Inorganica Chimica Acta, 14 (1975) 201–205
© Elsevier Sequoia S.A., Lausanne – Printed in Switzerland

Incidence and Distance Matrices

M. J. CLARK and S. F. A. KETTLE
School of Chemical Sciences, University of East Anglia, Norwich NOR 88C, U.K.
Received October 29, 1974

Introduction

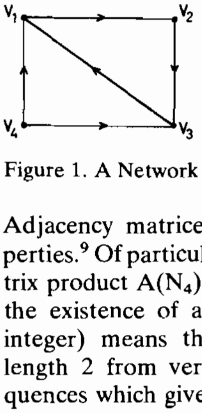
It is now recognised that many aspects of chemistry are, in some considerable measure, topologically determined. For example, graph theory has been applied to problems in Valence Theory – an aspect which has been explored by several authors.[1,2,3,4] Perhaps the earliest application of formal graph theory to a chemistry problem was Cayley's work in determining the isomer counts of the paraffinic hydrocarbons.[5] A much more recent application also concerns isomers and is the subject of the present communication.

Fluxional (stereochemically non-rigid) molecules, undergo rapid molecular rearrangements. All the isomers have the same equilibrium geometry and differ only in atomic permutations, but on a nuclear magnetic resonance time scale the properties of individual atoms may be a time average of those associated with mote than one position. Although is not always possible to associate a unique permutation with a unique permutational mechanism each species has a definite number of permutational isomers[6,7] each permutational mechanism leading to a characteristic pattern of interconversions between these isomers. The various mechanisms are conveniently distinguished by the different shortest path sequences by which any given isomer is converted into another. The shortest paths are detailed in a so-called distance matrix of the system, examples of which have been given by several authors. However there appears to be in the chemical literature no details of the derivation of such matrices. It is the purpose of the present communication to remedy this ommision and to propose some innovations that will increase the information content of them.

Network Matrices

A network, N_n, is an array of vertices $(v_1, v_2, v_3, \ldots v_n)$ containing paths between the various vertices which may or may not be directed (a graph or digraph). A path $v_i \rightarrow v_j$ is represented by a line (edge), which if one way only, is distinguished by an arrow in the direction $i \rightarrow j$. A permutational isomerisation can be represented by a network if each permutational isomer is assigned to a unique vertex and the permutation mechanism is shown by joining the vertices with directed lines. Distinguishable mechanisms will have different connections between the vertices. In a network representing a fluxional mechanism, a directed line from vertex *a* to vertex *b* implies that there will also be a directed line from *b* to *a* (microscopic reversibility). This reversibility would not be a feature of a network appropriate to a system in which a net chemical reaction, accompanied by large energy changes, occurs.

The vertices and lines of a network, N_n, containing n vertices can be represented by an adjacency matrix $A(N_n)$. In this matrix the rows and columns are labelled with the indices of the vertices. The matrix element $a_{ij} = 1$ if there is a directed line connecting vertex i to j; $a_{ij} = 0$ if there is no such path. In Fig. 1 we show a simple network and the corresponding adjacency matrix (this network does *not* correspond to a permutational isomerisation).



$$A(N_4) = \begin{array}{c} \\ V_1 \\ V_2 \\ V_3 \\ V_4 \end{array} \begin{array}{c} \begin{array}{cccc} V_1 & V_2 & V_3 & V_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Figure 1. A Network and its Adjacency Matrix.

Adjacency matrices have a number of valuable properties.[9] Of particular importance is the fact that the matrix product $A(N_4) \cdot A(N_4) = A^2(N_4)$ is a matrix where the existence of an element $a_{ij} = m$ (where m is an integer) means that there are m directed paths of length 2 from vertex i to vertex j. The detailed sequences which give rise to these paths is implicit in the matrix multiplication of $A(N_4)$. Thus, in the example shown in Figure 1, the matrix $A^2(N_4)$ is

$$A(N_4) \cdot A(N_4) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

where the $a_{21}{}^2$ element is given by

$$a_{21}{}^2 = a_{21}a_{11} + a_{22}a_{21} + a_{23}a_{31} + a_{24}a_{41} = 0.0 + 0.1 + 1.1 + 0.1 = 1$$

In this multiplication the existence of both abortive and succesful pathways can be readily seen. Furthermore, it shows that the sole path of length 2 to vertex 2 from vertex 1 is given by the term $a_{23} \cdot a_{31}$, *i.e.*, the sequence $v_2 \rightarrow v_3 \rightarrow v_1$. In a similar way the products $A^3(N_4)$, $A^4(N_4)$.....$A^c(N_4)$.....$A^d(N_4)$ indicate those vertices that are connected by directed paths of length three, four ...c...d. Successive multiplications to higher order for a small network will not give more information about connected vertices because paths become duplicated. It is convenient to define a partial distance matrix $d^c(N_n)$ where each element in $d^c(N_n)$ is set equal to the c value in $A^c(N_n)$. For instance, in the above example

$$d^2(N_n) = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \end{pmatrix}$$

That is, each element of this matrix represents a distance. The least distance matrix $D(N_n)$ is given by

$$D(N_n) = d(N_n) \oplus d^2(N_n) \oplus d^3(N_n) \ldots\ldots..$$

In this equation the matrices on the r.h.s. are added in strict numerological sequence. The element D_{ij} of the distance matrix is either zero or is equal to the first non-zero ijth element encountered in the r.h.s. matrices. That is the least distance matrix gives the shortest distances between two vertices (non-connected vertices are represented by zero). Because the distance between a vertex and itself is zero, all diagonal elements of the least distance matrix are zero. Thus, in the above example,

$$D(N_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \end{pmatrix} \oplus$$

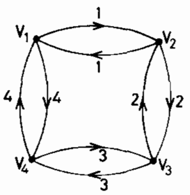
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \end{pmatrix} \oplus \ldots\ldots \quad = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 0 \end{pmatrix}$$

This process is readily computerised and can be used to calculate the distance matrices of networks when inspection methods are not practicable.

The above definition and derivation of adjacency and distance matrices uses binary arithmetic in networks – if two vertices are connected we represent this by the number one; if there is no connection we represent this by the number zero. Distance matrices thus defined suffer from two limitations. Firstly, it is assumed that there is a fixed distance between connected vertices (for a chemical system this would imply constant potential energy barriers). Secondly, there is no indication of the number of different paths (of the same length) which interconnect two vertices. We now suggest a simple method by which these limitations can be overcome.

An Augmented Adjacency Matrix

We require a matrix algebra that will allow characterisation of both connectivity and distance. In Figure 2 we show a network N_4 in which the distance between vertices varies from 1 to 4 units. Although, again, this matrix does not correspond to any particular fluxional mechanism, it is a simple model of a fluxional system in which a variety of mechanisms, characterised by different energy barriers, all operate. In the present context it is convenient to represent the different energy barriers by different distances (all of which we take to be simply related). We define an "adjacency-distance matrix" $X(N_4)$ which contains these distances as elements.



$$X(N_4) = \begin{pmatrix} 0 & 1 & 0 & 4 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 4 & 0 & 3 & 0 \end{pmatrix}$$

Figure 2. A Network and its $X(N_n)$ Matrix.

The partial distance matrices can be found by defining the law of combination as addition in the normal matrix multiplication sequence, but listing each contribution to the matrix element separately (we note that contracted representations will often be possible). That is, $b_{ij} = [(a_{ik} + a_{kj}), (a_{il}, + a_{lj}) + ..]$ with the condition that $a_{ik} + a_{kj} = 0$ if $a_{ik}, a_{kj} = 0$ (zero terms will not be included in the matrix). So, for the network of Figure 2 we have

$$X^2(N_4) = \begin{pmatrix} (2,8) & (0) & (3,7) & (0) \\ (0) & (2,4) & (0) & (5,5) \\ (3,7) & (0) & (4,6) & (0) \\ (0) & (5,5) & (0) & (8,6) \end{pmatrix}$$

For $X^3(N_4)$.... *etc.* the same method is followed except that when evaluating the sum $(a_{ik} + a_{kj})$ it is necessary to include all combinations of terms between a_{ik} and a_{kj} should each contain more than one. Thus, $(2,8) + (3,7) = (5,9,11,15)$.

The least distance matrix is computed in a manner similar to that given earlier except that in the expression

$$D'(N_4) = X(N_4) \oplus X^2(N_4) \oplus X^3(N_4) + \ldots$$

it is necessary to scan *all* terms in the r.h.s. expression in order to determine the least distance between two vertices. In the present example we have

$$D'(N_4) = \begin{pmatrix} (0) & (1) & (3) & (4) \\ (1) & (0) & (2) & 2(5) \\ (3) & (2) & (0) & (3) \\ (4) & 2(5) & (3) & (0) \end{pmatrix}$$

In deriving this expression we have recognised the existence of two independent and distinct paths of length 5 between vertices 2 and 4. Such replication of least distances is much more common in situations in which only one permutational mechanism is considered and is, evidently, very pertinent when considering the accessibility of one vertex from another. An unmodified distance matrix may, in such situations, give a misleading impression of accessibility. Indeed, for such cases, it is generally the case that the more remote that one vertex is from another, the greater the number of pathways of least length. As an illustration of this we describe the possible fluxions of the trigonal bipyramidal molecule, R_2NPF_4. This molecule has six permutational isomers (if we assume that the fluorines undergo rapid exchange while the amino group is static in an equatorial position[10]). The exchange of fluorines gives 4! permutations but some of the permutations are equivalent because of the point symmetry of the molecule (assumed to be C_{2v})* so that the number is reduced to $4!/4 = 6$. These six isomers of R_2NPF_4 are shown in Figure 3.

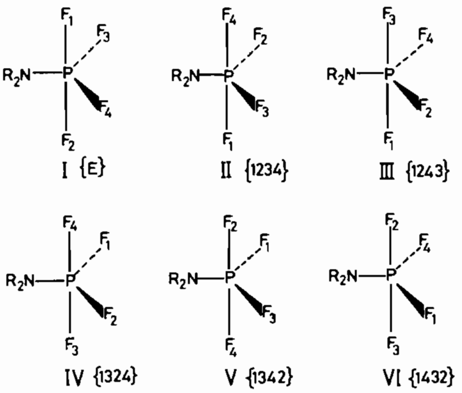
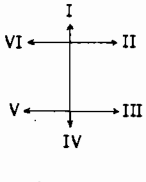


Figure 3. The permutational isomers of R_2NF_4.

Above the coalescence temperature there is rapid exchange between equatorial and axial fluorines on the NMR time scale. There are a number of possible mechanisms by which these fluorines could exchange: 1) the simultaneous exchange of all four fluorines, 2) simultaneous exchange of three fluorines, 3) simultaneous exchange of two fluorines (we recognise that these are not all in accord with the experimental data, but the present argument is illustrative only). The $X(N_6)$ matrices are shown below.

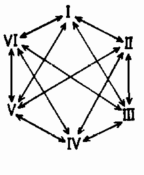
In constructing these matrices we have assumed that the four atom exchange is the Berry mechanism,[11] the

$$\begin{array}{c|cccccc} & I & II & III & IV & V & VI \\ \hline I & 0 & 0 & 0 & 1 & 0 & 0 \\ II & 0 & 0 & 0 & 0 & 0 & 1 \\ III & 0 & 0 & 0 & 0 & 1 & 0 \\ IV & 1 & 0 & 0 & 0 & 0 & 0 \\ V & 0 & 0 & 1 & 0 & 0 & 0 \\ VI & 0 & 1 & 0 & 0 & 0 & 0 \end{array}$$



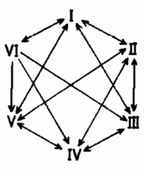
Mechanism 1

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$



Mechanism 2

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$



Mechanism 3

---

* This point group is appropriate to the N.M.R. time scale.

three atom process is an adapted turnstile mechanism,[12] and the two atom process is a straightforward interchange of an axial and equatorial fluorine. Although different mechanisms will generally lead to different matrices it is to be noted that cases 2 and 3 are equivalenty topologically. The least distance matrices can be found by the prescribed method and are shown below, where mechanisms 2 and 3 again give identical matrices.

$$\begin{pmatrix} 0 & 0 & 0 & 1(1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1(1) \\ 0 & 0 & 0 & 0 & 1(1) & 0 \\ 1(1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1(1) & 0 & 0 & 0 \\ 0 & 1(1) & 0 & 0 & 0 & 0 \end{pmatrix}$$

Mechanism 1

$$\begin{pmatrix} 0 & 1(1) & 1(1) & 4(2) & 1(1) & 1(1) \\ 1(1) & 0 & 1(1) & 1(1) & 1(1) & 4(2) \\ 1(1) & 1(1) & 0 & 1(1) & 4(2) & 1(1) \\ 4(2) & 1(1) & 1(1) & 0 & 1(1) & 1(1) \\ 1(1) & 1(1) & 4(2) & 1(1) & 0 & 1(1) \\ 1(1) & 4(2) & 1(1) & 1(1) & 1(1) & 0 \end{pmatrix}$$

Mechanism 2 and 3

TABLE I. Isomer Labelling.

Isomer Number	Ligand (Positions 1 and 5 are axial)				
	A	B	C	D	E
1	1	2	3	4	5
2	1	2	4	3	5
3	1	5	3	4	2
4	1	5	4	3	2
5	1	3	5	4	2
6	1	4	5	3	2
7	1	4	2	5	3
8	1	3	2	5	4
9	4	1	5	3	2
10	3	1	5	4	2
11	4	1	2	5	3
12	3	1	2	5	4
13	2	1	3	4	5
14	2	1	4	3	5
15	2	3	1	5	4
16	2	4	1	5	3
17	4	2	1	3	5
18	3	2	1	4	5
19	2	3	4	1	5
20	2	4	3	1	5

The additional information provided by our method is most significant for complicated systems. For example, the possible networks for PF_5 (which has twenty permutational isomers) are not trivial to obtain although individual steps are readily listed. Fortunately, only the latter are needed. As an example the $X(N_{20})$ matrix appropriate to a Berry mechanism is (the isomer labelling is as given in Table I):

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0
2	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0
5	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0
6	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1
7	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0
8	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0
9	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
10	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
11	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
12	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
13	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
14	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0
15	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
16	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
17	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
18	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
19	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
20	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0

from which the augmented distance matrix can be computed:

0	12(5)	2(3)	1(2)	1(2)	2(3)	2(3)	1(2)	4(4)	1(1)	4(4)	1(1)	2(3)	1(2)	1(1)	4(4)	1(2)	2(3)	2(3)	1(2)
12(5)	0	1(2)	2(3)	2(3)	1(2)	1(2)	2(3)	1(1)	4(4)	1(1)	4(4)	1(2)	2(3)	4(4)	1(1)	2(3)	1(2)	1(2)	2(3)
2(3)	1(2)	0	12(5)	2(3)	1(2)	1(2)	2(3)	2(3)	1(2)	2(3)	1(2)	1(2)	2(3)	4(4)	1(1)	1(1)	4(4)	4(4)	1(1)
1(2)	2(3)	12(5)	0	1(2)	2(3)	2(3)	1(2)	1(2)	2(3)	1(2)	2(3)	2(3)	1(2)	1(1)	4(4)	4(4)	1(1)	1(1)	4(4)
1(2)	2(3)	2(3)	1(2)	0	12(5)	2(3)	1(2)	1(2)	2(3)	4(4)	1(1)	1(1)	4(4)	2(3)	1(2)	1(2)	2(3)	1(1)	4(4)
2(3)	1(2)	1(2)	2(3)	12(5)	0	1(2)	2(3)	2(3)	1(2)	1(1)	4(4)	4(4)	1(1)	1(2)	2(3)	2(3)	1(2)	4(4)	1(1)
2(3)	1(2)	1(2)	2(3)	2(3)	1(2)	0	12(5)	1(1)	4(4)	2(3)	1(2)	4(4)	1(1)	1(2)	2(3)	1(1)	4(4)	1(2)	2(3)
1(2)	2(3)	2(3)	1(2)	1(2)	2(3)	12(5)	0	4(4)	1(1)	1(2)	2(3)	1(1)	4(4)	2(3)	1(2)	4(4)	1(1)	2(3)	1(2)
4(4)	1(1)	2(3)	1(2)	1(2)	2(3)	1(1)	4(4)	0	12(5)	1(2)	2(3)	2(3)	1(2)	2(3)	1(2)	1(2)	2(3)	1(1)	4(4)
1(1)	4(4)	1(2)	2(3)	2(3)	1(2)	4(4)	1(1)	12(5)	0	2(3)	1(2)	1(2)	2(3)	1(2)	2(3)	2(3)	1(2)	4(4)	1(1)
4(4)	1(1)	2(3)	1(2)	4(4)	1(1)	2(3)	1(2)	1(2)	2(3)	0	12(5)	2(3)	1(2)	2(3)	1(2)	4(4)	1(1)	2(3)	1(2)
1(1)	4(4)	1(2)	2(3)	1(1)	4(4)	1(2)	2(3)	2(3)	1(2)	12(5)	0	1(2)	2(3)	1(2)	2(3)	1(1)	4(4)	1(2)	2(3)
2(3)	1(2)	1(2)	2(3)	1(1)	4(4)	4(4)	1(1)	2(3)	1(2)	2(3)	1(2)	0	12(5)	4(4)	1(1)	2(3)	1(2)	1(2)	2(3)
1(2)	2(3)	2(3)	1(2)	4(4)	1(1)	1(1)	4(4)	1(2)	2(3)	1(2)	2(3)	12(5)	0	1(1)	4(4)	1(2)	2(3)	2(3)	1(2)
1(1)	4(4)	4(4)	1(1)	2(3)	1(2)	1(2)	2(3)	2(3)	1(2)	2(3)	1(2)	4(4)	1(1)	0	12(5)	2(3)	1(2)	1(2)	2(3)
4(4)	1(1)	1(1)	4(4)	1(2)	2(3)	2(3)	1(2)	1(2)	2(3)	1(2)	2(3)	1(1)	4(4)	12(5)	0	1(2)	2(3)	2(3)	1(2)
1(2)	2(3)	1(1)	4(4)	1(2)	2(3)	1(1)	4(4)	1(2)	2(3)	4(4)	1(1)	2(3)	1(2)	2(3)	1(2)	0	12(5)	2(3)	1(2)
2(3)	4(4)	4(4)	1(1)	2(3)	1(2)	4(4)	1(1)	2(3)	1(2)	1(1)	4(4)	1(2)	2(3)	1(2)	2(3)	12(5)	0	1(2)	2(3)
2(3)	1(2)	4(4)	1(1)	1(1)	4(4)	1(2)	2(3)	1(1)	4(4)	2(3)	1(2)	1(2)	2(3)	1(2)	2(3)	2(3)	1(2)	0	12(5)
1(2)	2(3)	1(1)	4(4)	4(4)	1(1)	2(3)	1(2)	4(4)	1(1)	1(2)	2(3)	2(3)	1(2)	2(3)	1(2)	1(2)	2(3)	12(5)	0

This form is matrix is illuminating in that it reveals, for instance, that permutational isomers 1 and 2 are connected by twelve independent paths of length 5, whereas other isomers are connected by unique paths (*e.g.* isomers 1 and 5 are connected by one path of length 2).

The form of the network matrices described in this paper is general and is therefore applicable to a wide variety of chemical systems. Applications to date have included bonding theory and isomer counts. In this paper we have concentrated on the already established application to fluxional molecules where the connectivity index relates to the possible interchange between isomers, and a distance parameter which gives the path length between the isomers. The final form of the distance matrix gives the number of paths of minimum length by using an algebra in which matrix elements are added in the conventional multiplication sequence. With this development the applications are widened to cases where, for instance, the distance parameter can be related to an energy barrier.

Acknowledgement

One of us (M.J.C.) thanks the Science Research Council for a Fellowship.

References

1 K. Ruedenberg, *J. Chem. Phys., 1954, 22,* 1878.
2 H.H. Schmidtke, *J. Chem. Phys.,* 1966, *45,* 3920.
3 A. Graovac. I. Gutman, N. Trinajstic and T. Zivkovic, *Theoret. Chem. Acta,* 1972, *26,* 67.
4 D. Cvetkovic, I. Gutman and N. Trinajstic, *Chem. Phys. Lett.,* 1972, *16,* 614.
5 A. Cayley, *Chem. Ber.,* 1875, *8,* 614.
6 E.L. Muetterties, *J. Am. Chem. Soc.,* 1969, *91,* 1636.
7 W.G. Klemperer, *J. Am. Chem. Soc.,* 1972, *94,* 6940.
8 E.L. Muetterties, *Rec. Chem. Progr.,* 1970, *31,* 51.
9 F. Harary, R.Z. Norman and D. Cartwright, "Structural Models: An Introduction to the Theory of Directed Graphs" Wiley, New York, N.Y., 1965. For a more mathematical discussion see B.A. Carré, *J. Inst. Maths. Applics.,* 1971, 7, 273–294.
10 G.M. Whitesides and H.L. Mitchell, *J. Am. Chem. Soc.,* 1969, *91,* 5384.
11 R.S. Berry, *J. Chem. Phys.,* 1960, *32,* 933.
12 P. Gillespie *et al., Angew. Chem. Internat. Edit.,* 1971, *10,* 687.