can you study memory-sharing as a Gröbern base problem itself.

A Not yet!

Q You showed many pictures: how do *we* produce them.

A These pictures were produced via Greuel's Imaginary: you should install these. `surf.lib` is the ray-tracing starting point.

4.7 Primary Decomposition and Parallelization: Schönemann

This is a basic tool. Computers are not getting faster, just more cores. Hence a change of approach is necessary. Singular's memory-management is very suited to GB, but is not thread-safe. Hence more common to use communication rather than memory-sharing. Classically: GB via CRT. Can use multiple threads in one processor for matrix operations (F4-style), as the sub-division is well-understood.

MathicGB (Roune) computes GB's via a matrix of machine integers.

In general, have a multi-area scheme: some exclusive to a thread, some shared and lockable.

Wu-Ritt

Gianni–Trager–Zacharias Here we add a pre-rpocessing step. We need dimension hence need a GB. Actually use a factoring GB, since this will automatically contribute to the primary decomposition. In practice this gets us most of the way very often. However, this also imposes inequations. Therefore should treat lowest-degree factors first. Use work-stealing on the factors

Eisenbud–Huneke–Vasconcelos Theorem 21 *If* $I \subseteq R = K[x_1, \ldots, x_n]$ ...

We can't yet parallelise the characteristic sets algorithm.

Factorising Gröbner is implemented, but not in parallel.

A good discussion on factoring, and JHD mentioned experience from [Dav87].

4.8 Criteria for Gröbner Bases: Gao

G is a Gröbner basis for I iff every polynomial $h \in I$ is top-reducible by G. Not algorithmically testable, hence the S-polynomial criterion. Hence LCM criterion, [MMT92], then Faugère's F4. [EderFaugere2014]: survey paper.

$$H := \{(u_1, ldots, u_m) \in R^m : \sum u_i g_i = 0\}$$

is calle dthe *syzygy module* of $\{g_i\}$. Want a term order $\prec_2$ on R^m which is compatible with the order $\prec_1$ on R. [Fau02] the signature of v is $\min\{\mathrm{lm}(u) : u.g = 0\}$. Reduce pairs $(u, v) \in R^m \times R$ by reducing the R part and tracking the changes in R^m part. We only reduce $(u_1, 0)$ by terms of the form $(u_2, 0)$.

$$M := \{(U, v) \in R^m \times R : u.g = v\}$$

and have a concept of a string GB of M. Then the elements with zero R-part are a GB for the syzygy module, and the projection onto R are a GB for the original.

If both $v_1 v_2$ are non-zero then the J-pair consists of doing the S-polynomial computation on both components. Let $T = \max(t_1 \operatorname{lm}(u_1), t_2 \operatorname{lm}(u_2))$.

Theorem 22 (us) *The following criterai are equivalient.*

1. *G is a strong GB for M*
2. *Every J-pair of G is covered by G.*

[RouneStillman2012a] etc. all have rules about "rewritable" when $tlm(v_2) \prec lm(v_1$ and "added later tan" rule. The last is not mathematical.

Note that we only store the signature $lm(u)$ not the whole u.

4.9 Modular Techniques in Computational Algebraic Geometry:

Rational reconsruction is an old idea. Preimage under the Farey map. This works as long as N is large enough, and none of the primes are bad.

1. Input modulo p is not valid
2. Algorithm fails (e.g. matrix not invertible)
3. Computable invariant is wrong (e.g. Hilbert polynomial). These primes are usually Zariski-closed in Spec $\mathbf{Z}$.
4. Computable invariant with unknown value is wrong: only solution is majority voting.

All $(x, y) \in \Lambda$ which reconstruct correctly are in a straight line. If M (bad primes) are small enough, Gauss–Langange will find the shortest vactor, and its norm divides the bad primes (take these out).

Hence this algorithm.

1. Compute I_p
2. Reduce P according to majority vote on $LM(U(p))$
3. compute termwise CRT-lift
4. Lift U(N) to error-tolerant rational reconstruction
5. Test on a new random prime
6. Verify
7. If any stage fails, repeat

Example 22 *$I \to \sqrt{I + Jac(I)}$ where*

Normalisation: $\overline{A}$ is the integral close of A in its quotient field.

Example 23 *Curve* $I = \langle x^3 + x^2 - y^2 \rangle$

Hence this algorithm.

1. Start from $A_o = A$ and J

Theorem 23 (GrauertTemmert)

Theorem 24 (201) *Suppose* $J = \sqrt{Jac(I)} = P_1 \cap \cdots P + r$ *is a primary decomposition, and* $A \subset B_i \subset \overline{A}$ *is the ring given by normalising* P_i*. Then* $\overline{A} = \sum B_i$*.*

Hence adjoint ideals.

Q Worst error fraction?

A Never occurs in practice! We actually start at laregst prime and work down.

4.10 Computing Integral Bases of curves in small characteristic: Stillman

This is about embracing the bad primes. k is $\mathbf{F}_p$ or $\mathbf{F}_q$, wth p small. $f(x, y)$ is an equation of a plane curve (think irreducible). Monic in y of degree n. $L; = k(C) = k(x)[y]/(f)$. Assume separble. Let $\emptyset \subset L$ be the integral closure of $k[x]$ in L. Suppose $P(x) \in k[x]$ is irreducible let $\emptyset_p$ denote the integral closure of $k[x]_{(P)}$ in L.

$\emptyset$ is rank n free $k[x]$-module. $\emptyset_p$ is a rank n free $k[x]_{(P)}$ module. If P^2 doesn't divide the discriminant then $\emptyset_P$ is trivial.

Definition 6 *A partial basis of* $\emptyset_p$ *is a set* $B = \left\{1, \frac{g_1(x,y)}{pd_1}, \ldots\right\}$ *where the* g_i *are monic in* t *of degree* i*, the fraction is integral over* $k[x]$*,* $0 \geq d_1 \leq d_2 \leq \ldots$*. Let* $L(B)$ *be the* $k[x]_{(P)}$*-span of* B*.* B *is a full basis if also* $L(B) - \emptyset_P$*. The delta invariant at* P *is* $\delta_p = \sum d_i$*.*

Example 24 $F = y^9(y-1) + (x^3 + x^2 + 1)^2 y + (x^3 + x^2 + 1)^3$*. 4 points in singular locus (one over the base field,*

Trager computed $\emptyset$ via "round 2". ALso [vH94] uses Puiseux series for large enough characteristic. [LeonardPellikaan2003]. Montes algorithm.

If I can compute integral basis for $\emptyset_x$ for $f' \in k'[xy]$ with k' a finite extensin of k. ....

General idea $g \in \emptyset_x$ iff $g^p \in \emptyset_x$ since

$$\emptyset_x = \{g \in L | v(g) \geq 0 \text{ for all valuations centred at } x = 0\}.$$

1. Start with a partial basis B of $\emptyset_x$ which is Frobenius stable.

2. For the momemt, assume that $k = \mathbf{F}_p$

3. Let M be the $n \times n$ matrix pf σ w.r.t. B. Since $L(B)$ is Frobenius-stable, M is a matrix of polynomials

4. For $c_0, \ldots, c_{n-1} \in$
$$\sigma\left(\frac{c - 0b_0 + \cdots}{p}\right) \ldots$$

5. Do this until $L(B) = \left\{g \in \frac{1}{x}L(B) | g^p \in L(B)\right\}$: Frobenius-stable.

* Nearly there: may need one more x.

6. Compute W_1 the kernel of ...

A singular point above 0 always leads to a fraction: quite often we can compute these "for free". $f(x,y) = TS^2 + xSU + x^2V(x,y)$ where $S, T, U \in k[y]$ and S is square-free and irreducible factor of T of multiplicity ≥ 2 does not divide U, then we have precisely the right fraction.

4.11 SIAM AG Business Meeting

Jan Draisma chaired the meeting. Note that we have `http://wiki'siam.org/siag-ag/index.php/Main_Page`. Note that you need to be added to the Wiki by a member.

Conference numbers over the history: 307; 386; 350 (at this year). Biggest AGs are 1500, 500 is probably median. Smilar size to Discrete Math. Note SIAM fellows: 2014 Sturmfels and 2015 Charles Wampler.

4.11.1 AG2017: Anton Leykin (Georgia Tech)

Beginning of August 2016. Duration 4 or 5 days — previous meetings were $3\frac{1}{2}$ days, but this is 5. We have a big airport in Atlanta. 25 minute subway to campus in 25 minutes. Many hotels in walking distance with discounted rates.

Q Can we avoid competing with MAA MathFest: we have last two times.

A Note that SIAM Conference is also around this time, and we can't conflict with that.

Q–Ottaviani MEGA is a conflict. A show of hands said that almost everyone present had been at a MEGA. MEGA 2017 will be in Nice in June.

Alicia MEGA was even, but went to odd because of ISSAC.

JHD But ISSAC is now every three years in Europe. We can live with 2017, but should think about 2019 if AG2019 should be in Europe. Note ICIAM 2019 is in Valencia.

All Also overlap with CRYPTO in Santa Barbara in late August.

Sandra Possibly merge the two when AG is in Europe.

Sturmfels/JHD Motion for the leaders of AG and MEGA to discuss. Fortunately there is intersection! Notably Jan Draisma himself.

***** A poll showed that about half the room would be going to ICIAM 2015.

Note that we are *recommending* to SIAM that Georgia Tech be approved.

4.11.2 Also

2016 will be SIAM AM July 11-15 with AMS Invited Lecture by Sturmfels. SIAM also has funding for summer schools.

We have 90 student members and 148 non-student. %age students is sligtly lower than SIAM average. However student members (of the AG) have been declining, attributable in part to a change in SIAM software. Academia is 84%, whereas SIAM average is 75%. 75% maths departments, 10% CS and the rest "other/none". US is about 2/3 of the total.

4.11.3 SIAM J. Applied Algebra and Geometry

Note that it is not formally an AG journal. Proposalwokred on by FS and TT. Proposal goes to SIAM Trustees this weekend.xs

Q Open Access?

A SIAM model (which itself is evolving).

Q Can we add "Applied Topology"?

A We should certainly reach out to that community.

Q Editorial board, process, terms etc.

A Being discussed. SIAM has guidelines.

Chapter 5

7 August 2015

5.1 Progress Report on Geometric Complexity Theory: Mulmuley

[Ikenmeyeretal2015a] is the main reference.

Problem 2 (The permanent verses determinant problem) *The permanent of an $n \times n$ variable matrix X cannot be approximated infinitesimally closely by symbolic determinants of $m \times m$ matrices whose entris affice linear combinations of the entries of X if m is polynomial in n.* This strengthens Valiant's conjecture, but is reducible to looking for representation-theoretic obstructions.

Let $G = GL_l(\mathbf{C})$ wheren $l = m^2$. Let $\lambda : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l \geq 0$ be a partition and let $V_\lambda(G)$ denote the Weyl module (irreducible representation) of G.

An obstruction (for given m and n) is a Weyl module $V_\lambda(G)$ that occurs on a certain G-orbit closure associated with the (padded) permanent but not on the G-orbit closure associated with the determinant.

Conjecture 1 *An obstruction exists if $m = poly(n)$.* This implies the truth of the previous problem.

Let $H = GL_m(\mathbf{C}) \times GL_M(\mathbf{C})$. Let ρ denote the natural monomorphism from G to $G = GL(\mathbf{C}^m \otimes \mathbf{C}^m) = GL_l(\mathbf{C})$. Given partitions $\lambda\mu\pi$, the *Kronecker coefficient* $K^\pi_{\lambda\mu}$ is the multiplicity of the irreducible H-module $V_\mu(H) \otimes V_\pi(H)$ in he irreducible G-module $V_\lambda(G)$ considered as an H-module via ρ. If λ is an obstuction for given n and m then

1. $|\lambda|$ is divisible by m
2. the height (numbe rof non-zero parts) of λ is $\leq n^2$
3. The rectangular coefficient $k^\lambda_{\delta(\lambda),\delta(\lambda)}$ vanishes.

Note that these are necessary, not sufficient. But satisfying them is challenging enough.

[Kirwan1984] the Kronecker cone for a given m is a polyhedral cone. BUt [Burgisseretal] $(\lambda, \delta(\lambda), \delta(\lambda))$ lies in the Kronecker cone, so can't be proved zero this way. This shows we can't use asymptotic techniques of algebraic geometry and representation theory on the Kronecker cone used to show the existence of vanishing rectangular Kronecker coefficients.

If $V_\lambda(G)$ is an obstruction for given $m = n^a$ and n, then the partition trible $(\lambda, \delta(\lambda), \delta(\lambda))$ must be (ϵ, b)=exception in teh following sense: letting $\epsilon = 2/a$ and b large enough: wher $(, \mu, \pi)$ is exceptional iff

0 $k^\lambda_{\mu,\pi} = 0$

1. $\mu = \pi = \delta(\lambda)$
2. height$(\lambda \leq m^\epsilon$
3. etc.

Hence our intermediate goal is to show that superpolynomiallly many exceptional partition triples (λ, μ, π) exist, as $m \to \infty$, for any fixed $\epsilon > 0$, with a large enough constant b depending on ϵ.

Hard. Let's relax condition 1 and just insist that $\mu = \pi$, but not necessarily rectangular.

Theorem 25 *For an $0 < \epsilon \leq 1$ there exists $0 < a < 1$ such that, for all n, there exist $\Omega(2^{m^a})$ partition triples such that [various items which meet the relatxed goals above].*

The proof uses the theory of NP-completeness. It *explicitly* constructs the obstructions.

This theorem disproves the conjecture that KRONECKER is in P [unless P=NP, I suppose]. Note that deciding positivity of Littlewood–Richardson. coefficients is in strong P [Knutson;Tao].

Conjecture 2 (GCT6) *There is a #P formula for the Kronecker coefficients.*

This is the complexity-theoretic version of the clasical problem of finding a positive rule for the Kronecker coefficients. analogous to the positive Littlewood–Richardson. rule.

All positive rules known so far for restricted classes of Kronecker coefficients. (such as Littlewood–Richardson) are for subclasses of partition triplies of type P.

The following result provides the first known instance of a positive rule for Kronecker coefficients. for a subclass of type NP.

Theorem 26 *There is a #P formula ...*

The next aim is to extend this explicit proof strategy to ..., but the problem is the "GCT chasm": the existing EXPSPACE versus P gap in the complexity of derandomizing Noether's Normalisation Lemma for explicit varieties.

Note that, in fact we only need rectangular Kronecker coefficients.

Q Can you explain NNL?

A Another lecture, but note that if we could solve permanent/determinant, we should close the chasm.

5.2 Homotopy continuation versus Gröbner bases for parametric systems: Leykin

Given $\phi \in K[px]^m$ and $V \subset]A_p = Spec(K[p])$ such that, for a generic $p_* \in V$ the set of solutions $\phi^{-1}(p + *)$ is finite, "get" this set.

5.2.1 Gröbner Trace

This of a parametric or comprehensive GB of $I = \langle]phi(p, x)\rangle$. There is an open $U_p \subset V$ and $G \subset K[p, x]$ such that $P(p_*, x) \subset K[x]$ is a GB for all $p_* \in U_p$.

But actually computing a comprehensive GB, so we use Gröbner trace ideas. This is a procedure which evaluates the coefficients of $G(p_0, x) \subset K[x]$ for a given $p \in V$, which are rational functions in p_0. We only tae extensin when we "solve".

5.2.2 Parametric homotopy

Classically, V is the dense open subset of $\mathbf{A}_p$. Generally

1. Take a generic (codim(V)-plane $L \in \mathbf{A}_p$
2. Find a structured witness set
$$V(\phi) \cap (L \times \mathbf{A}_x) \subset \mathbf{A}_p \times \mathbf{A}_x$$
3. Given $p_0 \in V$, pick a general $L_0 \subset \mathbf{A}_p$ that contains p_0
4. Deform from L to L_0.

Example 25 (Vision) *X is projected by three calibrated cameras $R_1 = I, R_2, R_3 \in SO_3$, and centers $C_1 = 0, C_2, C_3$. There is related work (Kileel) on calibrated trifocal varieties. Use the Cayley parametrisation of SO_3, and two matrices means six parameters. Write down a map*

$$\mathrm{Proj}(K[X, C]) \times SO_3 \times SO_3 \to (\mathbf{P}^2 \times \mathbf{P}^2 \times \mathbf{P}^2)^4$$

in some affine charts. $\dim(C) = 23$, $\dim(V) = 24$, *$\partial' phi$ is of full rank, so we have a codimension 1 variety.*

A vision problem is minimal *if a generic fiber $\phi^{-1}(y)$ is finite and the number of points and cameras is optimal. [HoltNetravali1995] for four points in 3 views, $|]phi^{-1}(y)| = 1$ for a generic view.*

Homotopy Take a real line $L_1 \in \mathbf{A}_y$ containng y_1 witha known solution and moving it to a real L_2 containing y_2 *works often* if $Y_1, , y_2$ are "not too far apart". Complex L_1, l_2 (in fact $L_1||L_2$) work even better. If single-path tracking succeeds, homotopy takes ≈ 0.1 seconds. But using the whole witness set (4000 points) is slow, and evaluation of ϕ and its derivatives is the bottleneck: problem with Macaulay2 currently.

Gröbner Computation of $\langle \psi(y_0, r)$ rank for exacty_0 takes $\approx$ 1sec (unfeasible).

In Macaulay2 use `GroebnerBasis` by Roune, with `Strategy=>"F4"` option, which beats Faugère's F4.

5.3 Integral bases via localisationa nd Hensel Lifting: ?Lapaigne

Let $A = k[x_1, \ldots, x_n]/I$ where I is a prime ideal. $x \in Q(A)$ iff it satsifies a *monic* equation. $\overline{A}$ = set of integral elements in $Q(A)$.

Example 26 *On $y^2 - x^3$, y/x is integral, but y/x^2 is not — consider Puiseux series around $x = 0$.*

An integral basis is a se of $k[x]$-module generator sfor $\overline{A}$.

Example 27 $I = \langle y^3 = x^2 \rangle$. *IB* $1, y, \frac{y^2}{x}$.

Factorisation is ring of Puiseux series. Therefore valuations and integral exponents = $\min_{\text{places } \gamma_j} v_j(y)$.

So we localise at the various singular points. If P_i is a singular point, let $A^{(i)}_{P_i} = \overline{A_{P_i}}$ Then $\sum A^{(i)} = \overline{A}$.

Note that $f \in k[x][y]$ can be factored in $k[[x]][y]$. But need to use CRT to reconstruct the answer from the different f_i.

Q Are the bases triangular?

A At the places P_i — yes. And at the end.

5.4 Gröbner Bases for Algebraic Number Fields: Decker

JHD observes that this seems to be [BFDS15].

As I said yesterday, we are revisiting basic algorithms. Given $I \subseteq K[x_i]$ where $K = \mathbf{Q}[\alpha]$. We think of $\tilde{I} \subset \mathbf{Q}[X, t]$ with f added. We know about modular/CRT Grp̈bner bases.

However, we also choose primes such that f factorises, so we have two levels of CRT-ing. Let $S = \mathbf{Q}(\alpha)[X]$ and $T = \mathbf{Q}[X,t]$. Fix a global product order $\succ_K = (\succ_1, \succ_2)$. Given $H = \{g_i(X,\alpha)\}$.

Theorem 27 *Let $\tilde{G}$ be the reduced GB of $\tilde{I}$ wrt $\succ_K$. Then if $\tilde{I} \neq 1$ ...*

[Noro2006] observed that we get many $t^b X^a$ which slow down the process. He therefore went monic in $\mathbf{Q}(t)[X]$.

Open question for us: what is a "good" factorisation?

Definition 7 *Take p dividing no (numerator or denominator) coefficient in the input f. Also [presumably f sqfr]. Then p is* admissible.

Let f_{ip} be the factors of $f \pmod p$.

Definition 8 *We say that p is* admissible of type B *if the sizes of $\tilde{G}_{i,p}$ and $\tilde{G}_{j,p}$ coincide.*

Definition 9 *Let $\tilde{I}$ be as above. Then p is lucky iff* $\mathrm{LM}(\tilde{G}_p) = \mathrm{LM}(\tilde{G})$. *This can only be etsted a posteriori.*

Use majority voting [IPS11] and then use CRT and rational erconstruction to produce a $\tilde{G}$.

Randomly chooose an additional prime p and check that $\tilde{I}_p$ reduces correctly an vice versa.

Theorem 28 ([Arn03, Pfi07]) *If we are homogeneous, if $\tilde{I}$ reduces to zero w.r.t. $\tilde{G}$ and if $\tilde{G}$ is the reduced GB of $\langle \tilde{G}$, then I is the GB required.*

Table of experimental data. Magma is mostly slow/timeout (on his examples!). New algorithm (on 32 cores) is significantly faster than the sequential non-modular algorithm.

Q-JHD How many primes?

A 12 such that there are at least three factors of f modulo p, then 12 more such and so on.

5.5 Tropical Homotopy Continuation: Jensen

Goal: mixed volume computation. $a \oplus b = \max(a,b)$; $a \odot b = a + b$. Tropical polynomials are piecewise linear. For $f \in \mathbf{R}[x_1, \ldots, x_n]$, the "tropical hypersurface" is defined s $T(f) := \{\omega \in \mathbf{R}^n : \max_i(c_i + \langle a_i, \omega\rangle)\}$ where its attained twice.

Volume $\sum \lambda_i C_i)$ is a polynomial, an dthe coeficients of $\prod \lambda_i$ os the mixed volume. BKK: The number of solutions is ....

Why is this tropical? To compute the mixed volume of Newton polytopes we find a subdivision of the Minkowski sum using a suitable lift.

Description of homotopy: divergence, collison do we find isolated solution? ALl these problems arise tropically.

Input Polynomials

Output Isolated SOlutoins

Tropical variant

Input

Output

tart with knwon volumes and their tropical hypersufrtace, how does this evolve.

The regular subdivisons of $\text{conv}(A_1)+\text{conv}(A_2)$ correspond to Cayley configurations. Wrte $u \sim v$ if they induce the same regular triangulation of $\text{conv}(A)$. Inspired by Gröbner fan ideas.

Question: which inequalities guarantee the existence of a given cell. Can we treat each mixed cell separately? We track the cells as we change the coefficients. As we evolve, cells can split, and paths can collide. We use Reverse Search ideas [AvisFukuda].

Regeneration [HauensteinSommeseWampler2011]. First solve a linear system. Do this for enough linear l_1 and interpolate to solve $f_1 = l_2 = \cdots = l_n = 0$, and continue. This is what to tropicalise.

I know the mixed volume of simplices = tropical lines. Shows an evolution (note via discrete transformations) from this to desired state.

Note several competing ideas, including [Mal14] — first tropical method. Shows Cyclic n. Li 2007 and Li 2011 (both use random floating-point lifts) scale by $\times$8–9 as n increases by 1, Malajovich by 4–5, me 5–7. My parallel (16 threads) seems to be 4–11. To JHD, Malajovich's times looked generally the best.

Q–Malajovic? There are two algorithms in my paper: one guaranteed and one random. Which one?

A I'm not sure.

***** The times were actually from the "guaranteed" version of Malajovich.

5.6 Lattices over Polynomial Rings and Applications to Function Fields: Bauch

Let $C/\mathbf{F}_q$ be a smooth curve determined by $y^n + a_1(x)y^{n-1} + \cdots a_n(x) = 0$. Hence $Jac(C)$. For elliptic curves, $Jac(E) \equiv E$. In general, how do we do arithmetic in Jacobians?

Consider $F = F(c) := \mathbf{F}_q[x, y]/(f)$. f assumed irreducible and separable. $\mathbf{P}_F$ is the set of places of F/k. Let $P_0(F) = \mathbf{P}(F) \setminus \mathbf{P}_\infty(F)$. $D_F = \{\sum_{finite} \lambda_P P : P \in \mathbf{P}(F)\}$.

Denoting classes by $[D]$, then $[D_1] = [D_2]$ is $\dim(L(D_1 - D_2)) > 0$.

Let $\emptyset''_F = Cl(\mathbf{F}_q[x], F)$ the finite maximal order. Also $\emptyset_{F,\infty}$. Write a divisor as finite+infinite and represent a sum of two ideals $\prod_{Q\in\mathbf{P}_{F,0}} Q^{-a_Q}$ and infinite equivalent.

Define degree of a rtonal function $|a/b| = \deg a - \deg b$. A basis is reduced if $||a_1b_1 + \cdots|| = \max|a_ib_i|$ for all a_i. There is a reduction algorithm.

Define $||\cdot||_D : F \to -infty \cup \mathbf{Q}$ by $||z||_D = -/lim_{P\in\mathbf{P}_\infty(F)} \left\{ \frac{v_p)z)+v_P(D)}{e(P/\infty)} \right\}$.

Theorem 29 *$(I_D, ||\cdot||_D)$ is a lattice., and a reduced basis has ...*

Let $sm(D) := sm(D, ||\cdot||_D)$.

Definition 10 *D_1 and D_2 are isomentric if same sm. Hence isomery class.*

If I have an A-basis $B)0$ and an orthonormal B_∞ then ....

If $r \in \mathbf{R}$ and $D'-D+r(x)_\infty$ Then this corresponds to the lattice $x^r I_D, ||\cdot||_D - r)$. Theorem states number of arithmetic operations: $O(n^5(h(D) + n^2C_f)^2 + n^{5+\epsilon}C_f^{2+\epsilon}\log q$.

Theorem 30 *Let B be a set of F having n leements. It is a$\mathbf{F}_q[x^{-1}]_{(x^{-1})}$-integral pasis of $\emptyset_{F,\infty}$ iff F s w-semi-reduced.*

So if we make our sums w-semi-reduced we have good arithmetic.

Q Plane curves only?

A So far yes.

5.7 On the Existence of Semi-Regular Sequences: Hodges

These correspond to systems of equations that are the hardest to solve for GB. Experimentally, most systems are. But we cant prove anything that corresponds to this. Systems of equations over a finite fields (generally $\mathbf{F}_2$). $p_i(x_1, \ldots, x_n) = \beta_i$. In genrla the ones in crypto are not smei-regular. See [BardetFaugèreSalvyYung]. Intutively,the polynomials are as independent as possible. Assume $x_i^2 - x_i$‘$foralli$.

Definition 11 *Let $\lambda_1, \ldots, \lambda_m \in B$ be a squence of homogeneous elements of positive degrees d_i and $I = (\lambda_1, \ldots)$. The sequence is smei-regular over $\mathbf{F}_2$ if ro all $i = 1, 2, \ldots, m$*

$$\lambda_i : \left(\frac{B}{\ldots}\right) \mapsto$$

The truncation of a power series is immediately before the first *non-positive* coefficient.

Theorem 31 (Bardet...) *The sequence is semi-regular iff the Hilbert series is $\left[\frac{(1+z)^n}{\prod(1+z^{d_i})}\right]$.*

Data: 20 sets of homogenepus quadratics, for varous m, n. Most combinations had 100% semi-regular. Can't even prove they always exist.

Conjecture 3 *1. For fixed m, the semi-regular tend to 0 as $n \to \infty$*

2. *As $n \to \infty$ the proportoinof sequence so flength n that are semi-regular tends to 1.*

3. *As $n \to \infty$ the proportion of all sequences in n variables that are semi-regular tends to 1.*

Proved 1 and 3: no progress on 2.

Keep $m = 1$ and fix degree. Varies dramatically on whether $n - d$ is odd (often 100%) or even (sometimes 0%). If $d = n, n-1$ then all elements are semi-regular. Can prove there are no semi-regular of degree $d > n/3$. If $n - d = 2s$ and $\left(\begin{array}{c} n \\ s \end{array} \right)$ is odd, there are no semi-regular.

Conjecture 4 *If $n - d$ is odd, most sequences are semi-regular.*

Have a result on when elementary symmetric polynomials are semi-regular. In particular when $d = 2^k$ ....

Problem 3 *Let$\pi(n, d)$ be the proportion of sequences that are semi-regular. Show this $\to \infty$.*

Also, what about $\mathbf{F}_q$?

5.8 New Results in Linear Cryptanalysis of DES; Semaev

$E_K(P_i) = C_i$ is encryption. Assume a lot of P_i, C_i pairs.

Average number of sides in the final equation $s/approx 2^{\text{rank}(J)} \prod_i^N (-\alpha_i)$.

2^{42} plain/cipher blocks and some approximations. Success probability seems o be 0.89. Experimental verification for 8-round DES. Note that this isn't limited to DES. (But he doesn't discuss how to find approximations.)

5.9 Enumeration and Gröbner Bases Methods on Solving Generic Multivariate Polynomial Systems: Yang

$MQ(m, n, q)$ problem: find a solution to a system of m quadratic equations in n variables over $\mathbf{F}_q$. More precisely. For any probailistic Turing macjone A trying to solve a MQ systems with randomly-drawn coefficients where $m/n = c + o(1)$ and sub-expoential functions $\eta(n)$, the probability that A returns the correct answer in time $\eta(n)$ is negligible. [Patarinetal].

"If (name==Faugère) then use F5, else use F4 in Magma" (except that Jintai Ding claims MutantXL).

For $\mathbf{F}_2$, brute force is often the best, otherwise anymptotically XL with sparse solver is best and for large fields with $c > 1$ then XL with sparse solver is often best.

XL was first suggested [Laz83], rediscovered by [Courtoisetal2000Eurocrypt]. Let $T^{(D)} = \{\}$ and $T := |T^{(D)}|$. Multiply every equation by every monials as long as degree $\leq D$. Then solve linear equations.

If we expect 1 solution, we can use sparse solvers. Claims that you can throw away rows at random (to get a sparse system) without losing solutions [JHD: surely this is obvious: a row is an extra equation]. Claims that XL2 [CourtoisPatarin2002] — suppose that we only manage to eliminate the toplevel monmials then multiply prepeatedly by others.

If we assume usual regularity conditions then #monomials= $[t^D]\left((1-t^q)^n(1-t)^{-(n+1)}\right)$. Also #free monomials equation.

Note that he is talking about generic equations, so this isn't HFE etc.

Courtois+Pieprzyk overclaimed efficiency of ZXL in2002. [Bardetetal2004] derives $D = (0.0090 + o(1))n$ for F5. [Bettaleetal2008] suggest guessing 0.45 of the variables. [Bouillaguetetal2013] brute force attacks on F_2 run very well on GPU/FPGA. Claims that you may think GB overtakes enumeration at 200, but we think the hardware effects actually moves this closer to 400.

Record holder for MQ challenge III ($m/n = 2, \mathbf{F}_{31}$) are us, using XL with Wiedermann. 32 days on 64core AMD 6282SE (4 sockets) with 512GB RAM.

Conclusion: Brute Force is probably the best way even for quadratics, an certainly for higher degrees.

5.10 Hodge Theory for Combinatorial Geometries: Huh

Three fundamental ideas:

1. A matriod is apiecewise linear object — tropical linear space [Sturmfels].
2. Hodge structure on the cohooly of projective toic varieties produces fundamental combinatorial inequalities [Stanley]
3. g-conjecture for polytopes can be proved using the "flip connectivity" of simplicial polytopes of given dimension [McMullen]

So consider a graph (vertices and edges). $\chi_g(q)$ =#numbr of proper colourings of G with q colours. For a square we get $q^4 - 4q^3 + 6q^2 - 3q$. always

$$a_i^2 \geq a_{i-1}a_{i+1}. \tag{5.1}$$

Can build up graphs, which proves polynomiality etc. However, this doesn't explain (5.1), because this isn't preserved under addition.

Matroids id a set of sets, which are called independent.

1. every subset of an independent set is independent.
2. If A, B are independent, and $|A| > |B|$ then there is an element of A which adde to B keeps it independent.

$n+1$ size of M the ground set

$r+1$ rank =

Not obvious to construct. Let G be a graph, base set is edges, and a set is independent if it's not a circuit. Of let A be a finite set fo vectors in V. Then "independent" = "linear independent".

Fano matroid is realisable iff char(k)=2 and non-Fani matriod iff $\neq 2$. Non-Pappus matroid is not realisable over any field. So ask how many matroids are realisable over a field.

Conjecture 5 *0% of matroids (limit as $n \to \infty$ are realisable.* Stated as an easy exercise in an early book!

When $k = \mathbf{Q}$ realisability is Holbert's 10th problem.

Conjecture 6 (Rota) *Define $\chi_M(q) = \chi_{M\setminus e}(q) - \chi_{M/e}(q)$ for a matroid. (5.1) still holds.*

Theorem 32 *Any noncontant homogenepus polynomial h defeines a sequence of milnor numbers $\mu^0(h), \ldots, \mu^r(h)$ with the following properties:*

1. *$\mu^i(h)$ is the number of i-dimensional cells in a CW-model of the complement $D(h) := \{x \in \mathbf{P}^r | \ldots\}$*
2. *...*

Consider the n-dimensional permutohedron, the convex hull of an orbit of the symmetric group S_{n+1}.

In a recent work with ..., we obtained inequalities that demonstrate Rota's inequality. Let X be a smooth projective variety of dimension$r$, and $k \leq r/2$. Let $C^k(X)$ b ethe image of the cycle class map in $H^{2k}(X, \mathbf{Q})$. Then Grothendieck's conjectures say that

1. Hard Lefschetz: Any hyperplane class defines an isomorphism
2. Hodge–Riemann: any hyperplane class ℓ defines a definite form of sign $(-1)^k$ ...

$$PA^k(M)_{\mathbf{R}} \times PA^k(M)_{\mathbf{R}} \to A^r(M)_{\mathbf{R}} \sim \mathbf{R}$$

Any structure that has these is said to be "like a smooth projective variety".

The toric variety of Δ_M is in the realisable case, 'Chow Equivalnt' to a smaooth projective variety. Itis tempting tothink of this as a Chow homotopy (but when the base field is $\mathbf{C}$, we must remember that it isn't!).

For any two matroids on $[n]$ with the same rank, there is a diagram of "flip" from one to the other: each "flip" preserves the "Kahler package" above.

Define the cohomology ... [he said "don't read these slides!"].

Then his Main Theorem is indeed that Hard Lefschetz: and Hodge–Riemann: are valid. So why does this imply (5.1)?

Part II

ICIAM 2015

Chapter 6

10 August 2015

6.1 Opening Ceremony

Attended by Vice-President of People's Republic of China, Minister of Education, President of the Academy of Sciences and President of Council of Tsinghua University.

We are united in believing that our mathematics is applicable outside the world of mathematics. Applied Mathematics might once have fitted into a single syllabus, but those days are long gone. Our member societies include Optimisation, Mathematical Biology, and Computer Science.

Over 3100 delegates as of this morning.

6.2 Prize Ceremony

Prizes are $5000, contributed by various societies, including the UK's IMA,

Collatz Annalisa Buffa

Lagrange Andrew J. Majda

Maxwell Jean-Michel Coron

ICIAM Pioneer Björn Engquist

ICIAM Su Buchin Li Tatsien

6.2.1 Buffa by Volker Mehrmann

"The use of highly suphisticated mathematical techniques in computer simulations" is the citation. Since 2004 Research Director CNR (Pavia), also ERC Starting Grant.

First worked one electromagnetics: both analytic (Sobolev spaces) and a general framework for coupled problems such as magneto-elesticity.

Then isogeometric analysis. Can we bring the methods of splines and NURBS as primitives in the discretisation of PDEs? First isogeometric GPL-licensed code.

6.2.2 Majda by Felix Otto

I always went to his Courant lectures on turbulent convection. Went in three years from Assistant to Full Professor at UCLA (1976–78).

"Remarks on the Breakdown of Smooth Solutions for the 3-D Euler Equations". This paper asked the question "if something does wrong, what"? Local existence theory doesn't answer this. They show that $\nabla \times u$ has to blow up.

"Absorbing Boundary Conditions for the Numerical Simulations of waves" (with Engquist). Require the constions to be local and to lead to a well-posed systems. Need pseudo-differential calculus for the locality. This comes up with a clean communicable result: three conditions wih increasing angle of incidence.

6.2.3 Coron by Alastait Pitt

Prize is for "originality in applied mathematics". "Highly sophisticated and novel mathematical techniques." Control Theory began with Maxwell's own apper "On Governance" (1860).

1992: fundamental paper in control systems. In finite dimensions, most systems can be stabilised by time-varying feedback laws — "Coron's return laws". Proof of controllability of Euler and Navier-Stokes, despite the fact that linearised Euler is not controllable.

These controllers are now being used to regulate the Meuse in Belgium.

6.2.4 Engquist by Kako

Siminal qoek in numerical methods for wave propagation in unbounded regions by introducing the absorbing Boundary Condition (ABC) or Radiating Boundary Condition (RBC). The Essentially Non-Oscillatory (ENO) sheme is used in industrail problems. Also Heterogeneous Multiscale method (HMM).

6.2.5 Li Tatsien by Yang

Major textbook, regarded as a model in China, and first two volumes translated into English. Directs Chinese Undergraduate Contest in Modelling, which has had a major influence on curricula

6.3 Revisiting Term Rewriting in Algebra: William Sit

My co-authors ave used these ideas to characterize Rota-Baxter type poerators.

k is a commutative unity ring (usually, but not always, a field). An *algebra* is a free associative k-module. A *rewritingsystem* is a set v wirh a binary relation $\to$. A *rule* $a \to b$ is just a pair $(a, b0$. Wite $\to *$ for the transitive reflexive closure. Define $a \to *b_1; a \to *b_2$ as a *fork*, and if every fork

Theorem 33 (Newman's Lemma) *A terminating RS is confluent iff it is locally confluent.*

The *symmetric closure* of $\to$ is $\leftrightarrow$.

Usual stuff on term algebras (largely skipped).

Fix a k-basis W of V a free j-module. For $f \in V$ the *support* of f is the set of $w \in W$ appearng with nonzero coefficients in the basis-expression of f. We are concerned with RS w.r.t. a fixed basis W. We therefore think of $\to \subset W \times V$, i.e. only basis elements get reduced by rules. Let T be the set of elements of W that actually get rewritten. We extend the rewriting system from T to the whole of V as $\to_\Pi$.

We say that the sytem is simple if $t \dot{+} v$ for all $t \to_v$.

Example 28 $W_1 = \{xy, x, y\}$, $W_2 = \ldots$. *Let* $\Pi = \{x \to y\}$.

Lemma 2 *1.* $f \to_\Pi g$*;*

2. $(f - g) \ldots$

3. $\ldots$

Then certain inferences, bt counterexamples against all others.

Theorem 34 *The following are equivalent. provided* $\to_\Pi$ *is simle.*

1. $\to_\Pi$ *is confluent*

2. $\downarrow_\Pi$ *is transitive*

3. JHD couldn't get this (12 items in all)

4.

There are three key arrows that require "simple". (l) may not hold for non-simple.

But "joinable" is not actually transitive. f and g joinable to $g_1 x + z$, g and h to g_2, and g_1 and g_2 aren't joinable. But f and g are!.

A *local base-fork* is $(ct \to_\Pi cv_1; ct \to_\Pi cv_2)$.

Theorem 35 *If* Π *is lcolly base-confluent, it is base-confluent.*

A *minimal descebdant chain* is the shortest from fron f to t.

"Every time one introduces a new concept of standard bases, one neds new defintiins and theorems (but generally the same proofs) — Mora". This should simplify this problem.

6.4 New effective differential Nullstellensatz: Richard Gustavson

Is a system of polynomial partial diffential equations consistent?

Example 29 $u_x - v_y = 0; u_y - v_x = 0; (u_{xx}+u_{yy})^2+)v_{xx}+v_{yy})^2 = 1$. *Simply regarding this as a polynomial system is consistent, but not when we differentiate. How often?*

Differential ring, and order($\theta = \prod \partial_{x_i}^{i_i}$) is $\sum i_i$. Let $K\{y_i\}$ be the ring of differential polynomials. Concept of "differentially closed"

Theorem 36 (Weak Differential Nullstellensatz) *Let K be a differential field of characteristic 0. For $F \subseteq K\{y_i\}$, we have $1 \notin [F]$ iff. for all differentially closed $L \subseteq K$, there is $(a_i) \in L^n$ such that $f(a_1, \ldots) = 0$.*

Let F have derivatives of maximal order d and degree h. Let $F^{(b)}$ extend F by all its derivatives up to order b. So an effective Differential Nullstellensatz would be a bound $b(m, n, h, d)$.

Example 30 $F = \{y' - 1, y^d\}$ *needs to differentiate d times. So $b(11, 1, d) \geq d$*

Extensions: $b(m, n, d, h) \geq d^{mn}h$

Theorem 37 (Sadik1985) *A lower bound h^{2^r} where $h = r + r\left(\frac{r-1}{h-2}\right) + 8$.*

First upper bound is due to [Gri89]. Triple-exponential and first-order systems only with a single derivations/ [GKOS08] Ackermann-based. [DJS15]

Theorem 38

$$k \leq (n\alpha_{T-1}d)^{2^{O(n^3\alpha_T^3)}}$$

where $\alpha_T = \begin{pmatrix} \alpha + T \\ T \end{pmatrix}$ *and T is to be defined.*

Uses a lemma [Pierce2014]. [FS14] use this to produce a recursive construction for $T.t = t(mn(2^i h))$ as in Pierce. Define $T = T_h^{m,n}$ with $T_h^{mn} = 2^{t(m,n,(2^i h))}$.

$b(1, n, h, d) \leq (n(h+1)d)^{2^{O(n^3(h+2)^3)}}$ which is [DJS15] but allows for non-constant coefficients. Shows some enormous figures for $m = 2, 3$. $T = 2^{2^{2^{520}+520}} + \cdots$ was one example. Have some improvements on [FS14].

6.5 Solving Polynomial Systems ...

Gives an n-variable 0-dimensional system, the output should be isolating boxes, and a $T_t(X) \in \mathbf{Q}[x]$ which defines the x_i-values.

Example 31 *In two-D ...*

Lots of GB and RC methods.

LGP bivariate: do a shift such that t

Our system LUR requires a different shift. We produce a *root candidate box* by interval methods.

For multivariates, first reduce to 2-D by resultants, Then isolate in 2D, and produce candidate root intervals in$\mathbf{R}^3$, Shift such that two projections onto $\mathbf{R}^2$ are disjoint.

Q How do you construct a random sample with mutliple roots?

A Discriminants of surfaces (?).

Q-MMM How do you do $n > 3$?

A Complex description.

Q-MMM Regular Chains now has a C version.

A But doesn't that require radicals? Not necessarily. A RC is always radical, but can represent non-radical systems.

6.6 Computing Equilibria of semi-algebraic economies using triangular decomposition and real solution classification: Li Xiaoliang

[Joint work with Dongming Wang]

Multiple equilibria are a problem for classical theories of economics. $P_i(u_j, x_k) = 0; Q_i(u_j, x_k)\sigma 0$ where $\sigma \in \{<, \leq, >, \geq, \neq\}$. Numerical problems have drawbacsk: instability, and are infeasible for multiple equilibria. [KublerSchmedders2010] uses Shape Lemma, also papers that use numerical homotopy. "Our methods are not new: we are rewriting RUR to make it accessible to econimists. In practice redoing [KublerSchmedders2010] in regular chains rather than GB.

Example 32 $x^3 - 20y^2 = 0; y^2 - 2x - 1 = 0; x - y \neq 0; 2x - y \neq 0; y > 0.$

1. *Let $x < y$. Regular Chain for equalities (easy). T. Let $T^* = T_{x=x+y}$ then decompose and get a chain T_1 and T_2 where T_1 is univariate in x and T_2 is now linear in y. Can always get such a "quasi-linear" system.*

2. *Back-substituting the linear variables gives a set of constraints in x alone.*

3. *Use the modified Uspensky Algorithm to isolate the roots of all the inequalities.*

4. *Then test sanple points to know which intervals satisfy the inequalities, and then see where T_1 has zeros.*

But what abut parameters.

Example 33 *1. Let $u < x < y$. Get three regular chains now. $T_2 = \{ux, y^2 - 1\}$. T_1 is the "main branch", i,e, greatest dimension.*

2. Then again make a linear transformation.

3. Then need to define border polynomial *which divides the patrameters space into regions.*

4. Take sample points of parameters space, and test these in the full system as before.

Example 34 (Exchange Economy [KublerSchmedders2010]) *$u_{11}(c) = 9c -$ $\frac{1}{2}c^2$ $u_{12}(c) = \frac{29}{4}$ ... [two parameters]. The sqfr BP has degree 25 and 249 terms. There are three equilibria when R0 There is a small rgion in $\mathbf{R}^2$ where this happens: probably not found by chance.*

Example 35 (Duopoly) *Customer can buy from A, B or neither. Customers have identical preferences AssumeU$_A > U_B > U_0 > 0$ are the utility functions for owning A B etc.*

Proposition 1 *There is a* Cournot equilibrium *with $fracR_1R_2 < \frac{1}{3}\frac{U_A - U_B}{U_A - U_0}$.*

Original proof was opaque.

Q When did semi-algebraic equilibria occur?

A Papers cited, but in practice most are,

Q Have you encountered problems you can't solve?

A Lots!

6.7 Triangular Systems over Finite Fields: Mou

Triangular iff mvars are distinct. Saturated ideal $sat(T) := |langleT\rangle : (\prod_{i=1}^{r} I_i)^{\infty}$.

Definition 12 *A triangular set is* simple *iff forall $i = 2 \ldots, n$ and aassociated prime p of* $\mathrm{sat}_{i-1}(T)$*, the image of T_i in $(K[x_1, \ldots, x_{i-1}/p)[x_i]$ under the natural homomorphism is square-free.*

Note that $(K[x_1, \ldots, x_{i-1}/\mathrm{sat}_{i-1}(T))[x_i]$ is not necesarily a IFD. Nowever our triangular representtation is a good representaton of algebraic extensions. So, after decomposing F into triangulars, we decompose the triangulars into simple sets. Hence by induction we want the "square-free" part of a polynomial in $(\mathbf{F}_q(x_1, \ldots, x_{i-1}]/\mathrm{sat}(T_{i-1})[x_i]$. Quotes as not a UFD.

Note that sqfr over finite fields is harder [GT96].

In the 0-dim case we have "generalised sqfr decomposition". $Q\prod_{i=1}^{p_1} P^i$ where Q is a p-th power.

For the positive dimensional-case we turn u into a parameter. [Kal98]. But the computation of radicals in positive dimension and positive characteristic. This is hard.

Note [Sei74] and the Condition P requirement, See also [FGT02].

What is a squarefree decomosition over an unmixed product of field extensions? Iff all the images over the components are sqfr, and the components are fields. We have anew algorithm

1. square-free decomposition plus D5
2. pth power identification: multiple derivations (new)
3. pth root extractuon via linear systems and Condition P

This is a new algorithms for simple decomposition.

Q Complexity result?

A Always hard for triangular systems. Also uses D5.

MMM We should have written up our D5 results. [DMSX06].

Q

A

6.8 Computing Decomposition...

Let K be a field. $\overline{K}$ the algebraic closure. $K[x]/I$ a finite o-dim ideal.

Definition 13 *$Dec(I) = \{\sigma \in S_n | F(t_{\sigma(1)}, \ldots = F\}$ is the decomposition group.*

Example 36 *$F = (t_1 + t_3)(t_2 + t_4)$ has decomposition group F_4*

1. $Dec(]langleT\rangle)$ is up to the Galois group of F [Anaietal1996]. $O(n^4)$.
2. There is an $O(n^3)$ algorithms by increasing chains of groups
3. We give a new algorithm: no complexity.

Proposition 2 *Under the above conditions, dots*

Definition 14 *$Zero_{\overline{K}}(I) = \{P_1, \ldots, P_N\}$ with*

Definition 15 *$S_i, S_j \in S$ are I-equivalent if $S_i = S_j$. $\sim_I$ is an equivalence relation.*

Lemma 3 (4.5) *The map $\Psi : K[x]/I \to$ is an endomeophism.*

Lemma 4 *$S_i \sim_I S_j$ iff $g_i(\lambda) = g_j(\lambda)$.*

Theorem 39 *Let* $P_I = \{B_k \subset \{1 \ldots, n\} | k = 1, \ldots, s\}$ *with* $B_k = \{n_1, \ldots, n_s\}$ *Then* $Dec(I) = \prod_{k=1}^{s} Sym(F_k)$

New algorithm:

1. Compute a Gröbner basis of I
2. Compute each m_{x_i}
3. Compute the characteristic polynomial f_i of each x_{m_i}
4. Construct the decomosition
5. use above theorem.

Application. $Zero_{\overline{K}}(\psi_\sigma(T)) \subseteq) Zero(T)$. Triangular decompsition of Cyclic-5. Algorithms give 15 sets. We get a union o f three sets $\psi_{1,4}$ etc.

Q You use GB - any chance of using triangular sets here?

A Somewhat confused

6.9 Solving Parametric Polynomial Optimiation via Triangular Decompsoition: Changbo Chen

Applicaton: Ecological Driver Assistance System. The Model Predictive Control is basically solving lots of optimisation problems, but these are really one parametric problem.

Minimise f subject to equations f_i, inequalities $g_j \leq 0$. Note that it is possible for optima to be at infinity.

CADs — naturally described by a tree. A strong projection algorithm (Collins, Hong) may use too many polynomials, but a ewaker one (McCallum etc.) may fail.

Introduce a new variable z to denote the optimal value, Add equatoinal constraints $z - f(umx) = 0$ Eliminate $x > z > u$. Eliminate with equalities and inequalities. Output the cells with smallest z value in each u-cylinder. Note that this will tellus about caes where the minimum is not attained as well.

We also have [JHD missed this] to see whather the KKT condition is valid. We need to exploit the structure of the MPC problem and the KKt condition to combine this with RC-CAD.

Q–JHD Exploit structure?

A A lot of the conditions are linear, so could use Fourier–Motzkin.

Q–JHD But doesn't RC-CAD do qute well in the linear case?

A Yes, but still does more work than we would like.

6.10 Disovering Multiple Lyapunov Functions for Switched Hybrid Systems: She

Two critical problems are safety verification and stability anaysis. Stability is done by constructing transitions that are suually sued for safety verification. So we will look at asymptotic stability.

Definition 16 *A switched hybrid system has N subsystems (modes) For each mode i there is an ODE $\dot{\mathbf{x}} = f_i(\mathbf{x})$. The state space is $X \subset \mathbf{R}^n$. Also there are switching functions into different modes.*

We want to use RRC to verify the existence of a multiple Lyapunov function. The family $\{V_i(\mathbf{x}) : i \in M\}$ is called a multiple Lyapunov function. (each $v_i(0) = 0$ and ... [Standard Lyapunov definition?]).

1. For each ode i
   (a) Let $\mathcal{X}_i$ be $\bigcup_{j=1}^{t_i}\{\mathbf{x}|E_{i,j}\mathbf{x} \geq 0\}$ $E_{i,j}$ is an $n \times n$ matrix
   (b) Let V_i be a quadratic form and write $V_i(\mathbf{x}_=\mathbf{x}^T P\mathbf{x}$ ...
2. Piece together.

So we use real root classification to under-approximate the constraints in out theorem. Formulate these under-approximations as a semi-algebraic set.

The algorithm got very involved here. We have a set of semi-algebraic sets from the various conditions. There is atheorm that states that,n if one satsifies all these semi-algebraic sets, then it is a MLF.

Example 37 *3D Eachs subsystem is asymptotically stable. The linearisation has two eigenvalues with real part 0.. We gte anya nswer. MI can't apply becaus eof the eigenvalues, and SOS doesn't terminate in five hours. As well as efficiency, note that LMI and SOS use floating-point, so have inherent problems.*

Note that QE could be applied, but is doubly-exponential, while this method is an adaptive CAD where some variables can be eliminataed.

Q In RRC do you need to make a recursive call into the variety of the border polynomial. Since the RRC output is the truth *outside* the border polynomial. [led to a discussion in Chinese with Changbo]

A Apparently use the "finder" interface. [??]

Q

A

Chapter 7

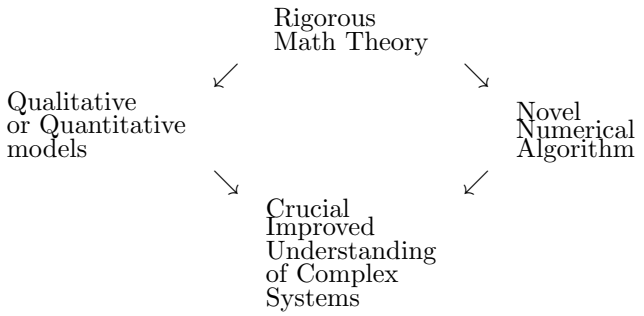
11 August 2015

7.1 : Majda

Climate Science is an extreme Complex System. Probably 10M or more unstable directions with a huge sate space. We need both statistical and applied mathematical skills together. We have to cope with model error: lack of physical understanding and inadequate resolution due to the curse of ensemble size. The computational cost of genrating even a small number of ensemble members is overwhelming.

Therefore we need uncertainty quantification (UQ) bounds for 1 and 2.

Therefore a new paradigm:



7.1.1 Ex 1: TBH

[MajdaTomoleyevPNAS2000] on the Truncated Burgers-Hopf (TBH) equations. Consider the finitie Galerkin truncation of inviscid Burgers equation. Statisticl predictons are equipartition of energy. correlation scaling law (large scales decorrelate more slowly), no separation of sclares. This is confirmed in sumulations with 40 odes. The Hamiltonian is actually $\int u_L ambda^3$, not the energy ($\int u_\Lambda^2$). We use thsi for dta assimilation as well.

7.1.2 Ex 2: Lorenz 96 model

$\frac{du_j}{dt} = (u_{j+1} - u_{j-2})u_{j-1} - u_j + F$. epeeninDepending on the forcng value F the system will exhibit completely different dynamic features. Can be weakly chaotic, strongly chaotic or turbulent. Miros mid-latitude baroclinic waves along midlatitude circle. nbergy of weather moves eastward but individual (Rossby) waves move westward. Been used for UQ modelling.

7.1.3 Ex 3: MMT equation

$iu_t = |\partial x|^{1/2}u + \lambda|u|^2u - iAu + F$. Consider focusing nonlinearity $\lambda = -1$. The instability of collapsing solitons radiate energy to large scales.

If you try to run with too few modes, you lose energy. But he has a trick (eddy terms) to restore this, and therefore beat the curse of ensemble size,

7.1.4 Stochastic Superparameterization

1. A general framework for stochasric subgridscale modelling with no scale separation.
2. Success in a difficult test problem.
3. ...

7.1.5 Extreme Events

[NeelinetalGRL2011] CO and CO_2 distribution in the atmosphere has bit fat tails compared with Gaussian. We want exactly solvable test models which display intermittency. [MajdaGershorinPhilRS2013]. This model shows the "exreme event" behaviour and fat tails that we observe.

[MajdaXinTongNonlineatirt2015] have a rigorous PDF which dislays intermittency. Thes eoccur when the random mean flow $U(t)$ gets close to a certain resonant set.

7.1.6 Information Theory

We can look at Shannon Entopy. Relative entropy quantifies the lack of information or model error in the statistics of u^M relative to that of u.

What we'd like to do is take the current climate and compute the response to forcing. Example of a perfect model and an imperfect one. They can predict the cimate perfectly but get the response to forcing completely wrong.

Equilibrium statistical fidelity is a necessray condition. Combine the information theory with linear response theory in improving predictive fidelity. Want a linear response operator calculated through correlation functions in the unperturbed climate.

7.1.7 Lessons for UQ and Failure of Polynomial Chaos

[MajdaBranickiDCDS2012] $\dot{u} = (-\gamma + \sigma_\gamma \xi)u + f(t)$ where parametric uncertainly is $\sigma_\gamma \xi$ .... It is easy to solve exactly.

Both PC with 120 coefficients and MC with 50,000 smaples will fail to predictthe variance with any accuracy.

7.1.8 Inverse Problems and Data Assimilation

Swows sample points in Atlantic. Firts rigorous math theory [NanChenMajdaNonlinearity2014JNLS2015] Inherent nonlinearity in measurement. Build exact closed analytic formulae for the optimal filter for the velocity field. Prove a man field limit at long tmes.

To recover incompressible flows need an exponential increase in the number of tracers for reducting the uncertainly by a fixed amount.

We have a rigorous mathematical model with comparable high skill in recovering GB modes ....

7.2 Filerting

A two-step porcess involving statistical prediction of the state variables through a forward operator followed by ....

Finite ensemblke Kalman filter (EnKF) often works well to estimate the mean when ensemble size is much smaller than phase space. Why?

There is a surprising pathology with catastrophic filter divergence. For filtering forced disspative systems such as L96, EnKF can explode to machine infinity in finite time. [HarlimMajda2008]. Wellposedness of EnKF is an issue.

We need *a priori* estimates for .... We look for energy principles inherited by the Kalman filtering scheme. We need modificatoin schemes for EnKF.

7.2.1 Madden–Julian Oscillation (MJO)

Starts in the Indian Ocean. Affects El Nino Australian and Asian monsoon tropical cyclones and midlatitude predictability. Rossby wave trains from this croos middle USA. This has slow eastward propagtion at 5 m/sec. Peculiar dispersive relation $\frac{d\omega}{dl} \approx 0$. MJO is actually an envelope of smaller-scale convection waves. GCMs typically don't adequately represent convectively coupled equatorial waves and the MJOs.

Needs Nonlinear Laplacian Spectral Analysis. We apply this to datasets 10^6 *in dimension*. Four ideas: lagged embedding; machine learning; adaptive weights; spectral entropy criteria.

Have a traiing period and predictive period for MJO1 and MJO2. hese both have exreme events and fat tails.

We have observed variables and hidden variables for stochastic damping and phase.

We would like (and see) that our ensemble spread captures the long-range forrecast uncertainty.

[MajdaStechmann2009PNAS] have a new model for the MJO, which caputes all three features: as above plus horizontal quadrupole structure. Neutrally stable interatcions between palnetary-scale lower-tropospheric moisture and ....

Minimal nonlinear osciallator model. Linearised primitive equations: equatorial long-wave scaling and Coriolis term: equatorial β-plane approximation. Ad dymanic equation for convective activity.

2011-12 massive effort to study MJO. There fore replace the $\partial_t a = \Gamma q a$ by a stochastic jump process. We get ntermittent egenration of MJO events plus organisation of MJO events into wave trains. We obsrve 39.7 days as average duration. and our skeleton model predicts 34.8 days.

There are squall lines at 200km sclare CCW at 2000 km and MJO at 20,000 km, Why? [Majdo2007JAS]. Paper son multicloud Model Dynamics. Good models run with 160km (v. coarse) resolution.

7.3 Grid and Grid Control Optimization in Europe — M2GI: Sax

Introduction: do you realise that the gas energy moves far more energy than the electricity grid.

Speaker: Gas represents 25% of Europe's energy/ 2/3 of this comes from Norway or Russia. These costs 1Meuro/km. There will be shortfalls from NL due to a recent court decision there. Shows pipeline network in Europe. Also extensions into Algeria via Sardinia. Tunisia via Sicily, Morocco via Spain etc. Also across the Black Sea from Ukraine to Turkey.

Open Gas Europe ... lots of statistics. Mentions NorthStream from Russia to Griefswald.

"yestrday, al their troubles seemed so far away" — gas was vertically integrated, and mathematically the optimisation problem was soluble, Inthe 1970s we wrote programs using physics, thermodynamics etc. Mathematically, these used tools like Reynolds numbers, Darcy–Weisbach equation etc.

But the EU did not like the huge profst from trading, and forced the companies to unbundle. This left "security of supply" no-one's problem. 2009–15 have demonstrated the issues this causes (e.g. 2011 there was zero gas flow at Waidhaus. which normally provides most of the gas for Germany). There were also shortage/low temperaturs problems in February 2012.

Gas storage provides no security. The gas providers use their storage to pursue profitable trading. Graph (in German) of the output of a porous reservoir storage. Above a critical withdrawal (50%), the efficiency declines seriously.

Until 2014 network planning was more-or-less "y hand". Scenarios were simulated by standard software, but this was limited. Gas notwork operators have obligations:

1. guarantee safe and reliable operation

2. non-discriminatory

3. transparent

4. at competitive prices

5. refusal of access requires a well-founded justification.

Therefore an entry/exit model. Offered capacities. booked capacities, nominated capacities. Thetransport company has to guarantee that any combination of entry/exit points is technically feasible. Even is one assumes this is reasonable, ths is not mathematically well-defined.

Therefore we have discrete decision, nonlinear equations giverning gas physics. The MATHEON project ealt with optimisation of gas transport and stable transient modelling and simulation of flow networls.

We had a large research project ForNE: 10 universities and 10 employees from Germany's operator OGE. Book "Evaluating Gas Network Capacities".

"With a little help from my friends" — our new target is a navigation system for control sstem decision in avolatile gas market gered to 24–48 hour time frame.

MODAL AG was led by ZIB to offer the gas industry computing kernels and a sustained researh effort. Funded by German Ministry of Educationa nd Research but a lot of industry support. This closes the gap between research and sueful deployment. ZIB's past track record was important here.

Part of the gas in the network is used to supplythe energy for transport, and therefore this should be minimised. A 10% reduction here is equal to one nuclear power plant!

Notethat a permanent risk is that a decision taken in the past will lead to a problem in the future. Such a solution [MODAL] would provide early warning of problems/ Algorithms would be able to find solutions for control problems that hmans can't find. This should increase the capacty of the network.

M2GI "More Mathematics in the Gas Industry" is the only way of maximising the provision of freely allocatable capacity, of optimising the grid and grid control to handle this ....

1. Research should be given the necessary time

2. Grass does not grow faster if you pull it

3. "a good start needs enthusiasm a good end discipline"

4. "Mathematics makes gas flow better"

7.4 Randomised ALgorithms in Linear Algebra: Kannan

This means "an algorithm can toss coins" or "the data tosses coins". i.e. average case analysis. The second is not our concern. We want results that work for *every* matrix. Examples:

- Quick Sort 1960s
- Primality testing 1970s — only recently deterministic
- Routing 1980s: randomness to avoid congestion
- Convex sets and volume in the 1990s
- Matrix Algorithms — this talk.

The simplest form is to compte witha small sample of rows/columns. Moderndata matrices can be massive. hence $O(1)$ access to an entry cannot be assumed.

We will prove error bounds on answers from a small sample. If fulldata is unavailable, only a sample may be available. Netflix has prefernces of 10^5 customers on 10^5 products.

A related question is distributed data. Communication is expensive, so the processors send sketches.

7.4.1 Setting

A is a large marix. How can I compute AA^T. and more generally AB. Then I might want SVD, Low Rank Approximation. Matrix Sketches. ensors: approximation by sum of rank 1 tensors.

- No free lunch: approximate answers only.
- But we will prove error bounds for all input matrices.

$||A||_F = \sqrt{\sum_{i,j} A_{i,j}}$, $||A||_2 =$ max eigenvalue norm.

If the rank is j, to solve these problems with error $\pm\epsilon||A||_F$, a sample of $f(k/\epsilon)$ rows/colums will suffice, provided that they are picked in i.i.d. trials, provided that the probability of picking a row/column is proprtional t its squared length. f is a small polynomial. [FiezeKannan196] did SVD and Low Rank Sampling. Many improvements.

Alternative Scheme, take a sample of entries, set others to zero, and compute faster because of sparsity. Note that this doesn't reduce matrix size.

Approximate AA^T in $O(n^2)$ time. Uniform smpling of rows is no good — what happens if all but one column of A are zeros. An unbiased estimator of AA^T $X = \frac{1}{p_j}$(column j)(rown j). This is why we need squared length. Then $E(||AA^T - est||_F) \leq \frac{||A||_F^2}{\sqrt{s}}$ with s samples. [DrineasKannaMahoney].

Can we do better woth the spectral norm? $E(||AA^T - est||_2)$. [Rudelson] $E(||AA^T - est||_2) \leq \frac{c||A||_F||A||_2}{\sqrt{s}}$. [Tropp] "User friendly tail bounds for ...".

Suppose P is a probability distribution on $\mathbf{R}^d$. We want the variance/covarince matrix of P $M_{i.j} = E_P(x_i x_j)$. We really want error bounds for sinite smaple size which depend on d alone. P might be log-concave like Gaussian, or uniform

on a convex subset. So how many samples should we take for relative error ϵ? $M \approx_\epsilon M'$. We want

$$M' : |x^T(M - M')x| \leq_\epsilon ||x||^2 \forall x.$$

Let B be the pseudo-left inverse of A. $VBA = I$ on the row space of A. Let p_j be proportional to the squared length of columns in BA. Draw s i.i.d. sample colums of A according to p_j and the W be the estimator of AA^T based on these columns. Then whp $|x^T A^T Ax - X^T Wx| \leq \frac{c\sqrt{r}}{\sqrt{s}} x^T A^T Ax^T$. We get relative error for every x provided $s?c \operatorname{rank} A$.

Graph has n vertices. Pick $O(n \log n)$ sample of weighted edges such that every cut has roughly(with ϵ) the same number of edges crossing it. Better is a spectral sparsifier: find a sma;; subset B of eighted colmns of A such that$AA^T \approx B^T B$. This is stringer than the cut sparsifier. THis can be solved by preconditioning, but that takes time. [SpielmanSr5ivatsava] says we can estimate the precoditioned probailities fast. Analogy is electriical resistance.

7.4.2 Matrix Sketching

Is a sample of rows sufficient? No.

Theorem 40 *Let A be any $m \times n$ matrix and CE = an $m \times s$ sample of s columns of A picked according to length squared. Ditto R but $\sqrt{s}$. Then their is a $s \times'$ sqrts matrix U such that $E(||A - CUR||_2^2) \leq \frac{c||A||_F^2}{\dots}$.*

SVD. Sample A to get $m \times s$ C. Find the top k eigenvalues of $C^T C$. Find the top k left singular vectors $U_1, \dots$ of C. A' = projection of A only the span of u_i. Then $E(||A - A'||_F^2) \leq$ best possible rank k approximination $+\dots$.

Data Handling – Pass Efficient Model. Sampling ALgorithms use a constant number of passes. Pass 1 computes lengthsquared probabilities, and pass 2 samples.

Traditional SVD can find best approximation A_k of rank k to A. Note that if A is a patient vs gene-expression matrix you say "principla component 1 is 3(patient 1)-4(10th patient) $+\cdots$" — would be better if the columns of the approximating matrix were actual columns.

Can get an approximationA' to A with error $< (1 + \epsilon)\times$ best possible. We want the probability of drawing an r-tuple of columsn withprobabiliyu proportional to squared volumn pf the siplex they span.

Lemma 5 (Johnson–Lindestrauss) *A fixed (notrandom) nit vector in $\mathbf{R}^d$. W a random $k \times d$ matriix. Then whp $|Wx| \approx \frac{\sqrt{K}}{\sqrt{d}}$. The probability of failure falls expoenntially in k. So o ensure this nolds simultaneously for N vectors, need $k > c \log N$.*

But ther are ony e^{cd} bectors in $\mathbf{R}^d$ of that norm, so with $k > cd$ we get whp $\forall x : |WAx|$'$approx|Ax|$.

7.4.3 Distributed data

Matrix spread over many servers. Do we need to communicat ethe same random projection to all servers? So instead use pseudo-random projectsions and just distribite the seeds. These need to be k-wise independent.

Suppose r servers. server t has a $n \times d$ matrix $A^{(t)}$ with $d > n$. Find a low rank approximation of $A^{(1)} + \cdots$ with communication being a scarce resourse. [emmeletal] have deterministic tight bounds. $O(\frac{nd}{\sqrt{r}})$. We want to allow randomness. We have $O(\frac{rdk}{\epsilon})$.

7.5 Numerical Solving for Parametric Polynomial Systems with Constraints: Wenyuan Wu

7.5.1 Computing Real Witness Points: Wenyuan Wu

. Note that lots of numerical work over $\mathbf{C}$, then the critical point approach started in symbolic computation [SafeyElDin]. We need a regularity assumption, that the Jacobian is of full rank. Sometimes the plane/distance approach will hit singular or ill-conditioned points. These are a compact set, so with high probability we will miss them. In dimension > 1 we also need to know the direction for path tracking. Also we need to determine the step size: need to avoid "jumping".

For a square system we need to estimatethe distance between two isolated points. Suppose $\max\{||\nabla J_{i,j}||_2\}$ on unit ball is $K(g)$.

Lemma 6 (Root Isolation) *Let σ'_{n+1} be smallest eigenvalue ...*

As far as direction determination is ocncerned, we need to incrase the smallest signular value. This lads to an optimisation problem. $\Delta x = \frac{Hc^t}{\sqrt{\cdots}}$ solves this optimisation problem.

Shows an example with ncreasding σ.

- Suppose $\dim V_]R(f)$ is $m = n-k > 1$ Let $g(x_0, \ldots, x_n) = \{f, (\sum x_i^2 - 1)/2\}$ when $K(g) = 1$ (rescaling).
- The diretcion is HC^t Left $J \oplus T$ ....
- ...

Define an expression for step size in terms of ρ. If Newton iteration converges to z_1 Then z_1 is on the same component as z_0 iff (?) $dist(z_1, z_0) :< \omega\delta$. If the convergence point z_2 is outside this ball, we may have had jumping. $\rho \approx 1.6$ is the appropriate value.

So we can ask now many prediction–corretion steps we need. Has a linear plot of this against $\log_{10} \sigma$.

7.5.2 Numerical Solving Parametric Systems

Many applications, but symbolic methods don't scale well (expression swell). Numerical methods can take advantage of sparsity. The goal is to answer these question for a 0-dim parametric system

1. Count how many connected cells in parameter space, and chooce a smaple poit in each cell.
2. Membership tests for these points
3. Construct a path from a given point to a sample point.

Suppose we have solved a square system off-line and have solutions S_p at p. then use real homotopy to follow a path from p to q. Consider the singular points ....

Conside $\mathbf{R}[a.b.x_1.....x_n]$. Suppose we are only intereste din ome (physcial) region of parameter space. Assuming convergence of Newton we areguaranteed to stay on the same component, and parmetric homotopy works.

7.6 Algebraic attack and algebraic Immunity of Boolean Functions: Lin

Compexity $O(\ldots)$.

Assume either that f has small degree, or
$existsg$:
$deg(fg)$ small.

Definition 17 *The* algebraic immunity *of* f *is* $\min_{g\neq 0} \deg g | fg = 0 or (f+1)g = 0\}$.

We want functions with maximal AI. There were Carlet–feng constructions. Need immunity against Fast Algebraic Attacks (FAA).

We can use LFSRl If $fg = h$ with $\deg(g)$ low use Berlekamp–Massey to eliminate h with $E \approx \left(\begin{array}{c} n \\ \deg(g) \end{array} \right)$ equations. A lot of symmetric function with high AI are vulnerable to FAA.

Theorem 41 (Curtois2003) *If* $e+d \geq n$ *then ther is a* $g \neq 0$ *with* $\deg(g) < e$ *and* $\deg(fg) < e$.

Want Perfect Algebraic Immune functions (PAI).

M. Liu et al. produce 2^k-variable Carlet–Feng functions.

There are various suggestions of PAI functions on nearly such, with no proofs only computer analysis.

Claims that algebraic attacks converts qualitative cryptanalysis into a quantitative approach. This is a unifirmmethod. So the problem is to construct such functions.

Q Any attacks based on sparsity, rather than degree?

A This could be a problem: there have been cases.

7.7 Davenport

See .

7.8 Extending Hybrid CSP with Porbability and Stochasticity: Shuling Wang

Given by someone else.

Example 38 *An aircraift*

- *flight pathis a sequence of line segments*
- *ideally should follow nomial path, but may deviate due to wind etc.*
- *If deviates, should fllow a correction heading*
- *The aircraft therfore acts as a continupus pkant, wirh stochastic influence, and the flight control system acts as a discrete controller.*

Hence we have a hybrid stochastic system.

There has been work on stochastic hybrid automata. Rachability anaysis is usually done by probabilistic model checking or simulation. This is not scalable. [Platzer] uses stochastic hybrid programs. Deductive-based verification, but concurrency and communication are not supported.

Let F be a σ-algebra on Ω and P is a probability measure on (Ω, F). Mapping $X : \Omega \to \mathbf{R}^n$ is an $\mathbf{R}^n$-valued random variable if for each $B \in \mathcal{B}$, we have $X^{-1}(B) \in F$. A stochastic process X is a function $X : \ldots$.

Use Hybrid CSP [HeZhou1994]. Adds timings consructs continuous evolution and interrupts. It inherits `ch!e` and `ch?e` from CSP. We have $P||Q$ for parallel composition.

Continuous evolution $\langle F(\dot{s}, s) = 0 \& B\rangle$ wher F is a differntal equation s is a vector of variables and B is a Boolean expression. Timeout: $langle F(\dot{s}, s) = 0 \& B\rangle \triangleright_d Q$ continues for d time units, then becomes Q. $P \sqcup_p Q$ is probabilistic choice: P with probability p and q with probability $1 - p$. Also adds a "commnication interrupt" $\triangleright|_{i \in I}$.

The semantics of SHCSP is defined by a set of transition relations. We can prove that this is well-defined, i.e. evolution doe sbnot look nto the future and evolution is a Markov process.

Example 39 (Continued) *Let* $\theta(t)$ *be* $\begin{cases} -\pi/4 & right \\ 0 & correct \\ \pi/4 & left \end{cases}$ *be the angle correction.*

Use $\{A; E\}P\{R; C\}$ *where* A *and* R *are the discrete pre/post-conditions, and* E *and* C *the continuus assumptions and conclusions.*

Ther are many inference rules. For example $\sqcup_p$-introduction (JHD wasn't sure of the details). The main one is stichastic continuous evolution.

Example 40 (continued) *We apply the SDE rule. Then we define the dangerous states. Our Booelan guard is* $f \geq 0 \wedge LF \leq 0$.

7.9 An Application of QE to Automatic Parallelization of Computer Programs: Marc MM

Supported by Chinese Academy of Sciences and IBM Centre for Advanced Systems: 2×CAS.

Our context is GPUs. Automatic generation may seem insne, but it makes sense for many of the kernels in scientific computation. (dense linear/polynomial algebra, stencil compilations). We focus on C→CUDA. Standard techniques (polyhedron model) are inear,, but parametric

- Old-fashioned parallelism: loops map to loops.
- Polyhedron parallelism:: performa "god" change of coordinates for the loops.

Dependence Analysis Transform the sequential object to a geometric object in index space. [Feautrier]. This talk responds to [GrosslingeretalJSC2006].

Parallelization Our real interest

Code Generation Important

Data is decomposed into segments, and the segment is given to a group of threads.

So we have serial code executing on the CPU ("host") and parallel parts ("kernels") executing on the GPU. Note that the threads are SIMD. Typically we have many more thred blocks (logical program threads) than physical processors. A threadblock has access to per-thread shared memory, and (slower) access to the GPU main memory. Cache/main memory is a good analogy.

We propose the MCM (Many Core Machine) model as an abstract machine model. [Haqueetal2015PARCO]. Let Z be the private memory size, $U$the data transfer time, ℓ be the number of threads per thread block etc.

Example 41 (DFT) *Two algorithms: CooleyTukey and Stckoham. Get exprssions for the ratio CT/S of work. span and paallelism overhead, e.g.*

$$\frac{W_{CT}}{W_S} = \frac{4n(47\log_2 n\ell + \cdots}{172n\log_2 n\ell + \cdots}$$

We should generate kernel code where ℓ etc. are parameters.

Example 42 (Dense Polynomial Multiplication) *Change coordinates to create concurrency: $p := i + j$. But this is not sufficient. Work is unenvenly distributed, and too many processors are implied. Hence rouping intothred blocks.*

Use `RegularChains:-QuantifierElimination` *on the system to eliminate i, j to give us a program in terms of thrad block and thread index.*

Most people believe that Fourier–Motzkin is doubly exponential $O(n^{2^d})$ coefficient operations, but in pratcice ideas from Linear Programming improve this.

Example 43 (Simplified LU) *The main loop is updating the kth column of L then the $n-k$ columns of U.*

INRIA had a MetaForl to CUDA translator for non-parametric code. We have a preliminary parametric version.

Various tables showing sppedups for various thread block sizes (which really matters). One problem isthat the CUDA compiler doesn't do common subexpresion elimination, which hurts his comparisons at the moment.

7.10 Modular Techniques for Efficient Computation of Ideal Operation: Yokoyama

Given by Marc MM.

Full Methods for recovering the true result from its modular images are necessary e.g. wanting a GB.

Partial Only do some of the computation based on modular techniques.

Cyclic-* took 14 seconds on $\mathbf{F}_{99981793}$ but 883 seconds over $\mathbf{Q}$. 50 primes of 27 bits are necessary for a candidiate.

1. Compute modular images
2. Glue these together
3. Verify the candidate solution.

hree styles.

- CRA
- Hensel

- hybrid, as in Gröbner trace

Marc MM shows his [not Y's] standard Euclidean/CRT approach with early discovery when degree doesn't change. [Arn03] does the same, using $h(d_0)$ as the compatibility test.

Then Y defines Pauer lucky and Hilbery lucky. Ned to check inclusion both ways: one is easy as it's reduction w.r.t. Gröbner basis.

Theorem 42 ([Arn03, Theorem 7.1]) *If the situation is homogeneous ...*

Proposes a trace-driven [Noro] for using modular information to determine useless syzygies.

F4+trace	630	verification 140
Buchberger+trace	900	verification: 160

There were many more examples.

7.11 From lexicographic Groebner bases to triangular sets: Dahan

See [Laz85]. Structure of a lex GB in two variables: exact division of the l.c. w.r.t. y (essentially GianniKalkbrener for two variables). This observation can in fact give us a triangular set.

What happens when $n > 2$? [Laz92] introduces LexTriangular via D5. Moeller also had a Groebner version in 1992. Note that this in in Singular, and has the advantage that it can handle non-radical ideals.

Lazard uses D5 to use "quasi-inverse" of the leading coefficients. Note that this needs g.c.d. computation.

My theorem

1. $\psi(g_t) \neq 0 \Leftrightarrow \psi(lc_{\leq k-1}(g_t) \neq 0$
2.
3.

[Bec94] proved 3, but not the crucial 1. Has a division result for several variables but requires zero-dimensionality and radicality, and is not as neat.

Example 44 *Vanishing ideal of a set of points.* $x^5, y^4, x^4y^2z, y^3z, x^4z^2, x^2yz^2 \ldots$.

$g_i = \sum_{\alpha \in A} L_\alpha)x, y)f_1(x, y)f_2(x, y, z)$ where the L are Lagrange interpolants.

Theorem 43 *Let* $g, g' \in$

1. ...
2. *If* $\deg_z(g) < \deg_z(g')$ *then by Theorem 1,* $lc_x(g)|lc_{x,y}(g)$ *and* $lc_x(g')|lc_{x,y}(g')$ ...

By induction we can suppose that we only have two polynomials in two variables (Triangular). But we might have many in three variables. Let $h := \frac{g_1}{lc_x(g)}$; $g_{l(2)} = q \underset{\cdots}{\cdots}$ etc.

Is it possible to have more than three variables? Yes, but it's tedious to state.

7.12 Characteristic Set Methods for Solvig Boolean Equations: Gao

Deciding whether a Boolean system has a solution is NP-complete. Butit has many applications. [Shannon1946] stated that a good cryptosystem was "equivalent to colving a certain systems pf simulatneous equations in many unknonws". Ther are logic approaches (SAT) and graph-theoretic (BDD) [Bryant1986]. Also Gröbner basis [Courtois2000]. Approximate algorithms [Has88].

Definition 18 *The r-approximation algorithms optputs $\widetilde{O}$ suich that $\frac{1}{r}O \leq \widetilde{O} \leq O$ where O is the true optimal.*

Various classes of NP-problems.

Any r-approx

r-approx beyond sone threshold

No r-approx

Most multivariate crypto systems are based on quadratic equations. [Hastad2001] shows it is NP-hard to approximate MAX-MQ in $\mathbf{F}_q$ for any ratio $q - \epsilon$. There is a polyomial-time algorithms with ratio $\frac{q^2}{\cdots}$.

Random assignment is a $q + \frac{q^2}{q^{n/2}-q}$-approximation algorithm. Hence q is basically the threshold.

Note [Wu 79] founding Mathematical Mechanisation. This has been applied for algebaric equations, differential equations (Riit, Kolchin), difference equations [Gaoetal2009JSC]. What about finite fields?

Any triangular set can be made monic. Note that a chain may not have zeros in $\mathbf{F}_q$: See $x^2 + 1$ in $\mathbf{F} - 3$.

Definition 19 *A* proper triangular set *has ...*

Theorem 44 $|Zero_q(P)| = \sum$ *triangular sets*

Theorem 45 *The bit-complexity of TDCS is $O(l^n) = O(2^{n \log l})$ where $l = |\mathbf{P}|$.*

Compare $O(P2^n$ for exhaustive serch, and $O(d^{2^n})$ for Gröbner basis.

Theorem 46 *Algorithm MFCS — Multiplication Free. uses MF well-ordering principle. Output* $Zero_q(P) = Zero_q(T) \cup \bigcup_i Zero_q(P_i)$ *where* T *is a [roper riangular set.*

Then the size of the polynomials is bounded by the size of the input. Bitwise complexity is $O(ln^{d+1} \sum_{P \in \mathbf{P}} term(P)$.

Theorem 47 ([HSL14]) *For a quadratic polynomial syste of* m *polynomials, the bitsize complexity of MFSC is [the same as exhaustive search].*

Examples of stream ciphers shows that MFCS ourperforms TDCS (always rund out of time) or GB (always runs out of memeory).

Problem 4 (COOK at SAT 2004) *If* $AB = I$ *as Boolean matrices does* $BA = I$*?*

For $n = 4$ we take 0.2 seconds and Magma 2363; $n = 5$ we took 10 and Magma overflowed. SAT took 800-2000. For $n - 6$ we take 166 seconds.

7.12.1 Aside

Theorem 48 *Let* $h_i \in K\{y_i\}$ *there is an algorithm to comoute satuated triangular sets* $A_q := \Psi_{q,1} \dots \Psi_{q,l_q}$ *such that*

1. *Zero(*h_i*)* $= \bigcup Zero(\Psi_i)$
2. *Complexity is (merely) triply-exponential*

Chapter 8

12 August 2015

8.1 Stabilization of control systems: from water clocks to rivers; Coron

Water clock (clepsydra) Hole at the bottom of a tank, and the height of the water reflects time. But as the water flows, the flow rate decreases, so it's nonlinear. Ctesibius (3rd century BC) apparently invented a regulator: none survive, but it is described in pseudo-Archimedes.

Watt's regulator 1788: se Science Museum in London. Showed Watt's original drawing. Maxwell (1868) "on goovernance" was the first publication. Shows the cart/inverted pendulum.

Definition 20 (Lyapunov) *Let y_e be an equilibrium*

Theorem 49 *If X is C^1 and y_e is an equilibrium point of $\dot{y} = X(y)$ If the eigenvalues of $X'(y_e)$ have ...*

Theorem 50 (something about stability)

Problem 5 *Can we ensure that we have stability, or asymptotic stability?*

Definition 21 (controllabilty) *Given states y^0 and y^1 can we move from one to the other? For non-linear systems,, we need to be more careful: "small-time local controllability" is the key idea.*

We know no N&S conditions for STLC. If the linearised system is controllable, then the nonlinear system is STLC by inverse mapping theorem.

Let $[X, Y](y) = Y'(y)X(y) - X'(y)Y(y)$. The Lie algebar rabk condition at 0 is satisfied if $\{h(0) : h \in Lie\{f_0, \ldots, f_m\}\} = \mathbf{R}^n$. Let P_M denote the characteristic polynomial of matrix M.

Example 45 (Baby stroller) $\dot{y}_1 = u_1 \cos y_3; \dot{y}_2 = u_1 \sin y_3 . \dot{y}_3 = \cdots$ *is controllable but does not satisfy the bracket condition.*

Example 46 (Satellite Attitude) *Need three actuators to make the linear system controllable, but with two it is STLC, Again it does not satisfy the bracket condition.*

Need to enlarge the system to allow feedback to depend on time. His 1992 theorem.

Definition 22 *The origina in* locally continuously reachable in small time *for the control system if for every positive T there is $\epsilon > 0$ such that ...*

Seems like a very strong property, but we know of no systems with asymptotic stability that doesn't have this.

If he is in dimension > 3 there is enough room for perturbations to avoid any crossing, so his theorem is proved: that we have asymptotic stability.

8.1.1 1D hyperboic PDE systems

Various gates on the Meuse. $V(t, x)$ is the velocity of the water at point s along the river. Conserve mass and momentum. "Théorie du mouvement non permanent des eaux ..." discovered when author was 74!. In genral $y_t + A(y)y_x = \cdots$ with boundary conditions on y. Note that there are many applications: see forthcoming book.

Let X be a Banach space of functions from $(0, 1)$ to $\mathbf{R}^n$. Let $\lambda_1 = \frac{4n}{4n+1}$ and $\lambda_2 = \frac{4n}{2n+1}$ which perturbs the stable case (1,2), we get instability. Hence we actually want *robust exponential stable*. given by a theorem of Silkowski, if $\rho_0(K) < 1$ then .... Chinese Theorem: If $\rho_\infty(G'(0)) < 1$ then dots.

For the nonlinear case, also need $|| \cdot ||_2$. $\rho_0 \leq \rho_2$. but for $n \in \{1, 2, 3, 4, 5\}$ we have equality (speaker; Voisin). Even for $n = 2$ they have eamples where $\rho_2 < 1$ but this isn't enough to guarantee exponential stability.

8.1.2 La Sambre

One checks that for $\eta \in (0, 1)$ there are nonlinear feedback laws such that $\rho_\infty(G'(0)) < 1$. Shows pictures, and states that they are currently working on La Meuse.

Note that when Maxwell wrote, there were probably 75000 Watt reguators in England. There was only one pendulum regulator, but that's what Maxwell wrote about. To get convergence, you need an integrator. This we are actually using on the rivers.

Q You mention robustness. What about robustness w.r.t. the model?

A Good question. Lyapunov theory doesn't really handle this. But it can handle any specified class of perturbations.

8.2 Computational Progress in Linear and Mixed Integer Programming: Bixby

Grötschel claims there are over 100K such programs running at any one time, affecting all aspects of our life: e.g. buying an airline ticket.

Will talk mostly about the second (MIP), but this relies crucially on LP.

8.2.1 Linear

1947 Dantzig invents simplex, and talks about it. First use was 120 man-days on desktop calculators.

1951 Used on computers at NBS

1960s commercially viable at oil companies, notably BP (still major users) etc.

1970s Interest flourished, but LP was hard.

1980s Thought this was as far as we can go. Airline model with 4420 constraints and 6711 variables was insoluble.

1981 IBM PC.

1984 [Kar84a, Kar84b] interior point methods.

1990s LP really took off. Simplex kept pace with primal/dual. 95% of problems are still solved with simplex, not interior point. Popular new applications.

401640 constraints; 1584000 variables. All numbers on same Pentium 4: 2GHz. My first CPLEX (1988) 29.8 days. CPLEX 5.0 (1997) 1.5 hours. CPLEX 0.9 (2003) was 59.1 seconds. The algorithm was Dantzig's primal simplex algorithm. Paper in *Journal OR*.

So today LP is considered a solved problem. Large (Millions) models can be solved robustly and quickly. There has been no real research in LP algorithms since 2004. The power industry still has big problems, and some mixed IP problems finding the LP a bottleneck

8.2.2 Mixed Integer

Definition 23 *The same (minimise $C^T x$ subject to $Ax = b$ and $l \leq x \leq u$) subject to constraints that some of the x_j must be integer.*

In 2012 we [Gurobi] sold to 200+ new customers acroos a range of industries (e.g. "ATM provisioning"; "sports betting").

Basic method is branch-and-bound. Relaxing the integer constraints leads to an LP problem. Then take a variable that should be integer but isn't, and try both $\lfloor x \rfloor$ and $\lceil x \rceil$ (and then do LP on one fewer variable!). Always keep track of upper bound (best solution we've found) and lower (LP solution). Difference is the "gap".

The last thing I want you to do is believe that Bixby has said we can solve these problems.

Example 47 (Schedule Generation Model) *157323 rows, 182812 columns, If we can solve the fleet assignment problem for a given schedule, why not optimise both? LP relaxation at root node was 18 hours. At 1710 nodes we found first feasible solution with a gap of 3.7%. Took 92 days!*

Example 48 (Real!) *44 constraints, 51 variables, maximisation. Immediately get a solutions at* -2186*. After 1.4 days, 5.5GB tree, 32M B&B nodes, made no progress.*

Example 49 (Toy) *Maximise* $x+y+z$ *subject to* $2x+2y \leq 1$*;* $z=0$ $x.y$ *free.*

Removing $z = 0$, or Euclidean reduction, do great simplification. Turning off presolve on all current codes will run forever.

Example 50 (Real supply chain) *Weekly model, daily buckets. Minimise end-of-day inventory. Production (single facility), inventory, shipping (trucks) to wholesellers (so demand is known). Initial modeling phase had a production grouping requirement, and a bizarre truck requirement (union rules). Couldn't get feasible solutions. So how did the humans do it? They fixed the producibles schedule first, and then solved in 1 hour (in 5.0; 4.0 wouldn't). Cplex 11 (2007) with Gomory fractional cuts took 0.63 seconds. So is the original problem soluble? Yes – 100seconds, and is 20% better than the soluton found with scheduler heuristics.*

1954 Dantzig/Fulkerson solves a 420-city TSP with LP and cutting planes.

1957 Gomory's cutting plane algorithms.

1960;1965 B&B formulated.

1969 BP does first commercial MIP.

1974 IBM's MPSX/370; and Sciconic. These were LP-based MIP.

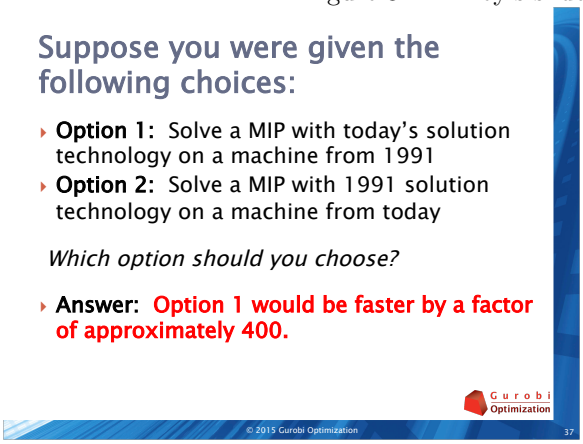
... Good old B&B remainded state of the art despite much theory (Padberg cutting planes, Balas disjunctive programming)

1998 Choice of branching variables (see TSP ideas: we had been very naïve here); lpsolve routines; cutting planes (Gomory's "lesser" ideas).

Our test set has 1852 real-world MIPs from customers. We used pure defaults. Best version→version improvements were 2.1→3.0 (Mature Dual Simplex LP: factor of 5.5) and 6.0→6.5 (Mined the theory backlog: factor of 9.5). After this we started solving real problems with "out of the box hits". Total improvement over 17 years 19990-2007 was $30,000\times$ in CPLEX. Since then in Gurobi we have seen a further $38.6\times$. Combining these two gives us $1.1M\times$, i.e. $\times 1.8$/year.

We see real problems with 1M rows/columns. 8% of models solved with gap $> 10\%$, 14% within $< 10\%$ and 75% optimal. Of the unsolved: 54 are blocked by LP, 16 were tunable, and 37 were "other".

Figure 8.1: Bixby's slide 37



Would you use today's technology on a 1991 machine or vice versa? [Most of the audience voted for modern computers and old software] Option 1 (today's software on 1991 machines) wins by 400×. See Figure 8.1.

JHD observes that 1.8× every year is 5.8× in three years, whereas the original Moore's Law [Moo65] is ×4 in three years. Clearly a win for the software/algorithms.

Q "Mixed Integer Rounding"?

A A simple idea for producing cutting planes.

Q What sort of time limit do people set in industry?

A We always customers what their criterion is. Often "overnight is fine". Other answers are "2 hours good; 5 minutes better".

Q How do you find the improvements? There was a slide showing version-on-version improvements. An older version of this slide was in [Bix10].

A 1998: literature. These days they come from practical problems, where we notice a feature, see if it shows up in the library, then implement it. This is a bag of tricks.

Q Thank you for your business model for helping researchers. What about quadratic?

A We can handle quadratic RHS. We don't see much of a demand for this.

8.3 On Convergence of the Multi-Block Alternating Direction Method of Multipliers: Yang

Subtitle: are there better methods for LP?

Minimise $C^T x$ subject to $Ax < b$ and $x \geq 0$. This is a data-driven model, which neds to be solved fast in practice. Geometrically, the constraints are a polyhedron, and the optimal plane has tobe found.

Markov decision processes provide a mathematical framework for sequential decision making where decision outcomes are partly random and partly decided by decision makers. At each time step, the process is in state i and the decision maker chooses an ation j, The process responds by moving to a state and producing a cost $c_{j,i}$ The probability of entering the next state in indepenednet of history. Hence we can ask for an optimal policy for teh decision maker.

Howard [1960] formulated this as "policy iteration method". ath-findng methods are $O9n\sqrt{n})$ But the im is to avoid matric inversion.

Eraly work was the vonNeumann projection (see als Freund).

Subgradient method [Renegar2014] transforms the problem assuming we know a feasible point with iteration complexity $O(L^{1/2}D^{/2})$. Also two-block ADMM.

onsider $min_{x\in\mathbf{R}^m} f_1(x_1) = \cdots$ such that $Ax = b$ We takethe augented Lagrangian function $L(x_1 \ldots, x_p, y)$ Do this or eahc x_i, then update y. Convergence was well-established when $p = 1$ or $p = 2$. What about $p > 2$? 2014 we would an example that can diverge when $p = 3$ $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$. Note that $\rho(A) > 1$. Perhaps we only update y by some $beta < 1$ of the true value. For $p = 1$. $0 < \beta < 2$ (Powell 1969). Is there a good problem-independent β? No!

Random-Permuted ADMM. Each round, use a random permutation for the order of updates. This seems ot owrk in practice,

Consider a square system of linear equations. After k rounds, we use a ...

Theorem 51 *If A is invertive, the expeted iterate ϕ^k converges ot the solution linearly for any $1 \leq p \leq n$.*

We can show that the "expected update matrix" $\frac{1}{n!}\sum_\sigma M_\sigma$ has radius < 1. Difficulty in proof as few tools for spectral radius of nonsymmetric matrix.

Showed examples of converges for large weakly Laplacian linear systems.

Consider the nonseparate quadratic problem to maximinse $t^T Hx + c^T x$. Then if each block converges, the whole converges.

So why multiblock? Consider the homogeneous and self-dual linear prgoramme to find $x., y, s)$ with $Ax - b\tau = 0$; $-A^T y - s + c\tau = 0$; $b^T y - c^T x - \kappa = 0$; $e^T x + \tau + e^T s + \kappa = 1$ $(x.\tau, s, \kappa) \geq 0$. Where the three blocks (x, τ), y, (s, κ) are alternately updated.

Also, consider the logarithmic barrier function as objective. Gradually reduce μ to 0 as in interior-point methods.

Note that ADMM is easily implementable on a distributed platform. Minimise $c^T \mathbf{x}$ Rather minimise $q_i(\mathbf{x}_i)$ independenently, then update $\mathbf{x}_0' := \max(\frac{1}{m}\sum \mathbf{x}_i, 0)$

1. Can we characterise the convergence rate? We have proved results about expectation.
2. can RP-ADMM convergence be "with high probability"?
3. Can we extend to more general convex optmisation?
4. So are there better LP algorithms out there?

8.4 Bounded-degree SOS Hierarchy for Polynomial Optimisation: Lasserre

LP- and SDP-certificates of positivity.

With $f'in'R[x]$ and $K := \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \geq 0.j = 1, \ldots, m\}$ being a compact semi-algebaric set. We are looking at the *global* minimisation problem. To prve this, we need to prove positivity of $f \geq f^*$. Can this be done effectively?

Real Algebraic Geometry helps. Such certificates exist, and are amenable to practical computation (note that Positivstellensätze for more general functions are not so amenable).

Theorem 52 (Putinar's Positivstellensätz) *If K is compact and satisfies a technical Archimedean assumption and $f > 0$ on K then*

$$f(\mathbf{x}) = \sigma_0(\mathbf{x}) + \sum_{j=1}^{m} \sigma_i(\mathbf{x}) g_j(\mathbf{x}) \tag{8.1}$$

where the σ_i are sums of squares.

Note that this theorem has no bounds on the degrees. Testing this (in bounded degree) is an SDP.

We can write K as $\{\mathbf{x} : g_i(\mathbf{x}) \geq 0; (1 - g_j(\mathbf{x})) \geq 0\}$

Theorem 53 (Krivine-etc.) *If K is compact and in that form, then*

$$f(\mathbf{x}) = \sum_{\alpha,\beta} \prod_{j=1}^{m} g_j(\mathbf{x})^{\alpha_i} (1 - g_j(\mathbf{x}))^{\beta_i} \tag{8.2}$$

This is solving an LP.

Many applications: generalised moment problem. This can be used to approximate set with quantifiers: $\{x \in \mathbf{B} : f(x, y) \leq 0 \forall y : (x, y) \in L\}$ for example.

We impose the constaint $\deg(\sigma_j g_j) \leq 2d$ in (52) to get element d in the SDP hierarchy, or same in (8.2).

This has been useful in problems of modest size, or larger if sparse. The SDP-hierarchy has been used in combinatorial complexity.

Note that our statement didn't distinguish between convex and nonconvex. and we can add Booleans by $x(1 - x) = 0$. However, the class of (easy) SOS-convex[1] problems is recognised as convergence occurs at the first level of the

[1] The Hessian factors as SOS.

hierarchy. For general convex problem convergence always happens. The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. best lower bounds) for hard 0/1 combinatorial problems, Note this doesn't occur for LP-hierarchy.

Theorem 54 (MarshallNie) *Let $\mathbf{x}^* \in K$ be a global miniimiser and assume*

1. *The gradients $\nabla g_i(\mathbf{x}^8)$ are linearly independent*
2. *struct complementarity holds*
3. *second-order sufficiency conditions hold at (x^*, λ^*)*

then

Therefore Putinar certificates are a generalisation. However, SDP-solvers have size constraints.

Can we do better. Assume that $g_j \leq 1$ on K (rescaling if necessary) and that $\{1, g_j\}$ generates $\mathbf{R}[\mathbf{x}]$. Then remember Lagrangian relaxation.

Q-EK I agree that you get a certificate with a good backward error. But is this meaningful? Only if the input problen is well-conditioned.

A Agreed the the certificate is not exact.

8.5 Smaller SDP for SOS Decomposition: Bican Xia

See [DX14]. It is known that SOS decomposition can be reduced to SDP, so in principle has a symbolic solution. We know that numerical SDP solvers can solve large SOS problems, and has available iplementations.

Definition 24 *SOSS(p,Q) means that p has support Q and is a sum of squares.*

This is equivalent to finding a positive semi-definite matrix M such that $p(x) = Q^T(x)MQ(x)$ where $Q(x)$ is the vector of monomials corresponding to Q. Let SOS(p.Q be any algorithms that solves this.

We find two classes of polynomials wher the original SOS proble can be transformed into smaller ones, and ways of detecting nn-SOS problems.

Between steps 1 and 2, we find check for evident non-SOSness, then check for a splitting.

Define the Newton polytope $N(p)$ to be the convex hull ....

Definition 25 *For a polynomial $p = \sum c_i \mathbf{x}^{alpha_i}$ and $T \subset \mathbf{R}^n$ denote by* Proj$(p.T)$ *the result of deleting*

Theorem 55 *If P is OS, then* Proj(pF) *is SOS for even face F of $N(p)$.*

Definition 26 *p is* convex cover polynomial *if there are some pairwise disjoint faces F_i of $N(p)$ such that $S(p) = \bigcup F_i$.*

In this case the problem decomposes.

In fact every convex cover polynomial is a split polynomial. If a polynomial. is split, then the SDP-matrix can be block-diagonalised.

States his Theorems 3 and 4, analogous to previous but for split polynomials.

Proposition 3 *Suppose Q has SOSS(p,Q) for a polynomial p. Then if p is SOS, the $\alpha \in Q + Q$ for all $\alpha \in S([)$.*

This gives us a quick negative check. After this check, we check for split polynomials, and solve them separately. Note that split polynomials may split further.

$SQR(k, n, d, t)$ is the sum of squares of k polynomials of n variables with t terms of degree at most d.

Various timing data: the check rejects all non-SOS polynomials very quickly.

Q–SMW Your non-SOS polynomials are very special. $g_1^2 + g_2 \sum_{i=1}^{n} x_i + 100 g_3^2 + 100$.

A We need the +100 to ensure the constantterms is not zero.

Q Which SDP solver?

A My student implemented this. The real point is the reduction.

8.6 Applications of homogenisation in SDP relaxations of polynomial optimisation: problems: Feng Guo

We want checkable conditions to veify non-negativity: NP-hard. For SDP programming the pimal problem: $\sup_W -Tr(CW)$ s.t. $Tr(AiW) = b_i$ and $W \geq 0$. $W^T - W$. Also dual problem. Spectrahedron $\{(\mathbf{x}) \in \mathbf{R}^n \mathbf{x}^T A.\mathbf{x} \geq 0$ and projected spectrahedron. Let Σ^2 be the set of sums of squares. Let $Q(X)$ ....

The Archimedean condition is equivalent to saying that S is compact. Shows an example of non-compact S with Putinar's Positivstellensätz failing.

So homogenise each generator to get $\widetilde{S}_>$ or $\widetilde{S}_\geq$ depending on what condition we impose on x_0. $f(x) \geq 0$ on S iff $\widetilde{f}(\widetilde{\mathbf{x}}) \geq 0$ on $cl(\widetilde{S})_>$ We say that S is *closed at infinity* if $cl(\widetilde{S}_>) = \widetilde{S}_>$. Detecting this is an open question.

8.6.1 Minimise a rational function

Mininise: $r^* + = \min \frac{p(\mathbf{x})}{q(\mathbf{x})}$. This is maximise r such that $p(\mathbf{x}) - rq(\mathbf{x}) > 0$.

Example 51 $\max \frac{1}{x_2^2+1}$ *which is not achievable, but is after homogenisation (i.e. at infinity).*

8.6.2 Semi-Infiite Polynomial Programming

$\min_{x \in X} f(x)$ subject to $G(x, u) \geq 0 \forall u \in U$. Problems when U is not compact.

8.6.3 Convex hulls of semialgebraic sets

Projected spectrahedron. Let P be the set of support hyperplanes of S and $M]\{(1.x) : x \in \mathbf{R}^m\}$. $\{1\} \times cl(co(S)) = P *^{\cap} M$.

Problems in the noncompact case. $\{x_1 \geq 0; x_1^2 - x_2^3 \geq 0\}$. Need a modified Lasserre's relaxation. We perspectively prject to $X_0 = 1$. We assue S is closed at ∞ and ....

8.7

Two tasks: compute zeros, and check whether we have an approximate zero.

Given a complex polynomial, we cantake real and imaginary parts.

Example 52 *$f_1 = x^2 - 2y$, $f_2 = y^2 - x$, $f_3 = x^2 - 2x_y^2 - 2y$. Then $P_0 = (00)$ is a zero of F_1, f_2, and we can get P_1 by Newton. But when we add F_3. We are going to consruct a square system containing all zeros of Σ, preserving reguar zeros: $f = \sum f_i^2$ and $\frac{\partial f}{\partial x}$ $\frac{\partial f}{\partial y}$: solve for the derivatives and $f - r$ in $\mathbf{Q}[x, y, r)$.*

For such a square system, we can use homotopy and *a posteriori* certificaton. Many deflation techniques for multiple zeros.

Theorem 56 *P is a simple zero of Σ iff $(P, 0)$ is a simple zero of the square system $\Delta = \{D_1(f), \ldots, D_n(f), f - r\}$.*

He gave a proof. Emphasised that, although the total degree has doubled, the computation is not much greater.

Theorem 57 *Given $\Sigma \subset \mathbf{Z}[X]$ and (P, r_0) is a zero of Δ within the root separation bound*

Showed an example with [Tsigaridasetal2010a] version of DMM bound of 10^{-138} while true answer is 2.5. Hence he claims that the certified simple zeros from Theorem 1 of Δ are with high probability the zeros of Σ.

Would like to look diretcly at multiple zeros in Theorem 1.

Q-Mourrain You could also consider $f_3 - r$ in the extended system.

A Not sure that the proof works.

Q You proof considers $A^T A$, so squares the condition number. Have you eplored this?

A No.

Q Why the slack variable?

A So that the Jacobian isn't zero.

8.8 Algebraic boundaries of convex sets: Sinn

Consider a polynomial optimisation: problem: minimise ℓ over $\{g_i(\mathbf{x}) \geq 0\}$. The optimal value function maps the coefficients of ℓ into the minimum value. By Tarski, thisis semi-algebraic.

A convex set is semi-algebraic iff its dual is. The dual of C is the set of supporting hyperplanes of C. $C = dual(dual(C))$. The *algebraic boundary* $\partial_a S$ of s is the Zariski closure in $'A^n$ of its boundary in the Euclidan topology. The optimal value function satisfies $\Phi(-\Psi(-alpha_1, \ldots, \alpha_n), -alpha_1, \ldots, \alpha_n) = 0$. hwer Φ is the defining polynmial of the algebaric boundary of the dual convex body.

Recall normal cone to $x \in \partial C$ is $N_C(x) = \{\ell \in (\mathbf{R}^n)^* : \forall y \in C \ell(y) \geq \ell(x)\}$.

There's also a dual variety: Zariski closure

$$\{[H] \in \mathbf{P}(V^*) : [H] \text{ is tangent to } X \text{ at a regular } P \in X_{reg}\}.$$

Biduality for *irreducible* varieties.

Definition 27 *An extreme point of C is a point x such that if $x = \frac{1}{2}(y+z)$ with $y, z \in C$ implies $y = z = x$.*

Theorem 58 *Let $Z \subset \partial_a C^0$ be an irreducible component. Then $\overline{Z}^*$ is an irreducible subvariety of $\overline{Ex_a(C)}$ the Zariski closure of the extreme points. $\overline{Z}^* \cap Ex(C)$ is Zariski dense in $\overline{Z}^*$. Assume $Y \subset Ex_a(C)$ and $Y \cap Ex(C)$ is Zariski dense in Y. Then $\overline{Y}^*$ is ...*

Theorem 59 *Let $C \subset' R^n$ be a convex compact semi-algebraic set with 0 in its interior. Suppose that every irreducible compact of $\partial_a C$ is smooth along ∂C. Let $Z \subset Ex_a(C)$ be an irreducible subvariety such that $\overline{Z}^*$ is an irreducible component of $\overline{\partial_a C}$.*

If $\operatorname{codim}(Z) = 1$ ...

Shows a semi-algebraic description of the "hard cases". See [Sinn2015 Researchinthe Mathematical Sciences 2(2015)]

8.9 Symbolic-numeric Methods for Linear and Integer Programming: Steffy

Linear Programming and MIP. 0/1 variables here. See Section 8.2. We use SCIP and SoPlex from ZIB — open souce and within an order of magnitude of commercial. Notethe problems of lfoating point, and we may only get "nearly feasible" solutions. Note that VLSI verification really cares about correctness. Ther are also poorly sclaed/conditioned problems, some biological systems.

Note that the dual solution provides an optimality certificate. Simplex also allows for fast re-optimisation (needed in MIP). `QSopt_ex` does nuerical simplex: checked exactly, then increase precision if necessary. Well-tried examples of

iterative refinement for linear systems, in mixed precision. an apply these ideas, but need to do the dual problem as well.

Near singularities. we keep adding nearly parallel planes. so ill conditined.

Q Mixed nonlinear systems?

A The solution may not be rational any more.

8.10 Problems on Symbolic Computation of Polynomial Equations in Wavelet ANalysis: Bin Han

For $\phi, \psi^1, \ldots, \psi^s \in L_2(\mathbf{R})$ define an affine system $AS(\phi, \psi^1, \ldots, \psi^s)$. Affine systems can have dilates and shifts.

Need to construct tight framelets from Filter Bank. Haar Orthonormal wavelet $\phi = \chi_{[0,1]}$ and $\psi = \chi_{[0,1/2]} - \chi_{[1/2,2]}$. Also Shannon and Daubchies wavelets.

Common characteristics are that we have linear equations from the approximations, and we also get total degree two equations.

Conjecture 7 (Since 1988) *Real-valued orthogonal filters a having arbitrary high linear-phase moments and smoothness $sm_2(a)$.*

Current stasus $lpm(a) = 5$; $sm(a) \approx 2.449$. Note that we would want $sm(a) > 2.5$ for a C^2 function ϕ.

A $d \times d$ matrix M is called a dilaton matrix if it is an integer ....

Problem 6 *Construct a finitely supported 1D real-valued orthonomral M-wavelet finter a such that $sm(a)$ is large and a is symmetric.*

Can prove it doesn't exist for $M = 2$. For $m = 3$ $sm(a) \approx 2.06$. For $M = 4$ $sm(a) \approx 2.53$.

If $M = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ then we have a problem with 36 unknowns — unsolved.

Also a sOS problem. For a given A, find filters $U_1, \ldots, u_s$ such that

$$|\hat{u}_1(\xi)||^2 + \cdots + |\hat{u}_s(\xi)||^2 = \hat{v}(\xi)$$

where $\hat{v} = \cdots$.

Chapter 9

13 August 2015

9.1 Without Mathematics and Supercomputing, no Effective Risk Reduction of Natural Disasters: Qing-Cun Zeng

Natural disasters cost many lives and much economic damage. 22% windstorm, flood 35% etc., 85% being meteorological in all. 53% of the deaths were caused by meteorological disasters.

In China, Great King Yu (21st C BC) led his people is controlling rivers. We should emphasise that there have been great progress in the last 50 years in risk reduction of natual disasters.

Typhoon/Hurricane Galveston (1900 Sept 8) kill6-8K people. 1992 Hurricane Andrew was tracjed by satellite, but Numerical Weather Preddiction could only given 24 hours warning. Sandy (2012) we had five days warning(better NWP, sending and assimilation). In China, storm Rammasun was monitored by Chinese satellite. NWP successfully prdicted the track. Landfall warning 36 hours in advance. No life was lost, but nevertheless direct econmic loss 26GYuan. NWP predicted heavy rain in Beijing oin2013, but several people wer killed. The NWP did not predict the quantity of rain (beat all records). Also the civil emergency system was not good enough.

Gave a six-step action plan (including verification and lessons learned). But the first stage, prediction, is key. Remote sensing needs to be inverted to find physical quantities, e.g. water vapour desnity from radiation. Thisis a Fredholm integral equation of the first kind. E/U of the solution requires $g(z'), W * z, z')$ to satisfy certain constraints. Also ill-conditioned.

Weather prediction is compliicated. There is mass conservation for atmospheric water (in three phases), cloud formation and evolution (very complicated), lower boundary conditions (kinematic, geometrcal and physical), and

upper boundary conditions (what is the upper boundary?) All fluzes (except radiation) $\to 0$ and $z \to \infty$. Then their initial conditions. Wellposedness was proved by Chinese scientists for linearised equations. For the nonlinear model we do *not* have stability. Meteorologists call t_c the "predictability" — the point after which prediction becomes impossible.

9.1.1 Computing Problems

1. Numerical Prediction. Note that von Neumann proposed numerical weather prediction, hence NWP and Computers have a common father. The two have grown up together! [Richardson: "Weather Presdiction by Numerical Processes" 1922] had a FEM formulation. This didn't succeed, because the equations were too simplistic, and computational resources inadequate. Shows a graph of ECMWF forecasting accuracy, and notes that storm warnngs from regional centres are 3-days, which is adequate.

   4D data assimilation was proposed by a French meteorologists in 1987. We can be proud of what we have done, but should do more to improve the disaster prediction. Need to iprove resolution (akss for a grod size of 500m-1000m!!), This would require morepowerful computers. We also need numerical (quantitaive) predictions of disasters, e.g. water flow. This requires a high-resolution ground model. Note that there are very irregular and complicated boundary conditions.

2. Visualisation. There is an optimisation problem: minimise (cost of actions) + (losses prevented). This requires real-time regulation, and the ability to explain to policy-makers.

Q How do we cope with the chaos caused by nonlinearity?

A Ensemble prediction.

Q–JHD Can we really achieve this incrase in resolution globally? Should we not be looking at local resolution?

A We need more powerful computers. [JHD fears he did not explain the question well enough]

9.2 Software and applications for polynonial homotopy continuation: Leykin

Q What is the meaning of "algebra" in Chinese?

A I have been told it means "substitution mathematics".

Q Ask your students what algebra means.

A In Arabic it means "the union of broken parts", or possibly "bone setting".

Example 53 *Have a target [polynomial] system F, and a simple system G with the same number of solutions, Then consider $H(\mathbf{x}, t) := (1-t)G(\mathbf{x}) + tF(\mathbf{x})$. Hence $\frac{d\mathbf{x}}{dt} = \left(\frac{\partial H}{\partial \mathbf{x}}\right)^{-1} \frac{\partial H}{\partial t}$. But there may be numerical issues, and problems with targeting singular solutions. There is also the issue of certification*[1]*.*

Software can even describe positive dimensional solutions (numerical algebaric geometry (PHCpack, Bertini, NAG4M2) but we won't talk much about this.

Let $K = \mathbf{C}$ (occasionally $'R$). Solve parametric problems for generic paraneters. Given $\Psi \in K[\mathbf{p}.\mathbf{x}]^m$ and $V \subset \mathbf{A}_P = K^{\#p}$ suchthat for a *generic* $\mathbf{p}_0]inV$. we have finitely many solutions.

If you wanted to solve this non-numerically, we could look at parametric or comprehensive Gröbner bases. This very expensive.

1. Take a generic codim(V) plane O
2. Find a structured witness set $V(\Psi) \cap (L \times \mathbf{A}_\mathbf{x}) \subset \mathbf{A}_\mathbf{p} \times \mathbf{A}_\mathbf{x}$.
3. Given $p_o \in V$ pick generic L_0
4. Deform L to L_0

Example 54 (Computer Vision) *Point X is projected on to three* calibrated *cameras, with local coordinate frames I, R_2, R_3. Use Cayley parametrisation of SO_3 (six parameters for two SO_3 instances). We have a rational map from the space of configurations (dimension 23) to V the space of views (dimension 24). n fact the pre-image has cardinality 1 [HoltNetravali1995]. Proof (uses [MorganSommese1989]): compute a Gröbner base in Macaulay (pre-M2) and prove the number of points in the fibre is constant, therefore 1.*

Notes that the GB solution is used in practice (Android 'phones). But takes 1second. With homotopies we have singularities, but when we are not too far from the reference solution, we are at 100ms, and could make 1ms. The bottleneck is the cmputation of Ψ and its derivatives, which is done via an SLP.

Example 55 (Definite Representations) *$f \in \mathbf{C}[x, y, z]$ of degree d. Determinental map $\Phi : M \to F$. [Nui1968] the set of hyperbolic polynomials is closed in F, contractible and path-connected [good news for homotopy!]. Follow Nui's paths (avoid singularities).*

Q How do you choose the parameters, and can you identify the bad ones.

A In practice it is hard [impossible for the vision problem] to identify the bad points, as you need to invert a rational map.

Q How big a problem can you solve?

A Millions of zeros if necessary. But note the vision problem, for example, only has one realistic solution.

[1] Claims that there is a difference between "verification" and "certification".

9.3 Bertini 2.0 and BertiniLab: Software for solving polynomial systems numerically: Bates

BertiniLab is a MatLab interface! But I'm not going to talk much about this.

1. This talk is numerical, but I believe in a mixed approach for many cases.
2. There's much other good software besides Bertini.
3. This is a software development talk, not algorithms.

In BertiniLand, we go from $t = 1$ to $t = 0$ [arguments about density of floating point numbers etc. can be made, but it's just a feature]. Bertini was written in the basement of the Students' Union at Notre Dame. Bertini how has some AI: autochanging other tolerances as you change one.

BertiniLab was written for a specific user.

`alphaCertified` will certify (in the sense of alpha theory) solutions.

`Paratopy` uses parameter homotopies and parallelism: using Bertini as the core engine. Does summary statistics.

`BertiniReal` — see Schost's ideas. Uses Morse-theoretic ideas to project critical points down, and look at the fibres over special points. It can produce MatLab `.fig` files and STL files [3D printers]: hence polynomial system → pretty solid.

We have money from "Advanced Cyber Infrastructure" branch of NSF for Bertini 2. 1.5 people → 10! More modular, GPL licencing, regression tests etc.

Q–EK LinBox took a long time to incorporate software from elsewhere: but eventually includes R+. What are your numerical analysis plans?

A We use Python for Numerical LA, and BOOST.

Q Also, it took us far too long to join SAGE.

A That's where I get the most flak about our licence. We're happy to join anyone.

9.4 Computing mixed volume ... in quermassintegral time: Malajovich

See `http://www.labma.ufrj.br/~gregorio/papers/beijing/pdf`

Mixed volume (Minkowsky). Take n convex objects in n dimensions. $V(A_1, \ldots, A_n) = \frac{1}{n!}\frac{\partial^n}{\partial t_i} Vil(t_1 A_1 +' cdots)$

Similarly Steiner formula: $Vol(A + \epsilon B^3) = Vol(A) + S\epsilon + \pi K \epsilon +^2 + \frac{4\pi}{3}\epsilon +^3$.

Quermassintegral: $V(A, A, B^3) = 3S$. $V(A, B^3, B^3) = 3\pi K$.

Theorem 60 (BKKh) *he generic number of roots in* $(\mathbf{C}^\times)^n$ *of* $f_1(x) = \cdots = 0$ *is* $n! V(Conv(A_1), \ldots, Conv(A_n))$.

To be contrasted with Bézout, which is for dense polynomials. Note that the "Bézout\BKKh" roots are in general degenerate, so bad news for homotopy.

Legendre transform of $\mathbf{a} \mapsto b_i(\mathbf{a})$ if $\xi \mapsto \lambda_i(\xi) = \max_{\mathbf{a}\in A} \mathbf{a}\xi - b_i(\mathbf{a})$. Tropical semi-ring: $(\mathbf{R} \cup \{\infty\}. + . \max)$ [That's what he wrote, but h said the other way round, which seems better]. Tropical limit: $\lim_{r\to\infty} e^{\tau\cdots} \sum \ldots$.

Mixed cells are *dual* to the solutions ξ of the tropical polynomial system.

Theorem 61 *With probability 1, the algorithm*[2] *computes the mixed volume and produces all the initial points in time bounded by $O(T + T')$ arithmetic operations, where [he suddenly changed into a software demo!] and*

1. *with probability 1. $v_d \leq n!V(A_1, \ldots, A_{d-1}, A, B^n, \ldots, B^n)$*
2. *...*

Shows a graph of T against measured time, which does look linear. Note that there is non-trivial numerical analysis involved here to translate the theorem into implementation. At times I need to move to quad computation.

9.5 Classifying Polynomial Systems Using the Canonical Form of a Graph: Yu

Polynomial $\sum_{\alpha\in A} c_\alpha \mathbf{x}^\alpha$. Homotopy from a binomial system to this. Shows PHC Web Interface. We have 16 cores of CPU and two K20c nVidia cards. Want to store various supports A to facilitate re-using stored results. But how do we store a support in a way that allows for commutativity etc.

Therefore want a unique key.

Definition 28 *Two polynomial systems are* isomorphic *if*

1. *they have the same dimension and number of equations*
2. *there is a permutation of variables and one of equations that takes one to the other.*

Hence think of variables as root vertices, pointing to powers, and hecne to monomials, Then can use standard graph tools.

Q What's the connection to graph isomorphism?

A In general graph isomorphism is hard, but ....

Q Look at Maple's technology: this very rapidly finds a match.

A Need one to be invariant over variable names.

JHD/Dan Roche Maple's is not invariant over variable names.

Q-Dan Do you solve the user's system or the one in the database?

unclear discussion.

[2] It is a randomised algorithm.

9.6 Labahn

Our solutions are invariant under symmetry groups. Why is there a set of dundamental invariants (?for the system, or for the solutionspace).

Action: $\mathcal{G} \times K^n \to K^n$. See [Gat90], [FaugereSvartz2013a], [HL12]. [HubertLabahnMathComp].

Definition 29 *$\mathcal{G}$ is diagonalisable if $\exists R : R \cdot \mathcal{G} \cdot R^{-1}$ is diagonal.*

Recall also the structure theorem for abelian groups. Also Hermite normal form for matrices. which we can make canonical.

Example 56 *Invariant under $x_1 \to x_2 \to x_3$. Write as $diag(\omega, \omega^2, 1)$.*

It's harder when the system is invariant, but not the individual polynomials.

We want to extend from abelian groups to soluble ones.

9.7 Arnold

Compares dense and sparse representations: dense has fast arithmetic. The *support* of a polynomial is the set of exponents of its non-zero sums.

Definition 30 *The sumset $A \oplus B = \{a + b : a \in A.b \in B\}$.*

Definition 31 *The* structural support *of $f \cdot g$ is $\operatorname{supp}(f) \oplus \operatorname{supp}(g)$, and the structural sparsity is $|\operatorname{supp}(f) \oplus \operatorname{supp}(g)|$.* That is, "ignoring cancellation"

Theorem 62 *Ther is a randomised algorithm that, with probability > 0.99 computes $h = f \cdot g$ on $\tilde{O}(Sn \log D + T \log C)$ whre C is a bound on the coeffcients and $D > \max_d \deg_{x_i}(f \cdot g)$.*

Note also [ColeHanharan2002] have a Las Vegas algorithm $\tilde{O}(T \log C \log^2 D)$.

Note that "grade school" is as ggos at it gets when there's no collisions. But squaring is a classical case where there is collision. Also composition of sparse polynomials.

Define $h^{\bmod p}$ to be $h \pmod{x^p - 1}$. Hence we can define "collision" of terms when we use $^{\bmod p}$. Say p is "good" if there are no collisions, and "OK" if less than half the terms collide. If we knew the number of terms in the sumset, we could produce probability estimates. So guess the size as $2, 4, 8, \ldots$, pikc p_G and $p_O K$ according to such estimates and [check for sanity].

See [AR15] for the details.

9.8 Computing Approximate GCRDs of Differential Operators: Giesbrecht

Note that we don't have unique factorisation. GCRD is the right question for joint solutions of differential equations.

Problem 7 *[Approximate GCD]Find* $\tilde{f}$ *and* $\tilde{g}$ *such that* $\deg \gcd(\tilde{f}, \tilde{g})) > 1$ *and* $\frac{||f-\tilde{f}||}{||f||}, \frac{||g-\tilde{g}|}{||g||}$ *small.*

Lots of alternatives, but this is the definition we shall generalise.

GCRD dates back to [Orr1933]. [li1997] hd a subresultant theory for generalised Ore polyomials. Aim was to merge these with [CGTW95].

However, it is not obvious that the question is even well-posed. Nee to define norms correctly, then can generalise Problem 7 precisely. Clear denominators and use[Kaltofenetal2006] to clear approximate contents.

Think of the differential Sylvester matrix S. The degree on the exact GCRD is the nullity of this matrix. Then inflate S to $\widehat{S}$ with numerical coefficients and do a reduced rank calculation here.

We can set this up as an optimisatoin problem, and then ask whether Φ *attains* its minimum. We can also ask whether Newton iteration is going to converge. Use ideas from [Kaltofenetal2007Unpubl].

Theorem 63 *Define the set of possible solutions: impose that* $\mathrm{lc}_x(\mathrm{lc}_\partial(h)) = 1$. *Then if the set is non-empty, the infimum of the error is attained.*

Also σ_v — the smallest singular value of the inflated Sylvester matrix — is some measure of[3] the condition number.

Q Does your unstructured perturbation take you to an inflated Sylvester matrix?

A Not necessarily, but near enough. Then the Newton iteration should take you back.

Q–EK Approximate factorisation?

A Future work.

9.9 European Research Funding: ERC and Mathematics

9.9.1 Bourguignon

ERC is a bottom-up individual-based pan-European comptition with host institutions in EU/Associated Countries. 15% of referees are outside the EU. There is an independent Scientific Council[4] with its own executive agency.

Starting 2–7 years[5] post PhD. Up to 1.5M+0.5(large facilities)

Consolidator 7–12 years post PhD. up to 2M+0.75

[3] Still needs more work.

[4] Just had an

[5] Throughout, women can automatically cliam 18 months extra/child: men if they can prove they had leave of absence.

Advanced Up to 2.5M+1

Proof of Concept Reserved for people who had already have an ERC grant.

ERC gets 17% of the EU Science Budget. 1.6Geuro this year, i.e. approximately 1000 grants. 2/3 of grants go to people between 30-40.

About 27% of the postdoc money used here goes to non-EU citizens. Note that these grants are portable (which ensures they are well-treated by hosts!). 8% of ERC grants go to PIs who are not European[6]. Success rate this year should be about 15% (up on historical).

During FP7 supported 4300 of which 237 in Mathematics. Last year 35/937 were mathematicians (decrease in proportion, which is slightly worrying). 2 Fields Medals and 3 Nobel Prizes went to ERC holders.

9.9.2 China

Ma [Sugaku Tushin 12(2007) 1]. Mathematics has a special fund (Tianyuan) and staus in China. But there are many funding schemes. Shows basic funding graph 1999-2009.

We are very happy with the improvements over the bast ten years. "National Centre for Mathematics and Interdisciplinary Sciences (CAS". Launched 24 November 2010 as part of Innovation 2020 initiative.

from Shandong But in China there is a problem for counting [evaluating numerically?] people, especially in universities. It tends to be numbers of papers and grant income.

Chan: President HKUST We have a separate system. I was also an AD at US NSF including mathematics.

NSF Funds 62% of the US mathematics basic research. Increased need to show societal relevance. Mathematicians do not participate as enthusiastically in the big inter-disciplinary programmes (Big Data etc.) as I think they should. Also there's more private money: Simons is about 25% of the size of NSF's DMS.

Shows graph of NSF's funcing" flat until 1984, climbing until 2002, then flat.

NSF does fund Oberwolfach for American participants.

HK UGC is 65%, RGC is 9%, rest private etc.

Institutes Mathematics is cheap, and institutes are cheaper than observatories.

Chinese NSF Shows 1986–2015 cumulative figures. Also figures for 2015. Tianyuan had 700 proposals but only 31 awards. In 1989 Tianyuan was 1M RMB/year, but now it is 25M/year. In the early days this was used

[6]Must spent at least half of the year in a (given) European Laboratory.

partly for grants, partly for "other matters important for the development of mathematcsi", but now all on this second task. There is an Academic Leading Group of the Tianyuan Fund. Administered by an office in CNSF.

9.9.3 Evaluation in ERC

There is a five-minute video ob ERC website which she recommends everyone to see.

Schemes ERC Grants: see J-PB.

Implementing Agreement Chinese researchers hwo are active holders on NSFC grants can be part of ERC teams for 6-12 months: NSERC covers international travel, ERC grants cover subsistence etc.

Enrique (Chair of Math Panel) Round 1 expects a 3:1 cut. Evaluates B1 and the CV. The Panel is 12–14 people. Four panelists will read your application. The key task is to envisage an ambitious major research theme, incorporating a team of several people. Diagrams about research group organisation help: you need both focus and to be understood by a large range of people.

Volker Mehrmann My grant is Maths/CS/Engineering, but there is no longer an interdisciplinary panel. This means that you have to make it in your own field first (see previous).

Maria Esteban: Chair Step 1 is only by panel members (but can ask other ERC panels). The second stage is external referees. At least four external reviews as well as the panelists. But the panel determines the ranking.

Carillo Starting and Consolidator have an interview as well. These are obtained before the interview. A lead reviewer is nominated for each candidate.

9.9.4 Past Grantholders

Annalisa Buffa I applied to the first round of ERC Starting Grants. "Innovative compatible discretizations for PDEs". This was a new and exciting topic. This was my chance to build a team. Complicated diagram of people, grants (one researcher got an ERC consolidator grant) and destinations.

Martin Hairer Consolidator Grant just before Fields Medal. So I had recently developed a theory of regularity structures to give meaning to stichastic PDEs that were previously thought to be ill-posed. *Ann Math.*; *Invent. Math.* etc. Hence the aim was to understand cross-over between regimes. Phase coexistence is one example.

As well as hiring a range of postdocs from different disciplines, I am organsing workshops, which is very easy at Warwick. I have been pleasantly surprised by the application process (and Warwick's research support staff). The interview (being shuffled from waiting room to waiting room) was the strangest part of the experience. The ongoing adminstration has been easier than I had feared, in particular changing the start date. Had problems with the University of Warwick's Housing system and the ERC's double charging rules.

Coron Spoke about nonlinear control, where the nonlinearity is important. I was at the Institut de France, with no teaching, but this was running out. How else to avoid teaching? Apply for an ERC Advanced Grant. More to the point, PhD students (2) and postdocs (which are very hard to get in France: these were my first). The administration is not very heavy.

Q Is there a right of respond to referees?

J-PB There is a formal "redress" process (about 2%), and we are always trying to make the reports of the panels helpful. The names of the referees are not revealed.

Q Suppose you have a string track record, but with to change area. The referees might not take account of this

J-PB This is something I stress in the briefing to panel members. The Panels do take risks like this.

Q Can an ERC grant holder and I (non-EU) write visits to each other into our grants?

A Yes.

Q Interviews are known to reinforce gender bias.

A We had 18% of women applying in Mathematics (which is above the EU average), but the success rate did not match this. In the past women did less well than men at stage 1, but better at stage 2. This year it was about even.

Chapter 10

14 August 2015

10.1 Applied Mathematics for Business Decision Making: the Next Frontiers: Kempf

Speaker is Chief Mathematician at Intel.

Every area has its core problems, which take years, decades, centuries to solve. Consider Weather Forecasting (Section 9.1) for example. This has interesting mathematics, and is important for humanity.

10.1.1 Background

The human brain has changed little since 200,000BC. First steps were fire, domestication of plans and animals, then the industrial revolution.[1] Observed that even in inflation-adjusted $, companies have grown bigger. 1900 Standard Oil was $1.4G ($70G in 2015), 1955 GM was $10G ($100G in 2015) but Sinopec is $500G.

Note that intuition lives in the earlier parts of the brain. Good intuition comes from structure, repetition and feedback. But business decisions tend not to fall into this category.

Intel has gone from 2300 transistors in 1971 to 6.5G today. Question: do computers belong on this timeline.

10.1.2 Problem

When the Chief Mathematician (speaker) goes into a room full of Vice-Presidents, they don't take his word for it: they want to apply their "business intuition".

1713: Nicholas Bernoulli and the Saint Petersburg Problem was the start of "perfect rationality" and the Expected Utility Hypothesis. But see Herbert

[1] JHD notes that writing, arithmetic etc. were omitted.

Simon's research (Nobel Prize in Economics) "Models of Bounded Rationality". Daniel Kahneman (2002 Nobel) claims that humans have biased bounded rationality. See his book "Thinking Fast and Slow".

> Overconfident professionals sincerely believe they have expertise act as experts, and look like experts. "You have to struggle to remind yourself that they may be victims of an illusion" — Kahneman.

But: how biased are we, and what can we do about it.

10.1.3 Towards a solution

Human techniques.

1. Heuristic search with paper/pencil
2. Heuristic Search with a Spreadsheet
3. Optimisation with a strong technique (CPLEX)
4. Automation with a strong technique

Solutions

1. Over-riding intuition
2. More recently, implicitly using intuition
3. Now, explicitly using intuition

Example 57 (A new (Intel) factory) *Building (clean rooms) costs \$2G; equipment \$6G*[2]*. If I give you the flow, characteristics of the equipment, can you decide how much equipment is needed? Note that there's re-entrant flow — machines used repeatedly in the process.*

Traditional methodology was divide and conquer: cost each machine type separately. Problem is that all the equipment is independent, but the flow is not. Hence we now do a discrete event simulation (warm it up for a year of simulated time, then run for two years). Typically end up with a set of equipment that costs less *but actually produces more.*

Example 58 *Lead time for equipment is growing: current 4Q-5Q. But forecasting is getting harder (1D–2Q). So what is your demand forecast? If we order equipment to hit the upper forecast, we risk using \$300M unfulfilled, If we hit the lower forecast, we risk unfulfilling \$3G of sales. So current strategy is to buy from the lower bound, with options (including paying for long-lead-time sub-assemblies) for more.*

[2] [Hsu21], admittedly six years later, claims \$5–20G, and quotes a single extreme ultraviolet lithography (EUV) machine that costs more than \$100 million.

I am confident telling *any* group at Intel that using our tools will halve your decision time (I normally get 5×–10×) and get a 5% better solution (I expect 10%–15%).

Then we should recall "the wisdom of crowds". Consider "Guess the weight" contests — the mean is generally closer than any individual forecast. Note that we hear 1:1 from our customers. Linear regression of forecast/actual purchase has r=0.78. There's a "Bass Model" for technology diffusion. This reduces our average forecast error by 25%.

Also set up a "prediction market" internally for our experts to buy "shares" in forecasts. 6 of 11 quarters are ±5%, 10 of 11 are ±10% — pretty good.

A large range of possible projects, with inter-relationships. Key concept is the "efficient frontier of non-dominated Portfolios". Then use "elimination by aspects". First one is budgets — in line with plans. Then look at "products/projects in all", in none etc. Then resources (by skill set) etc. Then market balance.

Executives are irrationally overconfident in their decisions. Application of analytics to exclude or employ intuition can yield better/faster decisions.

But what we are doing at Intel is only the tip of the iceberg! Call for young mathematicians to do more research here.

10.2 Developments in Computer Algebra Research and the Next Generation: Yokoyama

"Heuristic Counting of Kachisa–Schaefer–Scott curves": JSIAM Letters 6(2014) pp. 73-76.

Consider elliptic curve cryptography. Note that you can draw a curve in $\mathbf{R}^2$, but looked at over a finite field it's a set of dots. Given P and Q, can we compute n such that $Q = n \times P$.

Pairing-friendly curves. Supersingular. Miyaji-Nakabayaski-Takano, Barret-Naehig are the ones that interest us.

$$Q(y) = \frac{C}{\deg q^+ \deg r^+} \int_2^y \frac{1}{(\log c)^2} dx.$$

Use Hosten-Thomas' algorithm.

Computed various examples from `isl`.

10.3 Lattice-based Analysis and Their Applications in Public Key Cruptanalysis; Morozov

Note [Cop96] and [HerrmanMay2008] if we know some bits of p. [Bloemer-May2003] if some bits of d are known.

Other side-channel attacks. Suppose $N_i = p_i q_i$ and suppose p_i share bots. Once studied by Faugère.

Lemma 7 ([HG97]) *If* $||g(X_1 x_1, lgots, X_n x_n)|| > \frac{N}{\sqrt{\omega}}$ *then the root is exact.*

1. Collect polynomials with root x_0 modulo N^m
2. Construct a lattice with coefficents $g_i(xX)$ as basis vectors. Thel LL reduce
3. If $X < N^{1/3}$ this is an exact solution.

What we are doing: improve Sarkar-Maitra and revisit [Pengetal??].

10.3.1 SarkatMaitra

Suppose p_i share a certain number of MSB. Then $\gcd(N_1, N_2 + (p_1 - p_2)q_2, \ldots) = p_1$. We onbserve that $u_i^{(0)}$ contains a large prime q_1 determines by N_1.

Theorem 64 *Suppose* $p - 1$ *have* αn *bits (*q_i *have* $(1 - \alpha)n$ *bit)* p_i *share* γn *bits. Then* N_1, N_2 *can be factored in polynomial time if [condition on* α, γ*].*

10.3.2 Pengetal

Based on [MayRitzenhofen]. When $\gamma < \frac{k}{k-1}\alpha$ the reduced basis doesn't actually contain the required vector. They use [HerrmanMay2008]

Apparently these two produce the same bounds. This works in the case of balanced moduli ($\alpha = 0.5$). 512/512 bits needs 460 shared bits: lattice has dimension 105 and takes 2000 seconds.

10.4 Mansfield

See also work of Hubert. She really put moving frames on a rigorous basis for symbolic computation. I need a Lie group. $G \times M \to M$ is a regular free action. The elements of the group foliate the space, and there is a unique element of G that moves m to m'. $\rho : M \to G$ by

Solve $\phi_j(g \cdot z) = 0$ for $j = 1/ldots \dim G$. Solve this and invoke Implicit Function Theorem.

If $I(z_i)$ are the canonical invariants for $z = (z_1, ldots, z_i)$ and $F(z_1, \ldots, z_n)$ is an invariant, then we have a replacement rule

$$F(z_1, \ldots) = F(g \cdot z_1, \ldots) = \cdots.$$

Example 59 (difference) $u_{n+k} \mapsto \tilde{u}_{n+k} = \lambda u_{n+k} + \epsilon$ $\Phi : \times u_n = 1, \times u_{n+1} = 0$. $\lambda = -\frac{1}{u_{n+1} - u_n}$ *and* μ. *But* $I(u_{n+k})$ *has a fixed base point, which is not what I want. End up with a matrix in* λ, ϵ *independent of* n.

In anything, we end up with too many invariants. In the differential case, the components of the Maurer–Cartan matrices are (almost) generating.

The point is to be able to solve for the invariants without solving for the frame.

"multispace" is a manifold that contains the jet bundle, but also local lattice embeddings. Regard jet space as equivalent classes, and a function as equivalent to its Lagrange interpolation. Hence the points we interpolate have to be in general position. If points coalesce, we have points with multiplicity: ultimately interpolation becomes Taylor series.

10.5 Binomal Differnce Ideal and Toric Difference Variety: Yuan

In the algebraic case, these are well-studied. [EisenbudSturmfels1996].

A lattice is a mobule in $\mathbf{Z}[x]^n$. Note not a PID so may not have HNF. $(f_1, \ldots)$ is a GB iff a generalised Hermite form. (F, σ) is a difference field. Assume F algebarically closed. If $p = \sum c_i x^i$ write $a^p = \sum (\sigma^i a)^{c_i}$. Ratio of σ-monomials is a Laurent σ-monomial. Hence Laurent σ-binomial:.

We can define a partial character ona lattice L_p is a homomorphism onlt F^*. Let $I(\rho) = [\mathbf{Y}^f - \rho(f) | f \in L_p]$. Get a *difference ascending chain.*

Theorem 65 *f is a reduced GB with $[A] \neq [1]$ iff A is a reguar and coherent difference ascending chain iff A is a characteristic set of $[A]$.*

An ideal is *reflexive* of $\rho^x \in I \Rightarrow \rho \in I$. Also Perfec and prime. We can define $\mathbf{Z}$-saturated also x-saturated ($xf \in L \Rightarrow xinL$), and saturated if it is both.

Example 60 *$F = \mathbf{Q}(\sqrt{3})$ and $p = y^3 - 1$. $\{p\} = [p, y^{x-1} - 1]$ if $\sigma(\sqrt{-3}) = \sqrt{-3}$ and $\{p\} = [p, y^{x-2} - 1]$ otherwise.*

Theorem 66 *If $I(\rho)$ is perfect, then L_ρ is N-saturated; if L_ρ is x, M-saturated then ;dots.*

Can characterise reflexive closure of T and perfect closure.

Similar results in the non-Laurent case.

Torilc σ-deals and varieties. A toric variety is a σ-variety parameterized by σ-monomials.

Theorem 67 (equivalent) *1. $X \equiv Spec(\mathbf{Q}\{M\})$ wherne $M \subset \mathbf{Z}[X]$ is ...*

Theorem 68 *The σ-Chow form of X_α is the σ-sparse resultant with support α.*

Algorithms to compute the saturations (in both directions). [Gaoetal2014aArXiV]

10.6 Differential Algebar and the muduli space of products of elliptic curves: Freitag

This is really about conjectures coming from Diophantine Geometry. X is a family of sets (typically subvarieties of a fixed variety) and F a subset of special sets. F_a are special points. If $V \in X$ contains "many" special points implies $U \subset V$ is positive dimensional and special.

Example 61 *$V = \mathbf{C}^n$; X is $\{$ irreducible algebaric subsets of V $\}$. F is p_A where $p \in Tor(V)$. ...*

Also Manin–Mumford conjecture is of this form.

Recall j-function.

Definition 32 *a function f which is anlytic on some domain is automorphic if*

1. ...

Theorem 69 (Ax–Lindemann–Weierstrass) *If the a_i are linearly independent over $\mathbf{C}$, their exponentials are algebarically independent.*

For "algebraically dependent" if $a_i = g(a_j)$ we will say that$j(a_i), j(a_j)$ are modularly dependent. Note that j satisfies a third-order differential equation.

Theorem 70 (Pila) *Let $W \subset \mathbf{C}$ The a_i are modularly dependent iff the $3n$ functions are algebraiclly dependent.*

step 0 Let $f(t)]inMer(U)$ The Kolchin closure of $Iso(j(f(t)))$ is given by $\{x|\chi(x) = S_\delta(f(t))\}$.

Step1 By Nishioka's theorem; $j(f(t))$ is generic on this set.

Step 2 Shelah reflecting principle. Le $A \subset B \subset \mathcal{M}$ and $tp(a/b)$ be a forking extension of $tp(a/A)$. Thnen $cb(a/B) \subset acl(\{d_i\}_{i \in \mathbf{N}})$ where the d_i form an indiscernable sequence.

Step 3 By Pila, linearly dependent $\rightarrow$ algebraically dependent.

step 4 If we could assume $d_i = j(g_i t)$ for $g_i \in GL_2(\mathbf{R})$ we would be done.

Step 5 Siedenberg. Let $K = \mathbf{Q}(u_1, \ldots, u_n)$ be a differential field generated by n elements over $\mathbf{Q}$ and let$K_1 = K(v)$ be a simple differentual field extensin of N. SUppose $U \subset \mathbf{C}$ is an open ball and $\iota : K \rightarrow$ and ....

Suppose $Y \subset X_a \times X_b$ and Y cnnot prokect ontpo both X_a and X_b. then we say $X_a \perp X_b$.

10.7 Differential Chow Varieties Exist: Wei Li

Algebraic Chow Variety. Let $V = \sum s_i V_i$ be a d-cycle in $\mathbf{P}^n$. The Chow form of V is a aunique polynomial $F(u_{i,j}) = \sum_\omega c_\omega M_\omega(u_{i,j})$ such that $F(u_{i,j}) = 0 \Leftrightarrow V$ $bigcap_{i=1}^d \left(\sum_j u_{i,j} v - = 0\right) \neq 0$.

Example 62 *A line in* $\mathbf{P}^3$*. The Chow coordinates are then Plücker coordinates in this case.*

For an affine variety, the Chow variety may not be closed.

Differential Chow Form[GaoLiYuan2013]. Consider a sufficiently saturated $U \models DCF_0$ and $\mathbf{A}^n \equiv U^n$. Hence differential Cow coordinates.

Proved that these differential Chow varieties exist when $g = 1$. We will now show that these exost for all (d, h, g, m). Use various ingrediat=ents, especially differential characteristic sets. Need a prolongation sequence τ_l : functors for the category of algebraic varieties in $\mathbf{A}^n$ to the category of algebraic varieties in $\mathbf{A}^{n(l+1)}$. $\nabla_l : V \to \tau_l(V)$ — differential point.

A component of a differential variety with maximal Kolchin polynomial is called a *generic component*.

Need toprove results about what, relative to DCF_0, is definable in families. Butit is open whether primality of radical ideals is definable in families.

We proved that C_1 is a constructible set, with a 1-1 map to δ-Chow$_n(d, h, g, m)$.

Is there a more natural construction (this one used a lot of model theory!). Also the Ritt problem!

Q Do you know any one example?

A We know ones in dimension 1, not in higher dimension.

Q Is there any conjecture that can imply the Ritt problem?

A No!

10.7.1

10.7.2

Bibliography

[AR15] A. Arnold and D.S Roche. Output-Sensitive Algorithms for Sumset and Sparse Polynomial Multiplication. In D. Robertz, editor, *Proceedings ISSAC 2015*, pages 29–36, 2015.

[Arn03] E.A. Arnold. Modular algorithms for computing Gröbner bases. *J. Symbolic Comp.*, 35:403–419, 2003.

[BC11] P. Bürgisser and F. Cucker. On a problem posed by Steve Smale. *Annals of Mathematics*, 174:1785–1836, 2011.

[Bec94] T. Becker. On Gröbner Bases under Specialization. *AAECC*, 5:1–8, 1994.

[BFDS15] D.K. Boku, C. Fieker, W. Decker, and A. Steenpass. Gröbner Bases over Algebraic Number Fields. `http://arxiv.org/abs/1504.04564`, 2015.

[Bix10] R.E. Bixby. Mixed-Integer Programming: It works better than you may think. *www.ferc.gov/eventcalendar/Files/20100609110044-Bixby*, 2010.

[BKY09] M. Burr, F. Krahmer, and C. Yap. Continuous amortization: A non-probabilistic adaptive analysis technique. Technical Report TR09-136 Electronic Colloquium on Computational Complexity, 2009.

[BL95] W. Bosma and H.W. Lenstra. Complete systems of two addition laws for elliptic curves. *J. Number Theory*, 53:229–240, 1995.

[BPR06] S. Basu, R. Pollack, and M.-F. Roy. Algorithms in Real Algebraic Geometry, 2nd ed. *Springer*, 2006.

[CC86] D.V. Chudnovsky and G.V. Chudnovsky. Elliptic Functions and Algebraic Topology. *SCRATCHPAD II Newsletter 1(1985-6) 2*, pages 2–3, 1986.

[CGTW95] R.M. Corless, P. Gianni, B.M. Trager, and S.M. Watt. The singular value decomposition for polynomial systems. In A.H.M. Levelt, editor, *Proceedings ISSAC 1995*, pages 195–207, 1995.

[Cop96] D. Coppersmith. Finding a small root of a bivariate integer equation; factoring with high bits known. In *Proceedings EUROCRYPT '96*, pages 178–189, 1996.

[Dav87] J.H. Davenport. Looking at a set of equations (Technical Report 87-06, University of Bath Computer Science). `http://staff.bath.ac.uk/masjhd/TR87-06.pdf`, 1987.

[DJS15] L. D'Alfonso, G. Jeronimo, and P. Solernó. A decision method for the integrability of differential-algebraic Pfaffian systems. `http://arxiv.org/abs/1501.04941`, 2015.

[DMSX06] X. Dahan, M. Moreno Maza, É. Schost, and Y. Xie. On the complexity of the D5 principle. In J.-G. Dumas, editor, *Proceedings Transgressive Computing 2006*, pages 149–168, 2006.

[DX14] L. Dai and B. Xia. Smaller SDP for SOS Decomposition. `http://arxiv.org/abs/1407.2679`, 2014.

[Fau02] J.-C. Faugère. A New Efficient Algorithm for Computing Gröbner Bases Without Reduction to Zero (F_5). In T. Mora, editor, *Proceedings ISSAC 2002*, pages 75–83, 2002.

[FGT02] E. Fortuna, P. Gianni, and B. Trager. Derivations and Radicals of Polynomial Ideals over Fields of Arbitrary Characteristic. *J. Symbolic Comp.*, 33:609–625, 2002.

[FS14] J. Freitag and O.L. Sanchez. Effective uniform bounding in partial differential fields. `http://arxiv.org/abs/1411.0029`, 2014.

[Gat90] K. Gatermann. Symbolic Solutions of Polynomial Equation Systems with Symmetry. In S. Watanabe and M. Nagata, editors, *Proceedings ISSAC 1990*, pages 112–119, 1990.

[GKOS08] O. Golubitsky, M. Kondratieva, A. Ovchinnikov, and A. Szanto. A Bound for Orders in Differential Nullstellensatz. `http://arxiv.org/abs/0803.0160`, 2008.

[Gri89] D.Yu. Grigoriev. Complexity of quantifier elimination in the theory of ordinary differential equations. In *Proceedings EUROCAL 87*, pages 11–25, 1989.

[GT96] P. Gianni and B.M. Trager. Square–Free Algorithms in Finite Characteristic. *AAECC*, 7:1–14, 1996.

[GXD+14] Z. Guo, Q. Xia, Z. Du, L. Ji, and Z. Han. Research of critical ambient temperature of cylindrical fireworks and crackers. *Journal of Thermal Analysis and Calorimetry*, 115:1787–1792, 2014.

[Has88] B.J. Hastad. Solving Simultaneous Modular Equations of Low Degree. *SIAM J. Comp.*, 17:336–341, 1988.

[HG97] N.A. Howgrave-Graham. Finding Small Roots of Univariate Modular Equations Revisited. *Cryptography and Coding (Ed. M. Darnell)*, pages 131–142, 1997.

[HHS12] J.D. Hauenstein, N. Hein, and F. Sottile. Certifiable Numerical Computations in Schubert Calculus. `http://arxiv.org/abs/1212.3315`, 2012.

[Hı10] H. Hışıl. *Elliptic Curves, Group Law and Efficient Computation*. PhD thesis, Queensland University of Technology, 2010.

[HL12] E. Hubert and G. Labahn. Rational invariants of scalings from Hermite normal forms. In *Proceedings ISSAC 2012*, pages 219–226, 2012.

[HSL14] Z. Huang, Y. Sun, and D. Lin. On the Efficiency of Solving Boolean Polynomial Systems with the Characteristic Set Method. `http://arxiv.org/abs/1405.4596`, 2014.

[Hsu21] J. Hsu. The great chip crisis threatens the promise of Moore's Law. `https://www.technologyreview.com/2021/06/30/1026438/global-microchip-shortage-problem-m1-apple-tsmc-intel/`, 2021.

[HZ00] R. Hartley and A. Zisserman. See book: Multiple View Geometry in Computer Vision. *C.U.P.*, 2000.

[IPS11] I. Idrees, G. Pfister, and S. Steidel. Parallelization of Modular Algorithms. *J. Symbolic Comp.*, 46:672–684, 2011.

[JQ01] M. Joye and J.-J. Quisquater. On Rabin-Type Signatures. In B. Honary, editor, *Proceedings 8th. IMA Conf. Cryptography and Coding*, pages 99–113, 2001.

[Kal98] M. Kalkbrener. Algorithmic properties of polynomial rings. *J. Symbolic Comp.*, 26:525–581, 1998.

[Kar84a] N.K. Karmarkar. A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica*, 4:373–395, 1984.

[Kar84b] N.K. Karmarkar. A New Polynomial-Time Algorithm for Linear Programming. In *Proceedings 16th STOC*, pages 302–311, 1984.

[Laz83] D. Lazard. Gröbner Bases, Gaussian Elimination and Resolution of Systems of Algebraic Equations. In *Proceedings EUROCAL 83*, pages 146–157, 1983.

[Laz85] D. Lazard. Ideal Bases and Primary Decomposition: Case of Two Variables. *J. Symbolic Comp.*, 1:261–270, 1985.

[Laz92] D. Lazard. Solving Zero-dimensional Algebraic Systems. *J. Symbolic Comp.*, 13:117–131, 1992.

[Mal14] G. Malajovich. Computing mixed volume and all mixed cells in quermassintegral time. `http://arxiv.org/abs/1412.0480`, 2014.

[MMT92] H. Möller, T. Mora, and C. Traverso. Gröbner bases computation using syzygies. In *Proceedings ISSAC '92*, pages 320–328, 1992.

[Moo65] G.E. Moore. Cramming More Components onto Integrated Circuits. *Electronics*, pages 114–117, 1965.

[Pfi07] G. Pfister. On Modular Computation of Standard bases. *Analele Stiintifice ale Universitatii Ovidius Mathematical Series*, XV:129–137, 2007.

[Sei74] A. Seidenberg. Constructions in Algebra. *Trans. A.M.S.*, 197:273–313, 1974.

[SM15] M. Sagraloff and K. Mehlhorn. Computing real roots of real polynomials. *To appear in J. Symbolic Comp*, 2015.

[Stu95] B. Sturmfels. Gröbner Bases and Convex Polytopes. *Amer. Math. Sci.*, 1995.

[SVV10] F. Sottile, R. Vakil, and J. Verschelde. Solving Schubert Problems with Littlewood-Richardson Homotopies. In S.M. Watt, editor, *Proceedings ISSAC 2010*, pages 179–186, 2010.

[SY11] M. Sagraloff and C.K. Yap. A simple but exact and efficient algorithm for complex root isolation. In *Proceedings ISSAC 2011*, pages 353–360, 2011.

[SY12] V. Sharma and C.K. Yap. Near Optimal Tree Size Bounds on a Simple Real Root Isolation Algorithm. In *Proceedings ISSAC 2012*, pages 319–326, 2012.

[TB14] I. Tamo and A. Barg. A Family of Optimal Locally Recoverable Codes. *IEEE Trans. Information Theory*, 60:4661–4676, 2014.

[vH94] M. van Hoeij. An algorithm for computing an integral basis in an algebraic function field. *J. Symbolic Comp.*, 18:353–363, 1994.

[VY15] J. Verschelde and X. Yu. Accelerating Polynomial Homotopy Continuation on a Graphics Processing Unit with Double Double and Quad Double Arithmetic. `http://arxiv.org/abs/1501.06625`, 2015.

[Wu 79] Wu and Wen Tsün. On the Mechanization of Theorem Proving in Elementary Differential Geometry. *Sci. Sinica 1979 (Special Issue I on Mathematics)*, pages 94–102, 1979.