SOME EXAMPLES OF THE USE OF DISTANCES AS COORDINATES FOR EUCLIDEAN GEOMETRY

TIMOTHY F. HAVEL*

Abstract. Distance geometry provides us with an implicit characterization of the Euclidean metric in terms of a system of polynomial equations and inequalities. With the aid of computer algebra programs, these equations and inequalities in turn provide us with a coordinate-free approach proving theorems in Euclidean geometry. This paper contains a brief summary of the mathematical results on which this approach is based, together with some examples showing how it is applied. In particular, we show how it can be used to derive the topological structure of a simple linkage mechanism.

1. Introduction. Distance geometry may be defined as the classification and study of geometric spaces by means of the metrics that can be defined on them [Blumenthal, 1953, 1970]. It has been used to characterize Euclidean spaces [Menger, 1928, 1931; Schoenberg, 1937; Blumenthal, 1961], hyperbolic and elliptic spaces [Seidel, 1952, 1955], and Riemannian manifolds of constant curvature in general [Berger, 1981, 1985]. The Euclidean version also has a number of interesting applications to multidimensional scaling [Gower, 1982, 1985] and to molecular geometry [Crippen & Havel, 1988]. One of the key results on which all this work is based is Menger's intrinsic characterization of the Euclidean metric [Menger, 1928], which has the form of a system of polynomial equations and inequalities in the interpoint distances squared. If one understands the geometric interpretations of these polynomials, they can also be used to express a variety of common geometric conditions algebraically, and hence to use the *distances as coordinates* to prove theorems in Euclidean geometry. At least in principle, *all* of the theorems of Euclidean geometry can be derived in this way [Dress & Havel, 1987].

These polynomials can be written most succinctly as certain type of determinant. If $D(a_1, a_2)$ denotes the *squared* distance between a pair of Euclidean points labeled a_1, $a_2 \in A$, and $[b_1, ..., b_m]$, $[c_1, ..., c_m] \in A^m$ denote two m-element sequences of points, the *Cayley-Menger bideterminant* of these sequences is:

$$\begin{aligned} &D(b_1, ..., b_m;\ c_1, ..., c_m) \\ &:= 2\left(\frac{-1}{2}\right)^m \det \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & D(b_1, c_1) & D(b_1, c_2) & \dots & D(b_1, c_m) \\ 1 & D(b_2, c_1) & D(b_2, c_2) & \dots & D(b_2, c_m) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & D(b_m, c_1) & D(b_m, c_2) & \dots & D(b_m, c_m) \end{pmatrix} . \end{aligned} \tag{1}$$

*Institute of Science and Technology, University of Michigan, 2200 Bonisteel Blvd., Ann Arbor, MI 48109, U.S.A.

(The reason for the constant factor in this definition will become clear when the geometric meaning of these determinants is explained below). Since in many cases of interest the two sequences are the same, it will be convenient to abbreviate $D(a_1, ..., a_m;\ a_1, ..., a_m)$ by $D(a_1, ..., a_m)$, which is called simply a Cayley-Menger determinant. Observe that the Cayley-Menger determinant of a pair of elements a_1, $a_2 \in A$ is the same as the squared distance between them; thus our use of the symbol "$D(a_1, a_2)$" for both their squared distance and their Cayley-Menger determinant is consistent.

Let us now state Menger's characterization itself in two algebraically distinct but equivalent forms. Their geometric meanings are explained below.

THEOREM 0. *Let A be a set and $D : A \times A \to \mathbf{R}$ be a function such that for all $a, b, c \in A$:*

(i) $D(a, a) = 0$;
(ii) $D(b, c) \geq 0$;
(iii) $D(b, c) = D(c, b)$.

Then the following statements are equivalent.

(I) *There exists a function $\mathbf{p} : A \to \mathbf{R}^n$ such that $D(b, c) = \|\mathbf{p}(b) - \mathbf{p}(c)\|^2$ for all $b, c \in A$.*
(II) *For any positive integer m and sequence $[a_1, ..., a_m] \in A^m$ we have $D(a_1, ..., a_m) \geq 0$, and $D(a_1, ..., a_m) = 0$ whenever $m \geq n + 2$.*
(III) *For any positive integer m and two sequences $[b_1, ..., b_m], [c_1, ..., c_m] \in A^m$, we have $D^2(b_1, ..., b_m;\ c_1, ..., c_m) \leq D(b_1, ..., b_m)D(c_1, ..., c_m)$ with equality holding whenever $m = n + 1$.*

Hence $[D(a_i, a_j) \mid 1 \leq i, j \leq \#A]$ is a matrix of squared distances among a set of points in a Euclidean space if and only if either (II) or (III) (and hence both) are satisfied. For a proof of this theorem, the reader is referred to either [Blumenthal, 1970] or [Crippen & Havel, 1988].

To make these determinants seem more familiar, let us use the Pythagorian theorem for the plane $D(2, 3) = (x(2) - x(3))^2 + (y(2) - y(3))^2$ to translate $D(1, 2, 3)$ into Cartesian coordinates:

$$\begin{aligned} D(1,2,3) &= \frac{1}{4}\big(2\,D(1,2)D(1,3) + 2\,D(1,2)D(2,3) + 2\,D(1,3)D(2,3) \\ &\qquad - D^2(1,2) - D^2(1,3) - D^2(2,3)\big) \\ &= (x(1)y(2) - x(2)y(1) + x(3)y(1) - x(1)y(3) + x(2)y(3) - x(3)y(2))^2 \\ &= \det{}^2 \begin{pmatrix} 1 & 1 & 1 \\ x(1) & x(2) & x(3) \\ y(1) & y(2) & y(3) \end{pmatrix} \end{aligned} \tag{2}$$

Thus we see that $D(1, 2, 3)$ is four times the squared area of the triangle with side lengths $d(1, 2) := D^{\frac{1}{2}}(1, 2)$, $d(1, 3) := D^{\frac{1}{2}}(1, 3)$ and $d(2, 3) := D^{\frac{1}{2}}(2, 3)$. On performing this substitution in $D(1,2,3)$ and factorizing, we also find

$$\begin{aligned} D(1,2,3) \;=\; 1/4 \cdot &(d(1,2)+d(1,3)+d(2,3)) \cdot \\ &(d(1,2)+d(1,3)-d(2,3)) \cdot \\ &(d(1,2)-d(1,3)+d(2,3)) \cdot \\ &(-d(1,2)+d(1,3)+d(2,3)) \;, \end{aligned} \tag{3}$$

This is known as *Heron's formula* [Coxeter, 1969]. More generally, $D(a_1, ..., a_m)$ is $(m-1)!$ times the squared hypervolume of the simplex spanned by the points $a_1, ..., a_m$ of a Euclidean space.

Similarly, on applying the Pythagorian theorem to the Cayley-Menger bideterminant $D(1,2;\,1,3)$, we obtain

$$D(1,2;\,1,3) \;=\; (x(2)-x(1))\cdot(x(3)-x(1)) \;+\; (y(2)-y(1))\cdot(y(3)-y(1)) \tag{4}$$

i.e. the dot product of the vectors from 1 to the other two vertices of the triangle. In terms of the lengths of its sides, we get

$$D(1,2;\,1,3) \;=\; 1/2\,(D(1,2)+D(1,3)-D(2,3)) \;, \tag{5}$$

which is just the law of cosines for the dot-product. More generally, for four points $\{1,2,3,4\}$ in a Euclidean space $D(1,2;\,3,4)$ is the dot product of the vectors $\overrightarrow{12}$ and $\overrightarrow{34}$, and analogous geometric interpretations hold for the higher-order Cayley-Menger bideterminants [Havel & Dress, 1987; Crippen & Havel, 1988]. These interpretations are already enough to enable us to show that a rather wide variety of geometric conditions are equivalent to the vanishing of polynomials in the interpoint distances (squared). In Table 1, we provide a short list of common geometric conditions and their algebraic expression in terms of both planar generic Cartesian coordinates and distances.

Table 1

Geometric Condition	Cartesian Expression	Distance Expression
$\overline{12} \cong \overline{34}$ (*congruence*)	$(x(1)-x(2))^2+(y(1)-y(2))^2$ $-(x(3)-x(4))^2-(y(3)-y(4))^2$	$D(1,2)-D(3,4)$
$\overline{12} \perp \overline{13}$ (*perpendicularity*)	$(x(2)-x(1))(x(3)-x(1))$ $+(y(2)-y(1))(y(3)-y(1))$	$D(1,2)+D(1,3)-D(2,3)$

$((123))$ (*collinearity*)	$x(1)y(2) - x(2)y(1) + x(3)y(1)$ $- x(1)y(3) + x(2)y(3) - x(3)y(2)$	$2\,D(1,2)D(1,3) + 2\,D(1,2)D(2,3)$ $+ 2\,D(1,3)D(2,3) - D^2(1,2)$ $- D^2(1,3) - D^2(2,3)$
$\overline{12} \parallel \overline{34}$ (*parallelism*)	$(x(2) - x(1))(y(4) - y(3))$ $-(y(2) - y(1))(x(4) - x(3))$	$4\,D(1,2)\,D(3,4) - (D(1,3) +$ $D(2,4) - D(1,4) - D(2,3))^2$

It can be seen from the table that certain geometric conditions (e.g. congruence and perpendicularity) can be expressed more simply in terms of the distances than in terms of Cartesian coordinates. Of course, by a suitable choice of coordinate system the Cartesian expressions can be simplified substantially; for example, by placing point 1 at the origin and point 2 on the y-axis, perpendicularity becomes merely $y(2)y(3)$. When many polynomials in the Cartesian coordinates are needed to express a number of simultaneous geometric conditions, however, it is not always possible to find a single coordinate system that reduces them all to their simplest forms, especially in more than two dimensions. Moreover, when a polynomial describing a given geometric condition can depend upon our choice of coordinate system, it is more difficult to recognize it when it occurs among our results, and hence to understand the geometric meaning of those results. Since the distances are independent of our choice of coordinate system, they are also free of this particular problem. The use of distances as coordinates in fact offers the same advantages in Euclidean geometry that invariant formulations of geometric problems have more generally, as described in [Whiteley, 1989].

The drawback of using distances as coordinates is that the number of variables occurring in the polynomial equations is usually substantially higher than it is with Cartesian coordinates, because it is rarely possible to use only a small subset of the $\binom{N}{2}$ distances among N points in the course of a proof. In addition, because the vanishing of Cayley-Menger determinants of $n+2$ points is necessary to ensure that the configuration is n-dimensional, the total degree of the equations is at least $n+1$ in the squared distances. Nevertheless, it is reasonable to hope that there exists a canonical Gröbner basis for the ideals generated by these polynomials, as is known for the analogous ideals in projective geometry; by the results of [Sturmfels & White, 1989], reduction versus this projective Gröbner basis corresponds to the classical straightening algorithm for the Grassmann variety. If such a system of Gröbner bases can be found for Euclidean geometry, it is likely that in many cases the approach outlined here will become competitive with traditional approaches based on Cartesian coordinates. Because of its generality, it is also possible that the "distance geometry approach" would then provide a convenient framework in which to automate the proofs of theorems in Euclidean geometry, as has been done using Cartesian coordinates in e.g. [Chou, 1987; Kutzler, 1989; Kapur & Mundy, 1989]. We make no claim, however, that the proofs given in this paper are either automatic or automatable.

2. Examples. We now consider several examples which show how the distance geometry approach can actually be put into practice. In all of these examples, we have used the computer algebra program MAPLE [Char *et al.*, 1986] to perform the computations.*

2.1. ISOSCELES BISECTORS: As a first, very simple example, we show that the line between the odd vertex of an isosceles triangle is perpendicular to the base if (and only if) it bisects the base. If we number the vertices of the triangle $1, 2, 3$, where 1 is the odd vertex, and let a be the point which bisects the base, the hypotheses are:

(I)	$4\,D(1,2,3,a) = 0$	(*the triangle is coplanar with the point* a);
(II)	$D(1,2) - D(1,3) = 0$	(*the triangle is isosceles*);
(III)	$4\,D(2,3,a) = 0$	(*the point* a *is on the base* $\overline{23}$);
(IV)	$D(2,a) - D(3,a) = 0$	(*the point* a *bisects the base*).

By the Pythagorean theorem, the desired conclusion is:

$$\text{(V)} \qquad 2\,D(1,a;\ 2,a) = D(1,a) + D(2,a) - D(1,2) = 0 \qquad (\overline{1a} \perp \overline{2a}).$$

Figure 1.
Illustration of Example 2.1.

Note that, when the distance function $\sqrt{D}$ is non-Euclidean, the vanishing of the three-point Cayley-Menger determinant $D(2,3,a)$ does *not* imply the collinearity of $\{2,3,a\}$, since these may span a degenerate subspace (cf. [Snapper & Troyer, 1972]). Hence condition (I) is algebraically independent of (III). It can be derived from (III) only if we make use of the *in*equalities characteristic of the Euclidean metric (cf. Theorem 0). This is most easily done by using *Seidel's identity*:

$$(6) \qquad D(2,3)D(1,2,3,a) \;=\; D(1,2,3)D(2,3,a) - D^2(1,2,3;\ 2,3,a)$$

Since in a Euclidean space $D(2,3)D(1,2,3,a) \geq 0$, it follows that $D(2,3,a) = 0$ implies $D(1,2,3,a) = 0$ (unless $D(2,3) = 0$, i.e. the triangle is degenerate). Nevertheless, in applying these methods it is often very convenient to assume an elementary knowledge of Euclidean geometry, in order to avoid the more difficult task of arguing with inequalities.

To prove now our claim, we consider the Cayley-Menger determinant in hypothesis (I):

$$(7) \qquad 4\,D(1,2,3,a) \;=\; \frac{1}{2}\det\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & D(1,2) & D(1,3) & D(1,a) \\ 1 & D(1,2) & 0 & D(2,3) & D(2,a) \\ 1 & D(1,3) & D(2,3) & 0 & D(3,a) \\ 1 & D(1,a) & D(2,a) & D(3,a) & 0 \end{pmatrix}.$$

*This program and its documentation are available from the Symbolic Computation Group at the Univ. of Waterloo, Ontario. Those interested in a general introduction to the theory and applications of computer algebra programs are referred to [Davenport *et al.*, 1988].

Observe that by subtracting the third row from the fourth in this matrix and then expanding the determinant along the fourth row, it can be written in the following form:

$$D(1,2,3,a) \;=\; E_1 \cdot D(2,3) \;+\; E_2 \cdot (D(1,3) - D(1,2)) \;+\; E_3 \cdot (D(3,a) - D(2,a)) \;, \tag{8}$$

where E_1, E_2 and E_3 are polynomials. Hence if we use conditions (II) and (IV) to substitute $D(1,3)$ by $D(1,2)$ and $D(3,a)$ by $D(2,a)$ then, assuming the nondegeneracy condition $D(2,3) \neq 0$, we obtain an equation

$$\begin{aligned} E_1 \;&=\; \frac{1}{2\,D(2,3)} \det \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & D(1,2) & D(1,2) & D(1,a) \\ 1 & D(1,2) & 0 & D(2,3) & D(2,a) \\ 0 & 0 & D(2,3) & -D(2,3) & 0 \\ 1 & D(1,a) & D(2,a) & D(2,a) & 0 \end{pmatrix} \\ &= 4\,D(1,2,a) - D(2,3)\,D(1,a) \;=\; 0 \end{aligned} \tag{9}$$

that is *linear* in $D(2,3)$. Assuming $D(1,a) \neq 0$, this may be used to eliminate $D(2,3)$ from condition (III) by taking the pseudoremainder† of $4\,D(2,3,a)$ by E_1 with respect to $D(2,3)$, i.e.

$$prem(4\,D(2,3,a), E_1\,, D(2,3)) \;=\; 16\,D(1,2,a) \cdot D^2(1,a;\,2,a) \;=\; 0\;. \tag{10}$$

Thus, assuming the nondegeneracy condition $D(1,2,a) \neq 0$, the conclusion (V) follows.

To prove the converse (obtained by exchanging the conclusion (V) with the hypothesis (IV) above), we use (II) together with (V) to substitute for $D(1,3)$ and $D(1,2)$, respectively, and then take the resultant of (I) and (III) with respect to $D(2,3)$, obtaining after factorization:

$$D^2(2,a) \cdot (D(2,a) - D(3,a))^4 \;=\; 0\;. \tag{11}$$

The details are left to the reader.

2.2. The Congruence of Opposite Sides of a Parallelogram: As a slightly less trivial example demonstrating the use of Gröbner bases [Buchberger, 1985] in solving these problems, we prove that opposite sides of a nondegenerate planar parallelogram are pairwise equal in length. As hypotheses, we have:

(I)	$P(1,2;\,3,4) \;:=\; 4\,D(1,2)D(3,4) - 4\,D^2(1,2;\,3,4) = 0$	($\overline{12} \parallel \overline{34}$);
(II)	$P(1,4;\,2,3) \;:=\; 4\,D(1,4)D(2,3) - 4\,D^2(1,4;\,2,3) = 0$	($\overline{14} \parallel \overline{23}$);
(III)	$4\,D(1,2,3,4) = 0$	(*coplanarity*).

The conclusions we wish to derive are:

(IV) $D(1,2) - D(3,4) = 0$ ($\overline{12} \cong \overline{34}$)

†See [Chou, 1987] for a definition and examples of the use of the pseudoremainder function. In this simple case, it is the same as the resultant.

and

(V) $\quad D(2,3) - D(1,4) = 0 \qquad\qquad (\overline{23} \cong \overline{14}).$

Conditions (I) and (II) are the parallelism conditions given in Table 1. Once again, these conditions imply (III) whenever the metric is Euclidean, but the arguments required to establish this fact are relatively difficult.

Figure 2.
Illustration of Example 2.2.

In order to eliminate the "diagonal" squared distances $D(1,3)$ and $D(2,4)$, we compute the Gröbner basis of (I) — (III) with respect to the lexicographic monomial order induced by the variable ordering $[D(1,3), D(2,4), D(1,2), D(2,3), D(3,4), D(1,4)]$. The resultant Gröbner basis $\mathbf{G}$ contains 11 polynomials, one of which does not depend on either $D(1,3)$ or $D(2,4)$, as desired (for a detailed account of how one uses Gröbner basis computations to perform eliminations, see [Buchberger, 1985]). This polynomial $R(1,2,3,4)$, which is homogeneous of total degree 4 and has 35 terms, cannot itself be factored. If we define the polynomial map $D(i,j) \mapsto d^2(i,j)\ \forall\ 1 \le i < j \le 4$, however, $R(1,2,3,4)$ becomes a polynomial of total degree 8 in the $d(i,j)$ that factors into a product of linear terms:

$$
\begin{aligned}
r(1,2,3,4) \;:=\; & -(d(3,4) + d(1,2) + d(2,3) + d(1,4)) \cdot \\
& (d(3,4) - d(1,2) - d(2,3) - d(1,4)) \cdot \\
& (d(3,4) - d(1,2) + d(2,3) + d(1,4)) \cdot \\
& (d(3,4) + d(1,2) - d(2,3) + d(1,4)) \cdot \\
& (d(3,4) + d(1,2) + d(2,3) - d(1,4)) \cdot \\
& (d(3,4) + d(1,2) - d(2,3) - d(1,4)) \cdot \\
& (d(3,4) - d(1,2) - d(2,3) + d(1,4)) \cdot \\
& (d(3,4) - d(1,2) + d(2,3) - d(1,4))\ .
\end{aligned}
\tag{12}
$$

Interestingly enough, this polynomial is (up to sign) also obtained by taking the resultant of $4\,D(1,2,3)$ and $4\,D(1,3,4)$ with respect to $D(1,3)$. It vanishes if and only if at least one of the following holds:

(1) One or more of the first five factors f_1 through f_5 in the above equation vanishes;
(2) The sixth factor f_6 vanishes;
(3) Either seventh factor f_7, the eighth factor f_8 or both vanish.

Case (1) obviously implies that the parallelogram is degenerate (i.e. collinear or copunctual), and hence need not be further considered.

To take care of case (2), we consider another one of the polynomials $s(1,2,3,4)$ in our Gröbner basis $\mathbf{G}'$ after the transformation $D(i,j) \mapsto d^2(i,j)$ (which we have chosen simply because we are able to derive the desired conclusion from it). Since $f_6 = 0$, we may substitute $d(1,4) := d(1,2) + d(3,4) - d(2,3)$ in $s(1,2,3,4)$ to obtain

$$\begin{aligned} s'(1,2,3,4) &= (d(1,2)+d(3,4))^2 \cdot (d(2,3)-d(3,4)) \cdot (d(1,2)-d(2,3)) \cdot \\ &\quad (d(2,3)-d(3,4)-d(2,4))^2 \cdot (d(2,3)-d(3,4)+d(2,4))^2 \, . \end{aligned} \tag{13}$$

It follows that if the parallelogram is nondegenerate (i.e. noncollinear) then either $d(3,4) = d(2,3) \implies d(1,4) = d(1,2)$ or else $d(1,2) = d(2,3) \implies d(1,4) = d(3,4)$. If we substitute $d(3,4) := d(2,3)$ and $d(1,4) := d(1,2)$, however, then we find another polynomial $t(1,2,3,4)$ in our transformed Gröbner basis $\mathbf{G}'$ which becomes

$$t'(1,2,3,4) \;=\; -4\,d(2,4)^2 \cdot (d(1,2)+d(2,3))^2 \cdot (d(1,2)-d(2,3))^2 \, . \tag{14}$$

Hence our parallelogram is an equilateral quadrilateral, and in particular conclusions (IV) and (V) hold. The same result can be proved by an analogous argument if $d(1,2) = d(2,3)$ and $d(1,4) = d(3,4)$.

Finally, to handle case (3), suppose $f_8 = 0$. Then on making the substitution $d(1,4) := d(2,3) + d(3,4) - d(1,2)$ in $s(1,2,3,4)$ we get

$$\begin{aligned} s''(1,2,3,4) &= (d(2,3)+d(3,4)) \cdot (d(1,2)-d(2,3)) \cdot (d(1,2)-d(3,4))^2 \cdot \\ &\quad (d(2,3)+d(3,4)-d(2,4))^2 \cdot (d(2,3)+d(3,4)+d(2,4))^2 \, . \end{aligned} \tag{15}$$

Hence if the parallelogram is not degenerate either $d(1,2) = d(2,3) \implies d(1,4) = d(3,4)$ or else $d(1,2) = d(3,4) \implies d(1,4) = d(2,3)$. As shown in the previous paragraph, in the former case the parallelogram must be equilateral, whereas in the latter case opposite sides are congruent, as desired. The case $f_7 = 0$ is handled by an analogous argument.

It is worth noting that the well-known *Law of Parallelograms* is an immediate corollary of this result. For if we set $D(1,4) := D(2,3)$ and $D(3,4) := D(1,2)$, we find that every polynomial in our untransformed Gröbner basis $\mathbf{G}$ which does not vanish after this substitution has $2\,D(1,2) + 2\,D(2,3) - D(1,3) - D(2,4)$ as a factor; in particular we obtain

$$-8\,(D(1,2)-D(2,3))^2 \cdot (2\,D(1,2)+2\,D(2,3)-D(1,3)-D(2,4)) \;=\; 0 \tag{16}$$

Hence either the desired result holds or else the parallelogram is equilateral. In the latter case, however, on setting $D(1,2) := D(2,3) := D(3,4) := D(1,4)$ in our Gröbner basis, we obtain a polynomial

$$D^2(2,4)\,(D(1,3)+D(2,4)-4\,D(1,2)) \;=\; 0 \, , \tag{17}$$

which proves the same thing.

2.3. SIMSON'S THEOREM‡: Given three points $\{1,2,3\}$ in the plane together with a fourth point 4 which lies on their circumcircle, the feet of the perpendiculars a, b, c from the point 4 to the sides of the triangle $\overline{12}$, $\overline{13}$ and $\overline{23}$, respectively, are collinear.

‡After R. Simson; the theorem is actually due to W. Wallace, see [Johnson, 1929].

Figure 3.
Illustration of Example 2.3.

This theorem has become a favorite example for demonstrating automated proofs of geometric theorems (cf. [Chou, 1987]). The following proof, although it is not automatic, nevertheless provides a good illustration of how one translates geometry into algebra by using the distances as coordinates. To do this, we shall need one little-known fact about the Euclidean metric, together with its geometric interpretation: This is known as the *Ptolomeic inequality* [Johnson, 1929; Apostol, 1967], and states that given any four points in the plane $\{1,2,3,4\}$ and the distances $d(i,j)$ $(i,j=1,...,4)$ among them, we have

$$d(1,2)d(3,4) \;\le\; d(1,3)d(2,4)+d(1,4)d(2,3) \tag{18}$$

with equality if and only if $\{1,2,3,4\}$ are cocircular with $\{1,2\}$ separating $\{3,4\}$ on their mutual circumcircle. Since this is effectively a triangle inequality involving the products of pairs of distances, by substituting these products for the distances given in equation (3) we obtain a completely general and symmetric cocircularity condition:

$$\begin{aligned} 0 &= C(1,2,3,4) \\ &:= 2D(1,2)D(3,4)D(1,3)D(2,4)+2D(1,2)D(3,4)D(1,4)D(2,3)+ \\ &\quad 2D(1,3)D(2,4)D(1,4)D(2,3)-D^2(1,2)D^2(3,4)- \\ &\quad D^2(1,3)D^2(2,4)-D^2(1,4)D^2(2,3) \end{aligned} \tag{19}$$

The hypotheses of the theorem may now be formulated as follows:

(I)	$C(1,2,3,4) \;=\; 0$	(*points* $\{1,2,3,4\}$ *are cocircular*);
(II)	$D(1,4,a,b) \;=\; 0$	(*points* $\{1,4,a,b\}$ *are coplanar*);
(III)	$D(2,4,a,c) \;=\; 0$	(*points* $\{2,4,a,c\}$ *are coplanar*);
(IV)	$D(3,4,b,c) \;=\; 0$	(*points* $\{3,4,b,c\}$ *are coplanar*);
(V)	$D(1,2,a) \;=\; 0$	(*points* $\{1,2,a\}$ *are collinear*);
(VI)	$D(1,3,b) \;=\; 0$	(*points* $\{1,3,b\}$ *are collinear*);
(VII)	$D(2,3,c) \;=\; 0$	(*points* $\{2,3,c\}$ *are collinear*);
(VIII)	$D(1,a)+D(4,a)-D(1,4)=0$	$(\overline{1a}\perp\overline{4a})$;
(IX)	$D(2,a)+D(4,a)-D(2,4)=0$	$(\overline{2a}\perp\overline{4a})$;
(X)	$D(1,b)+D(4,b)-D(1,4)=0$	$(\overline{1b}\perp\overline{4b})$;
(XI)	$D(3,b)+D(4,b)-D(3,4)=0$	$(\overline{3b}\perp\overline{4b})$;
(XII)	$D(2,c)+D(4,c)-D(2,4)=0$	$(\overline{2c}\perp\overline{4c})$;
(XIII)	$D(3,c)+D(4,c)-D(3,4)=0$	$(\overline{3c}\perp\overline{4c})$.

The conclusion is simply:

(XIV) $D(a,b,c) \;=\; 0$ (*the points* $\{a,b,c\}$ *are collinear*).

Note that we do not need to make explicit use of all of the coplanarity conditions.

The proof will be simplified substantially if we make use of our knowledge of Euclidean geometry together with conditions (V) — (VII) to derive a preliminary result:

LEMMA. *Given a triangle* $\{1,2,3\}$ *and two points* a *and* b *on the lines* $\overline{12}$ *and* $\overline{13}$*, respectively:*

$$Q(1,2,3,a,b) \;:=\; 4\,D(1,2)D(1,3)D(1,a,b) - 4\,D(1,a)D(1,b)D(1,2,3) \;=\; 0 \tag{20}$$

PROOF: If the angle at the point 1 in the triangle $\{1,2,3\}$ is θ, the angle at the point 1 in the triangle $\{1,a,b\}$ is either θ or else $\pi-\theta$. Hence by the law of cosines, we have

$$\frac{D(1,2)+D(1,3)-D(2,3)}{\sqrt{D(1,2)D(1,3)}} \;=\; \pm\frac{D(1,a)+D(1,b)-D(a,b)}{\sqrt{D(1,a)D(1,b)}} \tag{21}$$

By rearranging this equation, squaring both sides and collecting terms appropriately, one obtains equation (20). (We note that, although we have made use of the Euclidean concept of angle in proving this Lemma, it is possible to prove it using only the cospatiality condition $D(1,2,3,a,b)=0$, the coplanarity conditions $D(1,2,3,a)=D(1,2,3,b)=0$ and the collinearity conditions $D(1,2,a)=D(1,3,b)=0$.)
In a similar fashion, one finds that

$$Q(2,3,1,c,a) \;:=\; 4\,D(1,2)D(2,3)D(2,a,c) - 4\,D(2,a)D(2,c)D(1,2,3) \;=\; 0 \tag{22}$$

and

$$Q(3,1,2,b,c) \;:=\; 4\,D(1,3)D(2,3)D(3,b,c) - 4\,D(3,b)D(3,c)D(1,2,3) \;=\; 0 \tag{23}$$

Proceeding now with the proof of the Theorem, we start by substituting for $D(4,a)$, $D(4,b)$ and $D(4,c)$ in $D(1,4,a,b)$, $D(2,4,a,c)$ and $D(3,4,b,c)$ using the perpendicularity conditions (VIII) through (XIII), obtaining polynomials which vanish by conditions (II) through (IV), and which may be factorized as:

$$P(1,4,a,b) \;:=\; 4D(1,4)D(1,a,b) - 4D(1,a)D(1,b)D(a,b) \;=\; 0 \tag{24}$$

$$P(2,4,a,c) \;:=\; 4D(2,4)D(2,a,c) - 4D(2,a)D(2,c)D(a,c) \;=\; 0 \tag{25}$$

$$P(3,4,b,c) \;:=\; 4D(3,4)D(3,b,c) - 4D(3,b)D(3,c)D(b,c) \;=\; 0\,. \tag{26}$$

Combining this result with equations (20), (22) and (23), we get:

$$\begin{aligned} 0 &= 4D(1,2,3)P(1,4,a,b) - 4D(a,b)Q(1,2,3,a,b) \\ &= 4D(1,a,b)\cdot(4D(1,4)D(1,2,3) - 4D(1,2)D(1,3)D(a,b)) \end{aligned} \tag{27}$$

$$\begin{aligned} 0 &= 4D(1,2,3)P(2,4,a,c) - 4D(a,c)Q(2,3,1,c,a) \\ &= 4D(2,a,c)\cdot(4D(2,4)D(1,2,3) - 4D(1,2)D(2,3)D(a,c)) \end{aligned} \tag{28}$$

$$\begin{aligned} 0 &= 4D(1,2,3)P(3,4,b,c) - 4D(b,c)Q(3,1,2,b,c) \\ &= 4D(3,b,c)\cdot(4D(3,4)D(1,2,3) - 4D(1,3)D(2,3)D(b,c))\,. \end{aligned} \tag{29}$$

Assuming the nondegeneracy conditions $D(1,a,b) \neq 0$, $D(2,a,c) \neq 0$ and $D(3,b,c) \neq 0$, we obtain:

$$
\begin{aligned}
&(30) \qquad R(1,2,3,4,a,b) := 4D(1,4)D(1,2,3) - 4D(1,2)D(1,3)D(a,b) &= 0\\
&(31) \qquad R(2,3,1,4,c,a) := 4D(2,4)D(1,2,3) - 4D(1,2)D(2,3)D(a,c) &= 0\\
&(32) \qquad R(3,1,2,4,b,c) := 4D(3,4)D(1,2,3) - 4D(1,3)D(2,3)D(b,c) &= 0\ .
\end{aligned}
$$

Finally, we solve these equations for $D(1,4)$, $D(2,4)$ and $D(3,4)$, respectively and use them to eliminate point 4 entirely from our cocircularity condition (I), obtaining:

$$(33) \qquad D^2(1,2) \cdot D^2(1,3) \cdot D^2(2,3) \cdot D^4(1,2,3) \cdot D(a,b,c) = 0\ .$$

Hence, assuming the nondegeneracy condition $D(1,2,3) \neq 0$, the Theorem follows.

2.4. TOPOLOGY OF THE EQUILATERAL PENTAGON LINKAGE: One of the most promising areas for the application of the distance geometry approach is to the study of *linkages*, i.e. mechanisms obtained by fastening together fixed length bars at flexible joints (and allowing the bars to pass through each other). Examples of this approach to the study of linkages may be found in [Schoenberg, 1969] and [Dress, 1982]. Here we shall present a new example in which we use distance geometry in conjunction with Morse theory to determine the topology of the configuration space of the linkage which is obtained by allowing the angles at the vertices of an equilateral planar pentagon to vary freely while preserving the lengths of its sides.

DEFINITIONS. The *configuration space* of the equilateral pentagon linkage consists of all five-point subsets $\{1,2,3,4,5\}$ of the Euclidean plane such that $d(i,i+1) = 1$ for all $i = 1,...,5$ ($i+1$ computed $\mod 5$). It can be defined analytically as the set of all possible Cartesian coordinates for such five-point subsets, with those members thereof which differ only by a translation and/or proper rotation identified. Let M be a smooth manifold (embedded in R^n, say) and let $f : \mathsf{M} \to \mathsf{R}$ be a smooth function. A *critical point* of f is any point of the manifold at which its gradient $\nabla f = 0$. The function f is called a *Morse function** if its Hessian $\nabla^2 f$ is nonsingular at all its critical points. The *index* of such a *nondegenerate* critical point is the number of -1's in the signature of its Hessian. A well-known result in Morse theory states that the Euler characteristic χ_M of the manifold is related to the number N_i of critical points of index i by the formula $\chi_M = \sum(-1)^i N_i$. In the case of a compact orientable two-dimensional manifold of genus G_M, the Euler characteristic is just $\chi_M = 2 - 2G_M$.

THEOREM. *The topological structure of the configuration space of the planar equilateral pentagon linkage is that of a compact, connected and orientable two-dimensional manifold of genus 4.*

*For a detailed account of Morse theory, the reader is referred to [Morse & Cairns, 1969].

Figure 4.
Illustration of Example 2.4.

To prove this theorem, we first show that the real algebraic variety that is obtained by setting the distances around the pentagon to unity is in fact a smooth manifold. To do this, let $\underline{\mathbf{p}} = [\mathbf{p}_1, ..., \mathbf{p}_5] \in \mathbf{R}^{10}$ denote a set of coordinates for the linkage, where $\mathbf{p}_i = [p_1{}^1, p_i{}^2] \in \mathbf{R}^2$ for $i = 1, ..., 5$, and denote by $\mathrm{vol}_{\underline{\mathbf{p}}}(i,j,k)$ the *oriented area* of the triangle $[\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k]$, which is given by

$$\frac{1}{2}\det\begin{pmatrix} 1 & 1 & 1 \\ p_i{}^1 & p_j{}^1 & p_k{}^1 \\ p_i{}^2 & p_j{}^2 & p_k{}^2 \end{pmatrix}. \tag{34}$$

We now define an atlas on the configuration space, whose coordinate patches are given by

$$\mathbf{P}_i(\underline{\eta}) := \{\underline{\mathbf{p}} \in \mathbf{R}^{10} \mid \mathrm{sign}(\mathrm{vol}_{\underline{\mathbf{p}}}(i, i+j, i+j+1)) = \eta_j\} \tag{35}$$

for $i = 1, ..., 5$ and $j = 1, 2, 3$, where $\underline{\eta} \in \{-1, +1\}^3$ and the index sums are computed $\mod 5$ as before. Since each coordinate patch is the inverse image of an open set in the range of a smooth function, it follows that each is open in $\mathbf{R}^{10}$ and hence is likewise open in the appropriate quotient topology. It is easily seen that these $2^3 \cdot 5 = 40$ open sets completely cover the configuration space. Using the method of triangulation, it can further be shown that the two squared distances $D(i, i+2)$ and $D(i, i+3)$ determine the configuration of the pentagon uniquely on each of the eight coordinate patches $\mathbf{P}_i(\underline{\eta})$ ($\underline{\eta} \in \{-1, +1\}^3$), and so constitute a local parametrization of each. From this, one sees that the configuration space is a compact, connected, two-dimensional manifold, and what remains is to determine its genus.

We shall do this by applying Morse theory. The exact choice of Morse function is not critical, but the oriented area $V(1, ..., 5)$ of the pentagon as a whole turns out to be computationally convenient. We recall that the oriented area of a polygon is the sum of the oriented areas of the triangles in any triangulation thereof [Klein, 1939], while the absolute area of a triangle $|V(a,b,c)|$ with vertices a, b, c is given in terms of the lengths of its sides by Heron's formula (Equation (3)), i.e. by $|V(a,b,c)| = 1/2\, D^{\frac{1}{2}}(a,b,c)$. Hence on a given coordinate patch $\mathbf{P}_i(\underline{\eta})$, the oriented area of our pentagon is given by

$$\begin{aligned} 2\,V(1,...,5) \;=\; & \eta_1\sqrt{D(i,i+1,i+2)} \\ & + \eta_2\sqrt{D(i,i+2,i+3)} + \eta_3\sqrt{D(i,i+3,i+4)}\,, \end{aligned} \tag{36}$$

where $D(i, i+j, i+j+1)$ denotes a three-point Cayley-Menger determinant and the index sums are computed $\bmod 5$ as always. In the case that $i = 1$, the derivatives of this expression are

$$\begin{aligned}
&\frac{\partial V(1,...,5)}{\partial D(1,3)} \\
&= \frac{\eta_1}{2D^{1/2}(1,2,3)} \cdot \frac{\partial D(1,2,3)}{\partial D(1,3)} + \frac{\eta_2}{2D^{1/2}(1,3,4)} \cdot \frac{\partial D(1,3,4)}{\partial D(1,3)} \\
&= \frac{\eta_1(D(1,3)-2)}{\sqrt{D^2(1,3)-4D(1,3)}} + \frac{\eta_2(D(1,3)-D(1,4)-1)}{\sqrt{1-2(D(1,3)+D(1,4))+(D(1,3)-D(1,4))^2}}
\end{aligned} \tag{37}$$

and

$$\begin{aligned}
&\frac{\partial V(1,...,5)}{\partial D(1,4)} \\
&= \frac{\eta_2}{2D^{1/2}(1,3,4)} \cdot \frac{\partial D(1,3,4)}{\partial D(1,4)} + \frac{\eta_3}{2D^{1/2}(1,4,5)} \cdot \frac{\partial D(1,4,5)}{\partial D(1,4)} \\
&= \frac{\eta_2(D(1,4)-D(1,3)-1)}{\sqrt{1-2(D(1,3)+D(1,4))+(D(1,3)-D(1,4))^2}} + \frac{\eta_3(D(1,4)-2)}{\sqrt{D^2(1,4)-4D(1,4)}} .
\end{aligned} \tag{38}$$

If we set these derivatives to zero, rearrange and square both sides, we get

$$(D(1,4)-1) \cdot (D^2(1,3) - 2D(1,3) - D(1,4) + 1) \;=\; 0 \tag{39}$$

and

$$(D(1,3)-1) \cdot (D^2(1,4) - 2D(1,4) - D(1,3) + 1) \;=\; 0 \tag{40}$$

respectively (note the signs η_i cancel on squaring). Thus, if $D(1,4) \neq 1$ and $D(1,3) \neq 1$, we have

$$D(1,3) \;\in\; \{1 - D^{1/2}(1,4),\; 1 + D^{1/2}(1,4)\} \tag{41}$$

and

$$D(1,4) \;\in\; \{1 - D^{1/2}(1,3),\; 1 + D^{1/2}(1,3)\} \tag{42}$$

respectively, so that the only nonzero simultaneous solutions of equations (39) and (40) are

$$\begin{aligned}
&D(1,3) = D(1,4) = \frac{\sqrt{5}+1}{2} , \\
&D(1,3) = D(1,4) = \frac{\sqrt{5}-1}{2} , \\
&\text{and} \quad D(1,3) = D(1,4) = 1 .
\end{aligned} \tag{43}$$

The first two solutions are the squared diagonal distances in the convex and inverted regular pentagons, respectively. Examination of the Hessian of V at these configurations establishes that for $\eta = -1$ and $+1$ the convex pentagon corresponds to critical points of index 0 and 2, respectively, which lie in all of the coordinate patches $\mathbf{P}_i(\eta, \eta, \eta)$ $(i = 1, ..., 5)$. Similarly, for $\eta = -1$ and $+1$ the inverted pentagon corresponds to critical points of index 2 and 0, respectively, which lie in all of the coordinate patches $\mathbf{P}_i(\eta, -\eta, \eta)$ $(i = 1, ..., 5)$. These observations show that $N_0 = N_2 = 2$. The last solution occurs in only two of the coordinate patches, namely $\mathbf{P}_1(+1, -1, +1)$ and $\mathbf{P}_1(-1, +1, -1)$. In this case, evaluation of the Hessian reveals that these are nondegenerate critical points of index 1. Since there is one such critical point in each coordinate patch of the form $\mathbf{P}_i(\eta, -\eta, \eta)$ $(\eta \in \{\pm 1\}, i = 1, ..., 5)$, the total number of such critical points is $N_1 = 10$.

We now plug these numbers into our equation for the Euler characteristic and get:

$$\chi_M \ = \ 2 - 10 + 2 \ = \ -6 \tag{44}$$

which corresponds to a manifold of genus 4 as claimed. Since the only nonorientable manifold of this genus has two cross-caps, orientability can be established by a symmetry argument.

Acknowledgements

The author thanks the editors, B. Sturmfels and N. White, for many useful discussions and suggestions during this work.

References

1. T. M. Apostol, *Ptolemy's inequality and the chordal metric*, Math. Mag. **40** (1967), 233–235.
2. M. Berger, *Une caractérisation purement métrique des variétés Riemanniennes à courbure constante*, in "E. B. Christoffel (ed. P. L. Butzer & F. Feheŕ)," Birkhäuser Verlag, Basel, 1981.
3. M. Berger, *La géométrie métrique des variétés Riemanniennes*, Soc. Math. Fr., Astérisque, hors série (1985), 9–66.
4. L. M. Blumenthal, "Theory and Applications of Distance Geometry," Cambridge Univ. Press, Cambridge, England, 1953 (Reprinted by Chelsea Publishing Co., Bronx NY (1970)).
5. L. M. Blumenthal, "A Modern View of Geometry," Dover Publ. Co., New York, NY, 1961.
6. Bruno Buchberger, *Gröbner bases: An algorithmic method in polynomial ideal theory*, In "Multidimensional Systems Theory", ed. N. K. Bose, D. Reidel Publ. Co., Dordrecht, Holland, (1985), 184–232.
7. B. W. Char, G. J. Fee, K. O. Geddes, G. N. Gonnet, M. B. Monagan, *A tutorial introduction to maple*, J. Symbolic Comput. **2** (1986), 179–200.
8. Shang-Ching Chou, "Mechanical Geometry Theorem Proving," Kluwer Academic Publishers, Hingham, MA, 1987.
9. G. M. Crippen & T. F. Havel, "Distance Geometry and Molecular Conformation," J. Wiley & Sons (Research Studies Press), New York, NY (Letchworth, UK), 1988.
10. Coxeter, H. S. M., "Introduction to Geometry," J. Wiley & Sons, New York, NY, 1969.
11. J. H. Davenport, Y. Siret and E. Tournier, "Computer Algebra," St. Edmundsbury Press, Ltd., Bury St. Edmunds, Suffolk, UK, 1988.
12. A. W. M. Dress, *Vorlesungen über kombinatorische Geometrie*, unpublished notes, Univ. of Bielefeld, W. Germany (1982).

13. A. W. M. Dress & T. F. Havel, *Fundamentals of the distance geometry approach to the problems of molecular conformation*, INRIA Workshop on Computer-Aided Geometric Reasoning, Sophia Antipolis, France (1987).
14. J. C. Gower, *Euclidean distance geometry*, Math. Scientist **7** (1982), 1–14.
15. J. C. Gower, *Properties of Euclidean and non-Euclidean distance matrices*, Linear Algebra Appl. **67** (1985), 81–97.
16. R. A. Johnson, "Modern Geometry," Houghton Mifflin, Boston, MA, 1929.
17. D. Kapur & J. L. Mundy, "Geometric Reasoning," MIT Press, Cambridge, MA, 1989.
18. F. Klein, "Elementary Mathematics from an Advanced Standpoint, vol. 2: Geometry," Dover Publications, London, U.K., 1939.
19. B. Kutzler, "Algebraic Approaches to Automated Geometry Theorem Proving," Ph.D. Dissertation, Johannes Kepler Univ., Linz, Austria, 1988.
20. K. Menger, *Untersuchungen über allgemeine Metrik*, Math. Ann. **100** (1928), 75–163.
21. K. Menger, *New foundation for Euclidean geometry*, Am. J. Math. **53** (1931), 721–745.
22. M. Morse & S. S. Cairns, "Critical Point Theory in Global Analysis and Differential Topology," Academic Press, New York, NY, 1969.
23. I. J. Schoenberg, *On certain metric spaces arising from Euclidean spaces by a change of metric and their imbeddings in Hilbert space*, Annals Math. **38** (1937), 787–793.
24. I. J. Schoenberg, *Linkages and distance geometry, I & II.*, Nederl. Akad. Wetensch. Proc. Ser. A (= Indag. Math.) **72 (= 31)** (1969), 43–63.
25. J. J. Seidel, *Distance-geometric development of 2-dimensional Euclidean, hyperbolic and spherical geometry. I & II*, Simon Stevin **29** (1952), 32–50,65–76.
26. J. J. Seidel, *Angles and distances in n-dimensional Euclidean and non-Euclidean geometry. I–III*, Indag. Math. **17, no. 3 & 4** (1955).
27. E. Snapper & R. J. Troyer, "Metric Affine Geometry," Academic Press, New York, NY, 1971.
28. W. Whiteley, *Invariant computations for analytic projective geometry*, J. Symbolic Comput., this issue.
29. B. Sturmfels & N. White, *Gröbner bases and invariant theory*, Adv. Math., in press..

Keywords. distance geometry, computer algebra, linkages

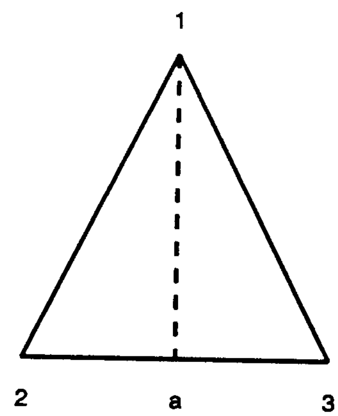

Figure 1

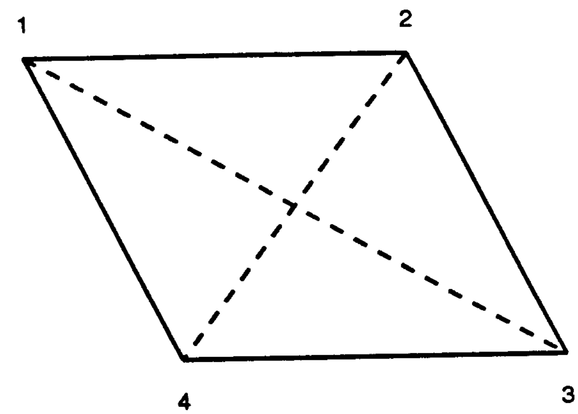

Figure 2

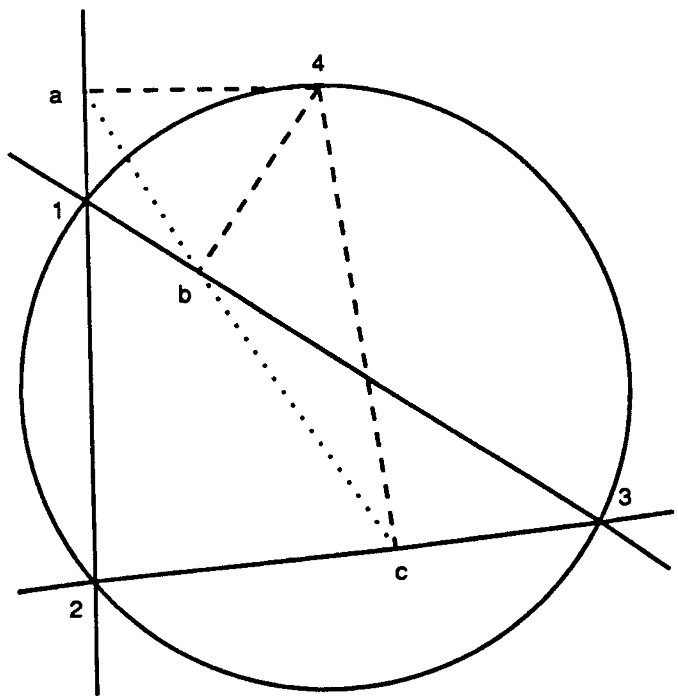

Figure 3

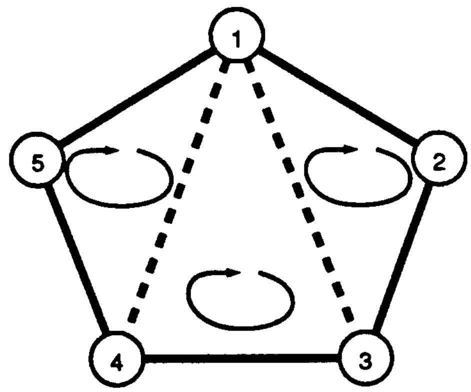

Figure 4