Discrete Applied Mathematics 19 (1988) 205–214
North-Holland

EMBEDDING AND CHARACTERIZATION OF QUANTUM CHEMICAL REACTION GRAPHS ON TWO-DIMENSIONAL ORIENTABLE SURFACES

Frank HARARY
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

Paul G. MEZEY
Department of Chemistry and Department of Mathematics, University of Saskatchewan, Saskatoon, S7N 0W0, Canada

Received 26 September 1985
Revised 20 February 1986

Quantum chemical reaction graphs defined on multidimensional potential energy hypersurfaces are embedded on two-dimensional orientable surfaces. Topological invariants of these graphs and those of the embedding two-dimensional surfaces can characterize various families of chemical reactions. Several examples are given for embeddings of reaction graphs of chemical conformational processes.

1. Introduction

In chemistry it has been a long established practice to represent molecules by graphs, with the constituent atoms and the interconnecting bonds taking the roles of vertices and edges, respectively. Whereas such representations contain a great amount of chemical information, and are of considerable value, nonetheless, they are in general insufficient to account for energy differences between various chemical species, and for the energies of activation required for a chemical reaction. Information on the energy variation of individual molecules during a chemical reaction is fundamentally important for a reliable prediction of favoured reaction mechanisms. In the last decade major advances have been made in the quantum chemical calculation of the energies of molecules, and now it is possible to evaluate the energy content of medium size molecules in any nuclear configuration occuring during a chemical change. Our purpose is to discuss a graph-theoretical representation of the energy function of molecules. With the recent spectacular advances in computer technology the quantum chemical calculation of molecular energy as the function of atomic arrangements is rapidly becoming a routine task, and the proposed graph-theoretical framework is expected to reveal some of the regularities and general trends within families of chemical reactions.

Graph theory is a powerful mathematical discipline for a concise description of the most essential interrelations among entities from which a given physical system is composed [for selected applications in the field of the natural sciences, see, e.g.,

0166-218X/88/$3.50 © 1988, Elsevier Science Publishers B.V. (North-Holland)

[1, 2, 5, 6, 9, 10, 11, 17, 19, 20]. Chemistry, more specifically, the interrelations among chemical reactions, is a field where typically a very large number of entities (chemical species) may be involved. The complexity arises in part from the large number of possible interconversions among chemical species, as well as from the complicated manner in which the individual transformations occur. Thus graph theory is an ideal tool for the analysis of chemical reactions.

The energy content of chemical species and the energy variation during the course of chemical reactions fundamentally influence the stability and the very existence of molecules. Energy is one of the most fundamental quantum mechanical observables, and contemporary analysis of chemical reactions involves, almost without exception, a study of energy relations and energy profiles. Within a semi-classical model the energy content of molecules depends on the mutual position of the nuclei that can be characterized either by a set of cartesian coordinates, or by a set of internal coordinates, such as bond distances, bond angles and torsional angles. (Note that there are three coordinates for each nucleus, that is, for N nuclei altogether $3N$ coordinates, however, in the absence of external electromagnetic or other fields, three translational and three rotational degrees of freedom of the molecule as a whole can be eliminated, hence one has only $3N-6$ dimensions for $N>2$). In general, the energy functional of a molecule of N nuclei can be represented as a $(3N-6)$-dimensional energy hypersurface $E(K)$ over a nuclear configuration space M defined in terms of a suitable set of 'internal' coordinates. Points along such potential energy hypersurfaces can be calculated using quantum chemical methods, and the analysis of such energy hypersurfaces leads to a detailed understanding of chemical processes and possibly to the discovery of new reactions (see, e.g., [3, 4, 12, 13, 14, 15, 16, 18, 21, 22, 23]).

Take a fixed collection of N nuclei and consider all possible molecules which can be composed from them. For simplicity, we may assume that all these molecules are in their most stable (lowest) electronic state, although this restriction is not essential. (For a recent review of the concepts of reaction topology, see [23].) Considering all possible relative positions of these nuclei, it is evident that all these molecules correspond to some domains of the lowest energy hypersurface $E(K)$, defined over a $(3N-6)$-dimensional nuclear configuration space M. Stable molecular arrangements, e.g., a single stable molecule of N nuclei, or two non-interacting stable molecules of N_1 and N_2 nuclei, where $N_1+N_2=N$, correspond to $(3N-6)$-dimensional basins called *catchment regions* of energy minima of the hypersurface. The ith catchment region and the ith minimum are denoted by $C(0,i)$ and $K(0,i)$, respectively. Stable molecules or molecular systems can be thought of as formal 'residents' of such multidimensional catchment regions. It is also evident, that all chemical reactions involving precisely this set of nuclei and no electronic excitation, correspond to transformations along the energy hypersurface $E(K)$, that is, to formal displacements from one domain of the hypersurface to another. Transition structures, also referred to as 'transition states', occurring in chemical transformations between two stable arrangements, correspond to $(3N-7)$-dimensional 'moun-

tain ridges', i.e., $(3N-7)$-dimensional catchment regions $C(1,i)$ of saddle points $K(1,i)$ of one negative canonical curvature. In the above notation the first index in the parentheses, generally denoted by λ, is the number of negative eigenvalues of the Hessian matrix at the critical point $K(\lambda,i)$ of $E(K)$, also called the critical point index, whereas i is an index of ordering. A general $(3N-6-\lambda)$-dimensional catchment region of a critical point $K(\lambda,i)$ of $E(K)$, is denoted by $C(\lambda,i)$. These catchment regions generate a *partitioning* and a *topology* on the potential energy hypersurface, and provides an assignment of chemical species to domains of the hypersurface, or equivalently, to those of the nuclear configuration space M.

Here M is a *metric space*, and a general *quantum chemical reaction graph* on $E(K)$ is defined as an intersection graph of the catchment regions and their closures, where the closure refers to the metric of M. The chemically most useful reaction graphs $G(\lambda',\lambda'')$, $\lambda'<\lambda<\lambda''$, are restricted to stable molecules and proper transition structures, i.e., to the cases of $\lambda=0$ and $\lambda=1$. (Note that in some related studies the $G(\lambda,i)$ notation has been used for a coordinate neighbourhood of catchment region $C(\lambda,i)$. In the present study letter G always represents a graph.) These catchment regions constitute the vertex set of the reaction graph $G(0,1)=(V,E)$,

$$V=\{C(0,i)\}\cup\{C(1,i)\} \tag{1}$$

whereas the edge set

$$E=\{[C(\lambda,i),C(\lambda',i')]:\mathcal{N}^{s}(i,i')=1\} \tag{2}$$

is defined in terms of the *strong neighbour relation* [13]:

$$\mathcal{N}^{s}(i,i')=\begin{cases}1 & \text{if } [C(\lambda,i)\cap\bar{C}(\lambda',i')]\cup[\bar{C}(\lambda,i)\cap C(\lambda',i')]\neq\emptyset,\\ 0 & \text{otherwise.}\end{cases} \tag{3}$$

The above definition is motivated by the direct involvement of formal transition structures in actual chemical transformations, and by the fact that two stable chemical species may undergo direct interconversions into each other by more than one reaction mechanism, that is, involving different transition structures [13].

Embeddings of the graph $G(0,1)$ on an orientable surface

Even those reactions of chemical interest which involve small molecules correspond to rearrangements of approximately five to ten atoms, hence the associated potential energy hypersurfaces have typical dimensions in the range of $3\times5-6=9$ to $3\times10-6=24$. A typical chemical reaction of organic chemistry may easily involve ten times more nuclei, and the dimensions of the relevant energy hypersurfaces may be several hundreds. It is of evident advantage to reduce the dimension of such problems, and embedding the corresponding reaction graphs on two dimensional surfaces can considerably simplify the analysis.

We may assume that the reaction graph $G(0,1)$ of an arbitrary potential energy

hypersurface $E(K)$ of n dimensions is drawn on a two-dimensional sphere S^2, possibly with some of its edges crossing one another. Following the method described on page 116 of [6], one can attach a handle to the sphere S^2 at each crossing and allow one of the crossing edges to go over the handle and the other under it. By this technique each such crossing can be eliminated and in the process the sphere S^2 is transformed into an orientable surface of genus g, where g is the number of handles required. If the embedding is carried out with the minimum number g of handles, then all faces of the graph on the orientable surface are simply connected. For this orientable surface the generalized Euler relation holds:

$$V-E+F=2-2g \tag{4}$$

where V, E, and F are the number of vertices, edges, and faces of the embedded reaction graph, respectively.

Characterization of embedded reaction graphs

There are several topological invariants of a reaction graph G embedded on orientable surfaces, and we shall consider four such invariants:

(i) the genus $g=g(G)$ described above,

(ii) the thickness $t=t(G)$, the minimum number of planar subgraphs whose union is G,

(iii) the coarseness $c=c(G)$, the maximum number of edge-disjoint nonplanar subgraphs of G,

(iv) the crossing number $x=x(G)$, the minimum number of pairwise edge intersections when G is drawn in the plane.

For a discussion of graph-theoretical properties of these invariants see [6, p. 116], whereas concerning invariants of toroidal graphs see [7] and [8].

Some of the topological invariants are determined fully by the topology of the orientable surface itself, e.g., the genus g, whereas the determination of some others (notably, the crossing number x) requires the graph itself and cannot be deduced from the topological properties of the surface.

These invariants characterize the complexity of the family of chemical reactions admitted by the potential energy hypersurface $E(K)$. Although the invariants themselves are rigorously defined, they do not correspond directly to common chemical concepts. Nonetheless, these invariants do indicate some of the qualitative features of the given family of chemical reactions.

A large crossing number x indicates that there are many interconversion processes involving relatively few molecules.

If a large crossing number x occurs in combination with a relatively small genus g, this is an indication that the assignment of interconversions to chemical species is not random, and interconversions admit a classification; the classes can be associated with individual handles on the orientable surface.

The thickness t can be taken as the number of (possibly overlapping) molecular families where within each family the number of interconversion processes is relatively small as compared to the number of molecules, and their assignment to molecule pairs is not very different from random.

The coarseness c, on the other hand, indicates the number of disjoint families of chemical species which are interrelated to one another within a family by a large number of different reactions. These families can then be taken as semi-autonomous reaction families on the original potential energy hypersurface $E(K)$.

Embeddings and invariants of conformational processes

Conformational processes can be regarded as special chemical reactions, which are

(i) monomolecular processes,

(ii) rearrangements in which no chemical bonds are formed or broken.

For a general potential energy hypersurface $E(K)$ conformational domains of the nuclear configuration space M can be confined to various *level sets*

$$F(A) = \{K : K \in M, E(K) < A\}, \tag{5}$$

i.e., to families of those points K of M where the energy is less than some energy bound A. For a conformational problem A is usually chosen below the lowest dissociation energy of a molecule. If the level set $F(A)$ is disconnected, then we consider its connected components separately. Each connected component is a separate conformational domain of M.

In most conformational problems it is relatively easy to assign the most important conformational rearrangements to chemically indentifiable internal coordinates of the molecule, e.g., to angles of bond rotations or to angles of pyramidal inversions, etc. This fact often leads to some simplifications.

For simplicity, we shall choose chemical examples where between each two stable conformers there exists at most one direct conformational transformation. Hence, instead of graphs $G(0, 1)$ it will be sufficient to consider the simplest reaction graphs $G(0)$ defined by the neighbour relation [13]

$$\mathcal{N}(i, i') = \begin{cases} 1 & \text{if } \bar{C}(0, i) \cap \bar{C}(0, i') \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

where $i \neq i'$.

Consider the example of *m*-xylene derivatives, where both methyl groups of the *m*-xylene molecule contain one large substituent L, and there is another substituent Z in ring position 2 (see Fig. 1). The *conformational polyhedron* of this problem, restricted to the case where energy bound A is lower than that required for free rotations of the substituted Me groups, has been discussed in [14].

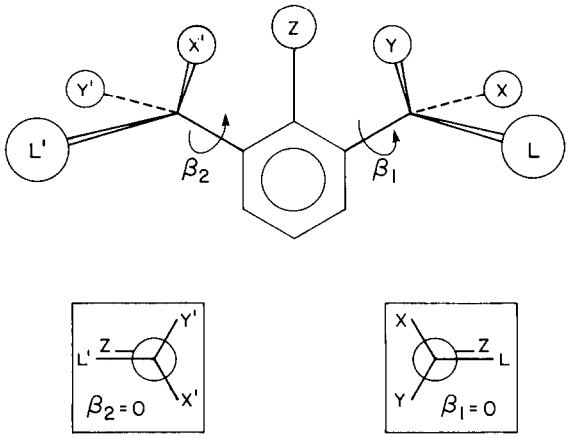

Fig. 1. The two-dimensional conformational problem of a family of substituted m-xylene derivatives. Full 2π rotations along angle coordinates β_1 and β_2 are hindered.

Here we shall use the same chemical example with some modifications to illustrate an embedding of a reaction graph and the use of invariants.

A schematic potential surface of this (essentially two-dimensional) conformational problem is shown in Fig. 2. If the energy bound A corresponds to the solid contour line, drawn with the heavy line, then the corresponding level set $F(A)$ contains five minima $m_1, \ldots, m_5$, four saddle points $s_5, \ldots, s_8$ and the corresponding parts of their respective catchment regions. Minimum m_5 is the lowest minimum. The energy maxima are denoted by capital M. For $F(A)$ one obtains the reaction graph $G(0)$ of edges shown by solid lines in Fig. 3. If one takes a somewhat higher energy bound A' shown by the dashed line in Fig. 2, then graph $G'(0)$, is obtained, which contains, in addition to the edges of $G(0)$, the edges shown by the dashed lines in Fig. 3. A further increase in the energy bound to the value A'' leads to the level set $F(A'')$, shown by the dotted contour line in Fig. 2. By periodicity of rotation angles β_1 and β_2 the opposite boundary line segments of $F(A'')$ are equivalent and one obtains graph $G''(0)$, containing all solid, dashed and dotted edges of Fig. 3.

The simplest of these graphs is the planar $G(0)$, a star, $K_{1,4}$, which belongs to energy bound A. For all planar graphs, the invariants are $g=0$, $x=0$, $t=1$ and $c=0$. Thus this holds for both the star $G(0)=K_{1,4}$ and the wheel $G'(0)=W_5$. However, the complete graph $G''(0)=K_5$ is nonplanar and the relevant invariants are $g=1$, $x=1$, $t=2$ and $c=1$.

The embeddings of graphs $G(0)=K_{1,4}$ and $G'(0)=W_5$ at energy bounds A and A', respectively, can be done on a sphere, but the embedding of $G''(0)=K_5$ requires at least one handle, i.e., a torus. A change from a planar graph to a nonplanar one as energy bound A' is increased to A'', brings about significant changes in the invariants.

Let us assume now that substituent L contains a group that is capable of a pyramidal inversion. We shall also assume that the two conformers arising when

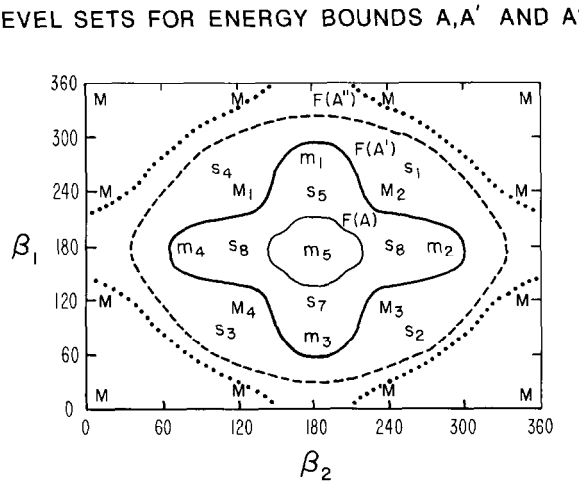

Fig. 2. Conformational level sets $F(A)$, $F(A')$ and $F(A'')$ of the accessible nuclear geometries of the *m*-xylene derivative problem at energy bounds A, A' and A'', respectively. Energy contours at the above energy bounds are denoted by heavy solid line, dashed line and dotted line, respectively. Minima, saddle points and maxima are denoted by m, s and M, respectively.

other conformational motions are frozen are of about equal stability. Furthermore, we assume that the energy barrier to this pyramidal inversion is high, and it can occur at some energy bound A''', slightly higher than A'', only if the two substituted methyl groups are in the vicinity of their lowest minimum conformation m_5.

Since the above conformational problem is essentially three-dimensional, energy being the fourth dimension, it is impossible to represent this potential hypersurface pictorially. However, the graph-theoretical description and embeddings of this problem are easily visualizable. The corresponding reaction graph $G'''(0)$ is shown with solid lines in Fig. 4, where the two K_5 components are drawn so as to give the

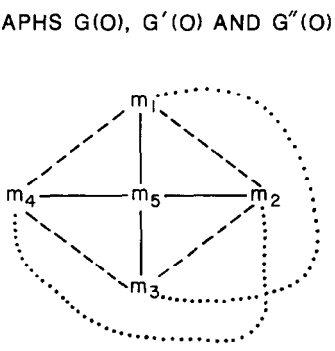

Fig. 3. Reaction graphs $G(0)$, $G'(0)$ and $G''(0)$ of level sets $F(A)$, $F(A')$ and $F(A'')$, are shown with solid, solid + dashed, and solid + dashed + dotted lines, respectively.

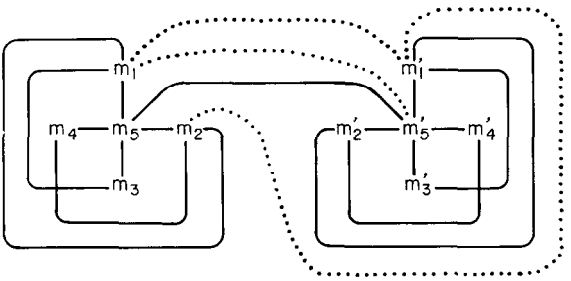

Fig. 4. If substituent L of the m-xylene derivative contains a group capable of pyramidal inversion, subject to some energy constraints (see text), then the conformational reaction graphs $G'''(0)$ and $G''''(0)$ are obtained, containing the solid edges and all edges, respectively.

smallest crossing number x for $G'''(0)$. The two subfamilies of minima associated with the two possible conformations of the L group are shown as the vertices of one K_5 component with no primes and those of the other with primes, respectively. The energy conditions on this family of conformational changes imply that the two subfamilies, associated with the two conformations of the L group, are joined only though the boundary of catchment regions $C(0,5)$ and $C(0,5')$ of minima m_5 and m_5', respectively. Elimination of the m_5m_5' edge leads to a chemically significant partitioning of $G'''(0)$ into two blocks. Elimination of edge m_5m_5' corresponds to an $A'''\rightarrow A''$ *lowering* of the energy bound, whereas the two K_5 blocks obtained at energy A'' correspond to two 'autonomous' conformational subsets of the potential energy hypersurface $E(K)$, one for each of the two possible conformations of group L. Evidently, the theorem of Battle, Harary, Kodoma and Youngs [6, p. 119] applies, and the genus $g(G'''(0))$ depends only on the genus $g(K_5)$ of its two K_5 blocks.

If at a higher energy bound A'''' the additional $m_1\leftrightarrow m_1'$, $m_5\leftrightarrow m_1'$ and $m_1\leftrightarrow m_5'$ direct conformational interconversions also become possible, then the graph $G''''(0)$ is obtained, that includes, in addition to the edges of $G'''(0)$ also those shown by dotted lines in Fig. 4. The $A'''\rightarrow A''''$ energy increase does not change the graph invariants, and $g=2$, $t=3$, $c=2$ and $x=2$ for $G''''(0)$ as well as for $G'''(0)$. If, however, the energy bound is further increased allowing for further direct conformational interconversions involving m_5 or m_5' (the energetically most favourable conformations), then this energy change inevitably leads to changes in the values of at least some of the invariants.

Summary

For the analysis of the most essential interrelations among reactions occuring along multidimensional potential energy hypersurfaces a simple graph-theoretical embedding technique is proposed. An energy criterion is used which restricts the embeddings and the analysis to the chemically important low energy regions of the

potential energy hypersurface. Such embeddings can always be accomplished on two-dimensional, orientable surfaces, and this reduction of dimension allows one to easily visualize relations among complicated reaction mechanisms. Topological invariants of the reaction graphs and their embedding surfaces provide characterization of various families of chemical reactions.

Acknowledgement

Financial support for this study has been provided by research grants from the National Science Foundation, U.S.A. and from the Natural Sciences and Engineering Research Council of Canada. The authors wish to express their thanks to Pano Rizzo for stimulating conversations and for his hospitality, while visiting his Institute.

References

[1] A.T. Balaban, ed., Chemical Applications of Graph Theory (Academic Press, New York, 1976).
[2] K. Balasubramanian, Computer generation of the characteristic polynomials of chemical graphs, J. Comput. Chem. 3 (1984) 387–394.
[3] K. Fukui, A formulation of the reaction coordinate, J. Phys. Chem. 74 (1970) 4161–4163.
[4] B.C. Garrett, D.G. Truhlar and R.S. Grev, Determination of the bottleneck regions of potential energy surfaces for atom transfer reactions by variational transition state theory, in: D.G. Truhlar, ed., Potential Energy Surfaces and Dynamics Calculations (Plenum Press, New York, 1981) 587–637.
[5] F. Harary, ed., Graph Theory and Theoretical Physics (Academic Press, London, 1967).
[6] F. Harary, Graph Theory (Addison - Wesley, Reading, MA, 1969).
[7] F. Harary, P.C. Kainen and A.J. Schwenk, Toroidal graphs with arbitrarily high crossing numbers, Nanta Math. 6 (1973) 58–67.
[8] F. Harary, P.C. Kainen, A.J. Schwenk and A.T. White, A maximal toroidal graph which is not a triangulation, Math. Scand. 33 (1973) 108–117.
[9] W.C. Herndon and J.E. Leonard, Canonical numbering, stereochemical descriptors, and unique linear notations for polyhedral clusters, Inorg. Chem. 22 (1983) 554–557.
[10] R.B. King, Chemical Applications of Topology and Graph Theory (Elsevier, Amsterdam, 1983).
[11] D.J. Klein, Ground state features of Heisenberg models, J. Chem. Phys. 77 (1982) 3098–3100.
[12] P.G. Mezey, Analysis of conformational energy hypersurfaces, Progr. Theor. Org. Chem. 2 (1977) 127–161.
[13] P.G. Mezey, Quantum chemical reaction networks, reaction graphs and the structure of potential energy hypersurfaces, Theor. Chim. Acta 60 (1982) 409–428.
[14] P.G. Mezey, The reaction polyhedron and group theory of reaction mechanisms, Internat. J. Quantum Chem. Symp. 19 (1986) 93–105, paper presented at the 1985 (25th Anniversary) Sanibel Symposium.
[15] P.G. Mezey, Reaction topology, in: V.H. Smith, Jr., H.F. Schaefer III, and K. Morokuma, eds., Applied Quantum Chemistry (Reidel, Dordrecht, 1986).
[16] W.H. Miller, N.C. Handy and J.E. Adams, Reaction path Hamiltonian for polyatomic molecules, J. Chem. Phys. 72 (1980) 99–112.

[17] N. Nakanishi, Graph Theory and Feynman Integrals (Gordon & Breach, New York, 1971).
[18] P. Pulay, Direct use of gradients for investigating molecular energy surfaces, in: H.F. Schaefer III, ed., Applications of Electronic Structure Theory (Plenum Press, New York, 1977) 153–185.
[19] M. Randic, On characterization of molecular branching, J. Am. Chem. Soc. 97 (1975) 6609–6615.
[20] N. Trinajstic, Chemical Graph Theory (CRC Press, Boca Raton, 1983).
[21] A. Tachibana and K. Fukui, Differential geometry of chemically reacting systems, Theor. Chim. Acta 49 (1978) 321–347.
[22] A. Tachibana, K. Yamashita, T. Yamabe and K. Fukui, Exact formula for the gradient of the CI potential energy surfaces, Chem. Phys. Lett. 59 (1978) 255–260.
[23] P.G. Mezey, Potential Energy Hypersurfaces (Elsevier, Amsterdam, 1987).