

# Symbolic and Numerical Computation for Artificial Intelligence

edited by

**Bruce Randall Donald**

Department of Computer Science  
Cornell University, USA

**Deepak Kapur**

Department of Computer Science  
State University of New York, USA

**Joseph L. Mundy**

AI Laboratory  
GE Corporate R&D, Schenectady, USA



**Academic Press**

*Harcourt Brace Jovanovich, Publishers*

London San Diego New York  
Boston Sydney Tokyo Toronto

ACADEMIC PRESS LIMITED  
24-28 Oval Road  
London NW1

*US edition published by*  
ACADEMIC PRESS INC.  
San Diego, CA 92101

Copyright © 1992 by  
ACADEMIC PRESS LIMITED

This book is printed on acid-free paper

*All Rights Reserved*

No part of this book may be reproduced in any form, by photostat, microfilm or any other means, without written permission from the publishers

A catalogue record for this book is available from the British Library

ISBN 0-12-220535-9

Printed and Bound in Great Britain by  
The University Press, Cambridge

## Chapter 3

# On the Solutions of a Set of Polynomial Equations

Philip S. Milne

*School of Mathematical Sciences*

*University of Bath*

*Claverton Down, Bath BA2 7AY, UK*

`psm@maths.bath.ac.uk`

---

Given  $n$  polynomials in  $n$  variables the paper suggests an algorithm for locating their zero-dimensional or “point-like” solutions. The technique differs from projection-based techniques by using the isolation strategy which is often used in the one dimensional case. Around each of the solutions the algorithm computes an arbitrarily small  $n$ -dimensional rectangle or *box* in which numerical techniques may be used to approximate the solution. The crux of the algorithm is a multivariate generalization of Sturm's theorem which may be computed without explicit generation of the symmetric functions of the system. To generate this sequence, a new construction, the *volume function*, is introduced and it is noted that standard elimination techniques may be used to compute it.

---

### 1. Introduction

What follows is a description of a technique for locating the real zero-dimensional or “point-like” solutions of systems of simultaneous polynomial equations.

In one dimension it is common to implement this procedure in two phases: an isolation phase and an approximation phase. The isolation phase produces a set of intervals sufficiently small for there to exist a single solution in each of them and is often implemented by recursively dividing a bounding interval given some strategy for counting the number of solutions inside an arbitrary interval. The second phase takes each of the intervals and uses numerical techniques to approximate the solution to some given tolerance.

It is possible to use this strategy in many dimensions as well, provided that an analogous technique for counting the number of solutions that lie within an  $n$ -dimensional rectangle or *box* is available. The initial bounding box may be computed by using any of the classical root bounds on each of the  $n$  univariate *eliminants* or projections onto the axes. This strategy differs from those of CAD and Gröbner Bases, firstly in that it is not projection-based and secondly in that it is only suitable for answering zero-dimensional or “point-like” problems.

This paper describes an extension of Sturm's theorem which is, in many ways, suitable for this sort of procedure. Hermite (1880) demonstrated the existence of a class of sequences which could be used to count roots inside boxes in  $\mathbf{R}^2$  and provided a method for computing them, in terms of the symmetric functions of the solutions. The construction described here is less general and allows the computation of just one element of the class of sequences that Hermite described. This particular sequence, however, has the property that it may be computed using elimination techniques alone. Like Hermite's sequences, each solution is counted just once regardless of its multiplicity which is a necessary property for an isolation algorithm.

Paul Pedersen's recent work (Pedersen, 1991b) has demonstrated how sequences similar to the construction Hermite (1852) describes may be extended to any number of dimensions, as well as providing generalizations to arbitrary shapes as well as boxes. Despite the fact that Pedersen's technique is able to generate not just a single sequence but a whole class of them, it is a surprising fact that the sequence described here is not amongst them. Unlike Pedersen's sequences each term in this sequence is in the ideal generated by the first two terms. The last term cannot therefore be a 1, as it is in Pedersen's construction, but is instead a perfect square. Pedersen (1991a) outlines the differences between these two techniques.

This algorithm is provided without complexity estimates. The crucial step of this algorithm is the computation of the volume function. A number of technologies are available for computing the volume function, amongst them: Buchberger's Algorithm, Macaulay Determinants, other results in Elimination Theory and, currently in two dimensions alone, the Subresultant Algorithm. Recent advances in the above techniques leave a choice of the most suitable technology unclear and hence appropriate complexity bounds difficult to state at the time of writing.

## 2. The Volume Function

We begin with a little motivation for the volume function which, like the  $u$ -resultant (van der Waerden, 1950), uniquely characterizes the point-like solutions of a system of  $n$  equations in  $n$  variables. A naïve analogue of the univariate factorization:

$$f(x) \propto \prod_{\alpha} (x - \alpha)$$

might, in two dimensions, be

$$f(x_1, x_2) \propto \prod_{\alpha} ((x_1 - \alpha_1)(x_2 - \alpha_2))$$

where  $\alpha$  is used to mean "has the same zero set as". But this, of course, suffers from the problem that the solutions  $(\alpha_1, \alpha_2)$ ,  $(\beta_1, \beta_2)$ , etc. are not distinguished from the spurious intersections of the factors of  $f$ , so that the points,  $(\alpha_1, \beta_2)$ ,  $(\gamma_1, \alpha_2)$ , etc. appear to be solutions as well. One way of including this information in the construction is to add another variable,  $u$ , to this product, so as to bind algebraically the components of the factors into their associated pairs.

DEFINITION 2.1. In  $\mathbb{R}^n$  we define<sup>†</sup> the volume function,  $V(u, \mathbf{x}; \mathbf{f})$ , of a system of  $n$  rational polynomials,  $f_i(\mathbf{x}) \in \mathbb{Q}[\mathbf{x}]$ , where  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ , in terms of the solutions,  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ , to the equation  $\mathbf{f}(\alpha) = \mathbf{0}$ ,

$$V(u, \mathbf{x}; \mathbf{f}) \propto \prod_{\alpha} (u + \prod_i (x_i - \alpha_i)).$$

When the number of solutions is not finite, the volume function is defined to be zero.

When the last argument is dropped, as in:  $V(u, \mathbf{x})$ , we assume that  $\mathbf{f}$  is given by the context.

Returning to the two dimensional case we note that, when  $u = 0$ , the volume function specializes to the “factorization” above. It also satisfies

$$\frac{V_u(0, x_1, x_2)}{V(0, x_1, x_2)} = \sum_{\alpha} \frac{1}{(x_1 - \alpha_1)(x_2 - \alpha_2)},$$

the numerator of which is 0 at the spurious intersection points mentioned above.

The function  $\frac{V_u(0, \mathbf{x})}{V(0, \mathbf{x})}$  is, in many ways, a good analog of the function  $\frac{f'}{f}$  in one dimension. It is trivial to show (Milne, 1990) that there is an analogue of Cauchy’s “root counting” integral in  $\mathbb{C}^n$  involving the above quotient and a volume integral. In  $\mathbb{C}^2$  we have,

$$\iint_{C_1 \times C_2} \frac{V_u(0, z_1, z_2)}{V(0, z_1, z_2)} dz_1 dz_2 = -4\pi^2 n,$$

where  $n$  is the number of roots inside both of the Jordan curves  $C_1$  and  $C_2$  but this is, of course, just a special case of the general theory of residues in many complex variables. It is also easy to show, by way of analogy with the Newton relations, that the symmetric functions of the system may be written in terms of the volume function,

$$\frac{V_u(0, \mathbf{x})}{V(0, \mathbf{x})} = \sum_{\mathbf{i} \in \mathbb{N}^n} \frac{1}{x^{\mathbf{i}}} \sum_{\alpha} \alpha^{\mathbf{i}-\mathbf{1}},$$

where  $x^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$  and  $\mathbb{N}$  is the set of all strictly positive integers.

### 3. Computation of the Volume Function

In essence, the volume function is computed by eliminating  $n$  new variables,  $a_i$ , from the  $n + 1$  equations:  $f_i(\mathbf{a})$  and  $u + (x_1 - a_1)(x_2 - a_2) \dots (x_n - a_n)$ . We define an *associate* of the volume function to be any polynomial with the same zero set and note that for the algorithms which follow it suffices to be able to compute any of the volume function’s associates.

Buchberger’s algorithm may be used to compute a reduced Gröbner basis for this system. If the solution set is zero-dimensional and a purely lexicographic ordering,  $a_i > a_{i+1} > u$ , is used then a single polynomial of degree zero in all of the elimination variables will be contained in the Gröbner basis (Buchberger, 1985b). Since this basis generates

<sup>†</sup> Currently, just up to a constant.

the same ideal as the input, this element must be an associate of the volume function and when there are not finitely many solutions there will be no such term. This statement follows trivially from the theory of ideals.

Denoting by  $G(p, \mathbf{x})$  the reduced Gröbner basis of the elements  $p_i$  with respect to the purely lexicographic ordering  $x_1 > x_2 > x_3 \dots$  we may write the volume function in terms of the Gröbner basis.

$$V(u, \mathbf{x}; \mathbf{f}) \propto \mathbf{R}[u, \mathbf{x}] \cap G((f_1(\mathbf{a}), f_2(\mathbf{a}), \dots, f_n(\mathbf{a}), f_{n+1}(u, \mathbf{x}, \mathbf{a})), (a_1, a_2, \dots, a_n, u)),$$

where  $f_{n+1}(u, \mathbf{x}, \mathbf{a}) = u + (x_1 - a_1)(x_2 - a_2) \dots (x_n - a_n)$ .

Since it is only the final eliminant which is required, computation of the full Gröbner basis seems very wasteful. By contrast, the techniques used in elimination theory (van der Waerden, 1950) may be used to generate the volume function on its own. In particular, Macaulay's construction allows any eliminant to be written as the quotient of two determinants containing the coefficients of  $\mathbf{f}$  (Macaulay, 1903). If  $\mathbf{x}$  is specialized before the elimination is computed, then techniques similar to those of Canny (1990) and Renegar (1989) allow the eliminant to be computed in singly exponential time even in the case when the determinants vanish identically.

In two dimensions the subresultant algorithm appears, in practice, to be the most efficient way to compute the volume function.

$$V(u, x_1, x_2) \propto \frac{\text{Res}_{a_2}(\text{Res}_{a_1}(f_1(a_1, a_2), f_3), \text{Res}_{a_1}(f_2(a_1, a_2), f_3))}{u^{\deg(f_1(x_1, 0)) \deg(f_2(x_1, 0))}}$$

where  $f_3(u, x_1, x_2, a_1, a_2) = u + (x_1 - a_1)(x_2 - a_2)$ †.

#### 4. Univariate Sturm Sequences

Throughout this text we make use of two new functions *var* and *per*† which respectively count the number of variations and permanencies of sign in a sequence of reals. As long as the convention is consistent, a zero may be taken as either positive or negative since it will be straddled by terms of opposite sign§. We note that:

$$\text{per}(S(x)) + \text{var}(S(x)) = l - 1, \tag{4.1}$$

where  $l$  is the length of the sequence. For completeness, we now give the following few theorems about univariate Sturm sequences.

Given a square-free polynomial  $p(x)$  we can construct a *Sturm* sequence of polynomials  $S_i = -\text{rem}(S_{i-2}(x), S_{i-1}(x))$ , where  $S_1(x) = p(x)$  and  $S_2(x) = p'(x)$ .

† The proof of this result is not very enlightening and has not been published.

‡ The per operator, which replaces the conventional var operator in these theorems, does nothing except introduce a minus sign throughout. This makes the evaluation of these sequences analogous to that of definite integration.

§ Except at the endpoints, where the choice defines closure.

**THEOREM 4.1. (STURM)** *The square-free polynomial  $p$ , with Sturm sequence  $S$ , has  $\text{per}(S(b)) - \text{per}(S(a))$  real roots in the interval  $[a, b]$ .*

**PROOF.** The property is clearly true when  $a = b$ . As we increase  $b$  through the roots of  $p(x)$  we note the following:

- Through an upward going root,  $p(x)$  changes from negative to positive whilst  $p'(x)$  is positive.
- Through a downward going root,  $p(x)$  changes from positive to negative whilst  $p'(x)$  is negative.

We therefore introduce one permanency of sign at the head of the Sturm sequence as we pass each root. Of the remaining elements in the list we observe:

$$S_i(x) = -\text{rem}(S_{i-2}(x), S_{i-1}(x))$$

so,

$$S_i(x) = -q(x)S_{i-1}(x) - S_{i-2}(x).$$

The sequence cannot contain two consecutive zeros since the last term is a constant. As  $S_{i-1}(x)$  vanishes, we see from the above relation that the two terms  $S_i(x)$  and  $S_{i-2}(x)$ , which straddle  $S_{i-1}(x)$  in the sequence, are of opposite signs. A sign change in  $S_{i-1}(b)$  as  $b$  passes through the root therefore makes no change to the number of permanencies of sign in the sequence and  $\text{per}(S(a))$  is incremented when and only when we cross a root of  $p(x)$ . □

Now we write down exactly what  $\text{per}(S(x))$  is, rather than just how it changes. We will restrict ourselves to *normal* sequences: those in which the degree drop between consecutive terms is unity. We also observe that the (negated) subresultant algorithm (section 12), may be used to generate the remainder sequence since, in the normal sequence, multiplications in the coefficient domain are all by perfect squares and are therefore sign preserving.

**LEMMA 4.1.** *If  $S(x)$  is a normal subresultant Sturm sequence for the square-free polynomial,  $p(x)$ , then  $\text{per}(S(a))$  is the number of real roots of  $p$  which are less than  $a$  plus the number of pairs of complex roots of  $p$ .*

**PROOF.** The number of real roots of  $p$  in the interval  $[a, b]$  is given by Sturm's theorem as  $\text{per}(S(b)) - \text{per}(S(a))$ . Consequently,

$$\text{per}(S(\infty)) - \text{per}(S(-\infty)) = r \tag{4.2}$$

where  $r$  is the total number of real roots.

Now, because  $S$  is normal,  $\text{per}(S(-\infty))$  is the number of permanencies of sign in the leading coefficients after alternate terms have been negated; this is just the number of sign variations in  $S(\infty)$ , so:

$$\text{per}(S(-\infty)) = \text{var}(S(\infty)).$$

Summing (4.1) evaluated at  $x = \infty$  and the above we have,

$$\text{per}(S(\infty)) + \text{per}(S(-\infty)) = l - 1,$$

where  $l$  is the length of the sequence. But, because the sequence is normal,  $l - 1 = \text{deg}(p)$

and from the fundamental theorem of algebra,  $\deg(p) = n$ , where  $n$  is the number of roots of  $p$ . So,

$$\text{per}(S(\infty)) + \text{per}(S(-\infty)) = n.$$

Subtracting (4.2) from the above gives,

$$2\text{per}(S(-\infty)) = n - r = c$$

where  $c$  is the number of complex roots. Since, for a polynomial with real coefficients, each complex root has a conjugate,  $\text{per}(S(-\infty))$  counts the number of such conjugate pairs.

Using Sturm's theorem we have that:  $\text{per}(S(a)) - \text{per}(S(-\infty))$  is the number of real roots less than  $a$ .  $\text{per}(S(a))$  is therefore the number of real roots less than  $a$  plus the number of pairs of complex roots.  $\square$

**COROLLARY 4.1.** *Provided the subresultant Sturm sequence,  $S$ , of a square-free polynomial  $p$  is normal,  $\text{per}(S(0))$  is equal to the number of negative real roots of  $p$  plus the number of pairs of complex roots.*

In all of the above statements about the Sturm sequence, we may replace the restriction that the polynomial should be square-free with the restriction that the last term should not vanish when the sequence is evaluated. We can do this because the polynomial  $g(x) = \gcd(p(x), p'(x))$  divides each element of the sequence to give precisely the sequence pertinent to the square-free part of  $p$ . Multiplication of each element of the sequence by a constant does not change the number of permanencies of sign unless that constant is zero. In particular, the above corollary transforms to a statement about the distinct negative roots of the polynomial  $p$ .

**COROLLARY 4.2.** *Provided the subresultant Sturm sequence,  $S$ , of a polynomial  $p$  is normal and the last term in the sequence,  $S_i(x)$ , is non zero at both  $a$  and  $b$ ,  $\text{per}(S(0))$  is the number of distinct negative real roots of  $p$  plus the number of distinct pairs of complex roots.*

## 5. An Evaluation Function for the Sequence

We now define an evaluation function  $E(M, I)$  which given a sequence of multivariate polynomials  $M$  and a coordinate aligned box,  $I_i = [a_i, b_i], 1 \leq i \leq n$ , evaluates the sequence at the corners of the region and uses the per operator to return a single integer.

In  $\mathbf{R}$ ,

$$E(M, I) = \text{per}(M(b)) - \text{per}(M(a)).$$

In  $\mathbf{R}^2$ ,

$$E(M, I) = \frac{1}{2}(\text{per}(M(b_1, b_2)) + \text{per}(M(a_1, a_2)) - \text{per}(M(b_1, a_2)) - \text{per}(M(a_1, b_2))).$$

In  $\mathbf{R}^n$  we define the function recursively,

$$E(M, I) = E_n(M, I),$$



$$E_i(M, I) = \frac{1}{2}(E_{i-1}(M(x_i \leftarrow b_i), I) - E_{i-1}(M(x_i \leftarrow a_i), I)), \tag{5.1}$$

where  $f(x \leftarrow a)$  is used to denote specialization of  $f$  with the substitution  $x = a$ ,

$$E_0(M, I) = 2\text{per}(M).$$

### 6. The Generic Sequence

Given a system of  $n$  polynomials,  $f_i$ , we may compute the associated volume function, which will be non-zero if and only if there are finitely many solutions, and in this case,

$$V(u, \mathbf{x}; \mathbf{f}) \propto \prod_{\alpha} (u + \prod_i (x_i - \alpha_i)).$$

Treating  $u$  as the main variable of  $V$  we can now compute a multivariate sequence  $S(u, \mathbf{x})$  of  $V(u, \mathbf{x})$  by using the negated subresultant algorithm on  $V$  and  $V_u^\dagger$ ,

$$S(u, \mathbf{x}) = \text{nprs}_u(V(u, \mathbf{x}), V_u(u, \mathbf{x})).$$

Define  $M$  to be the value of the sequence at  $u = 0$ ,

$$M(\mathbf{x}) = S(0, \mathbf{x}).$$

**THEOREM 6.1.** *The number of distinct simultaneous real roots of a system of  $n$  polynomials,  $\mathbf{f}$ , with normal sequence  $M(\mathbf{x}; \mathbf{f})$  in any coordinate aligned box  $I_i = [a_i, b_i]$  in  $\mathbf{R}^n$  is precisely  $E(M, I)$  provided that the last term of the sequence does not vanish at any of the vertices of the box.*

**PROOF.** We have from corollary 4.2 that if  $S$  is a normal subresultant sequence then  $\text{per}(S(0))$  is the number of distinct negative real roots plus the number of distinct pairs of complex roots.  $M(\mathbf{x})$  is just such a sequence, for the volume function evaluated at  $u = 0$ .  $\text{per}(M(\mathbf{x}))$  is therefore the number of distinct roots,  $\alpha$ , of the system for which  $-(x_1 - \alpha_1)(x_2 - \alpha_2) \cdots (x_n - \alpha_n)$  is distinct, negative and real plus half the number of distinct roots for which this expression is complex.

First, we note that any complex root,  $\alpha \notin \mathbf{R}^n$ , for which  $-(x_1 - \alpha_1)(x_2 - \alpha_2) \cdots (x_n - \alpha_n) \notin \mathbf{R}$  at any of the corners of the box, makes the same contribution at each corner and, in combination with the evaluation function  $E(M, I)$ , contributes zero to the final root count.

Second, we consider  $\alpha \notin \mathbf{R}^n$  and  $-(x_1 - \alpha_1)(x_2 - \alpha_2) \cdots (x_n - \alpha_n) \in \mathbf{R}$ . Since the polynomial map has real coefficients, any complex root,  $\alpha \notin \mathbf{R}^n$ , has a conjugate  $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_1, \dots, \bar{\alpha}_n)$  which is also a root of the system. So, for the conjugate  $-(x_1 - \bar{\alpha}_1)(x_2 - \bar{\alpha}_2) \cdots (x_n - \bar{\alpha}_n) = \overline{-(x_1 - \alpha_1)(x_2 - \alpha_2) \cdots (x_n - \alpha_n)} = -(x_1 - \alpha_1)(x_2 - \alpha_2) \cdots (x_n - \alpha_n) \in \mathbf{R}$ . This, therefore, corresponds to a double root of the volume function and the last element of  $M$  will be zero.

† Since the leading coefficients of both  $V$  and  $V_u$  are integers, multiplications in the coefficient domain are all by perfect squares of polynomials in  $\mathbf{R}[\mathbf{x}]$  (section 12) unless the sequence is *abnormal*. It seems likely that there is an algorithm similar to the subresultant algorithm which can generate precisely those terms of section 8 regardless of anomalies in the performance of the Euclidean algorithm.

For the real roots then,  $\alpha \in \mathbb{R}^n$ , we break the case up into two parts: when the roots are inside and outside of the box. Firstly, if  $\alpha \notin I$ , we may, without loss of generality, take  $\alpha_1 \notin I_1$ .

The recurrence relation (5.1) gives  $E_1(M, I)$  in terms of  $\text{per}(M)$ . The sequences given to  $\text{per}$  are the Sturm sequences of the two functions;

$$\prod_{\alpha} (u + (b_1 - \alpha_1)(x_2 - \alpha_2)(x_3 - \alpha_3) \cdots (x_n - \alpha_n))$$

and

$$\prod_{\alpha} (u + (a_1 - \alpha_1)(x_2 - \alpha_2)(x_3 - \alpha_3) \cdots (x_n - \alpha_n)),$$

where each  $x_i$  is consistently instantiated to either  $b_i$  or  $a_i$ . Since  $\text{sgn}(b_1 - \alpha_1) = \text{sgn}(a_1 - \alpha_1)$  the contribution  $\alpha$  makes to the number of negative real roots in the first function is equal to its contribution to the second. Thus  $\alpha$  contributes 0 to  $E_1(M, I)$  and hence 0 to  $E(M, I)$ .

Last then, we consider  $\alpha \in I$  and observe that such a root contributes to each vertex with which it makes positive volume. Since there are  $2^{n-1}$  of these, this root will contribute 1 to  $E(M, I)$ .

Thus  $E(M, I)$  computes the number of roots which lie in the box. □

## 7. An Example

As a simple example, let us take the two polynomials  $P(x, y) = x^2 + y^2 - 2$  and  $Q(x, y) = x - y$ . This we might think of as a circle and a line, intersecting at the points  $(1, 1)$  and  $(-1, -1)$ . To compute the volume function we need to eliminate two variables,  $a$  and  $b$  say, from the following system

$$\{P(a, b), Q(a, b), u + (x - a)(y - b)\}.$$

This is,

$$\{a^2 + b^2 - 2, a - b, u + (x - a)(y - b)\}$$

so we can use the second polynomial to substitute  $a$  for  $b$  in the other two and this gives,

$$\{2a^2 - 2, u + (x - a)(y - a)\}$$

or

$$\{a^2 - 1, u + xy - a(x + y) + a^2\}.$$

Now use the first term to eliminate  $a^2$  in the second,

$$\{a^2 - 1, u + xy + 1 - a(x + y)\}.$$

then multiply the first term by  $(x + y)$ , the second by  $a$  and add them,

$$\{u + xy + 1 - a(x + y), a(u + xy + 1) - (x + y)\}.$$

Finally multiply the first term by  $(u + xy + 1)$ , the second by  $(x + y)$  and sum to give,

$$\{(u + xy + 1)^2 - (x + y)^2\},$$

the volume function.

Its derivative with respect to  $u$  is:  $2(u + xy + 1)$ , and minus the remainder of the previous two terms is  $4(x + y)^2$ . Evaluating at  $u = 0$  and removing numeric contents gives the sequence as:

$$((x^2 - 1)(y^2 - 1), (xy + 1), (x + y)^2).$$

We may not evaluate the sequence anywhere along the line  $-x = y$  but anywhere else it will serve to count the number of roots.

For example, the region  $[-3, 3] \times [-2, 2]$  has corners at,  $(-3, -2)$ ,  $(-3, 2)$ ,  $(3, -2)$  and  $(3, 2)$  where the sequence evaluates to  $(24, 7, 25)$ ,  $(24, -5, 1)$ ,  $(24, -5, 1)$  and  $(24, 7, 25)$  respectively. The number of permanencies of sign in these sequences are 2, 0, 0 and 2 respectively so that the number of roots in the region is computed as two.

### 8. A Sturm Sequence in Terms of the Roots

Sylvester (1852) gave each of the  $l$  terms of the Sturm sequence in terms of the roots of  $S_0$ . We sacrifice the ability to represent the generic term here by using nineteenth century notation, with the convention that  $\sum_{\alpha\beta} \theta(\alpha, \beta) \equiv \sum_{i=1}^n \sum_{j=i+1}^n \theta(\alpha_i, \alpha_j)$  etc.

$$\begin{aligned}
 S_1 &\propto \prod_{\alpha} (x - \alpha) \\
 \frac{S_2}{S_1} &\propto \sum_{\alpha} \frac{1}{(x - \alpha)} \\
 \frac{S_3}{S_1} &\propto \sum_{\alpha\beta} \frac{(\alpha - \beta)^2}{(x - \alpha)(x - \beta)} \\
 \frac{S_4}{S_1} &\propto \sum_{\alpha\beta\gamma} \frac{(\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2}{(x - \alpha)(x - \beta)(x - \gamma)} \\
 &\vdots \\
 S_l &\propto \prod_{\alpha\beta} (\alpha - \beta)^2
 \end{aligned}$$

### 9. This Sequence in Terms of the Roots

Making substitutions:

$$x \leftarrow 0, \quad \alpha \leftarrow -(x_1 - \alpha_1)(x_2 - \alpha_2), \quad \beta \leftarrow -(x_1 - \beta_1)(x_1 - \beta_2), \quad \dots$$

into the expressions above yield the expressions for this sequence, in two dimensions, as a function of the Cartesian components of roots.

$$\begin{aligned}
 M_1 &\propto \prod_{\alpha} (x_1 - \alpha_1)(x_2 - \alpha_2) \\
 \frac{M_2}{M_1} &\propto \sum_{\alpha} \frac{1}{(x_1 - \alpha_1)(x_2 - \alpha_2)} \\
 \frac{M_3}{M_1} &\propto \sum_{\alpha\beta} \frac{((x_1 - \alpha_1)(x_2 - \alpha_2) - (x_1 - \beta_1)(x_2 - \beta_2))^2}{(x_1 - \alpha_1)(x_2 - \alpha_2)(x_1 - \beta_1)(x_2 - \beta_2)} \\
 &\vdots \\
 M_l &\propto \prod_{\alpha\beta} ((x_1 - \alpha_1)(x_2 - \alpha_2) - (x_1 - \beta_1)(x_2 - \beta_2))^2
 \end{aligned}$$

These terms are readily seen to be a special case of the general formulae Hermite (1853) provides.

## 10. Closure

The phrase "roots in the box" has been used to describe, rather loosely, how each root in  $\mathbf{R}^n$  contributes to the count, ignoring the possibility of a root on the boundary of the region of interest. First, we note that the evaluation function,  $E(M, I)$ , is *additive* so that, however *per* is defined, any two neighboring boxes must join invisibly along their common boundary.

We have already exempted the case of a zero appearing at the end of the sequence. We have also mentioned that internal zeros will be straddled by terms of opposite sign so that the number of permanencies of sign is not affected by the convention for these zeros. There remains the question of how to deal with zeros at the beginning of the sequence. We know that if the first two terms of the sequence are zero then the entire sequence vanishes identically and that this case has already been exempted. The final case is when the first term is zero and the second is non-zero and it is the convention adopted in this case which defines the extent to which roots on the boundaries are counted.

In the following section we adopt the convention that this zero is deleted, i.e. that  $\text{per}(0, 1) = \text{per}(0, -1) = 0$ . The statements that follow use the fact that, in one dimension, a Sturm sequence which uses this convention is left-closed, right-open. This means that zero does not contribute to  $\text{per}(S(0))$ : the count of the negative real roots of the volume function.

In  $n$  dimensions there are  $2^n$  vertices. Label the ones that make a positive volume with respect to the centroid: *positive vertices* and the others *negative vertices*.  $2^{n-1}E(M, I)$  is the number of roots which show a positive volume to the positive vertices minus the number which show a positive volume to the negative vertices. A root on a surface of the box, of co-dimension  $c$ ,  $c < n$ , contributes zero volume to all of the positive vertices which have a Cartesian component in common with the root. There remain  $2^{n-c-1}$  positive vertices which do not have a Cartesian component in common with the root.  $E(M, I)$ ,

the computed count, is this number divided by the normalizing factor  $2^{n-1} = 2^{-c}$ . If  $c = n$  then only the diagonally opposite vertex, which has no Cartesian components in common with the root, will contribute to the count. If this is a negative vertex then the root contributes zero to  $E(M, I)$  otherwise the contribution is  $2^{1-n}$ .

It is also possible to use the convention that  $\text{per}(0, 1) = \text{per}(0, -1) = \frac{1}{2}$ . A similar argument to the above applies except at the points of co-dimension  $n$  (the vertices of the box) where the contribution is the same regardless of whether the opposite vertex is positive or negative. If the vertex is positive, then it will contribute once to the count, but there will be an unmatched negative vertex which contributes  $-\frac{1}{2}$  to the count. If the opposite vertex is negative, then it contributes zero to the count but there will be an unmatched positive vertex which will contribute  $\frac{1}{2}$ . Either way this root contributes  $\frac{1}{2}$  to the count and we have the simple result that roots on the boundaries of co-dimension  $c$  always contribute  $2^{-c}$  to  $E(M, I)$ .

## 11. Future Work

### 11.1. EVALUATION MECHANISMS

In a practical implementation it is necessary to specify  $\mathbf{x}$  before the volume function is computed, instead of generating the volume function in its generic form and specializing the multivariate sequence,  $M(\mathbf{x})$ , afterwards. A number of new problems arise in this scheme. In particular, it is necessary to know the total number of multiplicities of the system. It is easy to show, by analogy with the univariate case that this number is precisely:  $\text{deg}_u(\text{gcd}(V(u, \mathbf{x}), V_u(u, \mathbf{x})))$ , for generic  $\mathbf{x}$ , but this is of little use if it is impractical to compute the generic volume function in the first place.

Dealing with the specialized instances of the volume function at each corner is made easier, however, by the fact that they become univariate. Conventional Sturm sequence technology may be used to compute the number of negative real roots of these functions, allowing the restriction to normal sequences to be dropped. In fact, Sturm sequences can be dropped altogether, since any scheme which counts the number of negative real roots of these functions may be used to implement the evaluation function  $E(M, I)$ .<sup>†</sup>

### 11.2. DEFINITENESS OF A SINGLE REAL POLYNOMIAL IN A BOX

Deciding the definiteness of a single polynomial  $p(\mathbf{x})$  in a box is the application for which this work was undertaken. The volume function, although useful, is not the only function which may be used in the elimination phase of this algorithm. More generally, we may find the number of solutions to  $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$  which make an arbitrary *query* polynomial,  $q(\boldsymbol{\alpha})$ , negative in the same way; by eliminating the  $a_i$  from the  $n + 1$  equations:  $f_i(\mathbf{a})$  and  $u - q(\mathbf{a})$ .

We may modify a count of the number of simultaneous roots of the partial derivatives

<sup>†</sup> This follows from the last half of theorem 6.1 and assumes repeated roots in the function are dealt with in the same way that they are by the Sturm sequence.

of  $p$  to count  $+1$  or  $-1$  depending on the sign of  $p$  at these points. This is done by using  $f_i(\mathbf{a}) = \frac{\partial p}{\partial x_i}(\mathbf{a})$  and the query polynomial  $q = -(x_1 - a_1)(x_2 - a_2) \cdots (x_n - a_n)p(\mathbf{a})$ .

The strategy above detects "bubbles" inside the box. Definiteness follows from a study of the definiteness of the function on the surface of the box: a problem in  $n-1$  dimensions. Both this question and the problem of finding efficient evaluation mechanisms have been given only rather superficial study; the details are currently being considered.

## 12. The Algorithm

First, we give a procedure for producing a negated polynomial remainder sequence from two multivariate inputs  $p$  and  $q$  in a main variable  $u$ . It is Collins' subresultant algorithm (Collins, 1967) simplified for the special case when the degree drop between consecutive terms is unity. There is no check for this condition in the code.

```

procedure nprs(p, q);
  if degree(q, u) = 0 then
    if q = 0 then list p else list(p, q)
  else p . nprs(q, - rem(lc(q)^2*p, q) / lc(p)^2);

```

Second, we give a procedure to generate a multivariate Sturm sequence in terms of a list of polynomials  $f$  and a list of variables  $vars$ .

```

procedure M(f: set, vars: list);
begin local v, s, p, ans;
  p := for each var in vars product (var' - var);
  v := last(groebner(union(f, u + p), append(vars, list(u))));
  s := nprs(v, differentiate(v, u));
  ans := substitute(0, u, s);
  substitute(vars, vars', ans);
end;

```

## 13. Historical Note

It is of little virtue that this sequence was developed quite independently to the inspired works of Hermite. The author would like, nevertheless, to mention the paper from which this work was actually derived.

It was a short paper by Pinkert (1976) which deals with the problem of finding the number of complex roots of a polynomial inside a rectangle in  $\mathbb{C}$ . The paper contains a wonderful idea which Pinkert attributes to Bobby Caviness in the acknowledgments. The simple idea is to produce a new polynomial whose roots are *geometrically* related to those in the original problem. In Pinkert's paper, this is a polynomial whose roots are the squares of the roots in the original problem and a count of the number of these roots which have positive imaginary parts gives exactly the root counting primitive that is produced by the volume function. Indeed, it was within hours of reading this paper that the idea of the volume function presented itself and it is with much gratitude that Caviness and Pinkert's idea is acknowledged.

## Acknowledgments

Thanks to my supervisor James Davenport, to Geoff Smith and to Dan Richardson all of whom have made invaluable contributions to this work. The author would also like to thank the SERC's ACME Directorate and the IBM UK Scientific Centre, who jointly funded the Computer Algebra and Solid Modeling project in the Schools of Mathematical Sciences and Mechanical Engineering at Bath.

## References

- B. Buchberger (1985a), "A survey on the method of Gröbner bases for solving problems in connection with systems of multi-variate polynomials", *Proc. 2nd RIKEN Symp. Symbolic and Algebraic Computation*, N. Inada and T. Soma, eds., World Scientific Publication, 69-83.
- B. Buchberger (1985b), "Gröbner bases, an algorithmic method in polynomial ideal theory", *Recent Trends in Multi-Dimensional System Theory*, N.K. Bose, ed., Reidel, NY, 184-232.
- J. Canny (1990), "Generalized characteristic polynomials", *J. Symbolic Computation*, 9, 241-250.
- G.E. Collins (1967), "Subresultants and reduced polynomial remainder sequences", *J. ACM*, 14, 128-142.
- C. Hermite (1852), "Sur l'extension du théorème de M.Sturm a un système d'équations simultanées", *Comptes Rendus des Séances de l'Académie des Sciences*, tome XXXV.
- C. Hermite (1853), "Remarques sur le théorème De M.Sturm", *Comptes Rendus des Séances de l'Académie des Sciences*, tome XXXVI.
- C. Hermite (1880), "Sur l'extension du théorème de M.Sturm a un système d'équations simultanées", *Œuvres de Charles Hermite*, tome III, Mémoire inédit.
- F.S. Macaulay (1903), "Some formulae in elimination", *Proc. London Math. Soc.*, 33(1), 3-27.
- P.S. Milne (1990), *On the Algorithms and Implementation of a Geometric Algebra System*, Ph.D. Thesis, Technical Report 90-40, School of Math. Sci., University of Bath, Claverton Down.
- P. Pedersen (1991a), *Counting Real Zeros*, Ph.D. Thesis, Technical Report, Courant Institute, New York University, NY.
- P. Pedersen (1991b), "Counting real zeros", *Proc. Conf. Algebraic Algorithms and Error Correcting Codes*, Lecture Notes in Comput. Sci. 539, Springer-Verlag, NY, 318-332.
- J.R. Pinkert (1976), "Finding the roots of a complex polynomial", *ACM Tran. Math. Software*, 2, 351-363.
- J. Renegar (1989), *On the Computational Complexity and Geometry of the First-Order Theory of the Reals*, parts I-III, Technical Reports 853, 854, 856, Cornell School of Oper. Res. and Industrial Eng., Ithaca, NY.
- J.J. Sylvester (1852), "On a theory of the syzygetical relations ...", *Sylvester's Collected Works*, article 57, Cambridge University Press, Cambridge. Also appeared in *Phil. Trans. Royal Soc.*, 1853, 407-576.
- B.L. van der Waerden (1950), *Modern Algebra*, F. Ungar Publishing Co., NY.