J. Symbolic Computation (1991) **12**, 127-141

Wu's Method and the Khovanskii Finiteness Theorem

DANIEL RICHARDSON
School of Mathematics, University of Bath, U.K.
dsr@uk.ac.bath.maths

(Received 18 December 1989)

The Khovanskii finiteness theorem calculates an upper bound for the number of connected components of the intersection of an algebraic set with a Pfaff manifold in R^n. This paper uses the algebraic methods of Wu Wen-tsun to give an elementary proof of Khovanskii's theorem. An extension of the Wu-Ritt zero structure theorem is also obtained.

Introduction

The Khovanskii finiteness theorem [1980] calculates an upper bound for the number of connected components of the intersection of an algebraic set with a Pfaff manifold in R^n. This implies, for example, that the zero set of a polynomial in $x_1, ..., x_n$ and $exp(x_1), ..., exp(x_n)$ is a finite union of connected components in R^n. The object of this paper is to use the algebraic methods of Wu Wen-tsun to give an elementary proof of Khovanskii's theorem. Unfortunately, although the bound is recursively calculated in this paper, it is not explicitly given by a formula, due to lack of knowledge about the computational complexity of Wu's method. (However, the situation is improving. See [Gallo and Mishra, 1990].) The relationship between this bound and the explicit bound given by Khovanskii is not clear.

In the first section below Pfaff manifolds are defined and the Wu-Ritt decomposition algorithm is defined. In the second section this is used to prove a version of the Khovanskii theorem.

1. Pfaff manifolds and Wu's method

Every set discussed below will be assumed to be a subset of one of the real Euclidean n-dimensional spaces, R^n. When the space is understood, polynomials

0747-7171/91/080127+15 $03.00/0

will mean polynomials with integral coefficients with variables corresponding to the coordinate axes of the space. A polynomial vector field, V, will be a vector field in R^n determined by a list of n polynomials.

Define a smooth path in a manifold to be a differentiable function, $p(t)$, with continuous and non zero derivative, from [0,1] into the manifold. We will be concerned with differentiable manifolds which are embedded in R^n in such a way that smooth paths in the manifolds are smooth paths in R^n, and a smooth path in R^n which happens to lie in one of these manifolds is also a smooth path in the manifold.

A smooth path $p(t)$ is said to be orthogonal to a polynomial vector field V if the scalor product, $p'(t) \bullet V$, is identically zero.

If M is a differentiable manifold and v is a point on M, and V is a polynomial vector field, we will say that point v is a singular point of V on M if every smooth path $p(t)$ which lies in M satisfies $p'(t) \bullet V = 0$ whenever $p(t) = v$. In other words, a singular point of V on M is one where the tangent plane of M is orthogonal to V.

Let M and Y be differentiable manifolds in R^n and let V be a polynomial vector field in R^n. We will say that M is orthogonal to V in Y if:

C1) M is contained in Y and M has codimension one in Y. No point of M is a singular point of V in Y.

C2) If $p(t)$ is a smooth path which lies in Y, and if $p(t)$ is orthogonal to V, and if there are no singular points of V in Y on $p(t)$, and if $p(t)$ intersects M, then $p(t)$ is contained in M.

C3) If $p(t)$ is a smooth path which is contained in Y, and if $p(t_1)$ and $p(t_2)$ are in M with $t_1 < t_2$, then there is a number t so that $t_1 \leq t \leq t_2$ and $p'(t) \bullet V = 0$.

Condition C2) says that M is closed under nonsingular smooth paths orthogonal to V. Condition C3) implies that smooth paths in M must be orthogonal to V, but it is much stronger. The intuitive idea is that M is the boundary of a domain in Y so that on a path leaving the domain the V vectors are pointing outwards and thus $p'(t) \bullet V \geq 0$, and on a path entering the domain $p'(t) \bullet V \leq 0$; since $p'(t) \bullet V$ is continuous, this implies that there is a point t where $p'(t) \bullet V = 0$.

Example 1) Let G be the graph of $y = exp(x)$ in R^2. Let V be the gradient of $y - exp(x)$ on the curve, i.e. $(-y, 1)$. Then G is orthogonal to V in R^2.

Example 2) Let p be a polynomial. Let M be the non singular zero set of p, i.e. the set of points where p is zero but has non zero gradient. Then M is orthogonal to the gradient of p in R^n.

We will say M is a Pfaff manifold if there is a sequence $M_0, ...M_k$ of differentiable manifolds with $M_0 = R^n$, and, for each $i < k$, M_{i+1} is contained in M_i, and $M_k = M$, and there is a corresponding sequence $V_1, ..., V_k$ of polynomial vector fields so that, for each $i < k$, M_{i+1} is orthogonal to V_{i+1} in M_i.

Example 3) The surface of the cylinder of radius one around the z axis is a Pfaff manifold in R^3. The curve $y = sin(x)$ on this cylinder surface is a Pfaff manifold if it is restricted to an open x interval on which $sin(x)$ is not equal to +1 or -1.

The sequence $(M_i, V_i), i \leq k$ will be called a construction of M_k, and k will be the length of the construction. If M has construction from R^n of length k, then the dimension of M is $n - k$, since by C1) the dimension goes down one at each step. It is imagined here that we could describe a Pfaff manifold by giving a construction, and a sample point on each connected component. However a method of verifying that a given arbitrary sequence of polynomial vector fields and a sample point determines a manifold satisfying the conditions above is not known at present.

If M is a Pfaff manifold or a subset of a Pfaff manifold in R^n and v is a point of M, we will say M is locally a graph at v if the coordinate variables of R^n can be divided into $k \geq 0$ independent variables $z_1, ..., z_k$ and $j = n - k$ dependent variables $w_1, ..., w_j$, and there exist j analytic functions $f_1, ...f_j$ of the independent variables so that in some R^n neighbourhood N of v we have that the points u in M and in N are exactly the points in N given by the condition $(w_1, ..., w_j) = (f_1, ..., f_j)(z_1, ..., z_k)$, for $z_1, ..., z_k$ in the projection onto the space of variables $(z_1, ..., z_k)$ of N . In this case k is the dimension of the Pfaff manifold, and $(z_1, ..., z_k)$ is called a local coordinate system for M at v.

In the following, we will always use X to mean the vector of coordinate variables $(x_1, ..., x_n)$ of R^n. If a local coordinate system is understood and one of its independent variables is z, then in all of the following $\partial X/\partial z$ will mean the vector of partial derivatives of the variables in X with respect to z .

Local Graph Proposition.

If M is a Pfaff manifold, and v is a point of M, then M is locally a graph at v .

Proof. The proof is by induction on the length of the construction of M, using the lemma below.

Local Graph Lemma

Suppose M and Y are differentiable manifolds in R^n, Y is locally a graph, V is a polynomial vector field in R^n, and M is orthogonal to V in Y. Then M is also locally a graph.

Proof. Let v be a point in M. There is a neighbourhood N of V in which Y has a coordinate system $(z_1, ..., z_{j+1})$. These are coordinates in R^n and the other variables are expressed in N as functions of these.

Now look at the partial derivatives in relation to V: $(\partial X/\partial z_q) \bullet V$, for $q =$ 1 up to $j + 1$. All of these terms can't be zero at v. If they were, then any smooth path in Y passing through v would be orthogonal to V at v, and v would be a singular point of V on Y. But condition C1) asserts that this does not happen. Let z be some z_q such that $(\partial X/\partial z_q) \bullet V$ is not zero at v. Renumber the other independent variables to be $(z_1, ..., z_j)$. It is claimed that near v this is a coordinate system for M.

Let $\mathcal{P}_j$ and $\mathcal{P}_{j+1}$ be the two projections of R^n onto the space of the independent variables, $(z_1, ..., z_j)$, and $(z_1, ..., z_j, z)$ respectively. These projections just forget the dependent coordinates of a point near v.

Pick N so small that for every point in $\mathcal{P}_{j+1}$(N), $(\partial X/\partial z) \bullet V$ is bounded away from zero.

We now have to define z as a function of $(z_1, ..., z_j)$ so that $(z_1, ..., z_j, z)$ is in $\mathcal{P}_{j+1}(M)$ for $(z_1, ..., z_j)$ in $\mathcal{P}_j(N)$. We assumed, of course that v was in M; and we know, by condition C2), that M is closed under smooth paths orthogonal to V.

Given $(z_1, ..., z_j)$ in the projection $\mathcal{P}_j(N)$, draw a straight line in this projection from this point to $\mathcal{P}_j(v)$, and define $z(t)$ on this line to satisfy $dX/dt \bullet V = 0$, where t parameterizes the line. This differential equation has a unique solution for t sufficiently close to 0, because it satisfies a Lipschitz condition in N.

It remains to show that every point in M in this neighbourhood N is given in this form. Suppose not. There would then be two distinct values of z so that $(z_1, ..., z_j, z)$ was in $\mathcal{P}_{j+1}$ applied to the intersection of M and N. Draw a straight line from one such point to the other in the projection. (Only the z coordinate varies in the projection, although in R^n a curve is described.) By condition C3) there must be a point on the line at which $(\partial X/\partial z) \bullet V = 0$. But N was chosen sufficiently small so that this can not happen.

That completes the proof of the local graph proposition.

We wish to study zero sets of polynomials on Pfaff manifolds. We will allow parameters in the polynomials, although not in the definition of the Pfaff manifolds. So we adjoin a list $(c_1, ..., c_j)$ of parameters to the list of variables (x_i), and from now on a polynomial with parameters will mean a polynomial with integral coefficients involving both the variables and the parameters.

Let PS be a finite list of polynomials with parameters, $p_1(X), ..., p_j(X)$, and let $(M_i, V_i)_{i \leq k}$ be a construction of a Pfaff manifold, and let $J(X)$ be another polynomial with parameters. (We will assume that no parameters occur in the vector fields V_i, so the Pfaff manifolds are independent of the parameters.) We are interested in conditions of the form $\Gamma =$

$(M_i, V_i)_{i \leq r}$
$PS = 0$
$J \neq 0$

which define the set of X such that $PS(X) = 0$, $J(X) \neq 0$, and X is in M_r. The defined set depends on the parameters, since there may be parameters in PS or J. Conditions in this form will be called Gamma conditions. If the number of vector fields, r, in the construction of a Gamma condition is zero, we will say that the Gamma condition is algebraic. Given an interpretation of the parameters as real numbers, a Gamma condition determines a set, which will be called a Gamma set.

Khovanskii has shown that such sets are necessarily finite unions of connected components. The intention here is to prove this using Wu's method. The whole construction can be done uniformly over the parameters. That is, the upper bound which is eventually obtained will hold no matter how the parameters are interpreted.

If S is a set closure(S) will mean the set of finite limit points of sequences from S. If S is a Gamma set, the boundary of S will mean closure(S) - S.

The main idea of this paper is to break Gamma conditions into uniform parts, as described below.

A Gamma condition, Γ, is said to be uniform of dimension k if associated with it are k of the coordinate variables $z_1, ..., z_k$ so that:

U1) For any interpretation of the parameters if S is the Gamma set defined by Γ and v is in S, then $(z_1, ..., z_k)$ can be used as the independent variables in a local coordinate system for S near v. (This means that not only is the set equidimensional, but also the selection of independent variables can be the same for all points.)

U2) For each z among $(z_1, ..., z_k)$, and for each variable w, we are able to construct a rational function $R(X)$, involving the variables and the parameters, so that $\partial w/\partial z = R(X)$ in all the local coordinate systems of 1). The denominator of $R(X)$ cannot be zero on any of the Gamma sets.

U3) If S is any of the Gamma sets defined by Γ, $J = 0$ on the boundary of S. (This means that $J = 0$ at least defines a superset of the boundary.)

The number of variables in a uniform local coordinate system for a Gamma condition will be called its dimension. Dimension zero will turn out to imply that in every interpretation every point in the defined set is isolated.

We next use Wu's method to break Gamma conditions into uniform parts. The result will be to extend the Wu-Ritt zero structure theorem.

Wu's Method

We have a list of parameters $(c_1, ..., c_j)$ and another list of variables $(x_1, ..., x_n)$, and we wish to order these by importance, by making all the parameters less important than all the variables, and within the lists ordering by subscript, e.g., we will say x_i is less important than x_j, or $x_i \prec x_j$, if $i < j$. For any polynomial $p(X)$, let class$(p(X))$ be the most important variable or, in case there aren't any variables, the most important parameter which actually occurs in it. We will say class$(p(X))$ is 0 if $p(X)$ has no variables or parameters, and we will say p has parameter class if it contains no variables.

For two polynomials p and q, we will say $p \prec q$ if class$(p) \prec$ class(q), or if they have the same non zero class, but the degree of p in the class variable or parameter is less than the degree of q in the class variable or parameter. Polynomials which are incomparable in this ordering have the same class and degree, or are both of class 0.

The degree of a polynomial p will be the degree of its class variable, or parameter.

If the class of p is x and the degree of p is d, then p can be put in the form $p = I * x^d +$ (lower degree terms in x), in which all the coefficients of powers of x are polynomials in variables or parameters below x. The coefficient I is called the initial of p.

If p and q are polynomials, p is said to be reduced with respect to q if both have class 0, or if the class(q) variable or parameter either does not occur in p or only occurs with degree below degree(q).

A finite sequence of non zero polynomials ASC $= (p_i)_{i \le r}$ is called an ascending set if either $r = 1$ and class(p_1)=0; or if $0 \prec class(p_1) \prec class(p_2) \prec \ldots \prec class(p_r)$ and each p_j is reduced with respect to all p_i with $i < j$.

Let F $= (p_i)_{i \le r}$ and G $= (q_i)_{i \le s}$ be two ascending sets of polynomials. We will say that F has lower order than G if for some $j \le min(r, s)$ we have p_i and q_i are incomparable for all i less than j, but $p_j \prec q_j$; or if $r > s$ and p_i and q_i are incomparable for all $i \le s$.

Any descending sequence of ascending sets is necessarily finite.

Given any ascending set, $(p_i)_{i \le r}$, and any polynomial $g(X)$ we get by successive division the following remainder formula

$I_1^{s_1} * \ldots * I_r^{s_r} * g(X) = \Sigma_i q_i * p_i + R$

in which I_i are the initials of p_i, s_i are non negative integers, and q_i and R are polynomials with R reduced with respect to all p_i. We can make R unique by choosing the s_i to be as small as possible and R is then called the remainder of $g(X)$ with respect to the ascending set $(p_i)_{i \le r}$.

All of these definitions are taken from Wu [1987].

Now, we define triangular conditions, which are like the characteristic conditions of Wu, and also uniform.

Let Γ be a Gamma condition.

$\Gamma =$

$(M_i, V_i)_{i \le r}$

$(p_i = 0)_{i \le s}$

$J \neq 0$.

Then we define Γ^* from Γ by replacing r by $r - 1$ if $r > 0$; by replacing s by $s - 1$ if $r = 0$ and $s > 0$; and if r and s are both 0 then $\Gamma^* = \Gamma$. In other words, Γ^* is obtained from Γ by dropping the top condition.

A Gamma condition Γ, together with a uniform local coordinate system, as above, is said to be in triangular form if

T1) Γ is uniform, and $(p_i = 0)_{i \le s}$ is an ascending set, as defined above. If Γ is algebraic the dependent variables are the class variables of the polynomials. All the initials of the polynomials and all the derivatives of the polynomials with respect to their class variables are factors of J.

T2) Γ^* is also in triangular form. Also the uniform local coordinate system for Γ is either the same as the uniform coordinate system for Γ^* (which can only happen in the nonalgebraic case); or is obtained from the uniform local coordinate system of Γ^* by dropping one variable, say z, which is then associated with the top condition of Γ.

T3) When the uniform local coordinate systems for Γ and Γ^* are not the same, then the variable z which was dropped from Γ^* was the most important independent variable possible. If another variable, say w, is more important than z but remains independent in Γ, it must happen that the top condition of Γ does not depend on w. This means that in any interpretation, after projection to the uniform coordinate system of Γ^*, the set defined by the top condition has to be part of a cylinder parallel to the w axis.

The intuitive idea is that at each stage a triangular Gamma condition is obtained by defining, if possible, the currently most important independent variable as a function of the others.

We can now use Wu's algorithm to rewrite any Gamma condition as a finite union of Gamma conditions in triangular form. The new Gamma conditions will be based on the same Pfaff manifold. That is, they will only differ algebraically.

Let Γ be

$(M_i, V_i)_{i\leq r}$

$PS = 0$, where PS is $(p_i)_{i\leq s}$

$J \neq 0$

Within the Pfaff manifold we expect the vectors $(V_i)_{i\leq r}$ to be linearly independent. This is required to satisfy condition C1). Thus the r by n matrix of the vectors in the construction should have rank r. Let DS be the sum of the squares of the determinates of r by r submatrices of this matrix of polynomial vectors. Our first step is to replace J in Γ by DS$*J$.

The algorithm continues by recursion on the complexity of Γ. That is, we first show how to break a purely algebraic condition into a union of triangular conditions. Then we show how to write Γ as a union of triangular conditions, assuming that Γ^* is already a triangular condition.

Algebraic lemma.

Let Γ be an algebraic Gamma condition. Then we can effectively rewrite Γ as a finite disjunction of algebraic Gamma conditions which are triangular.

Proof.

Remark: this lemma essentially shows that semialgebraic sets can be broken up into equidimensional semialgebraic characteristic subsets by Wu's method. The only important element which has been added to Wu [1987] is the idea of equidimensionality.

We suppose that r, the length of the Pfaff construction is 0. In this case Γ is:

$PS = 0$

$J \neq 0$.

Any ascending set of lowest order of polynomials chosen from PS will be called a basic set from PS. Now, as in Wu,

a) Choose basic set BS from PS

b) Find remainders of everything in PS - BS with respect to BS. If all are 0, then stop. Otherwise, add the remainders to PS and go back to step a).

The ascending sets generated in this way are steadily decreasing in order so the process eventually terminates. We will be left with

ASC = 0

$J \neq 0$.

where ASC is an ascending set, and every polynomial in the original PS has remainder 0 with respect to ASC. If ASC contains a non zero polynomial of class 0, then ASC is inconsistent and may be discarded. Otherwise, let I be the product of all the initials of ASC and all the partial derivatives of polynomials in ASC with respect to their class variables, if any, or otherwise with respect to their class parameters.

Consider

ASC = 0

$J * I \neq 0$.

If the remainder of $J*I$ with respect to ASC is 0, this condition is inconsistent and can be dropped.

Whether or not the condition is consistent, it is in triangular form. Call this Γ_1. For any interpretation of the parameters, any point which is in the set described by Γ_1 is also in the set described by the original Γ. On the other hand any point in the original set must have ASC=0. Note also that all the initials of ASC and all the partial derivatives with respect to class variables are factors of $J * I$.

Put Γ_1 at the root of a tree, which is to be constructed. The partial derivatives associated with the uniform coordinate system for Γ_1 are sums of products of ratios of partial derivatives of polynomials in ASC.

Let $(I_1, ..., I_q)$ list the factors of I, i.e. the initials and the partial derivatives of polynomials in ASC with respect to class variables or parameters. For each one of these factors, I_j, we form a lower order branch by adding $I_j = 0$ to PS, keeping the $J \neq 0$ condition as it was in Γ , and for each branch we iterate the above process. Continuing, we get a tree of conditions Γ_i, all in triangular form, and where each ascending set has lower order than the one before. Each downward path in the tree is finite, since a descending sequence of ascending sets must be finite. Also we have an upper bound on the number of branches at each node: twice the sum of the number of variables and the number of parameters. So, by Konig's lemma, the tree is finite. Therefore, the process eventually terminates. Γ is equivalent to the disjunction of the Γ_i.

If we wish the final set of triangular conditions to be disjoint, an alternative version of this technique can be used. Given $(I_1, ..., I_q)$, the factors of I as described at the beginning of the last paragraph, we form 2^q branches by considering all ways of dividing these factors into two sets, a zero set and a non zero set. Then a branch is formed by multiplying J by the non zero set and by adding the zero set to PS.

(Another possibility is to factorize all the polynomials in ASC and to branch on all the factors, as well as the factors of I. Note that if the polynomials are all irreducible, the dimension goes down every time a new zero condition is added, so no path in the tree could have length more than n.)

It is not clear which of these alternatives is better computationally, even if we are not interested in disjointness.

That completes the proof of the algebraic lemma.

Next suppose that r, the length of the Pfaff construction, is greater than 0. In this case recursively put Γ^* in the form of a disjunction of triangular forms. All have the same Pfaff manifold, obtained by dropping the last part of the construction for Γ.

We now put Γ into triangular form under the assumption the Γ^* is already in this form.

Inductive lemma.

Given Γ, with Γ^* in triangular form, we can effectively rewrite Γ into a finite disjunction of triangular forms.

Proof.

Let Γ be

$(M_i, V_i)_{i \leq r}$

ASC $= 0$, where ASC is $(p_i)_{i \leq s}$

$J \neq 0$

Since Γ^* is triangular, condition T1) implies that all the initials of ASC and all the partial derivatives of polynomials in ASC with respect to their class variables are factors of J.

Since Γ^* is triangular, it is uniform. Let $z_1, ..., z_k$ be a uniform local coordinate system for all the defined sets. The last vector field in the construction is V_r. For each z an independent variable among $z_1, ..., z_k$, consider $Dz = (\partial X/\partial z) \bullet V_r$. (These $\partial X/\partial z$ are the partial derivatives defined by the relation between the dependent and the independent variables. By condition U2) in the definition of uniformity, they are all rational in X.) There are now two cases.

Case 1. For all z in the coordinate system for Γ^*, Dz has remainder 0 with respect to ASC. Then any path in the coordinate system for Γ^*, i.e. any smooth path in $(z_1, ..., z_k)$ is orthogonal to V_r. Also, as long as J is non zero, such a path can't go through a singular point of V_r in M_{r-1}, since we started with a strong enough J condition to ensure that this could not happen. In this case, therefore, the same coordinate system, namely, $(z_1, ..., z_k)$ can be used for Γ as was used for Γ^*. Also for any values of the parameters, the boundary of a Γ definable set is also on the boundary of a Γ^* definable set, and thus must have $J = 0$. So Γ is triangular already. In this case we will say that the top condition of Γ is marginal. Regrettably, it is not clear whether or not the marginal conditions can be eliminated. Perhaps the marginal conditions are needed to distinguish different components of the defined set.

Case 2. The other possibility is that for some z in the coordinate system for Γ^*, Dz does not have remainder 0 with respect to ASC. We pick such z with maximum order.

(So if independent w is more important than z but not picked, it must be that Dw has remainder 0 with respect to ASC; in this case the top condition does not depend on w.)

We now break into two branches, a zero branch and a non zero branch. The zero branch is obtained by adding Dz = 0 to ASC. After adding Dz=0 to get the zero branch, we have to recalculate Γ^* in triangular form. The ascending sets which are obtained will all have lower order than ASC. So the zero branch eventually terminates. (Note that if the polynomials in ASC are kept irreducible by factorizing and branching on the factors, then every time a zero condition is added, either the dimension goes down or one of the vector fields becomes marginal, thus replacing a possibly transcendental condition with an algebraic one. In this case, no path in the tree can have length more than $2n$.)

The non zero branch is obtained by replacing J by J^*Dz. It is claimed that the new condition, which we will call Γ_1, is in triangular form, and that the new coordinate system is obtained by removing z from $(z_1, ..., z_k)$. Let $Y = (y_1, ..., y_{k-1})$ be this new coordinate system. We are now in the same situation as we were when proving the local graph proposition above, and the same arguement shows that $(y_1, ..., y_{k-1})$ can be used as a uniform local coordinate system. If y is any one of the variables in Y, $\partial z/\partial y = -((\partial X/\partial y) \bullet V_k)/((\partial X/\partial z) \bullet V_k)$.

There is also a boundary condition to verify, condition U3). We want to show that JDz becomes zero at the boundary in all cases. Fix any values of the parameters. Let S be the set defined by Γ_1. Suppose that v is a limit point of S, but that JDz is not zero at v. We want to show that v is in S. Let S* be the set defined by Γ^*, with the same parameters. Since Γ^* is triangular and J is not zero at v, it must be that v is in S*. So the coordinate system $(z_1, ..., z_k)$ is valid around v in some neighbourhood N. We may also assume that N is picked so small that J^*Dz is bounded away from 0 in N. Since v is a limit point of S, there are points of S in N.

Let v^* be the Y coordinates of v, obtained by projection. Take any point u of S in N and find its coordinates $u*$ in Y. Draw a straight line in the Y projection from $u*$ to $v*$, and define the z values appropriately, with the initial value of z agreeing with u at $u*$, so that the resulting path is orthogonal to V_r. Shrink N so that all the paths from points in the new N lie in the old N, in which J^*Dz is bounded away from 0. (This can be done because we can bound the partial derivatives of z in the old N.) All these paths must end up with the same z value, say $z*$ when they arrive at $v*$, since otherwise Dz = 0 would have a solution in the old N. Let $z(v)$ be the z coordinate of v. Around $v*$ we can define z as a function of Y, with $z(v*) = z*$, and this function is continuous at $v*$. Also a sequence on the graph tends to v. Thus $z(v)$ is the same as $z*$. This means that there are paths orthogonal to V_r which pass through v and also intersect S in N. The only way v could then fail to be in S would be if the path did not lie in M_r, and this could only happen if the path contained a singular point of V_r in M_{r-1}. The original non zero condition J was fixed at the beginning so that where J is not zero, all the vectors are linearly independent. However at a singular point, V_r is normal to the orthogonal space of the previous vectors, and so the whole list of vectors is dependent.

This finishes the description of the algorithm.

For any Gamma condition Γ, let Zero(Γ) be the set defined by it, depending on the parameters. We have the following version of the Wu-Ritt zero structure theorem.

Zero structure theorem

Given any Gamma condition Γ, we can effectively find a finite list $\Gamma_i, i = 1, ..., m$, which are all in triangular form and so that Zero(Γ) is the union of Zero(Γ_i), $i = 1, ..., m$.

Note that the Γ_i are found uniformly, for all parameters.

The Gamma sets have remarkable closure properties, as mentioned in Khovanskii [1983], and this theorem has a large number of applications. They will not be discussed any further here, but in the next section the theorem will be used to prove a version of the Khovanskii finiteness theorem.

2. Khovanskii Finiteness

Define a refinement of a Gamma condition to be another Gamma condition obtained by adding either more polynomials to the polynomial set PS, or more factors to the non zero condition J.

Sampling Theorem

For any Gamma condition, Γ, we can find a finite set of refinements (Γ_i) each of which is triangular form and of dimension 0 so that, for any interpretation, every connected component of Zero(Γ) contains at least one point from one of the refinements Zero(Γ_i).

Proof

By the zero structure theorem we may as well assume that Γ is in triangular form. Suppose the dimension of it is k.

For an induction hypothesis, assume that the sampling can be done for all triangular Gamma conditions with dimension less than k.

Suppose Γ is
$(M_i, V_i)_{i \le r}$
$PS = 0$

$J \neq 0$

Define eccentricity E to be the sum of the squares of the coordinate variables plus the square of the reciprocal of the non zero J polynomial.

$E = \sum_{i=1}^{n} x_i^2 + (1/J)^2$

All the variables are functions of the independent variables $(z_1, ..., z_k)$ in the uniform coordinate system, and the functions have partial derivatives which are known rational functions of the coordinate variables. So dE/dz_k is a known rational function of $x_1, ..., x_n$. Since Γ is triangular and uniform, J is zero on the boundary. Thus any path that leads to the boundary necessarily sends E to infinity. Consider moving just z_k, leaving the other independent variables constant. A path which never gets to a boundary but goes to infinity sends E to infinity anyway. Therefore there is at least one minimum point of E on a path which is formed by starting at any point on any component and moving z_k in both directions. Therefore there is at least one isolated root of dE/dz_k $=0$ on each such path. Take $dE/dz_k = 0$ to be the algebraic condition which is used to refine Γ. This is done by first multiplying by some factor in order to clear the denominators and thus obtaining a polynomial DE. DE can't have remainder 0 relative to the ascending set PS. If it did E would be constant, which is impossible. Now form PS(E) by adding DE to PS. Let Γ(E) be the new Gamma condition. This new condition can now be broken down into triangular parts. We also have Zero(Γ(E)) intersects every connected component of Zero(Γ).

The triangular conditions obtained by breaking down Γ(E) all have lower order than the original PS, so this already proves the sampling theorem, by induction on order.

It is possible also to show directly that the triangular parts of Γ(E) all have dimension less than k, so the theorem can also be proved by induction on k, rather than by induction on order of ascending sets. Let γ be one of the triangular parts of Γ(E). Suppose, to get a contradiction, that this γ has dimension $\geq k$. Of course Zero(γ) is contained in Zero (Γ), and the bigger manifold has dimension k. A smooth path in the small manifold which happens to lie in the bigger manifold is also a smooth path in the bigger manifold, since they are locally graphs. Thus at each point the tangent plane of the small manifold is contained in the tangent plane of the bigger one. However we supposed that the smaller manifold had dimension at least k. So the tangent planes are equal. It follows that a smooth path in the large manifold which intersected the small manifold would be contained in the small manifold. It would follow that the path obtained by moving z_k as discussed earlier would have to have DE identically 0. However E tends to infinity eventually, so it is not possible for E to be constant.

So the triangular parts of Γ(E) have dimension less than k. By induction hypothesis, it follows that they can all be refined into triangular sets with dimension zero.

The Khovanskii finiteness theorem can now be proved.

Khovanskii's Finiteness Theorem

Given Gamma condition, Γ, which may depend upon parameters, we can find a number N independent of the parameters so that Zero(Γ) can have no more than N connected components.

Proof

The proof is by induction on the length, r, of the construction of Γ.

First suppose $r = 0$. By the sampling theorem we can find a list of Gamma conditions $\Gamma_i, i = 1, ..., k$, each of which is in triangular form and has dimension 0 so that every component of Zero(Γ) contains at least one of the points in Zero(Γ_i). However for each Γ_i, we can find an upper bound on the number of points which it can describe. One such upper bound would be the product of the degrees of the polynomials in PS_i, the polynomial set belonging to Γ_i Sharper upper bounds could be found using the Descarte rule of signs.

Next suppose that r is greater than zero but that the theorem has been proved for constructions of length less than r.

We can, by the sampling theorem, effectively find a finite list Γ_i of Gamma conditions with dimension 0 in triangular form so that every component of Zero(Γ) contains at least one of the sample points. It is sufficient, therefore, to show that we can calculate N in case Γ itself is in triangular form and has dimension 0.

As above, let Γ^* be obtained from Γ by dropping the top condition. By definition of triangular form Γ^* is also triangular, and it must have either dimension 0 or 1. By induction hypothesis, we can calculate an upper bound, say N_1 for the number of components of Zero(Γ^*). If Γ^* has dimension 0, we are finished, since N_1 is then also an upper bound for the number of points in Zero(Γ).

Suppose the dimension is 1, and x is a uniform coordinate. Each of the components of Zero(Γ^*) is a smooth path $p(x)$. Let $p'(x)$ be the derivative of $p(x)$ at x. Since Γ^* is uniform, $p'(t)$ is a vector of rational functions in the coordinate variables. Γ^* is also simpler than Γ, so by induction we can find an upper bound N_2 for the number of solutions of $p'(x) \bullet V_r{=}0$, where V_r is the vector field used in the last step of the construction of Γ. Take the N of the theorem to be $N_1 + N_2$.

Suppose there were more than N elements in Zero(Γ). They must be distributed among the at most N_1 paths. There must be more than N_2 adjacent intersections of a path $p(t)$ with M_r, the last manifold of the construction. By assumption C3) about Pfaff manifolds, there would have to be more than N_2 solutions of $p'(x) \bullet V_r{=}0$, which is not possible.

That completes the proof of the theorem.

Example 4) A polynomial in several variables $x_1, ..., x_m$ and the exponentials $exp(x_1), ..., exp(x_m)$ has a zero set which is necessarily a finite union of connected components.

Example 5) Let $f(x)$, defined for all $x > 0$, be a function from positive reals into R^n, whose graph is also a described by some Gamma condition. Then $f(x)$ is eventually monotone in all its coordinates. Thus the limit of $f(x)$, as x tends to infinity, exists (although some of the coordinates of the limit may be plus or minus infinity).

Remarks

The upper bound which is asserted to exist above is actually found only after doing the decomposition of the Gamma condition into triangular parts.

I don't know how to bound, in advance, the complexity of the triangular parts produced by the Wu algorithm. However, see [Gallo and Mishra, 1990].

In order to get a bound comparable to Khovanskii's, it might be desirable to begin the whole construction with a random change of coordinate system. In particular it might be possible, after a random change of coordinate system, to find a 0 dimensional sample of a k dimensional set directly by taking the critical points of the eccentricity; rather than by step by step reduction as is done here.

A technique related to the sampling technique used here has been developed independently by A. J. Wilkie, and used to prove model completeness results for restricted Pfaffian functions.

The sampling technique given here allows, in the algebraic case, reduction of the problem of whether a given set of polynomial equations implies another polynomial equation over the reals to the problem of whether a zero dimensional given complex variety defined by rational polynomials has a real element; and Philip Milne at Bath and Paul Pedersen in New York have recently independently reported remarkable progress in the solution of this later problem by multidimensional Sturm sequences.

We would like to solve triangular conditions of dimension zero, i.e. to approximate the finite sets they define. Even in the algebraic case, this is not easy computationally. The results of van den Dries, however, give hope that, at least in theory and at least for bounded regions, this can be done for all Pfaff manifolds.

Another unsolved problem is the adjacency problem. This would be, after decomposition of a Gamma set into triangular pieces, to decide which pairs of these pieces were connected in the sense that one had points which were limit points of sequences from the other. A related problem is to construct, if possible, the boundary of a triangular set as a finite union of triangular sets.

Note that the problem of deciding whether or not a Gamma condition can be satisfied generalizes the problem of deciding whether or not a polynomial is positive definite, and that all these problems have been reduced to questions about existence of solutions of triangular conditions of dimension zero.

References

Arnon, D.S., and B. Buchberger, Algorithms in real Algebraic Geometry, Academic Press, 1988

Bochnak, J, M. Coste, and M-F. Roy, Geometrie Algebraic Reele, Springer Verlag, Ergebnisse der Mathematik und ihrer Grenzgebiete, 1980

Denef,J., and L. van den Dries, P-adic and real subanalytic sets, Annals of Mathematics, 128 (1988) 79-138

van den Dries, L., On the elementary theory of restricted elementary functions, The Journal of Symbolic Logic, vol 53, Number 3, Sept 1988, 796-808

Gallo,Giovanni and Bhubaneswar Mishra, Efficient Algorithms and Bounds for Wu-Ritt Characteristic Sets, MEGA conference, 1990

Khovanskii, A.G., On a class of systems of transcendental equations. Soviet Math. Dokl. 22 (1980), 762-765

Khovanskii, A.G., Fewnomials and Pfaff manifolds. In: Proc. Int. Congress of Mathematicians, Warsaw 1983, pp 549-564

Milne, P.,On the Algorithms and implementation of a geometric algebra system, Ph.D. thesis, Bath University 1990

Pedersen, P., Generalizing Sturm's theorem to N dimensions, New York University, Department of Computer Science, April 1990

Richardson, D., Non standard models of the theory of elementary functions of a real variable, Zeitschr. f. math. Logik und Grundlagen d. Math, Bd 34, S 355-372 (1988)

Richardson, D., Finding roots of equations involving functions defined by algebraic differential equations, Proceeding of Conference on Effective Methods in Algebraic Geometry (MEGA), 1990.

Richardson, D., Computing (in a bounded part of the plane) he topology of a real curve defined by solutions of algebraic differential equations, submitted for publication 1991

Spivak, M., Calculus on Manifolds, W.A. Benjamin, 1965.

Tarski, A., A decision method for Elementary Algebra and Geometry, 2nd edition, revised. Berkeley and Los Angeles 1951

van der Waerden,B.L., Algebra, Springer Verlag, Berlin-Gottingen-Heidelberg, 1960

Wilkie, A.J., Some Model Completeness results for expansions of the ordered field of real numbers by Pfaffian functions, manuscript , Oxford 1989

Wu Wen-tsun, A zero structure theorem for polynomial equations solving, Institute of Systems Science, Academia Sinica 1987