match no. 21 pp. 133-186 1986

THE WIENER NUMBER OF GRAPHS.
I. GENERAL THEORY AND CHANGES DUE TO SOME GRAPH OPERATIONS

Oskar E. Polansky[a] and Danail Bonchev[b]

[a]Max-Planck-Institut für Strahlenchemie, D-4330 Mülheim a. d. Ruhr, F. R. Germany

[b]Higher School of Chemical Technology, Department of Physical Chemistry, BU-8010 Burgas, Bulgaria

(Received: November 1986)

ABSTRACT

A new formalism for the study of Wiener numbers is developed which makes use of distances and distance numbers of the vertices of a graph. The notion of "distance tree of a vertex" is briefly reviewed.

By means of this formalism the change in the distance and Wiener numbers upon some elementary graph operations (covering several graphs upon a common induced subgraph, joining graphs by edges, and subdivision of an edge by a vertex) are studied. Various general formulae for distance and Wiener numbers of the graphs resulting after these operations are given. They provide deep and new insights into the modelling of distance and Wiener numbers and their interplay.

1. Introduction

Almost 40 years ago H. Wiener [1-3] introduced an important parameter, W(G), defined as "the sum of the distances between any two carbon atoms in the molecule in terms of carbon-carbon bonds" for correlations with physico-chemical properties of paraffins. Hosoya [4] pointed out the relation between the *Wiener number* W and the distance matrix of a graph. He also generalized the definition of this index from trees to any kind of undirected connected graphs. These pioneering works prompted the research of the distance matrix of chemical graphs. As a result, the Wiener index has found various applications [5,6], including here the detailed studies on molecular branching and cyclicity [7-15]. This earlier work indicated the existence of some limitations originating from the fact that essential structural information is lost when the Wiener number is constructed as the sum of $n(n-1)/2$ distance matrix entries.

In an attempt to reduce the loss of such structural information we propose to use other quantities, based on the distance matrix of a graph, which are intermediate in magnitude when compared with both the Wiener number and the individual distances. The sum of the distances from a certain vertex r to all other vertices of a simple connected graph G specifies such a quantity which is called here the *distance number* $d(r|G)$ of vertex r. The distance numbers are essential characteristics of the graph vertices providing a deeper insight into the topological nature of the Wiener number and its changes. It will be shown in this paper that general properties and relationships can be specified on this basis for the Wiener number. The applications of the formalism developed here will be reported elsewhere [16].

Naturally, distances, distance and Wiener numbers have been also a subject for pure graph theoretical considerations [17-19] originally initiated by some psychometric studies; in [18] $d(r|G)$ and $W(G)$ are termed "distance of vertex r" and "distance of graph G", respectively. We refer the more differentiating terms introduced above which have been developed during an early stage of this work where, in unawareness of [17-19], we studied the properties of these quantities in some details.

The metric properties of graphs have permanently growing interest for chemistry, due to their relevance to physico-chemical properties of chemical compounds [1-15,20-27]. The Wiener number and its information-theoretic analogues [26] are frequently used in the modelling of such properties. Meanwhile, some pitfalls of this approach were detected. They originate from the global nature of the Wiener number which does not preserve a great deal of the structural information contained in its summands, the individual distance matrix entries.

Aiming to overcome these difficulties we advocate the use of the distance numbers. They are algebraic quantities intermediate between the distances and the Wiener numbers and, hence, they preserve more of the information on the graph structure than the Wiener numbers. They can be used as topological indices [22,23,25, 26] and their future application to chemical reactivity studies, to NMR-chemical shifts in molecules, etc., may be anticipated.

Moreover, as a consequence of their intermediate character, the use of distance numbers enables one to treat topological features and their influence on Wiener numbers in a quite transparent manner [27]. The present paper is devoted to that subject. In Section 2 the basic formalism is exposed. Then we study the changes in the distance and Wiener numbers due to certain graph operations such as covering of several graphs upon an induced subgraph in common (Sect. 3) joining of graphs by the addition of edges (Sect.4), and subdivision of edges by a vertex (Sect. 5). The results obtained there offer a new and deep insight into the modelling of distance and Wiener numbers and their interplay.

We restrict our study to simple connected graphs which are of use in representing chemical structures. The graph-theoretical terminology used corresponds to that of F. Harary [28].

2. Basic Formalism

2.1. Definitions: In order to present our results in a systematic and rigorous way we start with a number of definitions.

Let G be a simple connected graph having n vertices and k edges, the vertices being denoted by a,b,c,..., etc.

Definition 1: The length of the shortest path, the so-called geodesic, $W_m(u,v)$ between vertices u and v of graph G is termed the *distance* d(uv) between these two vertices:

$$d(uv) = |W_m(uv)| \; ; \qquad (1)$$

thus, d(u,v) equals the number of edges of $W_m(u,v)$.

If necessary the symbol $d(uv|G)$ will be used instead of d(uv) in order to indicate that the distance is taken in graph G.

The distances obey the axioms of metric. The number of individual distances between all n vertices of G is n(n-1)/2. The distances may be obtained via the powers of the adjacency matrix $\underline{\underline{A}}$ of G using the formula:

$$d(uv) = \min\{\nu \mid [\underline{\underline{A}}^{\nu}]_{uv} \neq 0 \ ; \ 0 < \nu \leq n-1\}$$

Definition 2: The *distance matrix* $\underline{\underline{D}}$ of graph G is such a square matrix of order n whose elements are

$$[\underline{\underline{D}}]_{uv} = d(uv) \quad . \tag{2}$$

As a result of Definition 1 all off-diagonal $\underline{\underline{D}}$-entries are positive integers, d(uu) = 0, and $\underline{\underline{D}}$ is a symmetric matrix.

Definition 3: The r-th column of $\underline{\underline{D}}$ is called the *distance vector* $\vec{\underline{\underline{d}}}_r$ of vertex r; its entries are the distances from vertex r to all other vertices of graph G.

Definition 4: The sum of all $\vec{\underline{\underline{d}}}_r$ entries is called the *distance number* d(r|G) of vertex r:

$$d(r|G) = \sum_p [\underline{\underline{D}}]_{pr} = \sum_p d(rp) \quad . \tag{3}$$

Definition 5: The sum of all upper (lower) triangular submatrix entries of $\underline{\underline{D}}$ specifies the *Wiener number* W of graph G:

$$W = \sum_{p<r} [\underline{\underline{D}}]_{pr} = \frac{1}{2} \sum_r d(r|G) \quad . \tag{4}$$

Comparison of eqs. (3) and (4) reveals the intermediate character of the distance numbers, as compared with the Wiener number and the individual distances.

In order to make possible the next definition all the spanning trees of G, {T(G)}, and their distance matrices, {$\underline{\underline{D}}$(T(G))}, have to be considered.

Definition 6: A *distance tree* $T_u(G)$ of vertex $u \in G$ is such a spanning tree of G whose distance vector $\vec{\underline{\underline{d}}}_u(T_u(G))$ of vertex u is identical with the distance vector of this vertex in G, $\vec{\underline{\underline{d}}}_u(G)$:

$$\vec{\underline{\underline{d}}}_u(T_u(G)) \equiv \vec{\underline{\underline{d}}}_u(G) \quad .$$

This implies that the distances of vertex u to all vertices in G and in $T_u(G)$ are the same, hence, according to eq. (3) vertex u has the same distance numbers in G and in $T_u(G)$:

$$d(uv|T_u(G)) = d(uv|G) \quad ,$$
$$d(u|T_u(G)) = d(u|G) \quad . \tag{5}$$

Without relevance is whether more than one spanning tree of G meets the requirements of eq. (5); for convenience the most simple one may be used.

A method for the construction of a distance tree and how to use it for a fast calculation of distance and Wiener numbers will be found in [27]. The notion of distance trees may be traced back to Ore [29]; some authors [17] call it isometric tree.

2.2. Some Properties of Distance Numbers have been proved in [18]; they are briefly reviewed in this subsection.

2.2.1. Distance Numbers of Orbit Equivalent Vertices:

Proposition 1: If two vertices s and s' of a graph can be automorphically mapped onto each other, then they have equal distance number:

$$d(s|G) = d(s'|G) \quad .$$

Note, in general, however, the opposite statement is not true.

2.2.2. Distance Numbers of Centroid Vertices of a Tree:

Proposition 2: The vertices forming the centroid of a tree possess minimum distance number [30].

2.2.3. Distance Numbers of Adjacent Vertices: Let G be a simple connected graph; further let u and v be two adjacent vertices of G, i.e. $d(uv|G) = 1$.

Lemma 1: With respect to their distances to the adjacent vertices u and v, all vertices of G can unambiguously be classified such that they belong to one of the following three vertex subsets:

$$\begin{aligned} p &= \{p_i | d(p_i u) < d(p_i v)\} \quad ; \\ q &= \{q_j | d(q_j u) = d(q_j v)\} \quad ; \\ r &= \{r_k | d(r_k u) > d(r_k v)\} \quad . \end{aligned} \tag{6}$$

Set p comprises all those vertices which are closer to vertex u than to vertex v, while the opposite holds for set r. Evidently, u and v belong to p and r, respectively.

$$u \in p \; , \qquad v \in r \quad . \tag{6a}$$

Thus, the sets p and r are never empty, but the set q is non-empty if and only if the edge (uv) belongs to a cycle of odd length. It can be easily shown that the vertices of p and r belong to respective connected subgraphs of G; this is not necessarily true for the vertices of q.

Lemma 2: Let u and v be adjacent vertices and t be any other vertex of G. Depending on the subset to which t belongs the following equations hold:

$$t \in p: \quad d(tv) = d(tu) + 1 \quad ; \tag{7a}$$

$$t \in q: \quad d(tv) = d(tu) \quad ; \tag{7b}$$

$$t \in r: \quad d(tv) = d(tu) - 1 \quad . \tag{7c}$$

Proposition 3: If u and v are adjacent vertices of G their distance numbers satisfy the equation:

$$d(u|G) = d(v|G) + |r| - |p| \quad . \tag{8}$$

2.2.4. Distance Numbers of Non-Adjacent Vertices: Let G be a simple connected graph. Let $u,v \in G$ be two non-adjacent vertices with the distance $d(uv|G) = d > 1$, and let t be an abitrary vertex of G.

Lemma 3: Depending on their distance to the vertices $u,v \in G$, $d(uv) = d > 1$, all vertices $t \in G$ are uniquely partitioned into one of the following 2d+1 vertex subsets:

$$V(\delta) = \{t | d(tu) - d(tv) = \delta\} \quad ,$$
$$-d \leqq \delta \leqq +d \quad .$$

The sets $V(d-2j)$, $j = 0,1,\ldots,d$, are never empty. The sets $V(d-2j\pm1)$ or some of them are non-empty if and only if the geodesic $W_m(uv)$ or a part of it belongs to a cycle of odd length.

Proposition 4: If the distance between vertices u and v of a simple connected graph G is $d(uv) = d > 1$, the distance numbers of these vertices satisfy the following equation:

$$d(u|G) = d(v|G) + \sum_{\delta=-d}^{+d} \delta \cdot |V(\delta)| \quad . \tag{9}$$

Proposition 4 is a generalization of Proposition 3 and can be proved analogously. Note, for $d(uv) = 1$ eq. (9) turns into eq. (8).

2.2.5. The Distance and Wiener Numbers of Some Selected Graphs are given in the Appendix.

2.3. General Formalism

The quantities described above may be used in order to express the Wiener number of a graph in terms of the metric properties of its adequately chosen subgraphs. In such a way the influence of particular substructures and their mutual arrangements on the modelling of distance and Wiener numbers may be studied. For such a purpose the graph under study is considered as made up from another one by either a well-defined graph operation or the addition of vertices and/or edges; examples for such procedures are given in Sections 3-5. Thereby, it is not necessary that the graph, from which the graph into consideration is formed, is connected; but, certainly, the resulting graph must be connected in order to obtain finite distance and Wiener numbers.

There are no limits for the imagination of such procedures, but they are of use only if the metric properties of the graph(s) used at the start are either not altered by the procedure applied or their changes are under control. This demand drastically reduces the number of feasible procedures and, as shown in the following Sections, it sometimes even limits the area within which they can be applied.

A completely different approach to Wiener numbers is offered by eq. (8) provided the distance number of one vertex of G is known. This is exemplified for the graph shown in Figure 1a. Only for the vertices a and h a non-branched distance tree (P_{14}) can be constructed; thus, as seen from the Appendix and eq. (5) we have (Fig. 1b)

$$d(a|G) = d(h|G) = 49.$$

The stepwise application of eq. (8), starting with u = a, is illustrated by Fig. 1c - h; therein the adjacent vertices u and v are denoted by a double circle (⊙), the vertices belonging to p and r by full circles (•), and those from q by open circles (o), respectively. Although this method is very useful for a fast calculation of distance and Wiener numbers by hand it does not bring the insights into the modelling of these numbers as the approach outlined above.

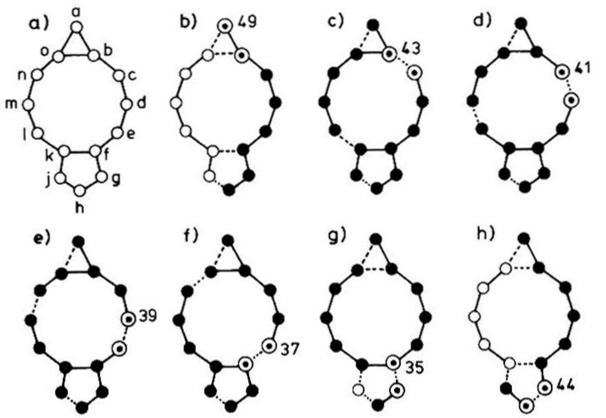

Fig. 1: Illustration of a fast calculation method by means of eq. (8).

3. Distance and Wiener Numbers of Graphs Generated by Covering

3.1. The Covering Procedure

In order to specify the term "covering" which is used here not strictly in the same sense as in [28] we call in mind the definition of an induced subgraph.

Definition 7: Let G be a graph with the vertex set V(G) and let G_o be a subgraph of G. Then, G_o is called an induced subgraph of G if its vertex set $V(G_o)$ is a subset of V(G) and two vertices are adjacent in G_o if and only if they are adjacent in G.

We now explain the covering procedure as follows: Let $G = \bigcup G_j$ be a graph consisting of J components $\{G_j | j = 1,2,...,J\}$ and let G_o be an induced subgraph of each component G_j. Let G_o have the (important) property that for any pair of vertices of G_o in each G_j at least one geodesic exists which consists of vertices of G_o only. Let the vertices belonging to $G_o \subseteq G_j$ be denoted by u_o, v_o,..., etc. and the corresponding vertices of G_j be denoted by u_j, v_j,... . Obviously the vertices u_o, v_o,... can be mapped into the vertices of any G_j according to the relations

$$u_o \to u_j \quad ,$$

$$v_o \to v_j \quad ,$$

etc., for all j.

The covering of the graphs G_j upon their common subgraph G_o, leading to a connected graph, H, means that the vertices $u_o, u_1, ...$ $..., u_J$ as well as $v_o, v_1 ..., v_J$, etc. are identified as single vertices u,v,..., respectively, in the resulting graph H. In the same manner the edges $(u_o v_o)$, $(u_1 v_1), \ldots, (u_J v_J)$, etc. are identified as a single edge $(uv) \in H$. The vertices not contained in G_o, namely $p_j \in G_j$, $p_j \notin G_o \subseteq G_j$, will be transferred unchanged into graph H: $p_j \in H$. This covering procedure is symbolically expressed by

$$H = G \operatorname{cov}(G_o) \quad .$$

Graph H obtained in such a manner has the following properties:

1. Each initial graph G_j is an induced subgraph of H; the same holds for G_o.

2. If one of the initial graphs G_j is removed from H, then H is decomposed into J-1 components, according to

$$H \setminus G_j = \biguplus (G_k \setminus G_o) \quad , \quad k \neq j \quad . \tag{10a}$$

3. If G_o is removed from H, then H is decomposed into J components, $G_j^c = G_j \setminus G_o$ according to

$$H \setminus G_o = \biguplus (G_k \setminus G_o) \quad . \tag{10b}$$

Obviously, the vertices of G_o act on H as a vertex cut set.

4. According to that decomposition one obtains the number n of vertices of H as follows:

$$n = \sum n_j - n_o(J - 1) \quad , \tag{11}$$

where n_j and n_o denote the number of vertices of G_j and G_o, respectively.

5. As a consequence of the property of G_o assumed above, the following equality

$$d(uv|G_o) = d(uv|G_j) = d(uv|H) \tag{12}$$

holds for all pairs of vertices which belong to G_o, $u,v \in G_o$, and for all j.

Proposition 5: Let graph H be obtained by means of a covering of graphs G_j, j = 1,2,...,J upon their common induced subgraph G_o which meets the requirements of eq. (12), then the distance number of vertex $u \in G_o$ in graph H is specified by

$$d(u|H) = \sum_{j=1}^{J} d(u|G_j) - (J-1)d(u|G_o) \quad . \tag{13}$$

Proof: a) Let the complement of G_o in G_j be denoted by $G_j^c = G_j \backslash G_o$. b) The distance vector of vertex u in G_j can be decomposed into two components as follows:

$$\vec{\underline{\underline{d}}}_u(G_j) = (\vec{\underline{\underline{d}}}_u(G_o)^T \; \vec{\underline{\underline{d}}}_u(G_j^c)^T)^T \quad .$$

The sum of the entries of $\vec{\underline{\underline{d}}}_u(G_j^c)$ is obviously $d(u|G_j) - d(u|G_o)$.
c) Due to eq. (12) the components $\vec{\underline{\underline{d}}}_u(G_o)$ are equal for all $\vec{\underline{\underline{d}}}_u(G_j)$.
d) The distance between any vertex $u \in G_o$ and any vertex $x_j \in G_j^c$ of the complements remains unchanged by the covering procedure, hence, one has $d(ux_j|H) = d(ux_j|G_j)$. Thus we can express $\vec{\underline{\underline{d}}}_u(H)$ as follows:

$$\vec{\underline{\underline{d}}}_u(H) = (\vec{\underline{\underline{d}}}_u(G_o)^T \; \vec{\underline{\underline{d}}}_u(G_1^c)^T \; \vec{\underline{\underline{d}}}_u(G_2^c)^T \ldots \vec{\underline{\underline{d}}}_u(G_J^c)^T)^T \quad .$$

e) Having in mind that the sum of the entries of $\vec{\underline{\underline{d}}}_u(G_j^c)$ equals $d(u|G_j) - d(u|G_o)$, the application of eq. (3) leads now directly to eq. (13). ◻

By means of eq. (13) the distance numbers of those vertices of H which belong to G_o are given, provided G_o has the property assumed above and, hence, eq. (12) is satisfied. For the distance numbers of the other vertices of H (which belong to the complements G_j^c) such a general formula can not be derived, because the distance between such two vertices, say $x_j \in G_j^c$ and $y_k \in G_k^c$, strongly depends on which of the vertices of G_o are passed through by the geodesic $W_m(x_j y_k | H)$. Hence, expressions for distance numbers like $d(x_j | H)$ can be derived only for particular G_o which must be defined in all their structural details.

In the next subsections we present such derivations of distance and Wiener numbers for the most simple subgraphs G_o. There we always will start from a disconnected graph G composed of J components G_j with n_j vertices, $1 \leqq j \leqq J$, $n_j > n_o$ and a particular choice of G_o. If the requirement of eq. (12) implies some constraints on the structure of G_j, they will be listed. It will always be assumed that all informations about G_j (e.g. distances, distance and Wiener numbers) are available.

3.2. Covering of Several Graphs upon a Vertex

The most simple case for the formation of graph H from several graphs G_j, $1 \leqq j \leqq J$, by the covering procedure is achieved if G_o consists of a single vertex only, i.e. $G_o = K_1$. Here we treat this case, $H = G\ \mathrm{cov}(K_1)$.

Let $u_j \in G_j$ denote the vertex upon which covering takes place, then we have

$$u_1 = u_2 = \ldots = u_j = \ldots = u_J = u \in H \quad .$$

Since $n_o = 1$, from eq. (12) one obtains

$$n = \sum n_j - (J - 1) \quad .$$

Let $t_j \in G_j$ be an arbitrary vertex of component G_j, different from u_j.

<u>Property 3.2.:</u> The distance numbers of the vertices in H, as well as the Wiener number of H, are specified by the equations:

$$d(u|H) = \sum d(u_j|G_j) \quad ; \tag{14}$$

$$d(t_j|H) = d(t_j|G_j) + (n-n_j)\cdot d(t_j u_j|G_j) + d(u|H) - d(u_j|G_j) \quad ; \tag{15}$$

$$W(H) = \sum W(G_j) + n\cdot d(u|H) - \sum n_j \cdot d(u_j|G_j) \quad . \tag{16}$$

<u>Proof</u>: a) Eq. (14) follows directly from eq. (13) because $d(u|G_o) = 0$. b) In the development of the proof of eq. (15) one proceeds from eq. (3) and obtains

$$d(t_j|H) = \sum_k \sum_{\{s_k\}} d(t_j s_k|H) \quad .$$

When $k = j$, then $d(t_j s_j|H) = d(t_j s_j|G_j)$ holds while for $k \neq j$, $d(t_j s_k|H) = d(t_j u_j|G_j) + d(u_k s_k|G_k)$ is valid because each path from a vertex of one component to a vertex of another one must pass through the covering vertex $u \in H$ which is a vertex cut set of cardinality 1. By substituting these expressions, as well as by summing over $\{s_k\}$ one obtains

$$d(t_j|H) = d(t_j|G_j) + \sum_{k\neq j} [(n_j-1)\cdot d(t_j u_j) + d(u_j|G_j)] \quad .$$

From this eq. (15) follows directly. c) From eq. (4) it follows that

$$2W(H) = d(u|H) + \sum_j \sum_{\{t_j \neq u_j\}} d(t_j|H) \quad .$$

Substituting eq. (15) into this formula one obtains eq. (16). ◻

The use of eqs. (14) - (16) is illustrated by Fig. 2: In Fig. 2a the resulting graph H is depicted. Fig. 2b shows the components G_j of G, j = 1,2,3,4, and the distance numbers of their respective vertices. The vertex u upon which the covering procedure takes place is marked by full circles (•). In Fig. 2c the distance numbers of the vertices of H are given; they may be obtained by means of eqs. (14) and (15).

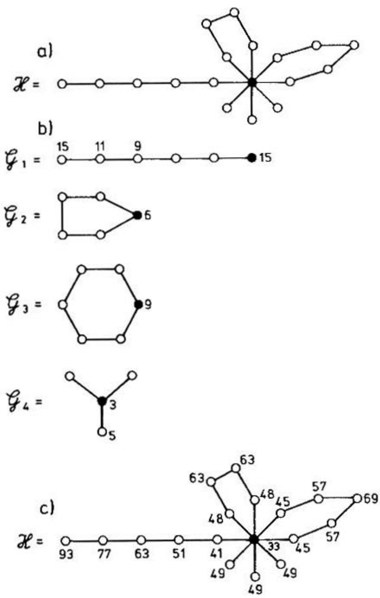

Fig. 2. Graph H formed according to H = G cov(K_1). For details see text.

Figure 2 also illustrates another aspect: Obviously the vertex u represents in H an articulation, i.e. a vertex cut set of cardinality 1. It is easily recognized that graphs G_j, j = 1,2,3,4, are related to the decomposition of H at the articulation u. Thus any graph H with an articulation can be considered as being formed according to H = G cov(K_1) which immediately leads to the components G_j of G. In this procedure, it is not necessary to decompose H into the largest possible number of subgraphs; e.g. G_4 in Figure 2 still has a vertex u such that the partitioning could be continued further. For convenience the decomposition should be carried out so as to arrive at a set of G_j's for which the values of $d(u|G_j)$ are known or can readily be calculated by means of the formulae given in the Appendix.

Let graph H be obtained by the covering of J isomorphic components G_j upon the equivalent vertices u_j, i.e. $G_1=G_2=\ldots=G_J=G'$ and for all pairs G_j and G_k, j≠k, let an isomorphic mapping exist which maps u_j onto u_k and vice versa. Then, the eqs. (14) - (16) take the following form:

$$d(u|H) = J \cdot d(u|G') \quad ; \qquad (14a)$$

$$d(t|H) = d(t|G') + (J-1)\cdot d(u|G') + (J-1)(n'-1)\cdot d(tu|G') \quad ; \qquad (15a)$$

$$W(H) = J \cdot W(G') + (n'-1)\cdot J(J-1)\, d(u|G') \quad ; \qquad (16a)$$

where $n'=n_1=\ldots=n_j$ is the number of vertices in the respective components.

As seen from eq. (16a), for a specified type and for a given number of components, the Wiener number W(H) depends linearly on the distance number of the vertex where the covering of components takes place. This linearity is illustrated for the graph given in Figure 3 by the equation:

$$W(H) = 92 + 12d(u|G') \quad .$$

The example shown in Figure 3 sheds some light on the branching problem: Since $d(v|G_j) > d(s|G_j) > d(u|G_j) > d(t|G_j) > d(r|G_j)$ the sequence $W(I) > W(II) > W(III) > W(IV) > W(V)$ results.

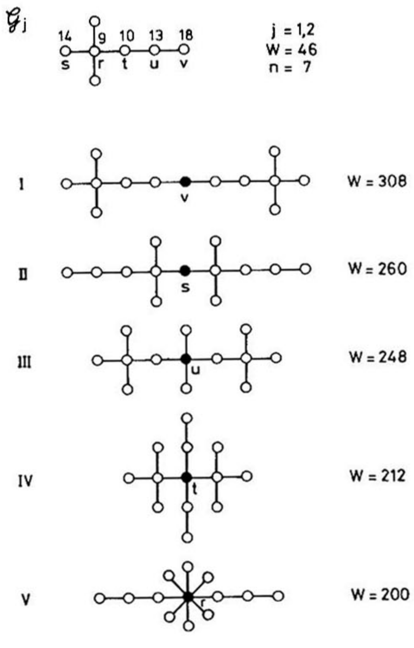

Fig. 3. Illustration of the application of eq. (16a) to the problem of branching: Due to relation $d(v|G_j) > d(s|G_j) > d(u|G_j) > d(t|G_j) > d(r|G_j)$ one obtains $W(I) > W(II) > W(III) > W(IV) > W(V)$.

3.3. Covering of Graphs upon an Edge

For the sake of simplicity we regard here the case of edge covering of a graph containing only two components.

Let G consist of components G_1 and G_2 having n_1 and n_2 vertices, respectively. Graph H is then formed by covering the edges $(a_1b_1) \in G_1$ and $(a_2b_2) \in G_2$, where $a_1=a_2=a \in H$ and $b_1=b_2=b \in H$. The number of vertices of H is $n = n_1+n_2-2$. Let $x \in G_1$ and $y \in G_2$ be arbitrary vertices in G_1 and G_2, respectively. By analogy with eqs. (6) and (7) we define:

$$\delta_1 = d(ax) - d(bx) \quad , \quad -1 \leqq \delta_1 \leqq +1 \quad ;$$

$$\delta_2 = d(ay) - d(by) \quad , \quad -1 \leqq \delta_2 \leqq +1 \quad .$$

The vertex sets of G_1 and G_2 are distributed on this basis among the following subsets:

$$G_1: \ p \ \ldots \ \delta_1 = -1 \quad , \qquad G_2: \ s \ \ldots \ \delta_2 = -1 \quad ,$$

$$q \ \ldots \ \delta_1 = 0 \quad , \qquad t \ \ldots \ \delta_2 = 0 \quad ,$$

$$r \ \ldots \ \delta_1 = +1 \quad , \qquad u \ \ldots \ \delta_2 = +1 \quad .$$

Denoting an arbitrary vertex of these six subsets by p, q, r, s, t, and u, respectively, we can specify the following

Property 3.3: The distance numbers of vertices in graph H, as well as its Wiener number are determined by the following equations:

$$d(a|H) = d(a|G_1) + d(a|G_2) - 1 \quad ,$$
$$d(b|H) = d(b|G_1) + d(b|G_2) - 1 \quad ; \qquad (17)$$

$$d(p|H) = d(p|G_1) + [d(a|G_2) + d(b|G_2) - |s| + |u|]/2 - 1 + (n_2-2)[d(pa|G_1) + d(pb|G_1) - 1]/2 \quad ,$$

$$d(q|H) = d(q|G_1) + [d(a|G_2) + d(b|G_2) - n_2 + |t|]/2 + (n_2-2)[d(qa|G_1) + d(qb|G_1)]/2 \quad ,$$

$$d(r|H) = d(r|G_1) + [d(a|G_2) + d(b|G_2) + |s| - |u|]/2 - 1 + (n_2-2)[d(ra|G_1) + d(rb|G_1) - 1]/2 \quad ,$$

$$d(s|H) = d(s|G_2) + [d(a|G_1) + d(b|G_1) - |p| + |r|]/2 - 1 + (n_1-2)[d(sa|G_2) + d(sb|G_2) - 1]/2 \quad , \qquad (18)$$

$$d(t|H) = d(t|G_2) + [d(a|G_1) + d(b|G_1) - n_1 + |q|]/2 + (n_1-2)[d(ta|G_2) + d(tb|G_2)]/2 \quad ,$$

$$d(u|H) = d(u|G_2) + [d(a|G_1) + d(b|G_1) + |p| - |r|]/2 - 1 + (n_1-2)[d(ua|G_2) + d(ub|G_2) - 1]/2 \quad ,$$

$$\begin{aligned} W(H) = W(G_1) + W(G_2) + 1 + \tfrac{1}{2}\{ &+|q|\cdot|t| - n_1 \cdot n_2 + \\ &+ (n_2-2)[d(a|G_1) + d(b|G_1)] + \\ &+ (n_1-2)[d(a|G_2) + d(b|G_2)] - \\ &- [d(a|G_1) - d(b|G_1)]\cdot[d(a|G_2) - d(b|G_2)]\} \quad . \end{aligned} \qquad (19)$$

<u>Proof:</u> a) Eq. (17) follows directly from eq. (13). b) From eq. (3) one obtains

$$d(p|H) = \sum_{\{x\}} d(px|H) + \sum_{\{y\}} d(py|H) - d(pa|G_1) - d(pb|G_1) \quad .$$

The last two terms are corrections, due to the fact that $a,b \in \{x\}$ and $a,b \in \{y\}$, i.e. the distances $d(pa)$ and $d(pb)$ are contained in both sums. The first sum yields $d(p|G_1)$ since $d(px|H) = d(px|G_1)$. The term in the second sum is $d(py|H) = d(pa|G_1) + d(ay|G_2)$, hence the second sum is $n_2 \cdot d(pa|G_1) + d(a|G_2)$. Taking into account the equality $d(pb|G_1) = d(pa|G_1) + 1$ one obtains

$$d(p|H) = d(p|G_1) + d(a|G_2) - 1 + (n_2-2) \cdot d(pa|G_1) \quad .$$

From $d(pa|G_1) = d(pb|G_1) - 1$ it follows that

$$d(pa|G_1) = [d(pa|G_1) + d(pb|G_1) - 1]/2 \quad .$$

Analogously: from eq. (8) $d(a|G_2) = d(b|G_2) - |\delta| + |u|$ one obtains

$$d(a|G_2) = [d(a|G_2) + d(b|G_2) - |\delta| + |u|]/2 \quad .$$

Substituting these two formulae into the intermediate result one arrives at eq. (18) which is thus symmetrical with respect to a and b, of G_1 and G_2, respectively. c) $d(r|H)$, $d(s|H)$, and $d(u|H)$ are derived in a similar manner. d) From eq. (3) we have

$$d(q|H) = \sum_{\{x\}} d(qx|H) + \sum_{\{y\}} d(qy|H) - d(qa|G_1) - d(qb|G_1)$$

The first sum yields $d(q|G_1)$. The terms in the second sum are: for $y \in \delta \cup t$, $d(qy|H) = d(qa|G_1) + d(ay|G_2)$; for $y \in u$, $d(qy|H) = d(qb|G_1) + d(by|G_2)$. Taking into account $d(qa|G_1) = d(qb|G_1)$ and $d(bu|G_2) = d(au|G_2) - 1$, one comes to the intermediate equation

$$d(q|H) = d(q|G_1) + d(a|G_2) - |u| + (n_2-2) \cdot d(qa|G_1) \quad .$$

Making use of substitutions given in (b), one obtains eq. (18) which once again is symmetrical with respect to a and b, of G_1 and G_2, respectively. e) $d(t|H)$ is obtained analogously. f) From eq. (4) we have

$$2W(H) = \sum_{\{x\}} d(x|H) + \sum_{\{y\}} d(y|H) - d(a|H) - d(b|H) \quad .$$

Hence, by means of eqs. (17) and (18) one obtains

$$\begin{aligned} 2W(H) = {} & 2W(G_1) + 2W(G_2) - d(a|H) - d(b|H) + \\ & + n_1[d(a|G_2) + d(b|G_2)]/2 - |p| - |r| + \\ & + n_2[d(a|G_1) + d(b|G_1)]/2 - |s| - |u| - \\ & - [|p|(|s| - |u|) + |q|(n_2-|t|) - |r|(|s| - |u|) + \\ & + |s|(|p| - |r|) + |t|(n_1-|q|) - |u|(|p| - |r|)]/2 + \\ & + (n_2-2)[d(a|G_1) + d(b|G_1) - |p| - |r|]/2 + \\ & + (n_1-2)[d(a|G_2) + d(b|G_2) - |s| - |u|]/2 \quad . \end{aligned}$$

Taking also

$$d(a|H) + d(b|H) = d(a|G_1) + d(b|G_1) + d(a|G_2) + d(b|G_2) - 2$$

we come to the intermediate result

$$\begin{aligned} 2W(H) = {} & 2W(G_1) + 2W(G_2) + 2 + \\ & + (n_2-2)[d(a|G_1) + d(b|G_1)] + \\ & + (n_1-2)[d(a|G_2) + d(b|G_2)] - A/2 \quad , \end{aligned}$$

where

$$\begin{aligned} A = {} & (n_1-|q|)(n_2+|t|) + (n_1+|q|)(n_2-|t|) + \\ & + 2(|p|-|r|)(|s|-|u|) \quad . \end{aligned}$$

From eq. (8) we have, however, $(|p|-|r|) = -d(a|G_1) + d(b|G_1)$ and $(|s|-|u|) = -d(a|G_2) + d(b|G_2)$. Taking into account these two equalities in calculating A one comes directly to eq. (17). □

When two graphs are covered upon different edges, the Wiener number W(H) of the resulting graph depends mainly on the distance numbers of the covered vertices while the influence of the other terms is negligible.

The last term in eq. (17) is of definite interest. When two graphs, G_1 and G_2, are to be covered on edges $(ab) \in G_1$ and $(cd) \in G_2$, and a and b, as well as c and d, are non-equivalent vertices then two different modes of covering are possible: In the first mode, generating H_1, one covers a on c and b on d; in the other mode, generating H_2, vertex a is covered on d and vertex b on c. In the study of topological effects on molecular orbitals (TEMO) [31-34] such pairs of graph like H_1 and H_2 play some role. As shown in [31,35-37] their eigenvalue spectra are particularly related to the so-called TEMO pattern. It is readily seen from eq. (17) that all terms except the last one are the same for the two modes of covering. Then the difference in the two Wiener numbers of the resulting graphs H_1 and H_2 is a simple expression:

$$W(H_1) - W(H_2) = [d(a|G_1) - d(b|G_1)][d(c|G_2) - d(d|G_2)]$$

Figure 4 presents an example of two such coverings. The resulted graphs are the C-graphs of 1,8-, and 1,5-naphthochinodimethane, respectively. This is an example of a pair of isomers formed by linking two equal fragments in two topologically different ways [31-37]. In the case that graphs H_1 and H_2 are isomorphic with the skeleton graphs of fully conjugated systems one has for the total π-electron energies [38,39] the inequality $E_\pi(H_1) \geqq E_\pi(H_2)$; this throws some light on the dependence of this quantity on molecular topology, as well as the reasons for the good correlation of E_π with Wiener numbers [40].

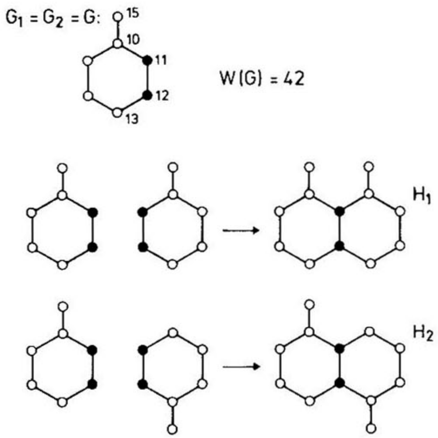

Fig. 4. Illustration of the importance of the last term in eq. (19): $W(H_1) = 175$, $W(H_2) = 176$.

Figure 4 also shows what type of graphs H may be considered as being formed according to $H = G\ \mathrm{cov}(P_2)$. The covering edge needs not to belong to a cycle of G_j.

If one has to cover more than two graphs upon an edge one treats the task in the same manner as outlined in the proof of eqs. (17) - (19).

3.4. Covering of Graphs upon a Path Graph P_m

To illustrate this procedure we treat here the most simple case where G_o represents a path graph P_m and all the G_j's are cycles $C_{k(j)}$. Due to the requirement of eq. (12) m must be chosen such that the relation $k_{(j)} \geqq 2m-2$ holds for all cycles to be covered. Then eq. (13) can be applied for all vertexes $t \in P_m$ and $d(t|G_o)$ may be taken from the Appendix. Then the vertex set of each cyclic component $C_{k(j)}$ is partitioned into 2m-1 subsets $V_j(\delta)$, $-m+1 \leqq \delta \leqq m-1$, according to Lemma 3. The further treatment is quite similar to that one used in the proof of property 3.3.

In the case most relevant to chemistry one covers two cyclic graphs (J=2). This is illustrated in Figure 5 where for $G_1 = C_6$, $d(u|G_1) = 9$, and for $G_2 = C_7$, $d(u|G_2) = 12$. In case of Figure 5a we choose $G_o = P_2$ (as already treated in the foregoing subsection) with the distance numbers (1,1); in case of Figure 5b $G_o = P_3$ with distance numbers (3,2,3) and a terminal vertex of P_3 is chosen as u. From eq. (13) one obtains the distance numbers $d(u|H_a) = 20$ and $d(u|H_b) = 18$, respectively. The Wiener numbers are $W(H_a) = 142$ and $W(H_b) = 105$, respectively.

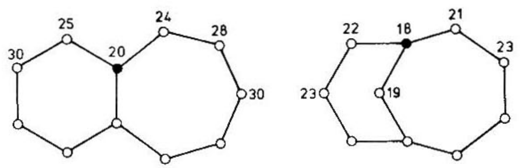

Fig. 5. Two graphs illustrating the covering upon a path graph P_m.

The requirement $k(j) \geqq 2m-2$ stated above does not limit the applicability of the covering procedure described here as it might seem at a first glance. Suppose we wish to generate graph H by covering J cycles of the lengths $k(1) < k(2) < \ldots < k(J)$ upon P_m. Because P_m is an induced subgraph of all those cycles, obviously $m \leqq k(1)$ holds in any case. Let us now consider graph H: It contains two vertices of degree J+1 which are linked by J+1 paths of the respective lengths m-1 and $k(1)+1-m < k(2)+1-m < \ldots < k(J)+1-m$. If $m-1 \leqq k(1)+1-m$ the requirement is met. Let us suppose, however, that it is failed by $m-1 > k(1)+1-m$. Observe in that case that H may be also formed by covering J cycles of the lengths k(1) and $[k(2)+k(1)+2-2m] < [k(3)+k(1)+2-2m] < \ldots < [k(J)+k(1)+2-2m]$ upon the path graph $P_{k(1)+2-m}$.

4. Linkage of Two Graphs by Means of One or More Disjoint Edges

The problems treated in this section strongly vary in their complexity which increases very fast with the number of added edges. The general equation for non-planar graphs is much more complex than that for planar graphs. Due to the fact that with only very few exceptions [41] chemical compounds may be represented by planar graphs we restrict our consideration exclusively to planar graphs.

In the following treatment we deal with a planar graph G of two components, $G = G_1 \cup G_2$, whose vertices are denoted by $x_j \in G_1$ and $y_k \in G_2$, respectively. Edges $\{(x_j y_k)\}$ are added to G to obtain the connected graph $H = G \cup \{(x_j y_k)\}$. Apparently, the set of added edges is a cut set of edges in H since its removal transforms

H once again into the two initial components. The added edges should be chosen such that the graph produced is planar.

4.1. Linkage of Graphs by Bridges

Let G_1 and G_2 be two simple graphs with n_1 and n_2 vertices, respectively and their union be denoted by $G = G_1 \cup G_2$. Let $x \in G_1$ and $y \in G_2$ be arbitrary vertices of the two components. Graph H is constructed as $H = G \cup \{(uv)\}$ where u G_1 and $v \in G_2$. Apparently, (uv) is a bridge in H.

Property 4.1(1): The distance numbers of vertices, as well as the Wiener number of H are specified as follows:

$$d(x|H) = d(x|G) + d(u|G) + n_2[d(ux) + 1] \quad ,$$
$$d(y|H) = d(y|G) + d(u|G) + n_1[d(vy) + 1] \quad ; \qquad (20)$$

$$d(u|H) = d(u|G) + d(v|G) + n_2 \quad ,$$
$$d(v|H) = d(v|G) + d(u|G) + n_1 \quad ; \qquad (21)$$

$$W(H) = W(G_1) + W(G_2) + n_1 n_2 +$$
$$+ n_2 d(u|G) + n_1 d(v|G) \quad . \qquad (22)$$

Due to the assumptions $x \in G_1$ etc., in eqs. (20) - (22) the indexing of components in $d(x|G_1)$, etc. are dropped.

Proof: a) Observe first that the added edge belongs to each geodesic $W_m(x_j y_k)$. b) From eq. (3) one generates equations for the individual distance numbers. Further, the sums over all vertices in H can be partitioned into two separate sums, the first one over all $x_j \in G_1$ and the second one over all $y_k \in G_2$. c) Thus one obtains

$d(x_i|H) = \sum_{G_1} d(x_i x_j) + \sum_{G_2} [d(x_i u) + 1 + d(vy)]$ from which eq. (20) follows directly. d) The proof of $d(y|H)$ is carried out analogously. e) From eq. (20) one obtains eq. (21) since $d(uu)=d(vv)=0$. f) An expression for $2W(H)$ follows from eq. (4). Further, the summing is carried out as described in (b) from which eq. (22) follows directly. □

The range of magnitude of $W(H)$ depends on the last two terms of eq. (20) for given two-component graphs, since $W(G_1)$, $W(G_2)$, n_1, and n_2 are constant. Therefore, $W(H)$ is minimal (maximal) when the two distance numbers $d(u|G)$ and $d(v|G)$ are minimal (maximal).

In the particular case where $G_2 = P_1$ contains a single vertex v only, one obtains $n_2 = 1$, $d(v|G) = 0$ and $W(G_2) = 0$. Hence, the following equations are derived from eqs. (20) - (22):

$$\begin{aligned} d(x|H) &= d(x|G) + d(xu) + 1 \quad , \\ d(u|H) &= d(u|G) + 1 \quad , \\ d(v|H) &= d(u|G) + n_1 \quad , \\ W(H) &= W(G_1) + n_1 + d(u|G) \quad . \end{aligned} \qquad (23)$$

These equations are useful when a star of graphs G_j is formed at vertex v (Fig. 6). They are applied to each graph $G_j \cup \{v_j\}$ forming G_j' and then, after joining these graphs G_j' by a covering on vertex v_j, eqs. (14) - (16) should be used.

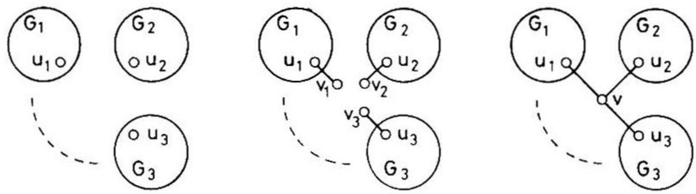

Fig. 6. Building of a star of graphs G_j around vertex v.

In the case of isomorphic components, $G_1 = G_2$, which correspond to the constitutional graph of a chemical compound, graph H obtained from G_1 and G_2, as described in the foregoing text, is the constitutional graph of the dimer of this compound. Analogously, one can calculate the vertex distance numbers and the Wiener number of the polymer constitutional graphs, assuming that the monomer units are linked regularly to each other by a single chemical bond. Two such types of structures can be constructed, known as Fascia- and Rota-graphs, respectively [42]. These two types are schematically represented by Figure 7. Their structure is uniquely specified by the structure G of the monomer units G_j, the number J of the monomer units, and the set of edges K linking the monomer units in a regular manner.

We consider first the distance properties of the Fascia-graphs $F = F(G,J,K)$. Let G be a graph composed of J isomorphic components G_j with n vertices which generally will be denoted by $p_j \in G_j$. Let u_j, $v_j \in G_j$ be the terminal vertices of the connecting edges $K_F = \{(u_{j+1}v_j) | 1 \leq j \leq J-1\}$. Evidently, $F = G \cup K_F$. The

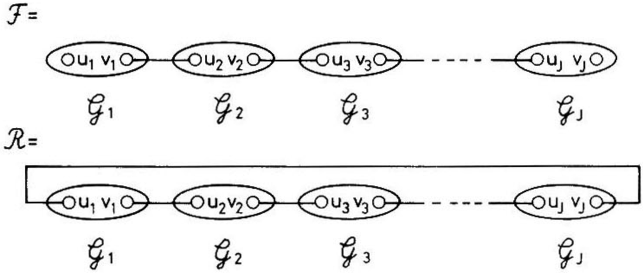

Fig. 7. Schematic representation of the Fascia-graphs $F(G,J,K_F)$ and Rota-graphs $R(G,J,K_R)$, respectively.
$K_F = \{(u_{j+1}v_j) \mid 1 \leq j \leq J-1\}$, $K_R = K_F \cup \{(u_1v_J)\}$.

distance between the vertices u_j and v_j in the component G_j, $d(u_jv_j)$, will be denoted (L-1) for short; usually (L-1) > 1.

Property 4.1(2): The vertex distance numbers and the Wiener number of F are specified by the following equations:

$$d(p_j|F) = d(p|G) + n(J-1) + nL[\binom{J-j}{2} + \binom{j-1}{2}] +$$
$$+ (J-j)d(u|G) + (j-1)d(v|G) +$$
$$+ n[(J-j)d(pv|G) + (j-1)d(pu|G)] \quad ; \qquad (24)$$
$$W(H) = JW(G) + n^2[\binom{J}{2} + L\binom{J}{3}] +$$
$$+ n\binom{J}{2}[d(u|G) + d(v|G)] \quad .$$

Proof: a) Let $x_k \in G_k$ be an arbitrary vertex in component G_k. According to eq. (3) one obtains for the distance number of vertex $p_j \in G_j$: $d(p_j|F) = \sum_{k=1}^{J} \sum_{\{x_k\}} d(p_jx_k|F)$, where the summation is taken over $k = 1,2,\ldots,J$, as well as over all vertices $x_k \in G_k$. b) Three cases $k < j$, $k = j$, and $k > j$ will be examined. In the first case x_k can be reached only via u_j while in the last case it can be reached only via v_j. c) When j and k are not neighbours, i.e. $|j-k| \geqq 2$, the components located between G_j and G_k will be crossed on the shortest path; the latter in each G_i is of length $d(u_i, v_i) = L-1$. In order to reach the next components (G_{i+1} or G_{i-1}) it is necessary to follow the edges which belong to K_F, so as to obtain $d(u_i u_{i-1}) = d(v_i v_{i-1}) = L$. d) Taking into account all these factors one obtains as an intermediate result for $d(p_j|F)$:

$$d(p_j|F) = \sum_{k=1}^{j-1} \sum_{\{x_k\}} [d(pu|G) + 1 + (j-k-1)L + d(vx|G)] +$$

$$+ \sum_{\{x_j\}} d(px|G) + \sum_{k=j+1}^{J} \sum_{\{x_k\}} [d(pv|G) + 1 + (k-j-1)L + d(ux|G)].$$

After some transformations this ressolves into eq. (24). e) The equation for W(F) is proved analogously. ◻

Rota-graphs $R = R(G,J,K_R)$, which differ from the respective Fascia-graphs $F = F(G,J,K_F)$ by the term $K_R = K_F \cup \{(u_1 v_J)\}$, cannot be characterized by equations as simple as eq. (24). The structure of G_j has a strong influence on the results. The parity of J ($J=2K$ or $J=2K+1$) is also of some importance.

Due to the complexity of the problem we only outline its treatment. We select an arbitrary vertex $p_k \in G_k$ in component G_k, and the same is done with the vertex $q_j \in G_j$ in component G_j, where $j < k$. As a consequence of the cyclic macrostructure of R, vertex q_j can be reached by two paths originating from p_k. The first one, $W_u(p_k q_j)$, proceeds via u_k and v_j while the other one, $W_v(p_k q_j)$, passes via v_k and u_j. Their lengths are as follows:

$$|W_u(p_k q_j)| = d(u_k p_k) + (k-j-1)L + 1 + d(v_j q_j) \quad ,$$

$$|W_v(p_k q_j)| = d(v_k p_k) + (J-k-1+j)L + 1 + d(u_j q_j) \quad ,$$

and, hence, their difference is

$$|W_u(p_k q_j)| - |W_v(p_k q_j)| = (2k-2j-J)L + \delta_p - \delta_q \quad , \tag{25}$$

where in accord with Lemma 3 δ_p and δ_q stand for

$$\delta_p = d(u_k p_k) - d(v_k p_k) \quad ,$$

$$\delta_q = d(u_j q_j) - d(v_j q_j) \quad .$$

$W_u(p_k q_j)$ will be the geodesic if the right hand side of eq. (25) is negativ or zero, otherwise $W_v(p_k q_j)$ is the geodesic. Due to the bounds given in Lemma 3, i.e. $-L+1 \leqq \delta_p, \delta_q \leqq L-1$, one has $|\delta_p - \delta_q| \leqq 2L-2$. With this in mind one concludes from eq. (25) that the sign of (2k-2j-J) selects the geodesic out of the two paths, provided $|2k-2j-J| > 2$; in the case of $|2k-2j-J| \leqq 2$ a detailed analysis is necessary.

Due to the cyclic macrostructure of the Rota-graphs, the equality $d(p_j|R) = d(p_k|R)$ holds for each pair of indices j and k. Hence, the Wiener number W(R) is (J/2) times the sum of the distance numbers of the vertices of one unit G_j.

4.2. Joining Two Graphs by Two Edges

Here one faces a complicated problem once again. We shall elucidate the condition which strongly simplifies the problem and present explicit formulae for this case only.

Let $G = G_1 \cup G_2$ be a graph composed of two components, G_1 and G_2. The connected graph $H = G \cup \{(ru),(sv)\}$ is obtained by adding two edges to G, namely (ru) and (sv), $r,s \in G_1$, $u,v \in G_2$. The distances between the pairs of vertices are $d(rs|G) = d_1$ and $d(uv|G) = d_2$. One may assume, without loss of generality, that $d_1 \geqq d_2$, i.e. $(d_1-d_2) \geqq 0$.

First of all we have to clarify whether the distances between two vertices of one and the same component could be shorter in H than in G; a consequence of this would be a reorganization of some geodesics by which the problem would become more complicated. Due to the condition $(d_1-d_2) \geqq 0$ assumed above such a shortening can only occur for vertices of G_1, say $p,q \in G_1$; obviously these vertices will be close to r and s, respectively. We will examine the possibility of $d(pq|H) < d(pq|G)$ a) for the vertices r and s which are end points of the added edges; then b) for a pair of vertices, say $a,b \in G_1$ in which case the shortest path $W_m(ab|G)$ should have λ edges, $0 < \lambda \leqq d_1$, in common with $W_m(rs|G)$; and finally c) for another pair, say $c,f \in G_1$, in the case where the shortest paths $W_m(cf|G)$ and $W_m(rs|G)$ have no edge ($\lambda=0$) in common. Particular examples of these are illustrated in Figure 8. In all these cases in H the path $W_1 = W_m(pq|G)$ is in competition with another path W_2 for minimal length; here $W_2 = W(p,\ldots,r,u,\ldots,v,s,\ldots,q)$ passes from p via r, u, s, and v to q and should be as short as possible. We shall call such paths *paths of relative minimal length*.

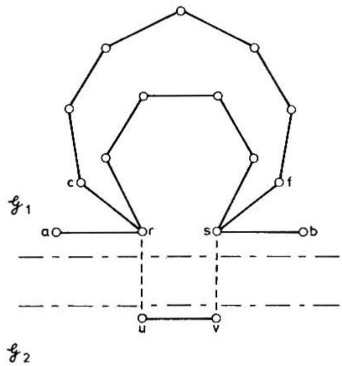

Fig. 8. Specific subgraph illustrating the shortening of distances when two components, G_1 and G_2, are linked by two edges, {ru} and {sv}, $d_1 = 5$, $d_2 = 1$, $d_1 - d_2 = 4 > 2$; $d(cf|G) = 6$, $d(cf|H) = 5$; $r,s \notin W_m(cf|G)$, $r,s \in W_m(cf|H)$.

a) Distance d(rs): In G we have $d(rs|G) = d_1$. The path W_2, which exists only in H, has a minimal length of $|W_2| = d_2 + 2$ and will be shorter than $W_1 = W_m(rs|G)$ if $|W_2| - |W_1| = (d_2+2-d_1) < 0$. Thus the condition for the shortening $d(rs|H) < d(rs|G)$ is expressed by

$$d_1 > (d_2 + 2) \quad . \tag{26a}$$

b) Distance d(ab) ($W_m(ab|G)$ and $W_m(rs|G)$ have λ edges in common, $0 < \lambda < d$): the path W_2 has the length

$$|W_2| = [d(ab|G) - \lambda] + [(d_1 - \lambda) + d_2 + 2] =$$
$$= d(ab|G) + [d_1 + d_2 + 2 - 2\lambda]$$

and will be shorter than $|W_1| = d(ab|G)$ if $[d_1+d_2+2-2\lambda] < 0$, i.e.:

$$2\lambda > (d_1 + d_2 + 2) \quad . \tag{26b}$$

If λ takes its maximal value $\lambda = d_1$, eq. (26b) converts into eq. (26a). In Figure 8 $\lambda = d_1$ is assumed for $W_m(ab|G)$.

c) Distance d(cf) ($W_m(cf|G)$ and $W_m(rs|G)$ have no edge (λ=0) in common): the path W_2 has the length

$$|W_2| = d(cr|G) + d_2 + 2 + d(fs|G)$$

and will be shorter than $|W_1| = d(cf|G)$ if

$$[d(cr|G) + d(fs|G) - d(cf|G) + d_2 + 2] < 0 \quad . \tag{26c}$$

Evidently, eq. (26c) can be realized only in situations where eq. (26a) is satisfied.

The essence of these considerations is as follows: (i) By $(d_1-d_2) > 2$, eq. (26a), is indicated that some of the geodesics established in G_1 are no longer geodesics when H is formed from $G_1 \cup G_2$; we use the term "reorganization of geodesics" to refer to that situation. (ii) However, if $(d_1-d_2) \leqq 2$, all the geodesics established in G_1 are also geodesics in H.

In the first case (i) no general expressions for the distance numbers of all vertices in H can be derived and an individual investigation of each class of graphs of this category is needed. In the second case (ii), however, general formulae for the distance and Wiener numbers can be derived. Thus, when two graphs are joined by two edges it has to be calculated first whether $d_2-d_1 > 2$ or $d_2-d_1 \leqq 2$.

We consider first the case (i) where eq. (26a) is satisfied. In this case one has to establish which vertex pairs of G_1 correspond to eq. (26b) and (26c), respectively. To do this one examines the vertices which belong to the spherical neighbourhoods of r and s by increasing the radii of the spheres stepwise by 1; in this manner all those vertices of G_1 are found some of whose distances are shorter in H than in G_1. For such a vertex, say t, one may conclude that $d(t|G_1)$ does not fully contribute to $d(t|H)$ as is the case for all the other vertices of G_1 and G_2.

Let us now turn to case (ii) where

$$(d_1 - d_2) \leqq 2 \tag{27}$$

holds. In this case the distances between any pair of vertices of one and the same component are not altered when the components are joined by two edges, i.e. $d(pq|G) = d(pq|H)$ for all $p,q \in G_1$ and $p,q \in G_2$. In this case one has only to determine whether the shortest path between the vertices $y \in G_1$ and $z \in G_2$ passes via the edge (ru) or (sv), Let us denote W^m_{ru} and W^m_{sv} the paths of relative minimal lengths passing from y to z via (ru) and (sv), respectively. Their lengths are $|W^m_{ru}| = [d(yr) + 1 + d(uz)]$ and $|W^m_{sv}| = [d(ys) + 1 + d(vz)]$, respectively. Hence, one obtains for their difference

$$|W^m_{ru}| - |W^m_{sv}| = [d(yr) - d(ys)] + [d(uz) - d(vz)] \quad . \tag{28}$$

When the vertex sets of G_1 and G_2 are partitioned according to Lemma 3 as follows

$$y(\delta_1) = \{y|d(yr) - d(ys) = \delta_1\}, \quad -d_1 \leqq \delta_1 \leqq +\delta_1$$

$$z(\delta_2) = \{z|d(zu) - d(zv) = \delta_2\}, \quad -d_2 \leqq \delta_2 \leqq +\delta_2$$

eq. (28) may be expressed as

$$|W^m_{ru}| - |W^m_{sv}| = \delta_1 + \delta_2 \quad . \tag{28a}$$

From this one can draw the following conclusions concerning the distance between the vertices $y\in G_1$ and $z\in G_2$ in H:

$$(\delta_1 + \delta_2) \leqq 0: \quad d(yz|H) = d(yr) + 1 + d(uz) \quad ; \tag{28b}$$

$$(\delta_1 + \delta_2) \geqq 0: \quad d(yz|H) = d(ys) + 1 + d(vz) \quad ; \tag{28c}$$

$$y\in y(\delta_1)\in G_1 \quad , \quad z\in z(\delta_2)\in G_2 \quad .$$

For $(\delta_1+\delta_2) = 0$ the eqs. (28b) and (28c) are equivalent.

We are now in the position to derive expressions for $d(y|H)$ and $d(z|H)$. According to eq. (3) we have:

$$d(y|H) = \sum_{\{x\}\in G_1} d(yx|H) + \sum_{\{z\}\in G_2} d(yz|H) \quad .$$

Under the condition of eq. (27), the first term equals $d(y|G)$. In view of eqs. (28b) and (28c) the summation over all $z\in G_2$ in the second term must be carried out in two partial sums where the first one runs over all $z\in z(\delta_2)$, $\delta_2 \leqq -\delta_1$, and the second one runs over all $z\in z(\delta_2)$, $\delta_2 > -\delta_1$. Thus one obtains for the vertices belonging to $y(\delta_1)$, i.e.: $y\in y(\delta_1)\in G_1$:

$$d(y|H) = d(y|G) + \sum_{\delta_2=-d_2}^{-\delta_1} \sum_{\{z(\delta_2)\}} [d(yr) + 1 + d(uz)] +$$

$$+ \sum_{\delta_2=-\delta_1+1}^{-d_2} \sum_{\{z(\delta_2)\}} [d(ys) + 1 + d(vz)] \quad .$$

By the substitution

$$d(yr) = d(ys) + \delta_1 \quad ,$$

$$d(uz) = d(vz) + \delta_2 \quad ;$$

the above equation is simplified:

$$y \in y(\delta_1) \in G_1 :$$

$$d(y|H) = d(y|G) + \sum_{\{z\}} [d(ys) + 1 + d(vz)] +$$

$$+ \sum_{\delta_2=-d_2}^{-\delta_1} \sum_{\{z(\delta_2)\}} (\delta_1 + \delta_2) \quad .$$

From this one finally arrives at:

$$y \in y(\delta_1) \in G_1 :$$

$$d(y|H) = d(y|G) + d(v|G) + n_2[d(ys|G) + 1] +$$

$$+ \sum_{\delta_2=-d_2}^{-\delta_1} |\{z(\delta_2)\}| \cdot (\delta_1 + \delta_2) \quad . \tag{29a}$$

Quite analogously one obtains

$$z \in z(\delta_2) \in G_2 :$$

$$d(z|H) = d(z|G) + d(s|G) + n_1[d(vz|G) + 1] +$$

$$+ \sum_{\delta_1=-d_1}^{-\delta_2} |\{y(\delta_1)\}| \cdot (\delta_1 + \delta_2) \quad . \tag{29b}$$

It should be noted that $(\delta_1+\delta_2) \leqq 0$ holds for both eqs. (29a) and (29b).

Making use of eq. (4) one can derive the Wiener number of H:

$$2W(H) = \sum_{\{y\}} d(y|H) + \sum_{\{z\}} d(z|H) \quad .$$

By inserting eqs. (29a) and (29b) into this formula one obtains

$$2W(H) = \sum_{\{y\}} \{d(y|G) + d(v|G) + n_2[d(ys|G) + 1]\} +$$

$$+ \sum_{\{z\}} \{d(z|G) + d(s|G) + n_1[d(vz|G) + 1]\} +$$

$$+ \sum_{\{y\}} \sum_{-d_2}^{-\delta_1} |\{z(\delta_2)\}| \cdot (\delta_1 + \delta_2) +$$

$$+ \sum_{\{z\}} \sum_{-d_1}^{-\delta_2} |\{y(\delta_1)\}| \cdot (\delta_1 + \delta_2) \quad .$$

The last two terms will be denoted by 2Q for short. When the summation is performed one arrives at:

$$2W(H) = 2[W(G_1) + W(G_2) + n_2 d(s|G) + n_1 d(v|G) + n_1 n_2] + 2Q \ , \qquad (30a)$$

wherein

$$2Q = \sum_{\delta_1=-d_1}^{+d_1} \sum_{\delta_2=-d_2}^{-\delta_1} |\{y(\delta_1)\}| \cdot |\{z(\delta_2)\}| \cdot (\delta_1 + \delta_2) +$$

$$+ \sum_{\delta_2=-d_2}^{+d_2} \sum_{\delta_1=-d_1}^{-\delta_2} |\{y(\delta_1)\}| \cdot |\{z(\delta_2)\}| \cdot (\delta_1 + \delta_2) \quad . \qquad (30b)$$

On inspecting these expressions one notes that:

1. The summation in the two terms runs only over pairs of δ_1 and δ_2 values that obey $(\delta_1+\delta_2) \leqq 0$;

2. The two term entries are the same function of the finite discrete variable parameter δ_1 and δ_2, namely

$$A(\delta_1,\delta_2) = |\{y(\delta_1)\}| \cdot |\{z(\delta_2)\}| \cdot (\delta_1 + \delta_2) \quad ;$$

3. $A(\delta_1,\delta_2) = 0$ for $(\delta_1 + \delta_2) = 0$.

4. The two terms differ only formally by the instruction as to how

the summation has to be carried out. In both terms of eq. (30b), however, exactly the same $A(\delta_1,\delta_2)$'s are summed up, hence, the two terms are equal.

The last result is not as obvious as the first three ones but it is easily proved. Hence, we can write down for Q:

$$Q = \sum_{\delta_1=-d_1}^{+d_2} \sum_{\delta_2=-d_2}^{-\delta_1} |\{y(\delta_1)\}| \cdot |z(\delta_2)\}| \cdot (\delta_1 + \delta_2) \quad . \tag{30c}$$

The explicit formula for W(H) is obtained from eqs. (30a) and (30c). It is not symmetrical with respect to vertices $r,s \in G_1$ and $u,v \in G_2$. If a symmetrization is needed, it could be achieved by substituting eq. (9), thus obtaining:

$$d(s|G) = [d(r|G) + d(s|G) - \sum_{\delta_1=-d_1}^{+d_1} |\{y(\delta_1)\}| \cdot \delta_1]/2 \quad ,$$

$$d(v|G) = [d(u|G) + d(v|G) - \sum_{\delta_2=-d_2}^{+d_2} |\{z(\delta_2)\}| \cdot \delta_2]/2 \quad .$$

We summarize the results obtained above as follows:

<u>Property 4.2:</u> Let $G = G_1 \cup G_2$, $r,s \in G_1$, $u,v \in G_2$, $d(rs) = d_1$, $d(uv) = d_2$, $0 \leqq (d_1-d_2) \leqq 2$; let also the vertex sets $\{y\} \in G_1$, $|\{y\}| = n_1$, and $\{z\} \in G_2$, $|\{z\}\{ = n_2$, be completely decomposed according to Lemma 3 into the subsets $\{y(\delta_1)\}$, $-d_1 \leqq \delta_1 \leqq +d_1$, and $\{z(\delta_2)\}$, $-d_2 \leqq \delta_2 \leqq +d_2$, respectively. The connected graph H is formed as $H = G \cup \{(ru),(sv)\}$. The following equations hold for the distance numbers of vertices and the Wiener number of H:

$y \in y(\delta_1) \in G_1$:

$$d(y|H) = d(y|G) + d(v|G) + n_2[d(ys\ G) + 1] + \tag{31a}$$

$$+ \sum_{-d_2 \leq \delta_2 \leq -\delta_1} |\{z(\delta_2)\}| \cdot (\delta_1+\delta_2) \quad ,$$

$z \in z(\delta_2) \in G_2$:

$$d(z|H) = d(z|G) + d(s|G) + n_1[d(vz|G) + 1] + \sum_{-d_1 \leq \delta_1 \leq -\delta_2} |\{y(\delta_1)\}| \cdot (\delta_1+\delta_2) \quad ; \tag{31b}$$

$$W(H) = W(G_1) + W(G_2) + n_2 d(s|G) + n_1 d(v|G) + n_1 n_2 + \sum_{\delta_1=-d_1}^{+d_2} \sum_{\delta_2=-d_2}^{-\delta_1} |\{y(\delta_1)\}| \cdot |\{z(\delta_2)\}| \cdot (\delta_1+\delta_2) \quad . \tag{32}$$

<u>Proof:</u> Not necessary, because eqs. (31) and (32) have been derived above. □

The problem of connecting the two components of a graph by means of two edges, handled here, is close to another problem: the addition of an edge (sv) to a connected graph $\tilde{G}$ which contains one bridge (ru). Removing this bridge from $\tilde{G}$ one obtains a two-component graph, $\tilde{G}\setminus(ru) = G = G_1 \cup G_2$. This union, $G_1 \cup G_2$, has been considered as the starting point for both procedures described in this subsection. Hence, the problem expressed by $H = \tilde{G} \cup \{(sv)\}$ can be treated using eqs. (20) - (22) and (31) - (32) provided that $|d(rs|\tilde{G}) - d(uv|\tilde{G})| \leqq 2$.

An interesting application of eqs. (31) and (32) is found within series of graphs which are successively formed in a regular manner according to $H_{2h+1} = H_h \cup G \cup \{e_1, e_2\}$ where e_1 and e_2 denote the added edges. Such series of chemical interest are easily constructed from the C-graphs of polycyclic aromatic hydrocarbons (PAH); examples for such series are benzene, naphthalene, anthracene, tetracene, etc., or benzene, naphthalene, phenanthrene, chrysene, picene, etc. and others. All these series start with the C-graph of benzene, i.e. $H_1 = P_2 \cup P_4 \cup \{e_1, e_2\} = C_6$ and continue with $H_{h+1} = H_h \cup P_4 \cup \{e_1, e_2\}$, where h indicates the number of six-membered cycles of H_h. Since for various series of PAH's general formulae are already known [7,9,11,12,14,43] which express W(h) in terms of powers of h, it is true that by the use of eqs. (31) and (32) no new results but insight into the modelling of these expressions are gained.

In the study of topological effects on molecular orbitals (TEMO) [31-34] eqs. (31) and (32) are of particular interest. Assume G_1 and G_2 are isomorphic, the vertices r and s are non-equivalent and can be mapped isomorphically onto u and v, respectively. Then two topologically different graphs $H_1 = G_1 \cup G_2 \cup \{(ru), (sv)\}$ and $H_2 = G_1 \cup G_2 \cup \{(rv), (su)\}$ can be formed. Their eigenvalue spectra are related particularly to the so-called TEMO-pattern [31,35-37]. From eq. (32) one concludes that $W(H_1) > W(H_2)$. In the case that H_1 and H_2 represent the skeleton graphs of fully conjugated systems then the total π-electron energies satisfy $E_\pi(H_1) \geqq E_\pi(H_2)$ [38]; this throws some light on the good correlation of E_π with Wiener numbers [40].

The procedure described here may be applied also to the Fasciagraphs of polymers [42] in which the monomeric units are linked by two bonds. The treatment is similar to that given in subsection 4.1.

4.3. Cyclisation of a Tree

The addition of an edge to a connected graph makes the number of independent cycles increase by 1. By means of the added edge new paths may be constructed for various pairs of vertices; some of them are shorter than the original ones. As a consequence of this a reorganization of the geodesic paths takes place and, hence, the estimation of the changes of distance and Wiener numbers usually is a very tricky problem which cannot be treated in full generality. The most simple and clear case is presented by the addition of an edge to a tree; in the following we shall present that case.

Let G be a tree, $u,v \in G$ two non-adjacent vertices and let d denote their distance $d = d(uv|G) > 1$. Let H be that graph which is formed from G by adding the edge (uv), i.e.: $H = G \cup \{(uv)\}$.

First we consider the reorganization of the geodesics due to the formation of H. For that purpose we apply Lemma 3 to the vertex set V(G) of G with respect to the distances of the vertices to u and v, respectively, i.e. V(G) is partitioned into the subsets $V(\delta)$, $-d \leqq \delta \leqq d$.

<u>Property 4.3 (1):</u> Let $x \in V(\delta_x)$ and $y \in V(\delta_y)$ be arbitrary vertices of G. When the graph H is formed from G according to $H = G \cup \{(uv)\}$ then the distances $d(xy|G)$ and $d(xy|H)$ are related as follows:

a) if $|\delta_y - \delta_x| \leqq (d+1)$:

$$d(xy|H) = d(xy|G) \quad ; \tag{33a}$$

b) if $|\delta_y - \delta_x| > (d+1)$:

$$d(xy|H) = d(xy|G) - [|\delta_y - \delta_x| - (d+1)] \quad . \tag{33b}$$

<u>Proof:</u> a) Observe first, in G there is only a single path $W_1(xy)$ for any pair of vertices. Provided $\delta_x \neq \delta_y$, in H a second path, $W_2(xy)$, is established which passes through the added edge, $(uv) \in W_2(xy)$. b) $W_1(xy)$ in G is the geodesic of the vertices x and

y, in H, however, $W_1(xy)$ and $W_2(xy)$ are in competition for that. The lengths of these paths are as follows:

$$|W_1(xy)| = d(xx') + d(x'y') + d(y'y) \quad ,$$

$$|W_2(xy)| = d(xx') + [(d+1) - d(x'y')] + d(y'y) \quad ,$$

where $x' \in V(\delta_x)$ and $y' \in V(\delta_y)$ denote those vertices of the respective subsets which also belong to $W_1(uv)$. c) Thus, the length of $W_2(xy)$ may be expressed as follows:

$$|W_2(xy)| = |W_1(xy)| + [(d+1) - 2d(x'y')] \quad . \qquad (33c)$$

Obviously, $W_2(xy)$ will be the geodesic if and only if $2d(x'y') > (d+1)$. d) Assume momentary $\delta_y > \delta_x$, then $d(x'y')$ may be expressed by one of the following equations:

$$d(x'y') = d(uy') - d(ux') \quad ,$$

$$d(x'y') = -d(vy') + d(vx') \quad .$$

This immediately leads to

$$2d(x'y') = [d(uy') - d(vy')] - [d(ux') - d(vx')] = \delta_y - \delta_x .$$

If $\delta_y < \delta_x$ is assumed one arrives at $2d(x'y') = \delta_x - \delta_y$. Both results are generalized to

$$2d(x'y') = |\delta_y - \delta_x| \quad .$$

e) The substitution of this in eq. (33c) leads to eqs. (33a) and (33b). f) The proof is completed by observing that in case of $\delta_x = \delta_y$ all vertices of the path $W_1(xy)$ belong to $V(\delta_x)$ and $W_1(xy)$ is also in H the single path connecting x and y. ⌑

Corollary to Property 4.3 (1): Let x belong to one of the subsets V(-1), V(0), or V(1), then the distances from x to any other vertex $y \in G$ are not altered when the edge (uv) is added to G; thus

$$d(xy|H) = d(xy|G) \tag{33d}$$

holds for all vertices y, provided $-1 \leqq \delta_x \leqq 1$.

Proof: a) Note the range of δ_y, i.e. $-d \leqq \delta_y \leqq d$. b) If $-1 \leqq \delta_x \leqq 1$ is assumed, then $|\delta_y - \delta_x| \leqq d+1$. Thus, eq. (33b) cannot be applied to any vertex of V(-1), V(0), and V(1), respectively. □

Note that G is a tree, hence, depending on the parity of d only either V(0) or V(-1) and V(1) are non-empty.

Now, we are sufficiently prepared for a consideration of the changes of distance and Wiener numbers due to $H = G \cup \{(uv)\}$.

Property 4.3 (2): Let H be formed according to $H = G \cup \{(uv)\}$ as described above, then for H the following distance and Wiener numbers result:

a) $-1 \leqq \delta_x \leqq 1$:

$$d(x|H) = d(x|G) \quad ; \tag{34a}$$

b) $-d \leqq \delta_x \leqq -2$:

$$d(x|H) = d(x|G) - \sum_{\delta_y=\delta_x+d+2}^{d} (\delta_y-\delta_x-d-1)\cdot|V(\delta_y)| \quad ; \tag{34b}$$

c) $2 \leqq \delta_x \leqq d$:

$$d(x|H) - d(x|G) - \sum_{\delta_y=-d}^{\delta_x-d-2} (\delta_x-\delta_y-d-1)\cdot|V(\delta_y)| \quad ; \tag{34c}$$

$$W(H) = W(G) - \frac{1}{2}\{\sum_{\delta_x=-d}^{-2} \sum_{\delta_y=\delta_x+d+2}^{d} (\delta_y-\delta_x-d-1)\cdot|V(\delta_x)|\cdot|V(\delta_y)| + \sum_{\delta_x=2}^{d} \sum_{\delta_y=-d}^{\delta_x-d-2} (\delta_x-\delta_y-d-1)\cdot|V(\delta_x)|\cdot|V(\delta_y)|\} \quad . \qquad (35)$$

Proof: a) Eq. (34a) follows directly from eq. (33d). b) From eq. (3) one obtains primarily $d(x|H) = \sum_{\delta_y=-d}^{d} \sum_{V(\delta_y)} d(xy|H)$.

If $-d \leqq \delta_x \leqq -2$ is assumed, $d(xy|H)$ is given by eq. (33a) for $-d \leqq \delta_y \leqq \delta_x+d+1$, but by eq. (33b) for $\delta_x+d+2 \leqq \delta_y \leqq d$. Thus, the summing up over $-d \leqq \delta_y \leqq d$ must be separately carried out for both these ranges, resulting in eq. (34b). c) Eq. (34c) is similarly proved. d) Eq. (35) follows from eqs. (4) and (34a) - (34c). Here again the summing up over $-d \leqq \delta_x \leqq d$ must be carried out in the intervals of δ_x, indicated in eqs. (34a) - (34c). ◻

Note, the two double sums in eq. (35) are equal; this is easily proved by comparison of the terms obtained for $\delta_x = a$, $\delta_y = b$ and $\delta_x = b$, $\delta_y = a$, respectively, $a \leqq b$.

From eq. (35) one concludes that the Wiener number of a tree is always reduced when an edge is added. Even for smallest acceptable value $d = 2$ one obtains from eq. (35) as follows:

$$W(H) = W(G) - |V(-2)|\cdot|V(2)| \quad .$$

4.4. Remark

In an earlier paper [14] analytical equations have been derived for the change in the Wiener number resulting from certain structural transformations. These include some particular cases of connecting two cycles by one or two edges, and is a subject of consideration in this section. The total number of vertices, the number, length and position of the bridges, etc., have been used as variable parameters in these equations [14]. The advantage of the procedure developed here lies in the provision of a treatment of such problems in a more general manner.

5. Subdivision of an Edge by a Vertex

Dividing an edge (uv)∈G by a vertex x increases the distance d(uv) from 1 to 2. All paths containing this edge will thus be lengthened by 1. When the edge (uv) is a bridge between two subgraphs, G_1 and G_2, then all paths from a vertex in G_1 to a vertex in G_2 will be lengthened analogously. In contrast to this, when (uv) is not a bridge, then the distances between some pairs of vertices increase due to the subdivision of (uv); these pairs are specified by the fact that all the shortest paths between them cannot avoid (uv). Such pairs of vertices can be detected by inspection of G. Let s be a vertex of such a pair and b(s) the number of geodesics which are enlarged due to the subdivision of the edge.

When H is a graph formed from G by dividing the edge (uv) by vertex x, the distance number of s in H is

$$d(s|H) = d(s|G) + b(s) + d(sx) \quad . \tag{36}$$

Vertex x could, however, be reached either via u or via v. Hence,

$$d(sx) = 1 + \min\{d(su|G), d(sv|G)\} \quad .$$

Now let the vertex set of G be decomposed according to eq. (6) into subsets {p}, {q}, and {r} which are distinct as to the differences of the distances to u and v. One thus obtains for the distance number of x

$$d(x|H) = d(u|G) + |p| + |q| = \\ = d(v|G) + |q| + |r| = \\ = [d(u|G) + d(v|G) + n + |q|]/2 \quad , \tag{37}$$

where $|p|$, $|q|$, and $|r|$ are the cardinalities of the respective vertex sets, and n is the number of vertices in G. The Wiener number is then presented by the equation

$$W(H) = W(G) + d(x|H) + [\sum_{s\in G} b(s)]/2 \quad . \tag{38}$$

In the case of (uv) being a bridge connecting the subgraphs G_1 and G_2, $u\in G_1$ and $v\in G_2$, let the number of vertices in the subgraphs be denoted by n_1 and n_2, respectively. Then, for each vertex $p\in G_1$ the quantity $b(p) = |r| = n_2$, as well as for each vertex $r\in G_2$, $b(r) = |p| = n_1$. Proceeding from eqs. (36) - (38) one obtains in the case of the subdivision of a bridge edge:

$$d(p|H) = d(p|G) + d(pu) + 1 + n_2 \quad ; \\ d(r|H) = d(r|G) + d(rv) + 1 + n_1 \quad ; \\ d(x|H) = d(u|G) + n_1 = d(v|G) + n_2 \quad ; \\ W(H) = W(G) + n_1 n_2 + [d(u|G) + d(v|G) + n_1 + n_2]/2 \quad . \tag{39}$$

Eqs. (36) - (38) can also be used to determine the distance numbers and the Wiener number of G from those of H when dealing with the inverse graph transformation, the collapse of two edges into a single one.

6. Conclusions

The theory of the Wiener number on the basis of the distance numbers is presented. Expressions for the changes of distance and Wiener numbers due to particular elementary graph operations are given. A variety of more complicated graph operations can be constructed by means of their combinations. As shown in subsection 4.3 the graph operation $H_{h+1} = H_h \cup G \cup \{e_1, e_2\}$ generates series of G-homologue graphs for which analytical expressions of the type $W(H_h) = f(h)$ are known [7,9,11,12,14,43]; thus, this earlier work can also be combined with the operations reported in Sections 3-5.

The distance numbers facilitate in a very convenient manner the calculation of the Wiener numbers (see Sect. 3-5). However, the real value and importance of the distance numbers lie in enabling one to express in a characteristic manner the changes in the Wiener numbers, resulting from some graph operations, as exemplified in Sections 3-5. The derivation of such expressions is very simple and straightforward in the case of some graph operations (see for example Sections 3 and 5) but rather complicated in other ones (e.g. Section 4). Difficulties always arise when the reorganisation of the geodesics due to the graph operation cannot be described with sufficient generality. Our experience with this subject matter may be summarized as follows: When the graph operation performed keeps the cyclomatic number (i.e. the number of independent cycles of the graph) constant, no serious difficulties occur. In contrast, the problem becomes very complex if a change of the cyclomatic number is involved. A verification of this behaviour is found in the subsections 4.1 and 4.2.

This observation is of some interest because it seems to reflect the effect of drastic changes in the topological structure associated with the graph operations under consideration. To make this point clear we will briefly refer to the changes in the topological structure upon the addition of an edge if a) this edge connects two components of a disconnected graph or b) this edge is added to a connected graph and closes an additional independent cycle. Let the topological structure of a graph be defined by the complete sets of spherical neighbourhoods of all graph vertices [44]. In the case of connecting two graphs by an edge the number of spherical neighbourhoods of each vertex increases and the same holds for the cardinalities of the vertex subsets which correspond to some of the spherical neighbourhoods. However, no vertex is transferred from one spherical neighbourhood to another. In contrast to that, the formation of an additional independent cycle by adding an edge to a connected graph is always accompanied by vertex transitions from outer spherical neighbourhoods into inner ones. Thus, an essential reorganization of the topological structure of the graph always occurs. The approaches to analysing such structural changes are briefly discussed in subsection 4.2.

Acknowledgements. Danail Bonchev is grateful to Deutsche Forschungsgemeinschaft and to Max-Planck-Gesellschaft for financial support during his stay in Mülheim in 1983, 1985, and 1986, respectively. The valuable assistance of Mrs. E. Currell and Mrs. I. Heuer in the preparation of the paper is appreciated.

APPENDIX

Distance and Wiener Numbers of some Graphs with n Vertices

1. Complete graph K_n:

$d(u|K_n) = n-1$, $W(K_n) = n(n-1)/2$.

2. Cycles C_n:

$d(u|C_{2m}) = n^2/4$, $W(C_{2m}) = n^3/8$;

$d(u|C_{2m+1}) = (n^2-1)/4$, $W(C_{2m+1}) = n(n^2-1)/8$.

3. Star graph $K_{1,n-1}$:

The center of $K_{1,n-1}$ is denoted by u.

$d(u|K_{1,n-1}) = n-1$,

$d(t|K_{1,n-1}) = 2n-3$, $W(K_{1,n-1}) = (n-1)^2$.

4. Path graph P_n:

From one end point to the other the vertices of P_n are labeled subsequently with the numbers 1,2,...,j,...n.

$d(j|P_n) = n(n+1)/2 - j(n+1-j)$,

$W(P_n) = n(n^2-1)/6$.

The central vertices of P_n have the following distance numbers:

$d(m|P_{2m}) = d(m+1|P_{2m}) = n^2/4$,

$d(m+1|P_{2m+1}) = (n^2-1)/4$.

References

[1] H. Wiener, J. Amer. Chem. Soc. 69, 17 (1947).

[2] H. Wiener, J. Amer. Chem. Soc. 69, 2636 (1947).

[3] H. Wiener, J. Phys. Chem. 52, 425 (1948).

[4] H. Hosoya, Bull. Chem. Soc. Jpn. 44, 2332 (1971).

[5] A.T. Balaban, I. Motoc, D. Bonchev, Ov. Mekenyan, Topics Curr. Chem. 114, 21 (1983).

[6] D. Bonchev, D. Sc. Thesis, Moscow 1984.

[7] D. Bonchev, N. Trinajstić, J. Chem. Phys. 67, 4517 (1977).

[8] D. Bonchev, N. Trinajstić, Intern. J. Quantum Chem.: Quantum Chem. Symp. 12, 293 (1978).

[9] D. Bonchev, Ov. Mekenyan, J.V. Knop, N. Trinajstić, Croat. Chem. Acta 52, 361 (1979).

[10] Ov. Mekenyan, D. Bonchev, N. Trinajstić, Math. Chem. (MATCH) 6, 93 (1979).

[11] D. Bonchev, Ov. Mekenyan, N. Trinajstić, Intern. J. Quantum Chem. 17, 845 (1980).

[12] Ov. Mekenyan, D. Bonchev, N. Trinajstić, Math. Chem. (MATCH) 11, 145 (1981).

[13] Ov. Mekenyan, D. Bonchev, N. Trinajstić, Intern. J. Quantum Chem. 18, 369 (1980).

[14] Ov. Mekenyan, D. Bonchev, N. Trinajstić, Intern. J. Quantum Chem. 19, 929 (1981).

[15] Ov. Mekenyan, D. Bonchev, N. Trinajstić, Croat. Chem. Acta 56, 237 (1983).

[16] O.E. Polansky, D. Bonchev, in preparation.

[17] J.G. Chartrand, M.J. Stewart, Lecture Notes Math. 186, 63 (Ed.: A. Dold, B. Eckmann), Springer-Verlag, Berlin, Heidelberg, New York 1971.

[18] R.C. Entringer, D.E. Jackson, D.A. Snyder, Czech. Math. J. 26 (101), 283 (1976).

[19] V.A. Skorobogatov, P.V. Khvorostov, Vychislitel'nye systemy 91, 3 (1981) (in Russian).

[20] J.R. Platt, J. Chem. Phys. 15, 419 (1947).

[21] J.R. Platt, J. Phys. Chem. 56, 328 (1952).

[22] A.T. Balaban, Chem. Phys. Lett. 89, 399 (1982).

[23] A.T. Balaban, THEOCHEM 21, 117 (1985).

[24] E.R. Canfield, R.W. Robinson, D.H. Rouvray, J. Comput. Chem. 6, 598 (1985).

[25] D. Bonchev, A.T. Balaban, Ov. Mekenyan, J. Chem. Inf. Comput. Sci. 20, 106 (1980).

[26] D. Bonchev, Information Theoretic Indices for Characterization of Chemical Structures, Research Studies Press, Wiley, Chichester, 1983.

[27] I. Gutman, O.E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin-Heidelberg 1986, pp. 124-127.

[28] F. Harary, Graph Theory, Addison-Wesley Publ. Comp., Reading (Mass.) 1972 (First printing: 1969; German edition: R. Oldenbourg, München, Wien 1974).

[29] O. Ore, Theory of Graphs, Amer. Math. Soc., Providence (Rhode Island) 1962.

[30] O.E. Polansky, D. Bonchev, Math. Chem. (MATCH), this issue.

[31] O.E. Polansky, M. Zander, J. Mol. Struct. 84, 361 (1982).

[32] O.E. Polansky, M. Zander, I. Motoc, Z. Naturforsch. 38a, 196 (1983).

[33] O.E. Polansky, J. Mol. Struct. 113, 281 (1984).

[34] M. Zander, O.E. Polansky, Naturwiss. 71, 623 (1984)

[35] I. Gutman, A. Graovac, O.E. Polansky, Chem. Phys. Lett. 116, 206 (1985).

[36] A. Graovac, I. Gutman, O.E. Polansky, J. Chem. Soc., Faraday Trans. II, 81, 1543 (1985).

[37] I. Gutman, A. Graovac, O.E. Polansky, Disc. Appl. Math., in press.

[38] A. Graovac, I. Gutman, O.E. Polansky, Mh. Chemie 115, 1 (1984).

[39] Zhang Fuji, Chen Zibo, Math. Chem. (MATCH) 21, this issue.

[40] D. Bonchev, Ov. Mekenyan, Z. Naturforsch. 35a, 739 (1980).

[41] The structural graph of the first Möbius compound recently synthesized by D.M. Walba, R.M. Richards, R.C. Haltiwanger, J. Amer. Chem. Soc. 104, 3219 (1982), is not planar.

[42] O.E. Polansky, N.N. Tyutyulkov, Math. Chem. (MATCH) 3, 149 (1977).

[43] I. Gutman, O.E. Polansky, Math. Chem. (MATCH) 20, 115 (1986).

[44] O.E. Polansky, Z. Naturforsch. 41a, 560 (1985).