Computer Science Department

TECHNICAL REPORT

On the 'Piano Movers' Problem
II. General Techniques for Computing
Topological Properties of Real
Algebraic Manifolds

BY

Jacob T. Schwartz

AND

Micha Sharir
February 1982 #41

NEW YORK UNIVERSITY

On the 'Piano Movers' Problem
II. General Techniques for Computing
Topological Properties of Real
Algebraic Manifolds

BY

Jacob T. Schwartz

AND

Micha Sharir

February 1982 $* 41$
Report No. 41

$C.1$

On the 'Piano Movers' Problem

II. General Techniques for Computing Topological Properties of Real

Algebraic Manifolds [*]

Jacob T. Schwartz
Computer Science Department
Courant Institute of Mathematical Sciences

and

Micha Sharir
Department of Mathematical Sciences
Tel Aviv University

ABSTRACT: This paper continues the discussion, begun in [SS], of the following problem, which arises in robotics: Given a collection of bodies B, which may be hinged, i.e. may allow internal motion around various joints, and given a region bounded by a collection of polyhedral or other simple walls, decide whether or not there exists a continuous motion connecting two given positions and orientations of the whole collection of bodies. We show that this problem can be handled by appropriate refinements of methods introduced by Tarski and Collins, which lead to algorithms for this problem which are polynomial in the geometric complexity of the problem for each fixed number of degrees of freedom (but exponential in the number of degrees of freedom.) Our method, which is also related to a technique outlined by Reif, also gives a general (but not polynomial-time) procedure for calculating all of the homology groups of an arbitrary real algebraic variety. Various algorithmic issues concerning computations with algebraic numbers, which are required in the algorithms presented in this paper, are also reviewed.

0. Introduction

The 'Piano Movers' problem (see [Re], [LPW], [IKP], [Ud], [SS]) is that of finding a continuous motion which will take a given body or bodies B from a given initial position to a desired final position, but which is subject to certain geometric constraints during the motion. These constraints forbid the bodies to come in contact with certain obstacles or 'walls', or to collide with each other. These walls can be curved, and the full collection of walls is not required be connected (see Fig. 0.1). The problem that we set out to solve is: Given two configurations (i.e. positions and orientations of all subparts) of the bodies B in which none of these bodies touches any walls, and in which none of the bodies B collide, find a continuous wall- and collision-avoiding motion of all the B between these two configurations, or establish that no such motion exists.

This paper will present a general, though not very efficient, method for deciding on the existence of such a path (and for constructing such a path if it exists). Specifically, we will show that this problem can be handled by a variant of Tarski's famous algorithm [Ta] for deciding statements in the quantified elementary theory of real numbers. Our approach is related to that outlined in an interesting paper of Reif [Re], and makes essential use of technical devices introduced by Collins [Co] and reviewed by Arnon [Ar]. As we shall see, these techniques also allow explicit, constructive calculation of the homology groups of an arbitrary real algebraic variety. In particular, the connectivity of such a variety can be calculated easily.

The paper is organized as follows. In Section 1 we begin to formulate the general mover's problem in which we are interested, as an abstract computational problem in algebraic topology. The algebraic mechinery required to handle this problem is then developed in Section
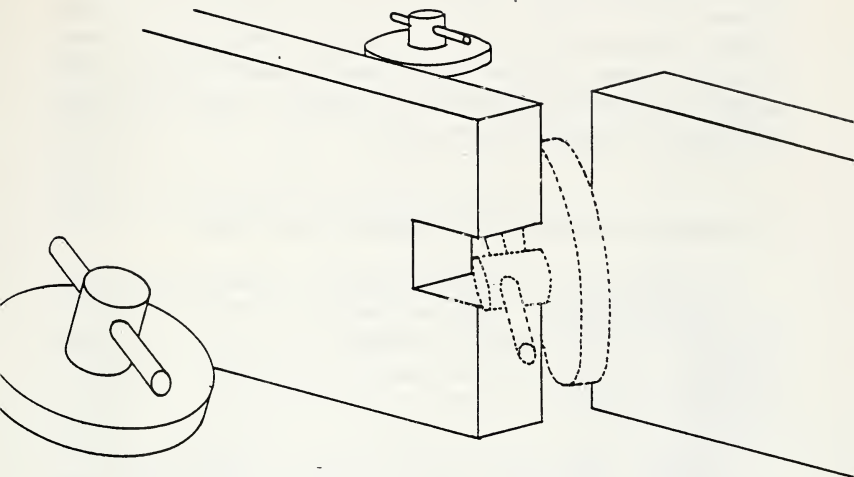
Fig. 0.1. <u>An instance of our case of the piano movers' problem.</u>
<u>The positions drawn in full are the initial and final positions</u>
<u>of B; the intermediate dotted positions describe a possible</u>
<u>motion of B between the initial and final positions.</u>

2. The results obtained in this section can be used to calculate more
general topological properties of algebraic varieties than those
required for the solution of the mover's problem, and Section 2 also
outlines some procedures for these calculations. Applications of the
general theory of Section 2 to the mover's problem are then given in
Section 3, yielding an algorithm which solves this problem in time
polynomial in the number of geometric constraints on the motion of the
body B, provided that the set of forbidden configurations for the body
B is closed.

For the sake of completeness and clarity, we include a simple
proof of the Collins construction on which our algorithms are based.
This is given in Appendix A. Appendix B reviews various efficient
techniques, some new, for exact calculations with algebraic numbers;

such calculations appear repeatedly in our algorithms. Finally, Appendix C gives technical details concerning the computations required to obtain the topological structure of the Collins decomposition. Note however that these, possible very expensive, computations are not required for the simpler task of determining the connectivity of the space of free configurations of the body B.

1. An Algebraic Formulation of the General Mover's Problem

In this section we reformulate the general motion-planning problem in abstract algebraic terms, and reduce it to the problem of decomposing certain algebraic varieties into their connected components. A solution to this abstract problem is then developed in subsequent sections.

Like Reif, whose work is to be presented more fully in a forthcoming paper, we study the space of all collision-free positions of one or more hinged bodies B. We assume each body B to consist of a finite number of rigid compact subparts $B_1, B_2, \ldots$, each bounded by various algebraic surfaces. These subparts can be connected to each other by various types of attachments, including the following:

(a) A point X on one part B_1 can be fastened to a point Y on another part B_2, in a manner which requires X and Y to be coincident but does not otherwise constrain the relative orientations of B_1 and B_2.

(b) The connection between X on B_1 and Y on B_2 can be a 'hinge', i.e. can constrain B_2 to revolve around an axis V fixed in the frame of B_1.

(c) The connection between B_1 and B_2 can permit B_2 to slide, or to slide and rotate, along an axis V fixed in the frame of B_1.

Various other forms of affixment might be envisaged and can be treated in much the same way that we will treat the more common fastenings (a-c). As noted above, we are willing to consider any number of disjoint hinged bodies of this kind, which are to move in a coordinated fashion throughout an empty space bounded by a finite collection of walls, which can themselves be arbitrary algebraic surfaces. We can regard the walls as part of the given system of bodies, with the additional properties that (a) they need not be compact; and (b) they are constrained not to move at all.)

Since we intend to proceed algebraically in what follows, our first task is to set up an appropriate algebraic parametrisation of a superspace of the set of all allowed positions of the hinged body B. It is convenient to proceed as follows. If $B = B_1$ is a single rigid body, we describe its position by giving a Euclidean motion T which takes B from some standard position to its given position. This transformation $Tx = Rx + X_0$ is defined by a pair $[X_0, R]$ consisting of a point X_0 in 3-dimensional Euclidean space E^3 and of a 3 by 3 rotation matrix R, and can therefore be regarded as a point in a smooth six-dimensional algebraic submanifold G of 12-dimensional Euclidean space E^{12}.

Next suppose that B is hinged, and that a second part B_2 of B is attached to B_1, say for the sake of definiteness in the manner (a). Then we can describe the overall position of the two parts B_1, B_2 of B as follows. As above, the position of B_1 is described by a Euclidean motion T which takes B_1 from a standard position to its actual position. By applying the inverse T^{-1} of T to both B_1 and B_2, we put B_1 into its standard position, and B_2 into a position which attaches a fixed one of its points X_2 to a point fixed on B_1. This position of B_2 is therefore defined by giving a Euclidean transformation T_2 such that $T_2 X_2 = X_2$. It is plain that the set of these transformations is in 1-1 correspondence with the set of rotation matrices R_2. Hence the overall position of B_1 and B_2 can be represented by a pair $[T, R_2]$, which once again varies over a smooth algebraic submanifold G_2 of a higher-dimensional Euclidean space E.

If instead B_2 is connected to B_1 in the manner (b), then much the same remarks apply, except that in this case the rotation matrix R_2 must satisfy $R_2 V = V$ for a certain 3-dimensional vector V. If B_2 can slide along an axis U fixed in B_1 but not rotate, its position is defined by a single real parameter u which defines the position of B_2 along this axis, etc. In all cases, the overall position of B_1 and B_2 is described by a pair $[T_1, T_2]$ of Euclidean motions, the first unconstrained, the second confined to some subgroup of the full Euclidean group. In all cases the allowed pairs form a smooth algebraic submanifold G_2 of some Euclidean space.

We can proceed similarly even if B consists of many parts hinged together in various ways. Suppose, for example, that B_2 is connected to B_1, and that a third part B_3 of B is connected to B_2. Then as above the overall position of B_1 and B_2 is defined by a pair $[T_1, T_2]$ of Euclidean transformations. T_1 maps B_1 from its standard to its actual position, and $T_1 T_2$ maps B_2 from its standard to its actual position. If we apply the inverse of $T_1 T_2$ to B_3, we put it into a position in which it is attached to a fixed point or axis of B_2 in one of the manners (a-c). Hence the actual position of B_3 is defined by a third Euclidean transformation T_3, belonging to a group of motions of one of the types we have already considered, and the mapping $T_1 T_2 T_3$ takes B_3 from its standard to its actual position.

These considerations make it clear that, irrespective of the manner in which the parts of a hinged body are connected together, the overall position of all its parts can always be defined by a point belonging to a smooth algebraic manifold G lying in a Euclidean space of some appropriate dimension.

Of course, the preceding considerations ignore all restrictions on the position of the parts of the bodies B imposed by the condition that none of these parts must collide. This point will be handled in Section 3 below, after the necessary algebraic machinery is introduced and developed in Section 2, which now follows.

2. Tarski Sentences and Sets; The Collins Decomposition

By a Tarski sentence we mean a sentence, possibly containing free variables, which can be formulated in the decidable quantified language studied by Tarski [Ta]. In this language, variables designate real numbers and are quantified over the set of all reals. The operators allowed in the language are +, −, *, and /, designating the usual real arithmetic operators. The allowed comparators are =, ≠, >, <, ≥, ≤, all of which have their standard meanings. In addition quantifiers and Boolean connectives are allowed.

A Tarski sentence $Q(x_1,...,x_n)$ containing the indicated n free variables and no others defines a subset of n-dimensional Euclidean space E^n, namely

(1) $\Sigma_Q = \{[x_1,...,x_n]: Q(x_1,...,x_n)\}$.

Sets of this form will be called Tarski sets (also known as semi-algebraic sets); the Q occuring in (1) is called the defining formula of Σ_Q. By a result given in the cited paper of Tarski, every Tarski set has a quantifier-free defining formula. A useful constructive proof of this result, involving a penetrating analysis of the geometric structure of Tarski sets, is given by Collins [Co] (see also Arnon [Ar]), and we will base our analysis of the general mover's problem on Collins' results, which substantially improve Tarski's earlier work. We will also prove certain topological properties of the 'cells' appearing in Collins' work, and will use these topological improvements to show that many standard topological properties (i.e. all the homology groups) of any algebraic variety are effectively calculable.

Using slightly different terminology than ours, Collins gives the following definitions and theorems:

Definition 1: For any subset X of Euclidean space, a decomposition of X is a finite collection K of disjoint connected subsets Y of X whose

union is X. Such a decomposition is a <u>Tarski decomposition</u> if each such subset Y is a Tarski set.

In what follows, E^r will denote the Euclidean space of r dimensions.

<u>Definition 2</u>: A <u>cylindrical algebraic decomposition</u> of E^r is defined as follows. For r=1 such a decomposition is just a partitioning of E^1 into a finite set of algebraic numbers and into the finite and infinite open intervals bounded by these numbers. For $r > 1$, a cylindrical algebraic decomposition of E^r is a decomposition K obtained recursively from some cylindrical algebraic decomposition K' of E^{r-1} as follows. Regard E^r as the Cartesian product of E^{r-1} and E^1, and accordingly represent each point p of E^r as a pair [x,y] with $x \in E^{r-1}$ and $y \in E^1$. Then K must be defined in terms of K' and an auxiliary polynomial $P = P(x,y)$ with rational coefficients, in the following way.

(i) For each $c \in K'$, let $c \times E^1$ designate the <u>cylinder over</u> c, i.e. the set of all [x,y] such that $x \in c$.

(ii) For each $c \in K'$ there must exist an integer n, such that for each $x \in c$ there are exactly n distinct real roots $f_1(x),...,f_n(x)$ of $P(x,y)$ (regarded as a polynomial in y), and hence these roots must vary continuously with x. We suppose in what follows that these roots have been enumerated in ascending order. Then each one of the cells of K which intersects $c \times E^1$ must have one of the following forms:

(ii.a) $\{[x,y]: x \in c, y < f_1(x)\}$

(lower semi-infinite 'segment' of $c \times E^1$).

(ii.b) $\{[x,f_i(x)]: x \in c\}$

('section' of $c \times E^1$)

(ii.c) $\{[x,y]: x \in c, f_i(x) < y < f_{i+1}(x)\}$

('segment' of $c \times E^1$)

(ii.d) $\{[x,y]: x \in c, f_n(x) < y\}$

(upper semi-infinite 'segment' of $c \times E^1$).

All these cells are said to have c as their <u>base cell</u> in K'; K' is said to be the <u>base decomposition</u>, and P the <u>base polynomial</u>, of K. It is convenient to put $f_0(x) = -\infty$ and $f_{n+1}(x) = +\infty$, and then to designate the cells (ii.a), (ii.b), (ii.c), and (ii.d) as c_0^*, c_i, c_i^*, and c_n^* respectively.

It obviously follows by induction that each of the sets constituting a cylindrical algebraic decomposition K of E^r is topologically equivalent to an open cell of some dimension $k < r$. We will therefore refer to the elements $c \in K$ as the (open) <u>Collins cells</u> of the decomposition K.

<u>Definition 3</u>: Let S be a set of functions of r variables, and K a cylindrical algebraic decomposition of E^r. Then K is said to be <u>S-invariant</u> if, for each c in K and each f in S, one of the following conditions holds uniformly for $x \in c$: either

(a) $f(x) = 0$ for all $x \in c$; or
(b) $f(x) < 0$ for all $x \in c$; or
(c) $f(x) > 0$ for all $x \in c$.

<u>Definition 4</u>: A point $p \in E^r$ is <u>algebraic</u> if each of its coordinates is a real algebraic number. A <u>defining polynomial</u> for p is a polynomial with rational coefficients whose set of roots includes all the coordinates of p.

<u>Theorem 1</u>: (Collins) Given any finite set S of polynomials with rational coefficients in r variables, we can effectively construct an S-invariant cylindrical algebraic decomposition K of E^r into Tarski sets such that each $c \in K$ contains an algebraic point. Moreover, defining polynomials for all these algebraic points, and quantifier-free defining formulae for each of the sets $c \in K$, can also be constructed effectively.

The proof of Collins' Theorem, which is not difficult, will be reviewed in Appendix A below.

In what follows we will find it useful to sharpen Collins' results in certain topological respects. For this, we have to impose an additional requirement on the decomposition; in certain unfavorable orientations of E^r, this condition can be false. However, as we shall show below, one can always restore this extra property by an appropriate rotation of the r-Euclidean space, and such a rotation can be easily calculated. This requirement is stated in the following

<u>Definition 5:</u> A Collins decomposition K is said to be <u>well-based</u> if the following condition holds. Let K' be the base decomposition and $P(b,x)$ the base polynomial of K. Then we require that $P(b,x)$ should not be identically zero for any $b \in E^{r-1}$. Moreover, we require that this same condition apply recursively to the base decomposition K'.

<u>Example:</u> Consider the polynomial

$$P(x,y,z) = (x^2+y^2)z + (x^2-y^2)$$

in 3-dimensional space E^3. Then, since $P(0,0,z)$ vanishes identically, no P-invariant Collins decomposition of E^3 whose final step projects E^3 onto E^2 in the z-direction is well-based.

For the correctness of the topological assertions that we are about to make, it is essential that the decomposition K be well-based. Later we shall see how to rotate the given Euclidean space so as to make the decomposition well-based. For the moment we assume a well-based decomposition K, and show that it has certain useful topological properties.

<u>Lemma 1:</u> Let $c \in K$. Then the closure of c is a union of cells of K.

<u>Proof:</u> To establish the lemma, we will use induction on the dimension r, and prove the following stronger

<u>Claim</u>: Let $r > 1$, and let K be a Collins decomposition of E^r. Then the closure of each cell in K is a union of cells of K. Moreover, for each cell c of K, each point z in the boundary of c, and each $\epsilon > 0$, the open ϵ-ball about z contains a (relative) neighborhood U of z in $c+\{z\}$ such that $U-\{z\}$ is a connected subset of c.

<u>Proof</u> <u>of</u> <u>Claim</u>: Our claim holds trivially for $r = 1$. Assume $r > 1$. With no loss of generality assume that c has either the form

(1) $\qquad c = c_j'^{*} = \{ [b,x] : b \in c', f_j(b) < x < f_{j+1}(b) \}$,

or

(2) $\qquad c = c_j' = \{ [b,f_j(b)] : b \in c' \}$,

for some $c' \in K'$ = the base decomposition of K. In either case the closure of c obviously contains c itself, and, in case (1), also the two 'sections' c_j' and c_{j+1}' bounding c from above and from below, all of which are of course cells in K. Any other point in the closure of c must be of the form $[b,x]$, where b belongs to the boundary of c', and where $x_1(b) < x < x_2(b)$; here we write

$$x_1(b) = \lim_{b' \in c', \ b' \to b} \inf f_j(b') \quad \text{and} \quad x_2(b) = \lim_{b' \in c', \ b' \to b} \sup f_k(b')$$

where $k = j+1$ in case (1), $k = j$ in case (2). Conversely, any such point clearly belongs to the closure of c. Let e be the set of all such points. By our induction hypothesis, the boundary of c' is a union of cells in K'. Let $c'' \in K$ be a cell of this boundary. Note that $x_1(b)$ and $x_2(b)$ are either roots of P, or are $-\infty$ or $+\infty$. We will show that these functions are (equal if $k = j$, and) continuous in the entire closure of c, and hence are continuous for $b \in c''$ (here we give the <u>extended</u> <u>real</u> <u>axis</u> R^*, including the points $-\infty$, $+\infty$, its standard topology, which makes R^* homeomorphic to the compact unit interval $[-1,+1]$.)

The asserted continuity is an easy consequence of the following auxiliary lemma.

<u>Lemma</u> 2: Let $c' \in K'$ be as above, and let $f_j(b)$ denote the j-th root of $P(b, \cdot)$ for $b \in c'$. Then there exists a unique continuous extension (in the sense of the extended topology of R^*) of f_j to the entire closure of the cell c'. Moreover, this extension, which we will also denote by f_j, is either infinite or else a root of P.

It will be shown below that Lemma 2 follows inductively from our other assertions. Assume for the moment that this has been proved. It then follows that the set

$$K(c'') = \{ [b,x] : b \in c'', x_1(b) < x < x_2(b) \}$$

must be a union of cells in K. We can show this, and even provide a more explicit characterization of these cells, as follows. Let $c' \in K'$ be as above, and let $f_j(b)$ be the j-th root (of m distinct real roots) of $P(b, \cdot)$ over c'. Let $c'' \in K'$ be any cell in the boundary of c'. Use the same symbol $f_j(b)$ to denote the continuous extension of f_j to $b \in c''$ which exists by Lemma 2. Then it follows from Lemma 2 that, for $b \in c''$, $f_j(b)$ is (either $-\infty$ or $+\infty$ or) one of the M roots of P over c''. In either case we can write $f_j(b) = F_J(b)$, where F_J is the J-th root of P over c'' (by our convention, this also includes the extreme cases J = 0,M+1); note that J is independent of b, because the roots F_J are isolated and f_j varies continuously over c''. Define a mapping

$$\rho(c',c'') : \{0..m+1\} \to \{0..M+1\}$$

by putting $\rho(c',c'')(j) = J$ if $f_j(b) = F_J(b)$ for some, hence for all b $\in c''$. Then the assertion we need to prove is contained in the following somewhat more detailed

<u>Lemma</u> <u>3:</u> Let $c' \in K'$ be as above. Then the cells of K which intersect the closure of the cell c_j' are c_j' itself, and all cells of the form c_J'', where c'' is contained in the closure of c' and $\rho(c',c'')(j) = J$.

The cells of K which intersect the closure of $c_j'^*$ are $c_j'^*$ itself, the two sections c_j' and c_{j+1}', and all cells of the form c_J'', $c_J''^*$, c_{J+1}'', $c_{J+1}''^*$,..., c_L'', where c'' is contained in the closure of c', $\rho(c',c'')(j) = J$, and $\rho(c',c'')(j+1) = L$. Moreover, all of these cells are contained in the closure of c_j' (resp. $c_j'^*$).

Proof: We prove only the second assertion, the first being even simpler. As noted in the paragraph immediately preceding the statement of Lemma 2, the closure of $c_j'^*$ consists of $c_j'^*$, c_j', c_{j+1}', and of the union of all the sets $K(c'')$ where c'' ranges over all the cells contained in the boundary of c'. By Lemma 2, and by the remarks preceding the present lemma, we have $x_1(b) = F_J(b)$, $x_2(b) = F_L(b)$, for $b \in c''$, where J, L are as in the statement of the present lemma. Thus $K(c'')$ is the union of all the cells c_J'',...,c_L''. It is then clear that all these cells, together with c_j' and c_{j+1}', are the only cells which can intersect (and hence be contained in) the closure of $c_j'^*$. Q.E.D.

We can now complete the proof of Lemma 1. Indeed, the first part of the claim to be proved is now immediate from Lemma 3. To prove the second part, let c, c', [b,x] $\in c'$, and let ε be as in the claim. Let d, d', be the base cells of c, c' respectively. Then either $c = d_j$ for some j or $c = d_j^*$ for some j.

Assume first that $c = d_j$, for some j. Project the ε-ball U about [b,x] onto a subset U' of E^{r-1}. By inductive hypothesis, U' contains a relative neighborhood V' of b in d+{x} such that V-{x} is connected. The continuity of $f_j(b)$ implies that for sufficiently small V' the set $\{[b,f_j(b)] : b \in V'\}$ is a connected neighborhood of [b,x] in c+{[b,x]} which is contained in U.

Next consider the case where $c=d_j^*$. Let J,L be as in Lemma 3, and let $0 < \alpha < 1$ such that $x = \alpha F_J(b) + (1-\alpha)F_L(b)$. Obtain U', V' as in the preceding paragraph, and consider the set

$$V = \{ [a, \beta f_j(a)+(1-\beta)f_{j+1}(a)] : a \in V', \ 0 < \beta < 1, \ |\beta - \alpha| < \delta \}$$

Note that V is relatively open in c. It also follows from the (uniform) continuity of f_j and f_{j+1} over V' that if V', δ are both sufficiently small, V will be a connected (relative) neighborhood of [b,x] in c+{[b,x]} which is contained in U.

Thus both assertions of our claim continue to hold in r dimensions, completing the inductive proof of Lemma 1. Q.E.D.

Next we return to finish the proof of Lemma 2, which will complete our whole interlocking set of inductions.

<u>Proof of Lemma 2:</u> It is sufficient to prove that f_j is uniformly continuous on any bounded subset of c'. Suppose the contrary. Then there exist $\delta > 0$ and two sequences b_n, $b_n' \in c'$, such that both sequences converge to a point b_0 on the boundary of c', but for all n we have $|f_j(b_n) - f_j(b_n')| > \delta$. By inductive hypothesis, for each integer m there exists a (relative) neighborhood N_m of b_0 in c'+{b_0} which is contained in the (1/m)-ball about b_0 in E^{r-1}, and such that N_m-{b_0} is connected. For all m > 1 define sets

$$R_m = \{ [b,f_j(b)] : b \in N_m-\{b_0\} \}.$$

and let R denote the intersection of the closures of these sets. R is obviously a compact set. We claim that it is connected. To see this we first observe that each of the sets R_m is connected, being the image of the connected set N_m-{b_0} under the continuous function $b \to$ [$b,f_j(b)$]. Next suppose that R is not connected. Then there exist two disjoint open sets U, V such that R is contained in the union of U and V and intersects both U,V. But this implies that for sufficiently large m R_m is also contained in the union of U and V, for otherwise there would exist a sequence z_m, such that $z_m \in R_m - U - V$ for all m, which, by compactness, must converge to some $z \in R - U - V$, which is impossible. Since R_m is connected, we can assume with no loss of generality that R_m is contained in U for all sufficiently large m. But

then R must also be contained in U, a contradiction which proves that R is connected. Since the projection of R into E^{r-1} consists of the single point b_0, R must be a vertical straight segment, all of whose points are obviously roots of $P(b_0, \cdot)$. However, since we have assumed that K is well-based, there are only finitely many such roots, and so R must consist of a single point $[b_0, x]$. Hence this point is the common limit of both sequences $\{[b_n, f_j(b_n)]\}$, $\{[b_n', f_j(b_n')]\}$, which is impossible. This shows that f_j is uniformly continuous over bounded subsets of c', and hence admits a unique continuous extension to the entire closure of c'. Q.E.D.

Remark: Lemma 2 is false if the decomposition is not well-based. Indeed consider the example given earlier. It is easily seen that $\{[0,0]\}$ is a base-cell. Hence, if c is a 2-dimensional base cell c containing $[0,0]$ in its boundary, the single zero of $P(x,y,\cdot)$ for $[x,y]$ ϵ c can never admit a continuous extension to $[0,0]$.

Lemma 2 has several consequences, which collectively show that a well-based Collins decomposition is free of local pathology.

Corollary: Assume K is well-based. Let c ϵ K', and let c' be a boundary cell of c. let $g_k(b)$ denote the k-th root of $P(b, \cdot)$ for b ϵ c'. For each b ϵ c' let $J(b)$ denote the set of all indices j such that $g_k(b)$ is the limit of some sequence $f_j(b_n')$, as b_n' ϵ c approaches b. Then $J(b)$ is a constant depending only on c', but not on the particular point b ϵ c'.

Lemma 4: Let c ϵ K, and let c' ϵ K be a boundary cell of c. Then for each point z ϵ c and each point z' ϵ c' there exists a continuous path connecting z to z', which, except for its endpoint z', lies wholly in c.

Proof: Proceed by induction on r. Let d, d' be the base cells of c, c' respectively. Let $z = [b, x]$, $z' = [b', x']$, where b ϵ d, b' ϵ d'. If d = d', then c = d_j^* for some j, and c' is its top or bottom face. Assume c' to be the bottom face d_j of c; then the desired curve is

constructed as follows. Let $q'(t)$ be a curve contained in d which connects b with b' (the recursive construction of Collins cells makes it quite easy to construct such a curve explicitly). Next let $0 < \alpha < 1$ be such that

$$x = \alpha f_j(b) + (1-\alpha)f_{j+1}(b).$$

Then the desired curve is simply

$$q(t) = \alpha(1-t)f_j(q'(t)) + [1-\alpha(1-t)]f_{j+1}(q'(t)), \quad 0 < t < 1$$

Otherwise, $d \neq d'$, and by inductive hypothesis there exists a continuous curve $q(t)$, $t \in [0,1]$ such that $q(0) = b$, $q(1) = b'$, and $q(t) \in d$ for all $t < 1$. First suppose that c' is a section

$$\{ [a, g_k(a)] : a \in d' \},$$

where $g_k(a)$ is the k-th root of $P(a, \cdot)$ over d', and that c is a section

$$\{ [a, f_j(a)] : a \in d \},$$

where f_j is the j-th root of $P(a, \cdot)$ over d. Extending f_j continuously to the whole closure of d, so that $f_j(a) = g_k(a)$ for $a \in d'$, the curve we want is simply

$$p(t) = [q(t), f_j(q(t))] , \quad t \in [0,1].$$

The other possible cases, i.e. those in which one or both of c and c' is a 'segment' rather than a section, can be handled in essentially the same manner; we leave details to the reader. Q.E.D.

Next we quote some standard definitions and results concerning finite cell complexes and their (singular) homology groups; see Cooke and Finney [CF].

Definition 6: A decomposition of a compact topological space S into finitely many disjoint sets $\{c_i\}$ is called a <u>cell complex</u> if

(a) Each c_i is homeomorphic to an open unit ball of some dimension d_i, which we call the dimension of c_i.

(b) For each integer dimension d, the union S_d of all the sets c_i of dimension $\leq$ d is closed.

(c) Each cell c_i of dimension d_i is open in the relative topology of S_{d_i}.

(d) For each cell c_i of dimension d_i, there exists a continuous mapping f_i of the unit closed ball B of dimension d_i onto the closure c_i^- of c_i, which maps the interior of B homeomorphically onto c_i.

Definition 7: The cell complex $\{c_i\}$ of the preceding definition is said to be <u>regular</u> if each of the mappings f_i is a homeomorphism of the closed ball B of dimension d_i onto the closure of the corresponding cell c_i.

Theorem 2: The cells of a (well-based) Collins decomposition of E^r form a regular cell complex.

We will, as usual, prove this theorem by induction on the dimension r, for which purpose the following easy lemma will be useful.

Lemma 5: Let B be the closed unit ball in E^r, and let f and g be two continuous real functions on B such that $g(b) < f(b)$ for each b in the interior of B. Put

$$B^* = \{[b,x] : b \ \varepsilon \ B, \ g(b) \leq x \leq f(b) \ \}.$$

Then B^* is homeomorphic to the closed unit ball B_1 of E^{r+1}.

Proof: By shifting and contracting B^* along the x-axis we can assume

that $g(b) = -f(b)$, and that $f(b) < 1/2$ for all $b \ \varepsilon \ B$. For each $b \neq 0$ in B, let $\theta(b) = b/|b|$, so that θ is continuous for all $b \neq 0$. Put $r(b) = (1-f^2(\theta(b)))^{1/2}$, and then let T map the point $[b,x]$ of B^* to

$$[br(b), \ x(1-|b|^2r^2(b))^{1/2}/f(b)]$$

(If $|b| = 1$ and $f(b) = 0$, then put $T(b,x) = [b,0]$.) Since $br(b)$ is continuous for all b, and since $f(b) \neq 0$ for $|b| < 1$, this mapping is plainly continuous for all $[b,x] \ \varepsilon \ B^*$ such that $|b| < 1$. T is also continuous when $|b| = 1$. This is plain for each b such that $|b| = 1$ and $f(b) > 0$; on the other hand, if $f(b) = 0$, then since $|x'/f(b')| \leqslant 1$ for $[b',x'] \ \varepsilon \ B^*$, it follows that $T[b',x'] \rightarrow [b,0]$ as $[b',x'] \rightarrow [b,0]$ from within B^*. Moreover, T is 1-1 on B^*. Indeed, if $T[b,x] = T[b',x']$ then plainly $\theta(b) = \theta(b')$, so that $r(b) = r(b')$, and hence $b = b'$, $x = x'$. Since T is continuous and 1-1 on the compact set B^*, it is a homeomorphism on B^*.

It is easily seen that the range $T(B^*)$ is the set B^+ of points $[a,y]$ in the closed unit ball $|a|^2 + y^2 \leqslant 1$ which satisfy the condition $|a| \leqslant r(a)$. The boundary of B^+ then consists of all points $[a,y]$ which either lie on the boundary of the closed unit ball B_1 of E^{r+1} and satisfy $|a| < r(a)$, or else are interior points of B_1 such that $|a| = r(a)$. It then follows that the boundary of B^+ has the property that each ray from the origin meets it in exactly one point. Hence B^+ is a star-shaped compact set, and it follows from Lemma 6 below that B^+, and hence also B^*, is homeomorphic to the closed unit ball B_1. Q.E.D.

Lemma 6: Let A be a compact set in E^r which contains the origin 0 of E^r in its interior, and which is star-shaped relative to 0 (i.e. each ray from 0 intersects the boundary A' of A in exactly one point.) Then A is homeomorphic to the closed unit ball B of E^r.

Proof: This is well known, but to prove it let θ designate an arbitrary point on the boundary of B, and let $r(\theta)$ be the length of the straight ray from 0 to A'. Then, since each such ray intersects A' in exactly one point, $r(\theta)$ obviously varies continuously with θ. The desired

homeomorphism simply maps each nonzero $b \in B$ to $br(\theta(b))$, where $\theta(b) = b/|b|$. Q.E.D.

We can now give the proof of Theorem 2: Suppose that Theorem 2 has been proven for a Collins decomposition K' of E^r, and let K be a Collins decomposition of E^{r+1} with base decomposition K'. Then any cell of K has either the form c_j or c_j^* for some $c \in K'$, and for some integer j. In the case of a section cell $c_j = \{[b, f_j(b)] : b \in c\}$, we can simply note that, since f_j is continuous on the closure c^- of c, the projection of c_j to c extends to a homeomorphism of the closure of c_j with c^-, which, by inductive assumption, is homeomorphic to a closed unit ball B of an appropriate dimension. Finally, using the same homeomorphism of c^- with B, the case of a segment cell c_j^* is covered by Lemma 5. Q.E.D.

Definition 8: If c and c' are Collins cells belonging to a decomposition K, and if c' is contained in the closure of c, then we say that c' is a face of c.

We continue our analysis by quoting another standard definition from Cooke and Finney.

Definition 8: Given the finite regular cell complex K, an incidence function α on K is a function assigning one of the integers $\{-1, 0, +1\}$ to each pair c, c' of cells of K, which satisfies the following conditions.

(i) $\alpha(c, c') \neq 0$ iff c' belongs to the boundary of c and has dimension exactly one less than the dimension of c.

(ii) If c is of dimension 1, and the two 0-dimensional cells (i.e. discrete points) constituting the endpoints of c are c_1 and c_2, we have $\alpha(c, c_1) + \alpha(c, c_2) = 0$.

(iii) If c'' belongs to the boundary of a cell c of dimension d, and has dimension exactly $d-2$, then

$$\sum \alpha(c,c')\alpha(c',c'') = 0,$$

where the sum extends over all cells c' of dimension d-1 whose closures are subsets of c^- and supersets of the closure of c''. (It is well known (see e.g. Cooke and Finney) that for any c, c'' there are precisely two cells c' of this kind, if the cell complex is regular.)

Every regular cell complex A admits an incidence function, and by a standard result proved at length in the cited work of Cooke and Finney, any such incidence function can be used to compute the homology groups of A in a purely combinatorial manner. For the Collins cell decompositions which we consider, incidence functions are easy to define; we simply proceed inductively on the dimension, and use a variant of the standard 'Cartesian product rule'. More specifically, suppose that K and K' are as in the proof of Theorem 2 above, and that an incidence function α' has already been defined for the cells of K'. Extend this to the cells of K by putting

(a) $\alpha(c_j, c_J') = \alpha'(c,c')$, if c' belongs to the boundary of c and $\rho(c,c')(j) = J$;

(b) $\alpha(c_j^*, c_j) = -1$, $\alpha(c_j^*, c_{j+1}) = +1$;

(c) $\alpha(c_j^*, c_k'^*) = -\alpha'(c,c')$, if c' belongs to the boundary of c and $\rho(c,c')(j) < k < \rho(c,c')(j+1)$;

and putting $\alpha(a,b) = 0$ in all other cases. It follows immediately from Lemma 3 that the function α defined this way satisfies condition (i) of Definition 8. It is trivial to verify that α also satisfies condition (ii) of that definition.

Next we verify that α satisfies condition (iii) of Definition 8. Let $c \in K$ be an s-dimensional cell, and let c'' be an (s-2)-dimensional cell contained in the boundary of c. Let d,d'' be the base cells of c, c'' respectively. Several cases are possible:

(i) Suppose $c = d_j$ for some j. Then it follows from Lemma 3 that c'' must have the form d_j''. Furthermore any (s-1)-dimensional face c' of c which contains c'' in its boundary must be of the form d_L', for some (s-1)-dimensional face $d' \in K$ of d which contains d'' in its boundary and is such that $\rho(d,d')(j) = L$ and $\rho(d',d'')(L) = J$. Then, by (a) above and by induction hypothesis, we have

$$\sum_{c'} \alpha(c,c')\alpha(c',c'') = \sum_{d'} \alpha'(d,d')\alpha'(d',d'') = 0$$

(ii) Next suppose that $c = d_j^*$. Then d must be (s-1)-dimensional. Once more Lemma 3 implies that c'' either is an (s-2)-dimensional face of d_j or d_{j+1} (both of which are (s-1)-dimensional), or c'' is d_k'' where d'' is (s-2)-dimensional, or c'' is $d_k''^*$ where d'' is (s-3)-dimensional. The first case is a special case of the second one, so we begin by assuming that the second case holds. Then it follows that d'' is a face of d, and that $\rho(d,d'')(j) = J < k < L = \rho(d,d'')(j+1)$. Assume first that k lies strictly between J and L; then the only possible cells c' that 'come between' c and c'' are $d_{k-1}''^*$ and $d_k''^*$. Hence, in this case, clause (b) of our definition of the incidence function α implies

$$\sum_{c'} \alpha(c,c')\alpha(c',c'') =$$
$$\alpha(d_j^*,d_{k-1}''^*)\alpha(d_{k-1}''^*,d_k'') + \alpha(d_j^*,d_k''^*)\alpha(d_k''^*,d_k'')$$

$$= -\alpha'(d,d'')\cdot(-1 + 1) = 0$$

Next suppose that $k = J < L$. Then the possible c' are the lower face d_j of c and the cell $d_J''^*$, and so we have

$$\sum_{c'} \alpha(c,c')\alpha(c',c'') =$$
$$\alpha(d_j^*,d_j)\alpha(d_j,d_J'') + \alpha(d_j^*,d_J{}^*)\alpha(d_J''^*,d_J'')$$

$$= (-1)\cdot\alpha'(d,d'') + (-\alpha'(d,d''))\cdot(-1) = 0.$$

A similar analysis covers the case $J < L = k$. Suppose finally that $k = J = L$. Then the possible c' are just the upper and lower faces d_j and d_{j+1} of c, and once again it is easy to verify (iii).

Next consider the case in which d'' is an $(s-3)$-dimensional cell, and $c'' = d_k''^*$. It is easily seen that in this case the only possible intermediate cells c' are of the form $d_\ell'^*$, where d' is a cell in K' intermediate between d and d'', and

$$\rho(d,d')(j) = J1 < \ell < J2 = \rho(d,d')(j+1);$$

$$\rho(d',d'')(\ell) = L_1 < k < L_2 = \rho(d',d'')(\ell+1).$$

However, it readily follows from Lemma 3 that for each intermediate cell d' there exists exactly one ℓ satisfying these inequalities. Hence, by inductive hypothesis, we have

$$\sum_{c'} \alpha(c,c')\alpha(c',c'') =$$
$$\sum_{d'} (-\alpha'(d,d'))(-\alpha'(d',d'')) = 0$$

Thus condition (iii) is established in all cases, so α is indeed an incidence function for K.

Since an incidence function for the cells of a Collins decomposition can be defined in this straightforward combinatorial manner, and since the homology groups of any finite regular cell complex can be computed combinatorially from such an incidence function, it follows that the homology groups can be computed by a purely finite procedure once we know what $(d-1)$-cells are faces of each given d-cell. A technique for determining this, which is based on Lemma 3 and its corollary, will be described below. Assuming this, we have the following result:

Theorem 3: For each j, the (singular) homology group $H_j(V)$ of the real

algebraic variety V defined by any set Π of polynomials equations $P(x_1,\ldots,x_n) = 0$ with rational coefficients, can be computed in a purely rational manner from the coefficients of the polynomials P.

In particular, since V is a regular cell complex the number of connected components of V can be formed simply by tracing sequences of cells which are faces of each other.

Collins gives estimates (which we reconstruct in Appendix A below) for the complexity of his cell decomposition procedure; in what follows we shall extend these to estimates of the work needed to determine whether one Collins cell c' is a face of another cell c, in the special case in which the dimension of c is the same as the dimension of the space E^r being decomposed. This will show that the connectivity analysis required to solve the general movers' problem can be performed in time polynomial in the total degree of the set Π of polynomials (but exponential in the number r of variables appearing in these polynomials.) Appendix C discusses the more difficult adjacency analysis in case c has a lower dimension than r. At the present moment we do not know whether this more complex task can also be handled also in polynomial time.

To complete the foregoing arguments we still need to show how a well-based Collins decomposition can be defined for any algebraic variety. Let P be an r-variate polynomial, for which we wish to construct a P-invariant Collins decomposition of E^r. Following Hironaka [Hi], we define a <u>good direction</u> to be a unit vector v (i.e. a point in the unit sphere S of E^r) such that P does not vanish identically on any line parallel to v. A simple technique for constructing a good direction, which is given by Lazard [La], is as follows. Let $Q(x_1,\ldots,x_r)$ be the homogeneous part of P of highest degree $= n$. If v is a direction on which Q does not vanish (i.e. $Q(v) \neq 0$), then v is a good direction for P. Indeed, for any $x \in E^r$ and for sufficiently large real t, $P(x+tv)$ behaves asymptotically as $Q(tv) = t^n Q(v) \neq 0$. By an easy Lemma of Schwartz [Sc], if $c > 1$, then the number of integer points in a cube of side $I > cn$ at which Q vanishes

is at most $c^{-1}I^r$. Taking c substantially larger than 1, a good direction v can be found quite rapidly simply by picking a v at random from such a cube, testing the condition $Q(v) \neq 0$, and repeating the choice until a v satisfying this condition is found. Once having found such a v, we rotate E^r so that v becomes the r-th axis, and we begin the Collins construction by projecting in the direction of v. As will be seen in Appendix A, this recursive construction generates a new polynomial Q in the remaining (r-1) variables from P and v; Q plays exactly the same role for the required base decomposition of E^{r-1} that P plays for the decomposition of E^r. This observation allows us to apply the above way of finding a good direction recursively, and in this way we build a well-based Collins decomposition.

Remark: The result of Hironaka [Hi] which asserts the triangulability of real algebraic varieties, follows immediately from what has gone before.

Next we must tackle the problem of deciding when a Collins cell c' of dimension d-1 forms part of the boundary of a d-dimensional Collins cell c. In analysing this question, we shall first handle the relatively simple case in which c is of maximal dimension, i.e. is of the same dimension as the Euclidean space E^r which is being decomposed; this is all that is needed for the 'movers problem' proper (see Section 3 for details). As usual, we proceed by induction on the dimension r, i.e. suppose that the decomposition K of E^r being considered has the base decomposition K', and that for each (r-2)-dimensional cell c' of K' we know the two (r-1)-dimensional cells c_1 and c_2 of K' of whose boundary c' forms part. We will also suppose that for each cell c ε K' considered, an algebraic point p(c) belonging to c is known, and that whenever c' forms part of the boundary of c, a vector v pointing from p(c') into c, i.e. a vector v such that $p(c') + \delta v \ \epsilon \ c$ for all sufficiently small δ, is known. We will shortly note that it is easy to carry the construction of such vectors forward inductively.

Any r-dimensional cell in E^r must have the form c_j^* for some $(r-1)$-dimensional cell $c \in K'$, and then, by Lemma 3, the $(r-1)$-dimensional cells lying in the boundary of c_j^* are its top and bottom cells c_j and c_{j+1}, together with all cells of the form $c_{j'}^*$, where c' belongs to the boundary of c and where $\rho(c,c')(j) < J < \rho(c,c')(j+1)$. The vector $[0,\ldots,0,-1]$ (resp. $[0,\ldots,0,+1]$) points from c_{j+1} (resp. from c_j) into c_j^*. Moreover, if $v = [v_1,\ldots,v_{r-1}]$ points from an $x \in c'$ into c, then $[v_1,\ldots,v_{r-1},0]$ points from $c_{j'}^*$ into c_j^* if the former is part of the boundary of the latter. Hence it is trivial to carry the necessary vectors v forward, and the boundary determination problem presently under consideration reduces to that of calculating the map $\rho(c,c')$ for an $(r-1)$-dimensional cell c and an $(r-2)$-dimensional cell c' of K'.

To make this calculation, we begin by observing that the corollary to Lemma 2 implies that to compute ρ it suffices to compute, for each root $f_j(b)$ over c', the number k of roots of P into which $f_j(b)$ splits as b moves slightly into c. Moreover, we can make this calculation for an arbitrary $b \in c'$.

Given $b = p(c')$, and the vector v which points from b into c, the points $b' = b + tv$ will lie in c for sufficiently small positive values of t. Moreover, the bivariate polynomial

$$Q(x,y) = P(b+xv, \ f_j(b)+y)$$

has coefficients which are algebraic since they depend algebraically on the (algebraic) coordinates of b, and this polynomial vanishes at the origin $[0,0]$. Let x vary in a sufficiently small interval $(0,x_0]$, and for each such x define the polynomial

$$R_x(y) = Q(x,y).$$

As is well known (see e.g. Van der Waerden [Wa2], Chap. 1), for sufficiently small x all the roots y of $R_x(y)$ which lie near the origin are expressible as a fractional power series of the form

$$y = c_1 \, x^{1/k} + c_2 \, x^{2/k} + \ldots =$$

$$= c_1 \, x^{1/k} + o(x^{1/k}),$$

where $k \leqslant n =$ the degree of R_x as a polynomial in y. Thus, if we substitute x^{n+1} for x and if we let x be small enough, it follows that each root y of $R_x n+1(\cdot)$ lying near the origin must belong to the interval $[-x,x]$, because the sum of the above fractional power series is $O(x^{(n+1/k)}) = o(x)$. It therefore suffices to determine the number N of zeroes of $R = R_x n+1$ in the neighborhood $[-x,x]$ of 0. To find N, we can use the Sturm technique (see Appendix B for a review of Sturm sequences), that is, compute the Sturm sequence of R at x and at $-x$, and then N is $S(-x) - S(x)$, where $S(a)$ is the number of sign changes in that sequence evaluated at a. (The Sturm sequence is obtained as the remainder sequence of R and its y-derivative R'.) Since R and R' depend polynomially on the parameter x, extra care must be taken to ensure that, as this sequence is being calculated by repeated divisions, no leading coefficient of any of the polynomials in that sequence vanishes for the values of x that we consider (if such a term did vanish, subsequent divisions might involve completely different polynomials). However, since all these coefficients of R are polynomials in x, and since there are only finitely many such coefficients, none of their zeroes will belong to the open interval $(0,x_0]$, if x_0 is chosen to be small enough.

Therefore we can find the number N of roots in which we are interested by calculating the Sturm sequence

$$(3) \qquad F_x^{(0)}(y),\ldots, F_x^{(k)}(y),$$

of $R_x n+1$ as a sequence of polynomials in y with coefficients which are rational functions in x; and since we can multiply all those coefficients by a common denominator, we can even assume that these coefficients are polynomials in x. After calculating the polynomials (3), we must find their signs at the points $-x$ and x, for x small enough. If we substitute $y = x$ in the sequence (3), we obtain a

sequence of polynomials $G_0(x),\ldots, G_k(x)$ in x, the signs of whose members are to be computed for x positive and sufficiently close to 0; and an exactly similar statement holds if we substitute $-x$ for y. Let $G(x)$ be any one of these polynomials; then the sign of $G(x)$ is either the sign of the coefficient in G of the nonzero term of the smallest degree, or, if all the terms of G vanish, is 0. This gives us N, i.e. the number of roots into which $f_j(b)$ splits as b moves into c. As already observed, by collecting these numbers for all the roots $f_j(b)$ of P for $b \in c'$, and by using Lemma 3 in a straightforward manner, we can reconstruct the map $\rho(c,c')$.

Next The analysis required to determine when a Collins cell c' of dimension $k-1$ is a face of a cell c of dimension $k < r$ is somewhat more difficult, for which reasons we prefer to present it in Appendix C below.

3. A General Algorithm for the Path-Planning Problem

As noted previously, the position of one of the hinged bodies B we consider is always described by a point T in a smooth algebraic manifold G, and the set of points occupied by a particular rigid part B_i of B is the range on B_i of a Euclidean transformation T_i whose coefficients can be expressed as polynomials in the components of T. We assume that each part B_i of each of our bodies is a compact Tarski set, and moreover that each face, edge, or other significant feature of such a part B_i is also a compact Tarski set (the walls however are not assumed to be compact but merely closed Tarski sets.) The set F of forbidden configurations for the body B is then defined as follows. First we note that if B_i and B_j are two parts of B which are hinged together in any of the ways specified above, then there exist some face or other closed feature B_i' (resp. B_j') of B_i (resp. B_j) which always touch each other. Any other touch between B_i and B_j at a point outside B_i' or B_j' is then considered to be forbidden (in particular, if B_i and B_j are not directly hinged to each other, then it is forbidden for them to intersect at all.) This defines a set F_0 of points of G which

represent forbidden configurations of B, and we define F to be the closure of F_0. F clearly constitutes a closed Tarski set, and hence a real closed semi-algebraic subset of G. (All this continue to apply in cases involving several disconnected, independently moving bodies, some of which can represent (immobile) walls.) Our path-planning problem is therefore that of deciding whether or not two points of G-F belong to the same component of G-F, where F is a real closed semi-algebraic subset of G, and, if so, construct a path that connects these points in G-F. (<u>Note</u>: If F is not required to be closed, the problem is still decidable, but the technique we are about to present may become less efficient; additional comment on this point is found below.)

To show how to decide this question, we can proceed as follows. Let Q be a quantifier-free defining formula for the set F of forbidden positions. Since the condition $a > b$ (resp. $a > b$) can be written as $a-b > 0$ (resp. $a-b > 0$ or $a-b = 0$) etc., we can suppose without loss of generality that Q is a boolean combination of clauses $P > 0$ and $P = 0$, P designating some arbitrary polynomial in the appropriate number r of variables. Let S designate the set of all polynomials appearing in Q or as defining equations of the manifold G, and use Collins' theorem to construct an S-invariant cylindrical algebraic decomposition K of E^r. Then it is clear that G, F, and G-F are all unions of collections of cells $c \in K$.

It follows from the general topological results presented in the preceding section that two points p and q of G-F can be connected by a continuous arc in G-F if and only if there exists a chain $c_1, \ldots, c_k$ of cells of G-F such that $p \in c_1$, $q \in c_k$, and such that for each j, either c_{j+1} is a face of c_j, having one dimension less than that of c_j, or vice-versa. (This follows from well-known properties of the homology group $H_0(G-F)$; see e.g. [CF].) Note that the recursive construction of the Collins cells gives us an effective way of connecting any two points in the same cell by a continuous arc lying wholly within that cell, while Lemma 4 gives us an equally explicit way of constructing an arc from any point p in a cell c' forming part of the boundary of another cell c to any point $q \in c$.

All in all, therefore, the results we have proved concerning the Collins decomposition give us a constructive way of determining whether the two given points p and q belong to the same (arcwise connected) component of G-F, and of finding a smooth arc connecting them when they do.

If F is not assumed to be closed, then to check for the existence of a chain of cells connecting the given points we may need to employ the more costly test for adjacency of Collins cells of arbitrary dimension described in Appendix C. At present we do not know whether this test can be carried out in time polynomial in the geometric complexity of the problem. By making the technical assumption that F is a closed subset of G we avoid this difficulty, since then connectivity in G-F can be determined by following chains of cells of codimension at most 1. To see this, assume for the moment that G can be represented as a whole Euclidean space E^k for some k (we will explain shortly how such a representation can be constructed.) Then, if F is closed it follows that G-F is a smooth manifold of dimension k in E^k, so that no submanifold of G-F of dimension k-2 or less can disconnect any connected component of G-F (for this well-known fact see, e.g., Lemma 1.9 of [SS]). Hence in this case two points in G-F are connected to one another if and only if the two cells containing them can be connected by a chain of cells all of which are either k-dimensional or (k-1)-dimensional. This can be done using the relatively efficient 'maximal dimension' adjacency testing method presented at the end of the preceding section.

To see that G can be represented in this manner, first note that by the general discussion of the algebraic representation of a collection of hinged bodies given in Section 1, G can be represented as the Cartesian product of a finite number of spaces, each of which is either a full Euclidean space E^k, or a rotation group of dimension 1,2 or 3. These groups can in turn be represented as the circle S^1, the 2-dimensional sphere S^2, and the 3-dimensional sphere S^3 respectively, with the third representation being double-valued, i.e. each rotation in 3-space is represented by two antipodal points in S^3 (for more

details concerning this representation, see e.g. [Ha]). Moreover, with the exception of one point, S^2 and S^3 can be mapped algebraically onto E^2 and E^3 respectively by appropriate stereographic projections. Omission of the exceptional points of one or more such stereographic mappings will not affect the connectivity of the open manifold in which we are interested since the points omitted all lie on submanifolds of G of codimension at least 2. Therefore, with no loss of generality, we can assume that G is represented as a product of a Euclidean space E^k by a finite number of circles S^1. Each such circle can be mapped algebraically (with the exception of one point) onto a line by a stereographic projection, but here the omitted point can affect the connectivity of the resulting image of G-F. To overcome this small technical difficulty, assume first that only one circle is involved, i.e. $G = E^k \times S^1$. Map S^1 onto a line R_1 by projecting it from a point $X_1 \in S^1$, and also onto another line R_2 from another point X_2. We then obtain two distinct representations G_1 and G_2 of G, and can construct corresponding Collins decompositions K_1 and K_2 for each of them. Next we can analyze the connectivity of G_1 (resp. G_2) by constructing an appropriate connectivity graph CG_1 (resp. CG_2) in the simplified manner described above, and finally we can merge these graphs into one graph by adding edges which connect a cell $c_1 \in CG_1$ to a cell $c_2 \in CG_2$ whenever c_1 and c_2 have a common point of G-F (this property of cells can be checked for easily if we take care to include the equation defining the subspace $E^k \times \{X_2\}$ (resp. $E^k \times \{X_1\}$) among the algebraic equations from which K_1 (resp. K_2) is generated). The connectivity of G-F can then be determined by analyzing chains of edges in this merged graph. Cases in which G is the product of E^k by more than one circle can be handled in a similar manner.

Note: Another minor technical point to be noted is that the representation of the full 3-dimensional rotation group as S^3 is bivalent, so that a path between two specified rotations R_1 and R_2 of some subpart B_i of B, can correspond either to a path between a point $\xi_1 \in S^3$ representing R_1 and a similar point ξ_2 representing R_2 or to a path between ξ_1 and $-\xi_2$. Thus in order to determine whether R_1 and R_2

can be connected we have to check for the existence of one of several paths.

Once a well-based Collins decomposition has been constructed, the connectivity analysis can proceed, as already noted, via a simple search through the underlined connectivity graph whose nodes represent the Collins cells of highest dimension, and whose edges indicate cell adjacency. The computation cost of such an analysis is plainly linear in the size of the Collins decomposition. Collins has shown (see also Appendix A below) that the number of cells in a cell decomposition K is $O((2n)^{3^{r+1}} \cdot m^{2^r})$, where m is the number of polynomials defining the sets G and F, and where n is the maximum degree of any one such polynomial. Note that n is related to the degree of any single geometric constraint, and that r is related to the number of degrees of freedom of the bodies B. If we fix r, it follows that the number of cells in K, as well as the number of adjacent pairs of cells, is polynomial in m and n, i.e. in the geometric complexity of the problem, that is, in the number of different walls, faces, and other features of the system B of bodies, and in their algebraic degrees. Moreover, the time required to construct the Collins decomposition, and to test for adjacency of cells of maximal dimension, can also be shown to involve a number of operations on algebraic numbers which is also polynomial in m and n. As is well known (see Appendix B for details), each such operation can be accomplished in time polynomial in the degree of the polynomials defining these algebraic numbers. Taking all this into account, we obtain the following result:

Theorem 4: The mover's problem for (algebraic) bodies having a fixed number of degrees of freedom whose set of forbidden configurations is closed can be solved in time polynomial in the number of geometric constraints present in the problem.

Remarks: (1) The computational cost of our solution of the movers problem is still exponential in the number of degrees of freedom of the bodies B. That this complexity growth is probably inherent is indicated by a theorem of Reif [Re], which asserts that the mover's problem for a

robot B with many jointed arms (all free to rotate around a common axis) is PSPACE-complete.

(2) A comparison of Theorem 4 and the discussion preceding it with the more elaborate technique used in [SS] to solve certain 2-dimensional cases of the motion-planning problem efficiently reveals a significant similarity between the two approaches. In both approaches the free space of configurations of B is partitioned into cells, and these cells are connected to each other whenever they are physically adjacent to each other. This imposes a combinatorial graph structure on these cells, whose connected components reflect the connected components of G-F. Moreover, in both cases these cells are constructed recursively by adding one dimension at a time. Also, the cells appearing in [SS] can be shown to consist each of a finite union of Collins cells in the associated decomposition. We will not pursue these observations in this paper, but they will reappear in a subsequent report on efficient algorithms for other special cases of the movers' problem.

Appendix A. The Collins Decomposition; Auxiliary Remarks

In this appendix we review the construction which leads to the proof of Collins' Theorem 1, and add various auxiliary observations. We begin with the following remark. Let $P_b(z)$ be a polynomial of fixed degree n whose complex coefficients depend continuously on a parameter b which varies in some connected set S. Suppose that the number of distinct roots of $P_b(z)$ is independent of b. Then these roots vary continuously with b. This follows immediately from the fact that the unique root of P_b lying in any small circle can be expressed by a Cauchy integral over this circle. Next suppose that the polynomial P_b also has real coefficients for each value of b. Then the number m of real roots of P_b is also independent of b, and for each $j \leq m$ the j-th largest real root of P_b depends continuously on b. To establish this, let S_k be the set of points b for which there exist exactly k real roots. Take a point b_0 in S_k, let $r_1 \ldots r_\ell$ be the distinct complex

roots of P_{b_0}, with $r_1 \ldots r_k$ real and the remaining roots non-real. Draw disjoint small circles C_j, $j = 1 \ldots \ell$, around these roots. Then for b sufficiently near b_0 each of those circles will contain exactly one of the roots of P_b, and each root of P_b will lie in one such circle. Since complex roots of P_b must occur in conjugate pairs, it follows that (if they are sufficiently small) the circles $C_1, \ldots, C_k$, and only these, contain real roots of P_b, which proves our assertion.

Next suppose, in addition to the assumptions made above, that $R_b(x)$ is a second polynomial with real coefficients depending continuously on $b \in S$, and that for each $b \in S$ all the zeroes of R_b are contained in the set of zeroes of P_b. Then for each j, R_b is nonzero and of constant sign in the open interval $I_j(b)$ between the j-th and the $(j+1)$-st largest real zeroes of P_b, and arguing by continuity and from the connectedness of S it is clear that the sign of R_b on the interval $I_j(b)$ is independent of b.

A simple variant of the Collins technique, sufficient for our purposes, but a bit less efficient t n the one developed by Collins, can be described as follows. Assume that we are given a (finite) collection $\{P_i(b,x)\}$ of polynomials in k+1 variables whose coefficients are all rational. (We continue to suppose that b designates a vector of k real variables, and that x designates the last of the k+1 variables on which P_i depends; accordingly, we will treat the multivariate polynomials P_i as polynomials in x with coefficients belonging to the ring of rational polynomials in the k other variables b.) Let P denote the product of all these polynomials. We can then construct a family of polynomials $\{Q(\cdot)\}$ in the k variables b with the property that for each (connected) k-dimensional set S over which each of the polynomials Q(b) maintains a constant sign (zero, positive or negative), the number of distinct real zeroes of $P(b, \cdot)$ is constant. Suppose for the moment that this has been done. Then the preceding remarks imply that the distinct real roots of $P(b, \cdot)$ over each such set S can be enumerated from smallest largest so that for each j the j-th root $f_j(b)$ varies continuously over S. On the other hand, once the family $\{Q(b)\}$ has been formed, we can partition E^k into connected

sign-invariant subsets S by a recursive application of the Collins technique to the collection {Q(b)}. The complexity of the total procedure then depends on the number of polynomials Q needed to ensure the invariance of the number of distinct real roots of P over each connected set on which they maintain a constant sign, and on their maximal degree.

To construct the required polynomials Q, we have only to use the following well-known observation: Let P' denote the x-derivative of P, and let R = GCD(P,P'). Let n be the degree of P, and m be the degree of R. Then P has n-m distinct roots. Hence it suffices to introduce enough polynomials Q(b) to ensure that the degrees of P and R are constant on any connected set on which the Q's are sign-invariant.

To do this we first recall some facts concerning resultants and subresultants of polynomials; for which see Brown and Traub [BT]: Let $A(x)$ and $B(x)$ be two polynomials in x, having degrees a and b respectively. Fix any $j > 0$, and consider the equation

$$(*) \qquad A(x)U_j(x) = B(x)V_j(x)$$

in two polynomials U_j, V_j, having degrees b-j-1 and a-j-1 respectively. The unique factorization theorem for polynomials implies that (*) has a nonzero solution if and only if A and B have j+1 common roots. By expanding (*) in terms of the coefficients of U_j and V_j, we obtain a system of a+b-j linear equations in a+b-2j unknowns. We prefer to reduce this system to a square system, to which end we use the following observation. Suppose that we already know that (*) admits a nonzero solution for all i = 0,...,j-1, so that A and B have at least j common roots. Replace (*) by the weaker condition

$$(**) \qquad A(x)U_j(x) - B(x)V_j(x) = C_j(x),$$

where $C_j(x)$ is an arbitrary polynomial whose degree is at most j-1. This system involves exactly as many equations as unknowns, so that equation (**) then has a nonzero solution if and only if $\psi_j(A,B) = 0$,

where $\psi_j(A,B)$ is the determinant of the $(a+b-2j)\times(a+b-2j)$ matrix of the homogeneous system of linear equations representing the condition that highest $(a+b-2j)$ powers of x in the left-hand side of (**) have zero coefficients. (The determinant $\psi_j(A,B)$ is known as the j-th <u>subresultant</u> of A and B; $\psi_0(A,B)$ is the <u>resultant</u> of these polynomials See Brown and Traub, op. cit. for more details.) If $\psi_j(A,B) \neq 0$, then (**), and hence also (*), has only trivial solutions, so that A and B have exactly j roots in common. On the other hand, if $\psi_j(A,B) = 0$, then there exist U_j, V_j and C_j satisfying (**). However, we already know that A and B have at least j roots in common. Hence $C_j(x)$ must be divisible by their product. But since $C_j(x)$ is of degree at most j-1, it must be identically 0, so that U_j and V_j also satisfy (*), and so have at least j+1 roots in common. Thus, given the two polynomials A and B, we can determine exactly how many roots they have in common by computing $\psi_j(A,B)$ for increasing j until $\psi_j(A,B)$ becomes nonzero. This establishes the following

<u>Lemma 1</u>: The number of common roots of two polynomials $A(x)$ and $B(x)$ is j, where j is the smallest integer such that $\psi_j(A,B) \neq 0$.

In particular, the degree of the polynomial $R(b,x)$ introduced above is the least j such that $\psi_j(P,P') \neq 0$. Note also that the process just described depends on the knowledge of the degree of A and B (more precisely, on the maximal degree of A and B). Hence if A and B also depend on some parameter b (as does happen in the case in which we are interested), these degrees may vary if the leading coefficients of these polynomials become zero. All these considerations lead us to the following

<u>Lemma 2</u>: Let $P(b,x)$ be of degree n in x. For each j =1,..,n let $P_j(b,x)$ denote the sum of terms of P whose degree in x is $\leq$ j, and let $Q_j(b)$ denote the leading coefficient of P_j. Also let $R_{jk}(b) = \psi_k(P_j,P_j')$, for k = 0,..,j-2. Let M be the collection of all polynomials $Q_j(b)$ and $R_{jk}(b)$. Then on each connected set S on which all polynomials in M maintain a constant sign, the number of distinct real roots of $P(b,\cdot)$ is constant.

To bound the computational cost of all this, Let $Q^*(b)$ denote the product of all polynomials in M. Then the degree of Q^*, as a polynomial in any of the components y of b, is easily seen to be $O(dn^3)$, where d is the degree of P in y. Indeed this product involves $O(n^2)$ polynomials, which are determinants of matrices of size 2n×2n at most, each element of which is of degree d in y.

Using the preceding remarks, the Collins decomposition can be built up in the following recursive manner. Let $S = \{P_i(b,x)\}$ be any set of polynomials in k+1 variables whose coefficients are all rational. Let P be the product of all the P_i, and let Q(b) be the product of all polynomials appearing in Lemma 2. Applying Collins' construction recursively, let K be a Q-invariant cylindrical algebraic decomposition of the Euclidean space E^k. Let c be any one of the cells of K. Then Lemma 2 implies that the number of distinct real roots of $P(b,\cdot)$ remains constant as b varies in c. Hence, if $f_1(b),..,f_m(b)$ designate the real roots of P over c in ascending order, then all the functions $f_j(b)$ are continuous in b for b ε c, and the collection of sets (ii.a) – (ii.d) of Section 2 partition the cylinder $c \times E^1$ in such a way so that the collection of all these sets over base sets c ε K define an S-invariant (k+1)-dimensional cylindrical algebraic decomposition.

Remark: Using the technique for selecting a good direction described in Section 2 we can reduce M somewhat. Let R denote the homogeneous portion of P of highest degree ℓ. Rotating axes, we can ensure that R has a nonzero term involving x^ℓ. In this case M need only include Q_ℓ and the subresultants $R_{\ell k}$, so that Q^* will be of degree $d\ell$ at most.

We omit the somewhat more refined argument, given by Collins, which shows how to find effectively quantifier-free defining formulae for the cells of the Collins decomposition.

As noted by Arnon [Ar], it is easy to write a Tarski statement which asserts that y is the j-th real root of P in ascending order. This is simply

$$P(y) = 0 \text{ \& (EXIST } y_1,\ldots,y_{j-1}) \mid$$
$$(y_1 < y_2 \text{ \& } y_2 < y_3 \text{ \& } \ldots \text{ \& } y_{j-2} < y_{j-1}$$
(2) $$\text{\& } P(y_1) = 0 \text{ \& } \ldots \text{ \& } P(y_{j-1}) = 0 \text{ \& } y_{j-1} < y$$
$$\text{\& (FORALL } x)(P(x) < 0 \text{ \& } x < y \text{ IMPLIES}$$
$$(x = y_1 \text{ v } x = y_2 \text{ v } \ldots \text{ v } x = y_{j-1}))$$

Collins also notes that it is easy to find an algebraic point in each cell in the decomposition K, in the following recursive way. Let K' be the base decomposition of K. Proceeding recursively, obtain such a point for each $c' \in K'$. Let $b \in c \in K$ be such a point. Then the points

$$[c,y_1-1], \quad [c,y_m+1],$$

$$[c,y_j], \qquad\qquad j=1,\ldots,m, \text{ and}$$

$$[c,\frac{y_j+y_{j+1}}{2}], \qquad j=1,\ldots,m-1,$$

are all algebraic and there is one such point in each cell intersecting $c \times E^1$.

Example: We illustrate the technique described above by finding a P-invariant decomposition of the 2-dimensional plane, where

$$P(x,y) = x^3+y^3-3xy$$

(This is one of the examples analyzed by Arnon [Ar2].) We begin by projecting E^2 onto E^1 in the y-direction. Since the leading coefficient of P (as a polynomial in y) is constant, in the first step of the Collins decomposition it is sufficient to construct the following polynomials in x (we delete the common factors of their coefficients):

$$\psi_0(P,P_y) = x^6-4x^3,$$

$$\psi_1(P,P_y) = x$$

Moreover, since the second polynomial is a factor of the first, only $Q(x) = \psi_0$ need be retained. The real roots of Q are 0 and $4^{1/3}$, so that the base decomposition K' of the decomposition we seek has 5 cells, namely

$$
\begin{aligned}
c_0^* &= (-\infty, 0) \\
c_1 &= \{0\} \\
c_1^* &= (0, 4^{1/3}) \\
c_2 &= \{4^{1/3}\} \\
c_2^* &= (4^{1/3}, +\infty).
\end{aligned}
$$

Next we determine how many dinstinct real roots $P(x,y)$ (as a polynomial in y) has over each of these cells. To do this we compute the Sturm sequence of P and P_y, which is

$$
\begin{aligned}
f_0(y) &= y^3 - 3xy + x^3 \\
f_1(y) &= y^2 - x \\
f_2(y) &= 2xy - x^3 \\
f_3(y) &= 4x - x^4
\end{aligned}
$$

From this sequence one easily finds out that $P(X, \cdot)$ has one root over c_0^*, one root over c_1, three roots over c_1^*, two roots over c_2, and one root over c_2^*.

Since the decomposition we have considered is well-based, it is a regular cell complex, and its topology will be completely determined once the ρ maps on its base cells are found. The maps ρ can be computed by using the technique described in Section 2. Omitting details, one finds that the map $\rho(c_0^*, c_1)$ maps 1 to 1; the map $\rho(c_1^*, c_1)$ maps all three roots 1,2,3 to 1; the map $\rho(c_1^*, c_2)$ maps the two upper roots 2,3 to the upper root 2, and the lower root 1 to the lower root 1; and, finally, the map $\rho(c_2^*, c_2)$ maps the single root 1 to 1.

Appendix B. On Exact Symbolic Computations with Algebraic Numbers.

This appendix addresses the problem f how to perform the exact calculations with algebraic numbers required for the algorithms described in this paper, fow which numerical approximate solutions are not acceptable, since such calculations may lead to incorrect conclusions, e.g. in comparing approximate quantities we may wind up putting them in an order which is different from the order of the original numbers, if these numbers are very close to each other. Of course, the algorithms to be described will never be able to give an 'exact' value of an algebraic number. Nevertheless they can be used whenever an answer to some discrete query involving algebraic numbers is needed, as in the Collins decomposition related technique sketched in this paper.

This kind of problem, i.e. how to perform exact calculations involving algebraic numbers, has been studied by many authors (see [Ak], [He], [CL], [Ru]). In this appendix we will review the methods used to perform calculations of this kind, describe various improvements of techniques that have appeared in the literature, and present a few additional techniques.

In the following discussion, we ignore all those (possibly substantial) computational costs which can (and will) arise from the growth in size of the integers with which the algorithms to be described must deal; that is, we will measure cost by assigning each operation on integers (and hence each elementary operation on rational numbers) a nominal cost of 1. (Note however that much prior research has concentrated on obtaining more realistic cost estimates for such algorithms, taking into account the possible growth of coefficients during certain operations on polynomials, such as computation of the GCD of two polynomials, the Sturm sequence of a polynomial, the sequence of derivatives of a polynomial, etc. (see [Br], [He], [CL]). These more refined estimates have shown that the extra cost incurred in such operations is still polynomial in the degree and the size of the

coefficients of the polynomial(s) involved. Our significantly more optimistic cost measure is like the one used by Aho, Hopcroft and Ullman [AHU].)

Some of the results presented below rely on the weak but useful lower bound on the smallest possible distance between two distinct real roots of a polynomial. This is the content of the result of Mahler [Ma] (see also Mignotte [Mi]) which the following definition and theorem summarize.

Definition 1: (a) Let P be a polynomial over the complex field. Then |P| is defined to be the sum of the absolute values of all the coefficients of P.

(b) The squarefree part P^* of P is the quotient of P by the greatest common divisor GCD(P,P') of P and its derivative P'.

As already observed, P^* and P have exactly the same roots, but all the roots of P^* are simple. If $P = P^*$, i.e. if P has simple roots only, then P is said to be squarefree. With the significant reservation noted above the squarefree part of a polynomial P of degree n can be calculated in time $O(n \log^2 n)$ by using fast techniques for the required GCD computation and division steps; see Aho, Hopcroft and Ullman [AHU] Chap. 8.

Theorem 1 (Mahler): The minimum distance between two distinct roots of a squarefree polynomial P of degree n with integer coefficients is bounded below by

$$\Delta(n, |P|) = \frac{1}{n^{(n+2)/2} |P|^{n-1}}$$

Theorem 1 is important in what follows, since it guarantees that sufficiently precise approximate calculations with algebraic numbers (of the type to be considered below) will yield entirely precise results. However, in most of the following algorithms we will not have

to compute the roots of a polynomial P to such a high degree of accuracy, unless roots of P actually happen to be that close to one another.

<u>Definition 2</u>: (a) Let P be a squarefree polynomial of degree n with integer coefficients, then a <u>P-isolating interval</u> for a real root r of P is an interval with rational endpoints, which contains r in its interior, but does not contain any other root of P.

(b) Let P be as in (a). A <u>P-separation</u> of the real line is a partition of R into a union of disjoint P-isolating intervals.

Theorem 1 yields a lower bound on the size of a maximal P-isolating interval for a root of P. However, assuming that the roots of P are randomly distributed, the size of P-isolating intervals can be expected to be much larger than $\Delta(n, |P|)$.

Following a convenient convention, we can represent an algebraic number x by a pair consisting of a squarefree polynomial with integer coefficients having x as a root and of a P-isolating interval for x. To proceed in this way, it is obviously important to be able to find isolating intervals for all the real roots of a squarefree P (i.e. to find a P-separation of the real line) rapidly. This <u>root isolation problem</u> is considered by Heindel [He], Akritas [Ak], and Collins and Loos [CL]. Akritas describes an isolation technique based upon systematic binary searching using Descartes rule of signs, which he indicates can solve this problem for a polynomial P of degree n in time $O(n^5)$. (However, the details of his efficiency estimate are not entirely clear.) The estimates of Heindel and of Collins and Loos give bounds $O(n^{10}+n^7\log^3|P|)$, which is also polynomial in n and $|P|$, although with a relatively high exponent. Below we will sketch a root-isolation algorithm, essentially an improved variant of the older technique suggested by Heindel [He], which uses a Sturm sequence-based technique and can accomplish root isolation using $O(n^3\log n)$ arithmetic operations (in our cost measure) in the worst case, but on the average will require only $O(n^2 \log^2 n)$ such operations.

We begin by reviewing the beautiful classical theory of Sturm sequences (see Marden [Mar], p. 130 ff.), which gives a very useful way of handling several of the problems that concern us. Let P be a univariate polynomial, and let P' be its derivative. The <u>Sturm sequence</u> of P is a sequence $\{f_i\}$ of polynomials such that $f_0 = P$, $f_1 = P'$, and such that for each $i > 1$, $-f_i$ is the remainder obtained by dividing f_{i-2} by f_{i-1}. For this sequence there plainly exists a sequence of quotient polynomials Q_{i-1} such that

$$f_{i-2} = Q_{i-1}f_{i-1} - f_i, \qquad i > 1,$$

where the degree of f_i is strictly smaller than that of f_{i-1}. Since P is assumed to be squarefree, this process must terminate with some constant function f_k. The Sturm sequence has the property that, for any interval $[a,b]$ the number of roots of P in this interval is $S(a) - S(b)$, where $S(x)$ is the number of sign changes in the sequence $[f_0(x), f_1(x),\ldots, f_k(x)]$.

Next we describe a fast procedure for the computation of the Sturm sequence of a given polynomial P of degree n. At a first glance it might seem that this task will require time at least $O(n^2)$, since that many coefficients appear in the polynomials constituting this sequence. However, by representing the sequence in a more economical way, we can reduce this time to $O(n \log^2 n)$. To do this, we note that the degree of Q_i is the difference of the degrees of f_{i-1} and f_i. Thus if m_i denotes the degree of the quotient Q_i, $i=1,\ldots,k-1$, the sum of all the m_i's is n. Hence we can represent the Sturm sequence by the sequence $[f_0(x), f_1(x), Q_1(x),\ldots, Q_{k-1}(x)]$, which involves only $O(n)$ coefficients. Once this representation is available, we can use it to evaluate the whole Sturm sequence at any given x, as well as the number of sign changes in the Sturm sequence, in time $O(n)$. To do this, we first compute $f_0(x)$, $f_1(x)$, and $Q_i(x)$ for $i=1,\ldots,k-1$ (which requires total time $O(n)$). Then, using the 'backward formulae'

$$f_i(x) = Q_i(x)f_{i-1}(x) - f_{i-2}(x), \qquad i = 2,\ldots,k,$$

the Sturm functions can be evaluated at x in $O(n)$ additional time.

For this more efficient evaluation, we simply need all the quotient polynomials Q_i. Up to a sign change, these are exactly the quotients obtained during calculation of the GCD of P and P'. To compute all these quotients efficiently, we can use the fast polynomial GCD procedure described in [AHU], Chap. 8. This procedure works as follows. Let $a(x)$, $b(x)$ be two polynomials of degree $\leq n$. Let the remainder sequence $\{r_i(x)\}$ and the quotient sequence $\{q_i(x)\}$ of $a(x)$ and $b(x)$ be defined so that $r_0(x) = a(x)$, $r_1(x) = b(x)$, $r_{i-1}(x) = q_i(x)r_i(x) + r_{i+1}(x)$, $i > 1$, and $\deg(r_{i+1}(x)) < \deg(r_i(x))$. For each i there exists a polynomial 2×2 matrix M_i such that

$$[r_i(x), r_{i+1}(x)] = [a(x), b(x)] M_i.$$

Furthermore, each M_i is the product of matrices of the form

$$N_j = \begin{bmatrix} -q_j(x) & 1 \\ 1 & 0 \end{bmatrix} \quad , j \leq i.$$

The fast GCD algorithm first computes the two middle elements $r_j(x)$ and $r_{j+1}(x)$ in the remainder sequence, and then calls itself recursively with $r_j(x)$ and $r_{j+1}(x)$ to process their remainder sequence, which of course coincides with the rest of the remainder sequence of a and b. To find r_j and r_{j+1}, the algorithm uses another recursive procedure HGCD which computes the matrix M_j. This second procedure uses the fact (see [AHU], Lemmas 8.6 and 8.7) that all the quotients $q_i(x)$, $i \leq j$ depend only on the most significant half of the polynomials a and b. HGCD thus discards the least significant halves of a and b, thereby obtaining polynomials a' and b' of degree at most n/2; it then calls itself recursively to compute the two middle elements $r_{j'}$ and $r_{j'+1}$ among the first j elements in the remainder sequence of a' and b', during which process it also calculates the matrix $M_{j'}$. The quotient $q_{j'}(x)$ and the matrix $N_{j'}$ are then calculated, after which HGCD calls itself once more with $r_{j'}$ and $r_{j'+1}$ as parameters to compute the product L of the

matrices $N_{j'+1},...,N_j$. The matrix M_j is then computed as $M_j \cdot N_j \cdot L$, and recursively returned. This description should make it plain that the algorithm sketched computes all quotients $q_i(x)$ appearing in the quotient sequence of $a(x)$ and $b(x)$. With minor modifications, it can therefore be used to obtain the desired efficient representation of the Sturm sequence of a given polynomial. Since the fast GCD algorithm runs in time $O(n \log^2 n)$, this is also the time required for the calculation of the Sturm sequence.

This gives us the following

Theorem 2: The number of distinct real zeroes of a polynomial with rational coefficients of degree n, lying in any given interval [a,b] can be found in $O(n)$ arithmetic operations, after preprocessing which requires $O(n \log^2 n)$ arithmetic operations.

We can now describe a root-isolation procedure as follows. We first compute the Sturm sequence of the given polynomial P, in the manner just described. Next we find an upper bound b and a lower bound a such that all real roots of P lie in the interval [a,b]. For example (see [Mar]), we can take

$$b = -a = \max \left\{ 1 + \frac{p_i}{p_n} : i = 0,...,n-1 \right\}$$

where p_i is the coefficient of the i-th power of x in P. Let $N = S(b) - S(a)$ denote the number of distinct real roots of P. We perform a binary search of the interval $I = [a,b]$ to find a point $c \in [a,b]$ which separates it into two subintervals each containing at least one root of P. It follows from Mahler's theorem that such a point will be found after at most $O(n \log n)$ bisections of I, at each of which we have to evaluate S, so that to find c will take $O(n^2 \log n)$ steps. (However, assuming random distribution of the roots of P, the expected number of required bisections will be $O(\log n)$, so that c will be found on the average after $O(n \log^2 n)$ steps.) We then apply the same process to each of the intervals [a,c] and [b,c], and continue in this manner until

isolating intervals for all the roots of P have been found. Obviously only N < n intervals will have to be processed, so that the whole procedure requires $O(n^3 \log n)$ steps in the worst case, and $O(n^2 \log^2 n)$ steps on the average; note that these time bounds also dominate the time required for the initial computation of the Sturm sequence.

An alternative, and possibly more attractive, technique for root isolation has been described by Collins and Loos [CL], and is based on Newton's approximation technique. Their technique proceeds inductively by first obtaining root-separating intervals for the derivative of P, making sure that none of these intervals contains a root of the second derivative of P. On any interval in the complement of the union of these intervals P' has constant sign, so that it is trivial to check whether such an interval contains a root of P. On any of the P'-isolating intervals I, P is either convex or concave throughout I. Hence Newton's technique will converge (very rapidly) to a root of P in I if such a root exists. This enables one to partition I rapidly into P-isolating intervals which do not contain a zero of P', thus allowing iteration of the process. More details can be found in [CL].

Once any root isolation procedure is available, we can use it to perform various exact computations in algebraic numbers. Before describing efficient procedures for such computations, we first note a simple but useful generalization of Sturm's theorem. Specifically, let $A(x)$ and $B(x)$ be two given polynomials, where A is squarefree. Form the generalized Sturm sequence of A and B (which, up to sign changes, coincides with the remainder sequence of A and B) as follows. Put $f_0(x) = A(x)$, $f_1(x) = B(x)$, and, for each $i > 1$,

$$f_{i-1}(x) = q_i(x)f_i(x) - f_{i+1}(x),$$

so that the last element in this sequence will be the GCD of A and B. Let $S(x)$ denote the number of sign changes in the sequence $[f_0(x),\ldots,f_k(x)]$. As in the case of standard Sturm sequences, it is easy to see that each time we cross a zero of some function $f_i(x)$, $i > 1$, $S(x)$ remains unchanged. However, each time we cross a zero x_0 of f_0

= A from left to right, $S(x)$ decreases by $\text{sign}(A'(x_0)B(x_0))$. (It is easy to check that this statement remains true even if x_0 is a zero of $B(x)$.) Hence, in any given interval $[a,b]$,

$$S(a) - S(b) = \sum \text{sign}(A'(x)B(x)),$$

where the summation extends over all distinct roots of A lying in (a,b). (In the case of a standard Sturm sequence, $B(x) = A'(x)$, so that the above sum is equal to the number of distinct real roots of A in (a,b).)

In particular, if the interval (a,b) is known to contain just one root r of $A(x)$, then $S(a) - S(b)$ is $0,1$ or -1, depending on the signs of $A'(r)$ and $B(r)$. Since A has been assumed to be squarefree, the sign of $A'(r)$ is easily calculable from the signs of $A(a)$ and $A(b)$. Hence the sign of $B(r)$ can also be calculated. That is, given an algebraic number r, represented as the i-th root of a squarefree polynomial A, for which an A-separation of the real line is available, and another polynomial B, we can find the sign of $B(r)$. (Note that (an efficient representation of) the generalized Sturm sequence can be computed and evaluated using precisely the same techniques prescribed for standard Sturm sequences.) This technique improves that described by Rump [Ru].

This generalized Sturm technique is applicable to a variety of other problems. For example, we can use it to determine the multiplicity of the real roots of a given polynomial P. To do this, we first compute the squarefree part P^* of P, and then isolate the real roots of P^* (i.e. the distinct real roots of P). For each root r of P, we use the above procedure to determine $\text{sign } P'(r)$. If this is nonzero, then r is a simple root. Otherwise r has multiplicity 2 at least; we then repeat our procedure to find the sign of $P''(r)$, and so on, until a nonzero sign is obtained, from which the multiplicity of r is immediately calculable. Using this technique, the multiplicities of all real roots of a polynomial of degree n can be found in time $O(n^2 \log^2 n)$ (assuming that a P-separation is available). Indeed, the total

number of roots is n, and to determine each multiplicity, an $O(n \log^2 n)$ procedure is applied.

A very similar procedure can be used to compare two real algebraic numbers a and b, given as roots of the squarefree polynomials P and Q respectively. To do this, obtain a P-separation and a Q-separation of the real axis, and merge them into one partitioning. Let $I = [c,d]$ be a P-isolating interval for a, and let $J = [e,f]$ be a Q-isolating interval for b. If I and J are disjoint, then the manner in which a and b compare is immediately obvious. Suppose then that I and J intersect. Several cases can arise; we will treat only the case in which $c < e < d < f$, since the other cases can be handled in a similar manner. Since P is squarefree, and since (c,d) contains only one root of P, $P(c)$ and $P(d)$ have different signs. By evaluating the sign of $P(e)$ we can determine whether a lies in the subinterval (c,e) or in the subinterval (e,d) of I. If a lies in (c,e) then we must have $a < b$. Similarly, if b lies in (d,f) we also have $a < b$. In the remaining case, both a and b (but no other root of P or Q) must lie in the interval (e,d). Using the generalized Sturm technique explained above, we can compute the sign of $P(b)$, and this shows at once how a and b compare. As before, this procedure takes $O(n \log^2 n)$ time, where n is the maximal degree of P and Q, provided that P- and Q-separations of the real axis are available.

Similar procedures can be used to perform various other exact computations with algebraic numbers. Suppose, for example, that we need to compare the sum (or product) of two algebraic numbers a and b to a third such number. One can of course compute a polynomial $R(x)$ having $a+b$ (or ab) as a root. However, it may be undesirable to do this since this can generate a polynomial whose degree is $\deg(A) \cdot \deg(B)$ (where $A(x)$ (resp. $B(x)$) is a polynomial having a (resp. b) as a root), and repeated computations of this sort may result in polynomials of extremely large degrees. To avoid this problem, a recursive symbolic representation of algebraic numbers might be more advantageous: We can specify an algebraic number r either as a polynomial $A(r_1, \ldots, r_k)$ in k other algebraic numbers, or as a root of a

polynomial $B(x)$, whose coefficients are algebraic numbers $r_1,\ldots,r_k$; where each of these numbers is in turn represented in this same fashion, until numbers explicitly representable by polynomials with rational coefficients are finally reached. Such semi-symbolic representations can be used to perform computations of the kind discussed above. Consider the typical problem of determining the sign of an algebraic number r specified in this recursive fashion. Suppose to be specific that r is specified as a polynomial $A(r_1,\ldots,r_k)$. Let $B(r_k) = 0$ be an equation for r_k, and isolate r_k as a root of B. Then use the technique explained above to determine the sign of $A(r_1,\ldots,r_k)$, which we regard as a polynomial in r_k with coefficients which are algebraic. This will require that we determine the sign of various polynomials in $r_1,\ldots,r_{k-1}$, which we can do by using the same technique recursively.

Fast algorithms for the computation of subresultants.

Schwartz [Sc] (see also Moenck [Mo]) describes a variant of the fast GCD algorithm described above which computes the resultant of two polynomials having degrees $\leq n$ in $O(n \log^2 n)$ time. Since the Collins decomposition technique involves numerous calculations of resultants and subresultants, it is of interest to note that the algorithm described in [Sc] can be generalized in a straightforward manner to yield a rather similar fast procedure for the computation of subresultants as well. For the convenience of the reader we will describe the necessary modifications in full detail.

First we recall the definition of a subresultant. Let

$$A(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_0,$$

and

$$B(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_0$$

be two polynomials of degrees m and n respectively. For $0 \leq j \leq \min(m,n)$, the j-th subresultant $\psi_j(A,B)$ of A and B is the $(m+n-2j) \times (m+n-2j)$ determinant

$$\begin{vmatrix} a_m & a_{m-1} & \cdots & a_{m-n+j+1} & \cdots & a_{2j-n+1} \\ 0 & a_m & a_{m-1} & & \cdots & a_{2j-n+2} \\ & & \cdots & & & \\ 0 & 0 & \cdots & a_m & \cdots & a_j \\ b_n & b_{n-1} & \cdots & b_{n-m+j+1} & \cdots & b_{2j-m+1} \\ 0 & b_n & b_{n-1} & & \cdots & b_{2j-m+2} \\ & & \cdots & & & \\ 0 & 0 & \cdots & b_n & \cdots & b_j \end{vmatrix}$$

the first $n-j$ rows of which involve coefficients of A, and the last $m-j$ rows of which involve coefficients of B. (Here we use the convention that $a_i = 0$ if $i < 0$.) As noted earlier, this is precisely the determinant of the linear transformation

$$T : [U(x), V(x)] \rightarrow (A(x)U(x) + B(x)V(x)) / x^j$$

where U is a polynomial of degree $n-j-1$, V is a polynomial of degree $m-j-1$, and where, as usual, the remainder after the indicated division is discarded. If $j=0$, $\psi_j(A,B)$ is just the resultant of A and B.

Let Q be the quotient obtained by dividing B by A, and let $R = B$ mod A be the corresponding remainder. Suppose that $n \geqslant m$, and let $k \leqslant n-m$. By subtracting an appropriate upper row from each lower row in (*) we see that $\psi_j(A,B) = \psi_j(A, B-x^k A)$. Using this last formula repeatedly, it follows that, if both sides of the following equation are considered as $(m+n-2j) \times (m+n-2j)$ determinants, we have

$$\psi_j(A,B) = \psi_j(A, B \bmod A)$$

Moreover, expanding the second determinant by minors of the first n-m+1 rows, we obtain

$$\psi_j(A,B) = L(A)^{n-m+1} \, \psi_j(A, B \bmod A)$$

where $L(A)$ denotes the leading coefficient of A, and where the determinant on the right-hand side is now an $(2m-1-2j) \times (2m-1-2j)$ determinant. However, since the remainder $B \bmod A$ can be of any degree k lower than m, a more appropriate reduction is

$$(1) \qquad \psi_j(A,B) = L(A)^{n-k} \, \psi_j(A, B \bmod A)$$

where this time $\psi_j(A, B \bmod A)$ is an $(m+k-2j) \times (m+k-2j)$ determinant. To rewrite this in the symmetric form which covers the case $m > n$, first note that

$$(2) \qquad \psi_j(A,B) = (-1)^{(m-j)(n-j)} \, \psi_j(B,A).$$

Hence, if $m > n$, and if k is the degree of $A \bmod B$, we have

$$(3) \qquad \psi_j(A,B) = (-1)^{(n-j)(m-k)} \, L(B)^{m-k} \, \psi_j(A \bmod B, B).$$

Finally, if $j = m < n$, then

$$(4) \qquad \psi_j(A,B) = L(A)^{n-m}.$$

It is also appropriate to put

$$(5) \qquad \psi_j(A,B) = 0.$$

if $\min(m,n) < j$. Note that the last equality is consistent with the reduction formulae (1) given above, in the sense that $\psi_j(A,B) = 0$ if either $\deg(A \bmod B) < j$ or $\deg(B \bmod A) < j$.

The technique for resultant calculation given in [Sc] depends only on the identities (1)-(5) and hence can be adapted to the calculation of subresultants. Fleshing out this summary remark, we shall now present an efficient technique for the simultaneous calculation of all subresultants of a given pair of polynomials P and Q. To this end, we make the following definition.

<u>Definition 3</u>: Let a pair of polynomials $w = [P,Q]$ of degree d, d' with coefficients in a field F be given, and let $d = \max(d,d')$. Write P mod Q for the remainder of P upon division by Q. Then the <u>RQ-sequence</u> $RQ(w)$ of w is the sequence t_i, $i = d, d-1,\ldots,0$, of quadruples

$$t_i = [[P_i,Q_i], a_i, b_i, M_i],$$

defined as follows:

(1) P_i, Q_i are polynomials, a_i is a quantity of F, b_i is always +1 or -1, and M_i is a 2x2 matrix of polynomials with coefficients in F.

(2) $t_d = [[P,Q],1,1,I]$, where I is the 2 × 2 identity matrix.

(3) $\max(\deg(P_i), \deg(Q_i)) \geq i > \min(\deg(P_i),\deg(Q_i))$ for $i > 0$.

(4a) If $\min(\deg(P_i),\deg(Q_i)) < i$, then $t_{i-1} = t_i$;

(4b) otherwise, if $\deg(P_i) = i$, then (dropping remainders in all polynomial divisions) we have

$$t_{i-1} = [[P_i, Q_i \bmod P_i], a_i e_i, b_i f_i, N_i M_i],$$

where $n = \deg(Q_i)$, $k = \deg(Q_i \bmod P_i)$, $f_i = 1$,

$$e_i = L(P_i)^{n-k},$$

and

$$N_i = \begin{bmatrix} 1 & 0 \\ -Q_i/P_i & 1 \end{bmatrix}$$

(4c) otherwise $\deg(Q_i) = i$, and then

$$t_{i-1} = [[P_i \bmod Q_i, Q_i], a_i e_i, b_i f_i, N_i M_i],$$

where $m = \deg(P_i)$, $k = \deg(P_i \bmod Q_i)$,

$$e_i = ((-1)^i L(Q_i))^{m-k},$$

$$f_i = (-1)^{m-k},$$

and

$$N_i = \begin{bmatrix} 1 & -P_i/Q_i \\ 0 & 1 \end{bmatrix}$$

The following lemma generalizes Lemma 3 of [Sc].

Lemma 3: Let $w = [P,Q]$, t_i, etc., be as in the preceding definition. Then the sequence $RQ(w)$ has the following properties:

(i) $[P_i, Q_i] = M_i[P,Q]$.

(ii) $\deg(M_i) < d - \max(\deg(P_i), \deg(Q_i)) \le d - i$.

(iii) For each $i < d$ and each $0 < j < \min(\deg(P_i), \deg(Q_i))$, $\psi_j(P,Q) = a_i \psi_j(P_i, Q_i)$ if j is even, and $\psi_j(P,Q) = a_i b_i \psi_j(P_i, Q_i)$ if j is odd.

Proof: All this is clear for $i = d$. A step from t_i to t_{i-1} via (4a) of the preceding definition clearly preser es the validity of (i)-(iii). Now suppose that rule (4b) applies to the step from t_i to t_{i-1}. Property (i) is clearly preserved. Moreover,

$$\deg(M_{i-1}) \; < \; \deg(Q_i) - \deg(P_i) + \deg(M_i)$$

$$< \; \deg(Q_i) - \deg(P_i) + d - \deg(Q_i)$$

$$= \; d - \deg(P_i) = d - \max(\deg(P_{i-1}), \deg(Q_{i-1})).$$

Concerning (iii), it follows by (1) and (2) that

$$\psi_j(P,Q) = a_i \psi_j(P_i, Q_i) =$$
$$= a_i(L(P_i))^{(\deg(Q_i) - \deg(Q_i \bmod P_i))} \cdot \psi_j(P_i, Q_i \bmod P_i)$$
$$= a_{i-1} \psi_j(P_{i-1}, Q_{i-1}),$$

if j is even, and similarly if j is odd.

Suppose finally that (4c) is used to obtain t_i from t_{i-1}. Again, property (i) is clearly preserved, and property (ii) follows by an argument symmetric to the one used above, in which P_i and Q_i are interchanged. Concerning (iii), using (1) and (3) we obtain, assuming j is even,

$$\psi_j(P,Q) = a_i \psi_j(P_i, Q_i) =$$
$$= ((-1)^i L(Q_i))^{(\deg(P_i) - \deg(P_i \bmod Q_i))} \cdot \psi(P_i \bmod Q_i, Q_i)$$
$$= a_{i-1} \psi_j(P_{i-1}, Q_{i-1}),$$

and similarly if j is odd. This proves the Lemma. Q.E.D.

<u>Lemma 4:</u> Let $w = [P,Q], w^* = [P^*, Q^*]$ be two pairs of polynomials. Suppose that $\max(\deg(P), \deg(Q), \deg(P^*), \deg(Q^*)) = d$, and suppose that the terms of order not less than $d - 2i$ in P, Q agree with the corresponding terms in P^*, Q^*. Then the first i+1 terms of the sequence $RQ(w) = [t_d, t_{d-1}, \ldots]$ have precisely the same components a_i, b_i, M_i as the corresponding terms a_i^*, b_i^*, M_i^* of the sequence $RQ(w^*) = [t_d^*, t_{d-1}^*, \ldots]$.

<u>Proof:</u> Except for the equality of b_i and b_i^*, the proof is completely

identical to that of Lemma 4 of [Sc]. The equality of b_i and b_i^* is also an easy consequence of that same proof. Q.E.D.

Lemmas 3 and 4 justify the following subresultant calculation algorithm, whose underlying idea is to compute the terms a_i, b_i and M_i by stepping through the sequence $[t_i]$ by steps of increasing length, each of length double that of the preceding step. Lemma 4 implies that each of these steps can use polynomials of substantially lower degree than that of the original polynomials. However, as soon as we have calculated a_i, b_i and M_i for some $i < d$ for whic $j = \deg(P_i) < \deg(Q_i)$, we can get $\psi_j(P,Q)$ directly. Indeed, using (4) and Lemma 3(iii), we have

$$\psi_j(P,Q) = a_i \psi_j(P_i,Q_i)$$
$$= a_i \cdot L(P_i)^{(\deg(Q_i)-j)},$$

if j is even, and similarly if j is odd (and of course analogous formulae are available if $j = \deg(Q_i) < \deg(P_i)$.) Note that if $t_i \neq t_{i'}$, then $\min(\deg(P_i),\deg(Q_i)) \neq \min(\deg(P_{i'}),\deg(Q_{i'}))$, so that for each j either there exists no i for which $j = \min(\deg(P_i,Q_i))$, or the i's for which this equality holds all have the same value t_i, and hence the above formula defines ψ_j unambiguously. Observe finally that, by (5), all ψ_j which do not appear in the above formula for any value of i are 0.

In view of these comments, the algorithm proceeds as follows.

Initialize four static lists, AL, BL, CL and DL to the null list each.

(1) Call an auxiliary routine HSBRSL(P, Q, 0). This routine will build up tne rirst half of the RQ-sequence $[t_i]$, for $d > i > d/2 = d'$, and will return the quantities $e_{d'}$, $f_{d'}$, and $M_{d'}$.

(2) Compute $[P_{d'},Q_{d'}] = M_{d'}[P_d,Q_d]$.

(3) Let $m = \deg(P_{d'})$, $n = \deg(Q_{d'})$. If $m < n$ then append the

quantities $e_{d'}$, $f_{d'}$, $L(P_{d'})^{n-m}$, m respectively to the end of the four lists AL, BL, CL, DL. Similarly, if m > n then append the quantities $e_{d'}$, $f_{d'}$, $L(Q_{d'})^{m-n}$, n, to the end of these four respective lists.

(4) Call the whole procedure recursively, passing $P_{d'}$, $Q_{d'}$ to it as inputs to complete the construction of the whole sequence.

The subprocedure HSBRSL(P,Q,ℓ) required is rather similar to the auxiliary procedure HGCD used in the fast polynomial GCD algorithm in [AHU]. It consists of the following steps.

(1) Discard the least significant half of the coefficients of P and Q, that is, let $P = P^*x^k + P'$, $Q = Q^*x^k + Q'$, where $k = \max(\deg(P),\deg(Q))/2$, and where $\deg(P')$, $\deg(Q') < k$.

(2) Apply HSBRSL to P^*, Q^*, and $\ell+k$, to obtain $e_{k'}$, $f_{k'}$ and $M_{k'}$ for $k' = 3k/2$ (by Lemma 4, these are the same as the corresponding quantities associated with the original P and Q.)

(3) Compute $[P_{k'},Q_{k'}] = M_{k'} [P,Q]$.

(4) As in step (3) of the main procedure, let $m = \deg(P_{k'})$, $n = \deg(Q_{k'})$. If m < n then append the quantities $e_{k'}$, $f_{k'}$, $L(P_{k'})^{n-m}$, $m+\ell$, to the end of the four lists AL, BL, CL, DL respectively. Similarly, if m > n then append the quantities $e_{k'}$, $f_{k'}$, $L(Q_{k'})^{m-n}$, $n+\ell$, to the end of these four respective lists.

(5) Let $P_{k'} = P^*_{k'}x^{k''} + P''$, $Q_{k'} = Q^*_{k'}x^{k''} + Q''$, where $k'' = k/2$, and where $\deg(P'')$, $\deg(Q'') < k''$.

(6) Call HSBRSL with $P^*_{k'}$, $Q^*_{k'}$ and $\ell+k''$, to obtain e^*_k, f^*_k and M^*_k.

(7) Return e^*_k, f^*_k, and $M^*_k \cdot M_{k'}$.

When the algorithm just sketched has terminated, we can perform one

final scan through the lists AL, BL, CL, and DL, to accumulate the scalar quantities a_i and b_i by repeated multiplication of the e_i's and the f_i's respectively. Using the following technique we can also calculate all the subresultants ψ_j. For each iteration step k, let γ_k, m_k be the k-th components of CL and DL respectively, and let a_{k-1} and b_{k-1} be the product of all preceding components in the lists AL and BL respectively. If m_k is even, put

$$\psi_{m_k} = a_{k-1}\gamma_k,$$

and if m_k is odd, put

$$\psi_{m_k} = a_{k-1}b_{k-1}\gamma_k.$$

At the end of this final iteration, all remaining undefined subresultants are set to 0.

The computational cost of the procedure just sketched is evidently the same as the original procedure given in [Sc]. To estimate this cost we note that multiplication and division of polynomials of order m can be accomplished in time $O(m \log m)$. During each call to the subprocedure HSBRSL(P,Q,ℓ) that we have just described polynomials of order at most $m=\max(\deg(P),\deg(Q))$ need to be multiplied and divided (at steps (3) and (7)). The total time $T(m)$ required to apply HSBRSL to two polynomials of degree m therefore satisfies $T(2m) = 2T(m) + O(m \log m)$ and hence has the bound $T(m) = O(m \log^2 m)$. Similar considerations show that the time used by the main procedure to build up the RQ-sequence has the same bound, so that this estimate also bounds the time required to calculate all the subresultants of two polynomials of maximal degree m.

Appendix C. Adjacency of Collins Cells of General Dimension.

In this appendix we complete our calculation of Collins cell adjacency by considering cells whose dimension is less than that of the whole space E^r being decomposed. The problem here is to determine when a Collins cell c' of dimension $k-1$ is a face of a cell c of dimension $k < r$. Our approach will again be based upon consideration of the base decomposition K' of K. However, the present case is somewhat more complex than the case $k = r$ considered previously, and this makes it necessary to use a bit more (largely classical) machinery drawn from the theory of algebraic curves. Accordingly, we recall the following definitions and lemmas, for which see Keller [Ke], Chap. 5.

Definition 1: (a) A fractional (Laurent) series is a formal series of the form

$$(1) \qquad y(x) = \sum_{i=m}^{\infty} a_i x^{i/D}$$

whose coefficients are complex numbers, and in which D is an integer. Such a series is said to be convergent if it converges in the neighborhood of $x = 0$. Note that we allow m to be negative. The series (1) is said to be truncated if it contains only finitely many terms.

(b) If $a_m \neq 0$, then m/D is called the leading exponent of the series y.

(c) If the leading exponent e of the series (1) satisfies $e \geqslant k$ (resp. $e > k$), we will write $y = O(x^k)$ (resp. $y = o(x^k)$).

It is easily seen that the standard definitions of addition, multiplication, etc. for power series make the collection of all fractional series into a field. This statement, like many others made in the next few paragraphs, is true irrespective of whether we consider convergent fractional series only, or allow arbitrary, nonconvergent series, and treat operations on them in a purely formal manner.

Moreover, this statement remains true even if the coefficients of the series (1) are required to lie in some subfield of the complex numbers.

In what follows we will designate the field of fractional power series by Fr.

Lemma 1: (see [Ke], Chap. V) The field of fractional series with complex (or with algebraic) coefficients is algebraically closed.

Thus if $P(y) = a_n y^n + \ldots + a_0$ is a polynomial in y with coefficients in the field Fr and if $a_n \neq 0$, P can be factored as

(2) $$P(y) = a_n(y-r_1) \ldots (y-r_n),$$

where $r_1, \ldots, r_n$ are themselves fractional series, namely the roots of $P(y) = 0$.

In what follows we will need to work in purely finite manner with fractional series representing the roots of various polynomials. Our ability to do so without ambiguity will rest upon various extensions of the following simple lemmas, which are also noted by Kung and Traub [KT].

Lemma 2: Let $P(x,y)$ be a polynomial with complex coefficients in two variables, of total degree $d>0$. Let its degree n in y be nonzero, and suppose that, when regarded as a polynomial $P_x(y)$ in y, P has no factor in common with its y-derivative $P_x'(y)$ and has a leading coefficient $a_n(x)$ which is not zero for $x=0$. Regard $P_x(y)$ as a polynomial in y with coefficients in Fr, and let $r_1(x), \ldots, r_n(x)$ be its roots (which are elements of Fr). Then

(a) If $r_i(x) - r_j(x) = O(x^{d(2n-1)})$, then $i = j$.

(b) We cannot have $P_x'(r_i(x)) = o(x^{d(2n-1)})$ for any root $r_i(x)$.

Proof: Consider the discriminant $D(x)$ of the polynomial $P_x(y)$, i.e. D

= $\mathrm{Res}(P_x, P_x')$ is the resultant of P_x and P_x'. This is a polynomial of degree $d(2n-1)$ in x, and since P_x and P_x' have no factor in common it is nonzero. Hence $D(x) = o(x^{d(2n-1)})$ is impossible. By well-known identities (see [Wa], Sections 30 and 31),

$$(3) \qquad D(x) = \prod_{i \neq j} (r_i(x) - r_j(x))$$

and also

$$(4) \qquad D(x) = \prod_{i \neq j} P_x'(r_i(x)).$$

Since the leading coefficient of y in $P_x(y)$ does not vanish at x=0, it is easy to see (by substitution) that no fractional series r(x) with negative leading exponent can satisfy $P_x(r(x)) = 0$. Hence $r_i(x) = O(1)$ for all i, and therefore if assertion (a) of our Lemma were violated, (3) would imply that $D(x) = o(x^{d(2n-1)})$, which is impossible. If we use (4) instead of (3) in this argument, (b) follows in the same way. Q.E.D.

<u>Lemma 3</u>: Let $P_x(y) = a_n(x)y^n + a_{n-1}(x)y^{n-1} + \ldots + a_0(x)$ be a polynomial in y, of degree n, with coefficients in Fr, and suppose that $a_n(x)$ does not vanish at x = 0. Let k ≥ 1, and let $y_0 = \sum \alpha_i x^{i/D}$ be an element of Fr such that $P_x(y_0(x)) = O(x^{nk})$. Then there exists a root $r_i(x)$ of $P_x(y) = 0$ such that $y_0(x) - r_i(x) = O(x^k)$.

<u>Proof:</u> Factor $P_x(y)$ in the manner (2). Then it is plain that if $y_0 - r_i$ has leading exponent less than k for all i, the leading exponent of $P_x(y_0(x))$ must be less than nk, contradicting our assumption. Q.E.D.

Lemma 3 asserts that any power series y_0 which comes close enough to making $P_x(y_0)$ equal to zero must lie quite close to a root of P_x. It is also worth noting that if $P_x(y)$ has no repeated roots, then, once $P_x(y_0)$ has been made small enough, any desired number of coefficients of a root of $P_x(y)=0$ can be calculated rapidly from the coefficients of y_0 by purely rational operations. A technique for calculating fractional power series which approximate each of the roots of $P_x(y)$

near x=0 to an arbitrary degree of precision is described by Kung and Traub [KT]. Their technique uses first the Newton's polygon method to obtain an initial (truncated) fractional series approximation for each of the roots of P_x, and then uses a variant of Newton iteration (similar to that described by Lipson [Li]) to extend each of these series to an arbitraru degree of precision. Kung and Traub give a complexity bound of $O(nN \log N)$ for their procedure, where n is the y-degree of the polynomial P_x, and where N is the number of terms sought in each fractional series. This bound holds asymptotically if n is held fixed and N increases, but it ignores the cost of applying the Newton polygon procedure (which indeed is independent of N) in order to obtain the initial number of terms required to ensure convergence of Newton's approximation method. At present our best estimates of this cost are still exponential, leading us to pose the following

Open Problem: Given a polynomial $P(x,y)$ with complex coefficients, of y-degree n and total degree d, does there exist a procedure for calculating truncated fractional series approximations for each of the roots of $P(x,y)$ (regarded as a polynomial in y) near x=0 up to terms of order $O(x^{d(2n-1)})$, in time polynomial in n and d?

As will become clear from the subsequent discussion, an affirmative answer to this problem would imply that adjacency of Collins cells of general dimension, and hence also the homology groups of any algebraic variety, can be calculated in time polynomial in the number and the maximal degree of polynomials defining a Collins decomposition (or an algebraic variety).

At any rate, once these fractional power series are available, we can use them to determine the adjacency of Collins cells. It is worth explaining what is involved by commenting briefly on the special case r = 3. Let K′ be the 2-dimensional base decomposition of a well-based Collins decomposition K of E^3 whose base polynomial is $P(x,y,z)$. Let c be a 1-dimensional curved cell in K′, and let $c′ = [x_0, y_0] \in K$ be an endpoint of c. Let $Q(x,y)$ be the base polynomial of K′. It follows from the preceding discussion that in the vicinity of c′ the curve c

can be specified uniquely by a truncated fractional series $y(x)$ which approximates the corresponding root of $Q(x,y)$ near (x_0,y_0). If enough leading terms of this fractional series are known, then, to determine the number of roots of $P(x,y,\cdot)$ into which a given root z_0 of $P(x_0,y_0,z) = 0$ splits as we move from (x_0,y_0) into c, we can substitute the curve $y(x)$ for y in P, thus obtaining a polynomial $R_x(z) = P(x,y(x),z)$ in z with coefficients in Fr. Then, using a Sturm-based technique of the sort described in Appendix B, we can determine the number of roots of R_x into which z_0 splits for sufficiently small positive x. We can also compute truncated fractional series approximating each of these roots, and this enables us to carry the same form of analysis inductively to cases involving more than three variables.

However, several technical difficulties must be overcome in following this conceptual approach. First of all, Lemma 2 is not immediately applicable to R_x, since the coefficients of R_x are not simply polynomial but lie in Fr; moreover R_x need not be squarefree. This makes it a little harder to state how many terms of a fractional series suffice to characterize a root of $R_x(z)$ uniquely. Also, the truncated fractional series $y(x)$ gives only an approximate representation of the root of Q which traverses the curve c, and at first glance it is not clear how many terms need be included in $y(x)$ to guarantee that the analysis we have outlined will carry forward inductively.

To overcome these purely technical difficulties, it is useful to consider systems of polynomial equations in several variables having the form specified in the following definition.

Definition 2: (a) A <u>triangular</u> <u>monic</u> <u>system</u> of polynomial equations is a system

$$
\begin{aligned}
&P_2(x_1,x_2) = 0 \\
(5) \qquad &P_3(x_1,x_2,x_3) = 0 \\
&\qquad \cdot \ \cdot \ \cdot
\end{aligned}
$$

$$P_n(x_1, x_2, \ldots, x_n) = 0$$

of polynomial equations of total degrees $d_2, \ldots, d_n$ with complex coefficients, whose j-th equation involves the variables $x_1, \ldots, x_j$ only, and which, regarded as a polynomial in x_j with coefficients polynomial in the remaining variables, has a constant leading term. The sequence $d_2, \ldots, d_n$ is the <u>degree</u> <u>sequence</u> of the triangular system (5).

(b) Given the triangular monic system (5), its <u>resultant system</u> is the sequence

$$(6) \qquad R_2(x_1, x_2), \ R_3(x_1, x_3), \ \ldots, \ R_n(x_1, x_n)$$

of bivariate polynomials defined as follows. R_2 is simply P_2. R_3 is obtained from P_3 and P_2 by regarding them both as polynomials in the variable x_2 and forming their resultant. More generally, to form $R_j(x_1, x_j)$ proceed as follows: Regard $P_j(x_1, \ldots, x_j)$ and $P_{j-1}(x_1, \ldots, x_{j-1})$ as polynomials in x_{j-1}, and form their resultant, thus obtaining a polynomial $Q_1(x_1, \ldots, x_{j-2}, x_j)$ from which the variable x_{j-1} has been eliminated. Then form the resultant Q_2 of Q_1 and P_{j-2} (as polynomials in x_{j-2}), obtaining a polynomial from which both variables x_{j-2} and x_{j-1} have been eliminated. Continue repeatedly in this way, thus finally obtaining the desired polynomial $R_j = Q_{j-2}$, from which all variables but x_1 and x_j will have been eliminated.

(c) The sequence $\delta_2, \ldots, \delta_n$ of degrees of the resultant polynomials (6) is called the <u>resultant degree sequence</u> of the triangular system (5).

<u>Definition</u> <u>3</u>: (a) A <u>solution curve</u> of the triangular monic system (5) is a continuous vector-valued function $f(t) = [f_2(t), \ldots, f_n(t)]$ with complex components defined for all sufficiently small nonnegative values of the parameter t, such that for all j=2,...,n we have $P_j(t, f_2(t), \ldots, f_j(t)) = 0$.

(b) Let P_j and f be as in (a), let $d = [d_2, \ldots, d_n]$ be a sequence of

rational numbers, and let $\alpha = [\alpha_2,\ldots,\alpha_n]$ be a sequence of truncated fractional series in the variable t, such that α_j contains terms of order $O(t^d j)$ at most. Then α is said to be an <u>order</u> <u>d</u> <u>descriptor</u> of the solution curve f if $f_j - \alpha_j = O(t^d j)$ for all $j=2,\ldots,n$.

Our aim is to generalize Lemmas 2 and 3 to general triangular monic systems of more than two variables. For this, we need the following straightforward technical lemma, whose main purpose is to enable us to bound the order of finite descriptors of solution curves of (5) needed to characterize such a solution uniquely.

<u>Lemma 4</u>: As in Definition 3, let $P_2,\ldots,P_n$ be a triangular monic system and let $R_2,\ldots,R_n$ be its resultant sequence. Then each R_j can be factored as

$$(7) \qquad R_j(x_1,x_j) = A(x_1)R_j^0(x_1,x_j)R_j^1(x_1,x_j),$$

where A, R_j^0 and R_j^1 are all polynomials in their respective variables, and where R_j^0 is the squarefree part of R_j (regarded as a polynomial in x_j) represented by a polynomial whose coefficients (which are polynomials in x_1) have no common factor. Moreover, we can arrange this factorization so that the leading coefficient $C(x_1)$ of R_j^0 does not vanish at $x_1 = 0$.

<u>Proof</u>: As in Definition 2, regard $P_j(x_1,\ldots,x_j)$ and $P_{j-1}(x_1,\ldots,x_{j-1})$ as polynomials in x_{j-1} with polynomial coefficients, and form their resultant Q_1. If we use the fact that P_{j-1} is monic in x_{j-1}, it follows from the fundamental theorem of resultant theory that a tuple $[x_1,\ldots,x_{j-2},x_j]$ is a solution of the system $P_1 = P_2 = \ldots = P_{j-2} = Q_1 = 0$ of equations if and only if there exists an x_{j-1} such that $[x_1,\ldots,x_{j-2},x_{j-1},x_j]$ is a solution of the system $P_i = 0$, $i=2,\ldots,j$. Arguing repeatedly in this way, we see that $[x_1,x_j]$ is a solution of the equation $R_j(x_1,x_j) = 0$ if and only if there exist $x_2,\ldots,x_{j-1}$ such that $[x_1,x_2,\ldots,x_j]$ is a zero of P_i for all $i=2,\ldots,j$. On the other hand, it follows since each P_i is monic that all the solutions of the system $P_i = 0$, $i=2,\ldots,j$ which lie over a small neighborhood of the

point $x_1 = 0$ remain bounded. Thus all the solutions of $R_j(x_1,x_j)=0$ over such a neighborhood have the same property.

Next regard R_j as a polynomial in x_j, take its (x_j-) derivative R_j', form $T^1 = GCD(R_j, R_j')$, and use it to factor R_j as $R_j = T^0T^1$. These are polynomials in x_j with coefficients rational in x_1, but then by multiplying through by an appropriate polynomial $C(x_1)$ we can write $CR_j = S^0S^1$, where now S^0 and S^1 have coefficients which are polynomials in x_1. Since the ring of polynomials in two variables is a unique factorization domain, every prime factor of C divides one of the polynomials on the right, which is to say, divides all its coefficients. This remark allows us to divide through by all these prime factors, and if we then collect all common factors of the coefficients of R^0 and R^1 in $A(x_1)$, we arrive at the factorization (7).

It is now clear that all the solutions $[x_1,x_j]$ of $R_j^0=0$ which lie over a sufficiently small neighborhood of $x_1=0$ have second components x_j which remain bounded. By Lemma 1, these solutions $r_1(x_1),\ldots,r_m(x_1)$ can be written as fractional power series $\sum_{i=m} \alpha_i x_1^{1/D}$, so that since they remain bounded near $x_1=0$ none of these fractional series contain any term with negative exponents. Suppose that the leading coefficient $a_m(x_1)$ of R_j^0 vanishes at $x_1=0$. Then since R_j^0 can be factored as

$$R_j^0(x_1,x_j) = a_m(x_1) \ \Pi(x_j-r_i(x_1)),$$

it would follow that all the coefficients of R_j^0 vanish at $x_1=0$. But then x_1 would be a common factor of all coefficients of R_j^0, contrary to the way in which R_j^0 has been defined. Q.E.D.

Corollary: Let the polynomials P_j, $j=2,\ldots,n$ form a triangular monic system, and let the resultant degree sequence of this triangular system be $[\delta_2,\ldots,\delta_n]$. Put

$$d = [\delta_2(2\delta_2-1), \ldots, \delta_n(2\delta_n-1)].$$

Then any two solution curves $f(t)$, $g(t)$ of this triangular system which

have identical order d descriptors are identically equal for all sufficiently small nonnegative values of t.

Proof: Let R_j be the resultant system of the triangular system P_j. Arguing as in the proof of Lemma 4, we can conclude that the components f_j of f satisfy $R_j(x,f_j(x)) = 0$, and similarly for the components of g. Our assertion is therefore an immediate consequence of Lemmas 2 and 4. Q.E.D.

Definition 4: Let P_j and δ_j be as in the preceding Corollary, let $d = [d_2,\ldots,d_n]$ be the degree sequence of the P_j, and define $d^* = \max\limits_{j=2,\ldots,n} d_2 d_3 \cdots d_j \delta_j (2\delta_j - 1)$. Put $d_j^* = d^*/(d_2 \cdots d_j)$ for $j=2,\ldots,n$, (so that d_j^* is a decreasing sequence of positive rational numbers). Suppose that for $j=2,\ldots,n$ $\alpha_j(x)$ is a truncated fractional series such that there exist auxiliary truncated fractional series $\alpha_j^*(x)$ (all of whose exponents can be assumed to be $\leq d_j^*$) such that $P_j(x,\alpha_2^*(x),\ldots,\alpha_j^*(x)) = O(x^{d_j^*-1})$ and $\alpha_j(x) - \alpha_j^*(x) = O(x^{d_j^*})$. Then $\alpha = [\alpha_2,\ldots,\alpha_n]$ is called an **adequate descriptor for a solution of the triangular system** $P_2 = \ldots = P_n = 0$ **of equations**.

Lemma 5: Let d, d^*, and δ be as in the preceding definition. Given any adequate descriptor $\alpha = [\alpha_2,\ldots,\alpha_n]$ for a solution of the triangular system (5), then for all sufficiently small t there exists a solution $f(t) = [f_2(t),\ldots,f_n(t)]$ of the system $P_j(t,f_2(t),\ldots,f_j(t)) = 0$, $j=2,\ldots,n$ such that

$$(8) \qquad f_j(t) - \alpha_j(t) = O(t^{\delta_j(2\delta_j-1)}).$$

Proof: Use the notations of Definition 4. We proceed by induction on j to establish the existence of a solution curve $f = [f_2,\ldots,f_n]$ which satisfies the formula

$$(8') \qquad f_j(t) - \alpha_j(t) = O(t^{d_j^*}),$$

for $j = 2,\ldots,n$. By definition of the d_j^* it is plain that $(8')$ implies (8). The base case $j=2$ for our induction is immediate from Lemma 3.

Suppose that we have already shown the existence of $f_2(t),\ldots,f_{j-1}(t)$ such that $P_i(t,f_2(t),\ldots,f_i(t)) = 0$ for $i=2,\ldots,j-1$, and such that (8′) holds for all such i. As in definition 4, there exist auxiliary fractional series $\alpha_i^*(t)$ such that $P_j(t,\alpha_2(t),\ldots,\alpha_j(t)) = O(t^{d_j^*}-1)$ and $\alpha_i^*(t) - \alpha_i(t) = O(t^{d_i^*}) < O(t^{d_j^*}-1)$ (since the d_i^* are decreasing). Thus by (8′) we have $\alpha_i^*(t) - f_i(t) = O(t^{d_j^*}-1)$, from which we see immediately that

$$P_j(t,f_2(t),\ldots,f_{j-1}(t),\alpha_j^*(t)) = O(t^{d_j^*}-1).$$

Hence by Lemma 3 there exists a root $f_j(t)$ of this same equation such that $f_j(t) - \alpha_j^*(t) = O(t^{(d_j^*-1/d_j)}) = O(t^{d_j^*})$ for small t, from which (8′) is plainly seen to hold for j too. Q.E.D.

<u>Corollary:</u> Let P_j, $f = [f_2,\ldots,f_n]$, and $\alpha = [\alpha_2,\ldots,\alpha_n]$ be as in Lemma 5. Suppose that all the polynomials P_j have real coefficients. Then f_j is real if and only if all the nonzero coefficients appearing in α_j are real.

<u>Proof:</u> Since all P_j are real, the whole situation being considered is invariant under complex conjugation. Thus the curve $\bar{f} = [\bar{f}_2,\ldots,\bar{f}_n]$ is also a solution curve of the system P_j, and $\bar{\alpha} = [\bar{\alpha}_2,\ldots,\bar{\alpha}_n]$ is an adequate descriptor for $\bar{f}$. It is plain from Definition 3 and from the corollary to Lemma 4 that $f = \bar{f}$ if and only if $\alpha = \bar{\alpha}$, from which the present corollary is immediate. Q.E.D.

Our preparation is now sufficient for computations of the topological relationships which interest us to be feasible. The following definition takes the next step in this direction.

<u>Definition 5:</u> Let K be a Collins decomposition of the Euclidean space E^r, let c be a d-dimensional cell in K, and let $c' \in K$ be a $(d-1)$-dimensional face of c. Then an <u>incidence</u> <u>curve</u> for the pair c, c' is a triple consisting of

(i) A triangular monic system $P_2,\ldots,P_m$ of m polynomials with real algebraic coefficients,

(ii) a continuous solution curve $f(t) = [f_1(t), f_2(t), \ldots, f_r(t)]$ all of whose components are real such that $f(t) \ \varepsilon \ c$ for all sufficiently small nonnegative t, and such that $f(0) \ \varepsilon \ c'$. All constant components of the vector f must be algebraic numbers, and its nonconstant components $f_{i_1}(t),\ldots,f_{i_m}(t)$, which for convenience we shall write in left-to-right order as $g_1(t),\ldots,g_m(t)$, must satisfy $P_j(g_1(t),\ldots,g_j(t)) = 0$ for $j=2,\ldots,m$. Moreover, $g_1(t)$ must have the form $g_1(t) = t+\zeta$, where ζ is an algebraic number.

(iii) An adequate descriptor α for the solution $g_2,\ldots,g_m$ of the triangular monic system $P_2 = \ldots = P_m = 0$. This (finite) descriptor will be used to represent the nonconstant components of the curve f in finite terms in a unique manner.

Our aim is to carry such a family of incidence curves through the whole inductive construction of the Collins decomposition, and at each stage to use them to determine which $(d-1)$-dimensional cells c' are faces of a given d-dimensional cell c. To this end, let K be a well-based decomposition of E^r, and let K' be the base decomposition of K. Suppose inductively that for each cell b of K' and each of its faces b' of one less dimension an incidence curve for the pair b, b' is available. Let c be a d-dimensional cell of K, and let c' be one of its faces. Then, by Lemma 3 of Section 2, either $c = b_j^*$ for some b ε K' and $c' = b_j$ or $c' = b_{j+1}$, or $c = b_j^*$ and $c' = b_j'^*$ where b' is a $(d-2)$-dimensional face of (the $(d-1)$-dimensional) cell b, or $c = b_j$ and $c' = b_j'$ where b' is a $(d-1)$-dimensional face of (the d-dimensional) cell b.

Let $P(x_1,\ldots,x_r)$ be the base polynomial of K; since K is well-based, the leading coefficient of P (regarded as a polynomial in x_r) is constant. If $c = b_j^*$ and $c' = b_j$, then we can take $[\zeta_1,\ldots,\zeta_{r-1}]$ to be any algebraic point in b, and then if ζ_r is the j-th root of $P(\zeta_1,\ldots,\zeta_{r-1},x_r)$ we can put $f(t) = [\zeta_1,\ldots,\zeta_{r-1},\zeta_r+t]$,

thereby defining an incidence curve. Of course, an entirely similar construction yields an incidence curve for the pair c_j^*, c_{j+1}.

Next suppose that $c = b_j^*$ and $c' = b_j'^*$, where b' is a $(d-2)$-dimensional face of b. Let $f' = [f_1(t),\ldots,f_{r-1}(t)]$ be an incidence curve for the pair b, b', let $g_1(t)=\zeta+t, g_2(t),\ldots,g_m(t)$ be the nonconstant components of this curve, enumerated in left-to-right order, and let $P_2,\ldots,P_m$ be the associated triangular monic system of polynomials. Put $\zeta_j=f_j(0)$ for $j=1,\ldots,r-1$, so that the point $Z = [\zeta_1,\ldots,\zeta_{r-1}]$ is an algebraic point in b'. Let P be as in the preceding paragraph, let η be the J-th real root of $P(\zeta_1,\ldots,\zeta_{r-1},y) = 0$ (or a sufficiently small rational number if $J=0$), and let η' be the $(J+1)$-st root of this same equation (or a sufficiently large rational number if this equation has only J real roots). Put $\zeta_r = \frac{1}{2}(\eta+\eta')$, and $f(t) = [f_1(t),\ldots,f_{r-1}(t),\zeta_r]$. This is plainly an incidence curve for the pair c, c', whose nonconstant components are plainly the same as those of f'. We have an adequate descriptor for the nonconstant components of f', and plainly this can also serve as an adequate descriptor for f.

Finally, suppose that $c = b_j$ and $c' = b_j'$, where b' is a $(d-1)$-dimensional face of b. Again, let $f'(t) = [f_1(t),\ldots,f_{r-1}(t)]$, $[P_2,\ldots,P_{m-1}]$, and $\alpha' = [\alpha_2,\ldots,\alpha_{m-1}]$ be an incidence curve for the pair b, b'. As in Definition 5, let $g_1,\ldots,g_{m-1}$ be the nonconstant components of f', so that $g_1(t) = t+\zeta$ where ζ is algebraic, and put $\alpha_1(t) = t+\zeta$ for convenience. Extend the truncated fractional series α_j, $j=2,\ldots,m-1$ to series α_j^* such that $g_j - \alpha_j^* = 0(t^D)$, where D will be chosen below. Let P be the same base polynomial as in the two preceding paragraphs, let d be its degree, and let P^* be the polynomial obtained from P by replacing each variable x_j in it for which $f_j(t) = \zeta_j$ is an algebraic constant by ζ_j. Clearly, the degree of P^* is at most d. Having P^*, we can then find all distinct real roots $y_1,\ldots,y_\ell$ of the equation $P^*(\alpha_1(0),\ldots,\alpha_{m-1}(0),y) = 0$ and arrange these solutions in increasing order. By Lemma 2 of Section 2, these roots are in 1-1 correspondence with the sequence of 'section' cells b_k' of K whose base cell is b'. Next, for each such y_k, we can find all the fractional

series solutions of the monic equation $P^*(\alpha_1^*(t),\ldots,\alpha_{m-1}^*(t),y) = 0$ whose constant term is y_k; all terms of these series up to those of order t^D must be constructed. Let $\alpha(t)$ designate any one of the truncated fractional series constructed in this way. Since $g_j - \alpha_j^* = O(t^D)$ for $j=1,\ldots,m-1$, we have $P^*(g_1(t),\ldots,g_{m-1}(t),\alpha(t)) = O(t^D)$ and thus by Lemma 3 there exists a solution g of $P^*(g_1(t),\ldots,g_{m-1}(t),g(t)) = 0$ such that $\alpha(t)-g(t) = O(t^{D/d})$. Append the polynomial P^* to the triangular monic system $P_2,\ldots,P_{m-1}$, thus obtaining a larger system $P_2,\ldots,P_m$, and let $R_2,\ldots,R_m$ and $\delta_2,\ldots,\delta_m$ be respectively the resultant system and the resultant degree sequence of this extended system. Arguing as in the first paragraph of the proof of Lemma 4, we see that $R_m(g_1(t),g(t)) = R_m(t+\zeta,g(t)) = 0$. Hence if we choose $D = d\,\delta_m(2\,\delta_m-1)$ it follows by Lemma 2 that this equation, together with the truncated fractional series α and the relationship $g-\alpha = O(t^{D/d})$, determine g uniquely, and in particular (arguing as in the proof of the Corollary to Lemma 5) that g is real for $t > 0$ if and only if all the coefficients of the fractional series α are real.

Dropping all those fractional series α which involve any nonreal coefficient, we can then go on to append any of the remaining α to $\alpha_1,\ldots,\alpha_{m-1}$, thereby obtaining an adequate descriptor for the nonconstant components of the curve $f = [f_1,\ldots,f_{r-1},g]$, where g is the unique (real) solution of $R_m(t+\zeta,g(t))$ represented by α. Let us agree to compare any two fractional series $\beta(t)$, $\gamma(t)$ with real coefficients by comparing coefficients of like terms lexicographically; this is equivalent to agreeing that $\beta < \gamma$ if $\beta(t) < \gamma(t)$ for all sufficiently small nonnegative t. Then it is plain that for each y_k the real solutions of $P^*(g_1(t),\ldots,g_{m-1}(t),g(t)) = 0$ are in 1-1 ordered correspondence with the truncated fractional series α which we retain and which satisfy $\alpha(0) = y_k$. The number of such series is therefore the number of distinct real roots into which the root y_k of $P(\zeta_1,\ldots,\zeta_{r-1},y) = 0$ splits as $f'(t)$ moves from the point $f'(0) = [\zeta_1,\ldots,\zeta_{r-1}]$ of b' to immediately neighboring points inside b. The sequences $P_2,\ldots,P_{m-1},P^*$ $g_1,\ldots,g_{m-1},g$, and $\alpha_1,\ldots,\alpha_{m-1},\alpha$ clearly define incidence curves for all pairs of cells having the form b_j, b_j', such that the latter is part of the boundary of the former.

.

This finishes our inductive construction of incidence curves, and also shows how the information on root splitting needed to determine the mappings $\rho(b,b')$, defined in the paragraph preceding Lemma 3 of Section 2, can be calculated; so that description of a technique for testing the adjacency of Collins cells of general dimension is now complete.

Remark: Given two Collins cells and Tarski sentences defining each of them, we can easily write a quantified Tarski sentence which is true if and only if they are adjacent. Thus in principle we can obtain an adjacency-testing algorithm using standard decision-procedures for Tarski sentences. However, since the number of quantified variables required to define adjacency of Collins cells in this way can be large (compare e.g. the explicit Tarski sentence defining a Collins cell given at the end of Appendix A), this technique will probably be much less efficient than that described in the present Appendix.

Acknowledgement

Bibliography

[AHU] A. Aho, J. Hopcroft and J. Ullman [1974] The Design and Analysis of Computer Algorithms. Addison-Wesley, Reading, Mass.

[Ak] A.G. Akritas [1980] The Fastest Exact Algorithm for the Isolation of the Real Roots of a Polynorial Equation. Computing v. 24, pp. 299-313.

[Ar] D.S. Arnon. [1981] Algorithms for the Geometry of Semi-Algebraic

Sets. Tech. Rept. 436, Computer Science Department, Univerity
of Wisconsin, Madison.

[Ar2] D.S. Arnon [1982] Automatic Analysis of Real Algebraic Curves.
Tech. Rept., Computer Science Department, Purdue University.

[BT] W.S. Brown and J.F. Traub [1971], On Euclid's Algorithm and the
Theory of Subresultants. JACM v.18, pp. 505-514.

[Ch] P.J. Cohen [1969] Decision Procedures for Real and p-adic Fields.
Comm. Pure and Appl. Math. v.22,pp.131-151

[Co] G. Collins [1975] Quantifier Elimination for Real Closed Fields by
Cylindrical Algebraic Decomposition. in: Second GI Conference
on Automata Theory and Formal Languages, Lecture Notes in
Computer Science v.33, Springer Verlag, Berlin. pp. 134-183.

[CL] G.E. Collins and R. Loos [1976] Polynomial Real Root Isolation
by Differentiation. in: SYMSAC 76 (Proc. of the 1976 ACM
Symp. on Symbolic and Algebraic Computations, Yorktown
Heights, N.Y.) pp. 15-25.

[CF] G.E. Cooke and R.L. Finney [1967] Homology of Cell Complexes.
Mathematical Notes, Princeton University Press, Princeton,
New Jersey.

[Ha] W.R. Hamilton [1969] Elements of Quaternions, Chelsea Pub. Co.
New York.

[He] L.E. Heindel [1971] Integer Arithmetic Algorithms for Polynomial
Real Zero Determination. JACM v. 22, pp. 533-549.

[Hi] H. Hironaka [1975] Triangulations of Algebraic Sets. in: Proc.
Symposia in Pure Math. v. 29, A   ican Math. Soc., Providence.
R.I., pp. 165-185.

[IKP] M.B. Ignat'yev, F.M. Kulakov and A.M. Pokrovskiy [1973] Robot
        Manipulator Control Algorithms. Report No. JPRS 59717, NTIS
        Springfield, Va.

[Ke] O.H. Keller [1974] Vorlesungen Uber Algebraische Geometrie.
        Akademische Verlagsgesellschaft, Geest & Portig, Leipzig.

[KT] H.T. Kung and J.F. Traub [1978] All Algebraic Functions can be
        Computed Fast. JACM v. 25, pp. 245-260.

[La] D. Lazard [1977] Algebre Lineaire sur $K[X_1,\ldots,X_n]$ et
        Elimination. Bull. Soc. Math. France, v. 105, pp. 165-190.

[Li] J. Lipson [1976] Newton's Method: A Great Algebraic Algorithm.
        in: SYMSAC 76 (proc. ACM Symp. on Symbolic and Algebraic
        Computations, Yorktown Heights, N.Y.), pp. 260-270.

[LPW] T. Lozano-Perez and M. Wesley [1979] An Algorithm for Planning
        Collision-Free Paths Among Polyhedral Obstacles. CACM v. 22,
        pp. 560-570.

[Ma] K. Mahler [1964] An Inequality for the Discriminant of a
        Polynomial. Michigan Math. J. v. 11, pp. 257-262.

[Mar] M. Marden [1949] The Geometry of Zeroes of a Polynomial in a
        Complex Variable. Mathematical Surveys v. 3, Amer. Math.
        Soc., Providence, R.I.

[Mi] M. Mignotte [1976] Some Problems About Polynomials. in: SYMSAC
        76 (Proc. 1976 ACM Symp. on Symbolic and Algebraic Computations,
        Yorktown Heights, N.Y.), pp. 227-228.

[Mo] R. Moenck [1973] Studies in Fast Algebraic Algorithms. Ph.D.
        Dissertation, University of Toronto.

[Re] J. Reif [1979] Complexity of the Mover's Problem and

Generalizations. Proc. 20th Symp. on the Foundations of
Computer Science, pp. 421-427.

[Ru] S. Rump [1976] On the Sign of a Real Algebraic Number. in:
SYMSAC 76 (Proc. 1976 ACM Symp. on Symbolic and Algebraic
Computations, Yorktown Heights, N.Y.), pp. 238-241.

[Sc] J.T. Schwartz [1980] Fast Probabilistic Algorithms for
Verification of Polynomial Identities, JACM v. 27, pp.
701-717.

[SS] J. Schwartz and M. Sharir [1981] On the 'Piano-Movers' Problem; I.
The case of a Two-Dimensional Rigid Polygonal Body Moving
Amidst Polygonal Barriers. Computer Science Technical Report
39, Courant Institute, New York University. To appear in
Comm. Pure and Appl. Math.

[Ta] A. Tarski [1951] A Decision Method for Elementary Algebra and
Geometry. (Second Edition, Rev.) University of California
Press.

[Ud] S. Udupa [1977] Collision Detection and Avoidance in Computer-
Controlled Manipulators, Ph.D. Dissertation, California Inst.
of Technology, Passadena, Ca.

[Wa] B.L. Van der Waerden [1939] Einfuhrung in die Algebraische
Geometrie, Springer Pub. Co., Berlin.

[Wa2] B.L. Van der Waerden [1960] Algebra (5th Edition) Springer
Pub. Co., Berlin.

This book may be kept

FOURTEEN DAYS

A fine will be charged for each day the book is kept overtime.

ILL 4·30·86			
GAYLORD 142			PRINTED IN U S A

N.Y.U. Courant Institute of
Mathematical Sciences
251 Mercer St.
New York, N. Y. 10012