Surface parametrisation without diagonalisation

Christiaan van de Woestijne
Institut für Mathematik B
Technische Universität Graz, Austria

ISSAC 2006, Genoa
10 July 2006

What are we trying to do?

Object of interest: rationally fibred surfaces over a field F.

By algebraic geometry: may assume the form

$$S : 0 = \sum_{i,j=1}^{3} a_{ij} X_i X_j, \qquad a_{ij} = a_{ji} \in F(t),$$

so the zero set of a *ternary quadratic form* over the function field $F(t)$.

This form can be obtained computationally but expensively; not topic of today.

By clearing denominators, we may assume

$$a_{ij} \in F[t].$$

What are we trying to do? (II)

Goal: find a *rational parametrisation* of the surface S, that is: dominant rational maps

$$\phi : \mathbb{A}^2(F) \to S, \qquad \psi : S \to \mathbb{A}^2(F)$$

that are inverses to each other.

That way, except possibly for a small "problematic" subset of S, each point of S is derived in an efficient and controllable way from a *unique* point in $\mathbb{A}^2$, the ordinary plane (in two coordinates x, y) over F.

Useful, for example, for *rendering* the surface S!

State of affairs

First point: rational parametrisation is usually impossible, depending on the *degrees* of the coefficients a_{ij}.

Second point: there exists an algorithm, due to Schicho (ISSAC '98, ISSAC '00), that

- decides the existence of a rational parametrisation;
- computes one if it exists;
- uses a polynomial number of operations in F;
- works in principle over any field of characteristic not 2.

Implementation depends on the possibility to solve conics over F efficiently. This is possible, more or less, for $\mathbb{Q}$, $\mathbb{R}$, finite fields...

State of affairs (II)

Third point: a paper by Van Hoeij (with Cremona, 2004) titled "Solving conics over $\mathbb{Q}(t_1, \ldots, t_k)$" develops the same method independently — the authors graciously propose to re-title it something like "an implementation of Schicho's algorithm".

Schicho's (Lagrange's?) method

Represent the form $f = \sum a_{ij} X_i X_j$ by a symmetric 3×3-matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}.$$

We define the discriminant Δ to be $\det A$. If $\Delta = 0$, the problem is trivial, so assume the opposite.

We need two things:

- removing all *redundant factors* from Δ;
- weighing the variables X_1, X_2, X_3 such that $f(X_1, X_2, X_3)$ is *almost homogeneous*.

Definitions

For weights $W = (w_1, w_2, w_3) \in \mathbb{Z}^3$, define

$$\begin{aligned} \deg_W(A, i, j) &= \deg a_{ij} + w_i + w_j; \\ \deg_W(A) &= \max_{i,j} \deg_W(A, i, j). \end{aligned}$$

In fact, $\deg_W(A, i, j)$ is the weighted degree of the term $a_{ij} X_i X_j$.

If f were (weighted!) homogeneous, all terms would have the same degree; and we would have

$$\deg \Delta = 3 \deg_W(A) - 2(w_1 + w_2 + w_3).$$

Now, we define the degree defect, which is nonnegative, as

$$\operatorname{def}_W(A) = 3 \deg_W(A) - 2 \sum w_i - \deg \Delta,$$

and call the form *almost homogeneous* if

$$\operatorname{def}_W(A) = 0 \text{ or } 1.$$

Schicho's (Lagrange's?) method (II)

For any weights of the X_i, Schicho defines the index of f as

$$\text{ind}_W(A) = 3\,\text{deg}_W(A) - 2\sum w_i,$$

so the "idealised" degree of the discriminant.

If the two things are done, we know:

there exists a rational parametrisation of f
if and only if
the index is at most 3.

This is an algebraic-geometric result of Iskovskikh. If the index is at most 3, it is easy to show that the form is either linear in one of the variables, or defined over the base field F.

Finding a good weight vector

If A is diagonal, finding suitable weights is easy; basically,

$$w_i = \left\lfloor \frac{\deg a_{ii}}{2} \right\rfloor,$$

with some minor subtleties.

So, Schicho starts by diagonalising the form f, using Gram-Schmidt orthogonalisation, to get

$$f = a_{11}X_1^2 + a_{22}X_2^2 + a_{33}X_3^2,$$

where we make the a_{ii} polynomials again by clearing denominators. We now have

$$\Delta = a_{11}a_{22}a_{33}.$$

On the other hand, for non-diagonal matrices, suitable weights may not exist.

Removing redundant discriminant factors

Theorem (Schicho 2000/Simon 2004) Let p be an irreducible factor of the discriminant Δ. We can find a transformation matrix T over $F[t]$ such that all entries of

$$T^t \, A \, T$$

are divisible by p; we have

$$\operatorname{disc}\left(\frac{T^t \, A \, T}{p}\right) = \frac{(\det T)^2 \Delta}{p^3},$$

so we have removed p from Δ if $\det T = p$.

If p^2 divides Δ, we can even find T of determinant p^2, such that $T^t \, A \, T$ is divisible by p^2, so we remove two factors p at the same time.

But: the transformed matrix need not be diagonal even if A is. So, how to keep a good weight vector?

Removing redundant discriminant factors (II)

Proof (sketch)

1. Reduce A modulo p; it is singular, compute the kernel. This is the *radical* of the form f modulo p.

2. Now f modulo p has essentially only two variables. See if the discriminant of this form is minus a square. If so, the corresponding space is a *hyperbolic plane* and f modulo p is a product of two linear factors.

The square root is taken in $F[t]/(p)$; for $F = \mathbb{Q}$, this is an algebraic number field, and this can give practical problems if $\deg p$ is large.

Removing redundant discriminant factors (III)

Take a new basis consisting of

- a radical vector (lifted to $F[t]$);
- a zero of f mod p that is independent of the first (also lifted);
- and any vector of $F[t]^n$ multiplied by p.

This, with some subtleties, gives the transformation matrix.

In ongoing work, I found a shortcut to this computation, that avoids the computation of the kernel of A mod p. The amount of field operations is roughly halved using this approach.

Reduced bases

Here basis reduction in polynomial lattices comes in. This has been practised by many authors:

Von zur Gathen (1984), A.K. Lenstra (1985), Paulus (1998), Mulders and Storjohann (2003), ...

under various names; in fact, the latter paper speaks about the Popov form, taking up a notion of Popov from 1969. Schicho still calls this process differently: it can also be described as computing a Gröbner basis of a 3-dimensional free module over $F[t]$, with respect to a term-over-position term ordering.

An essential feature of a reduced basis is that one of the basis vectors is a shortest vector of the generated module, where "shortest" is defined by means of the max-norm, so $\|v\| = \max_i \deg v_i$ for $v \in F[t]^m$.

Reduction and weights

Theorem (Schicho): If the columns of the transformation matrix T form a reduced basis of the column space, then it is easy to find a weight vector W' such that

$$\text{def}_{W'}(T^t\, A\, T/p) \leq \text{def}_W(A).$$

So we get the following algorithm:

(1) Diagonalise and find good weights W.

(2) For all factors p of the discriminant: try to remove p; if successful, apply reduction to the transformation matrix and update the weights.

Reduced quadratic forms

There is a *different notion* of reduction for modules over $F[t]$: if we have a quadratic form f on the free module $V = F[t]^m$, then the form is called reduced (in the sense of Hermite) if, with respect to some basis $v_1, \ldots, v_m$ of V,

v_i is shortest among vectors linearly independent of $v_1, \ldots, v_{i-1}$.

Let A be the Gram matrix with respect to the basis v_i. Now we know (Simon 2004/CvdW 2006):

- the form f is reduced if we have, for $1 \leq i < j \leq n$,

$$\deg a_{ii} < \deg a_{jj}, \qquad \text{and} \qquad \deg a_{ij} < \deg a_{ii}.$$

- a variant of the LLL algorithm transforms f into a reduced form, provided we do not encounter zeros of f in the process.

Reduction and weights (II)

Here is the main result:

Theorem (CvdW) If the form f is reduced in the sense of Hermite, then for the purpose of finding a good weight vector, we can treat it as being diagonal.

This arises because for Hermite-reduced forms, we have

$$\deg \Delta = \deg a_{11} + \deg a_{22} + \deg a_{33}$$

trivially, and, for appropriate weights W and $i < j$,

$$\deg_W(A, i, j) < \deg_W(A, i, i),$$

less trivially.

My algorithm

We get the following algorithm:

1. For all irreducible factors p of Δ: try to remove p.

2. Apply Hermite reduction to the resulting f. Either we are successful, or we find a zero of f, which is even better.

3. Compute a good weight vector for f and compute the index.

If desirable, we can apply basis reduction to the transformation matrices used in Step 1, to get smaller coefficients.

Reduced = reduced?

If we have $v_1, \ldots, v_m \in F[t]^n$, let V be the generated module; the usual dot product gives a quadratic form f on V.

Facts:

1. If the v_i form a reduced basis, f need not be reduced in the sense of Hermite.
2. If f is reduced in the sense of Hermite, the v_i need not form a reduced basis.
3. Not all quadratic forms f on free $F[t]$-modules arise in this way. E.g., take $f = a_{11}X_1^2$ over $\mathbb{Q}$ where $\deg a_{11}$ is odd.

This corrects my paper...

Why avoid diagonalisation?

Why is it proficient to avoid diagonalising the form f?

1. Row and column operations on the matrix A may introduce large degrees, and coefficient explosions. So use sparingly.

2. Diagonalising introduces denominators that have to be cleared. In fact, the factors

$$a_{11} \quad \text{and } a_{11}a_{22} - a_{12}^2$$

   occur doubly in the discriminant after clearing, and have to be removed again.

Example...

Questions...

1. How efficient is Hermite reduction?

Answer: it's easy to prove that it terminates fast, using the same proof as for LLL, and that it uses polynomially many field operations. But I do not have a bound on the occurring field elements...

2. How efficient is basis reduction?

Same answer, Mulders and Storjohann seem to give the tightest bounds. But can/do we avoid coefficient explosion over $\mathbb{Q}$?

Anybody?