MATHEMATICAL MODEL OF REORGANIZATION OF VALENCE ELECTRONS INVOLVING TWO ATOMS*

Jaroslav KOČA[a], Milan KRATOCHVÍL[b], Luděk MATYSKA[b] and Vladimír KVASNIČKA[c]

[a] *Department of Organic Chemistry,*
Faculty of Sciences, Purkyně University, 611 37 Brno,
[b] *Research Institute of Pure Chemicals, Lachema, 621 33 Brno and*
[c] *Department of Mathematics, Faculty of Chemistry,*
Slovak Institute of Technology, 812 37 Bratislava

Dedicated to Prof. J. Kováč on the occasion of his 60th birthday.

The algebraic description of states of atomic vectors and their interconversions in a form of 9--dimensional vectors is suggested. The elaborated algorithms LENGTH and PATH calculate the distance between two states of an atomic vector and construct all possible shortest paths between them, respectively. Illustrative applications demonstrate the chemical impact and interpretation of the theory.

This communication is a continuation of our recent work[1] devoted to the mathematical model of valence states of atoms and their interconversions. We have suggested simple mathematical formalism for the description of valence states of atoms. Following its simple geometric interpretation we are able to construct combinatorially all possible valence states restricted by a given number of valence electrons. The purpose of this article is to enlarge the above mentioned theory for pairs of atoms (called the atomic vectors). In particular, we take into account explicitly the assumption that all elementary chemical processes are running over pairs of atoms. In the mathematical model of Dugundji and Ugi[2] this very serious fact was partially covered by the notion atomic vector, which roughly speaking, corresponds to a pair of atoms that are connected by the bond in a studied molecule. We generalize this notion in such a way that under atomic vector we shall understand an arbitrary pair of atoms belonging to a given ensemble of molecules (EM), it will be denoted by (X, Y), where X and Y are atomic symbols. If the actual valence states are presented in an atomic vector, it will be called the state of atomic vector, *e.g.* —C≡N̄, ⟩C=N̄—, *etc.* We shall elaborate a proper formalism for description and conversions of states

---

* Part IX in the series Mathematical Model of Organic Chemistry; Part VIII: Čas. pěst. matem., in press.

of atomic vectors in two hierarchical levels depending on whether the suggested algebra is realized in E^3 or E^9 Euclidean spaces.

ALGEBRA OF STATES OF ATOMIC VECTORS IN E^3

This algebraic aproach is devoted to the description of states of atomic vectors, where the lone electrons and bonds between atoms from a given atomic vector are explicitly treated; the bonds between other atoms are now fully ignored.

Formal Description of States of Atomic Vectors and Their Conversions

We start our consideration by an abstract assumption that an atomic vector is an indiviual molecule and its state is described by the BE matrix[2] of the second order. Similarly, the conversion of an educt atomic vector onto product atomic vector is described by the reaction (R) matrix[2] of the second order. Since both these matrices are symmetric they may be reformulated in the form of 3-dimensional vectors from the E^3 Euclidean space. Formally,

$$\mathbf{v}_E + \mathbf{v}_R = \mathbf{v}_P, \qquad (1)$$

where $\mathbf{v}_{E(P)}$ is a 3-dimensional vector corresponding to the educt (product) and $\mathbf{v}_R$ is a 3-dimensional vector corresponding to a given R matrix.

Example 1: A nucleophile addition to carbonyl group can be initiated by the following step

$$\text{>C=}\overline{\text{O}} \quad \rightarrow \quad \text{>C}^{(+)}\text{---}\overline{\underline{\text{O}}}|^{(-)}.$$
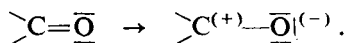
In the matrix formalism this process is described as follows
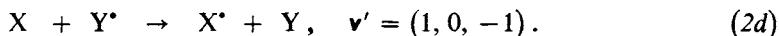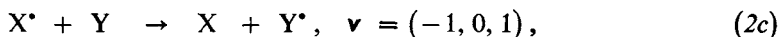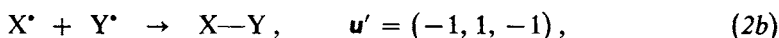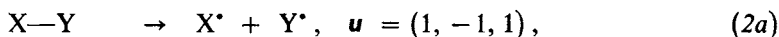
$$\begin{matrix} C \\ O \end{matrix} \begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 6 \end{pmatrix}.$$

Transforming this matrix equation in the vector form we arrive at

$$\underset{\mathbf{v}_E}{(0, 2, 4)} + \underset{\mathbf{v}_R}{(0, -1, 2)} = \underset{\mathbf{v}_P}{(0, 1, 6)},$$

where we have used only the matrix elements from upper-triangular parts of the corresponding matrices.

All the processes which are realizable simultaneously for both atoms of a given atomic vector we can decompose in the form of four elementary steps[3]:

$$X\!-\!Y \quad \rightarrow \quad X^\bullet + Y^\bullet, \quad \boldsymbol{u} = (1, -1, 1),\tag{2a}$$

$$X^\bullet + Y^\bullet \quad \rightarrow \quad X\!-\!Y, \quad \boldsymbol{u}' = (-1, 1, -1),\tag{2b}$$

$$X^\bullet + Y \quad \rightarrow \quad X + Y^\bullet, \quad \boldsymbol{v} = (-1, 0, 1),\tag{2c}$$

$$X + Y^\bullet \quad \rightarrow \quad X^\bullet + Y, \quad \boldsymbol{v}' = (1, 0, -1).\tag{2d}$$

In the right column the corresponding **R** vectors are presented; we emphasize that the processes $(2a)$, $(2b)$ and $(2c)$, $(2d)$ are mutually inverse, algebraically, $\boldsymbol{u}' = -\boldsymbol{u}$ and $\boldsymbol{v}' = -\boldsymbol{v}$.

It means thall all processes running simultaneously over both atoms of a given atomic vector may be formally expressed as the linear combination of the above introduced 3-dimensional **R** vectors $\boldsymbol{u}$ and $\boldsymbol{v}$,

$$\boldsymbol{v_R} = c_1\boldsymbol{u} + c_2\boldsymbol{v}.\tag{3}$$

Example 2: Let us consider the process

$$X\!\equiv\!Y \quad \rightarrow \quad \overline{X}\!-\!Y.$$
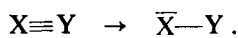
In the vector formalism the educt and product atomic vectors are represented by

$$(0, 3, 0) \rightarrow (4, 1, 0).$$

Then, following the Eq. (1), the assigned **R** vector is

$$\boldsymbol{v_R} = (4, 1, 0) - (0, 3, 0) = (4, -2, 0).$$

This vector is expressible by the linear combination of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ (see Eq. (3)),

$$\boldsymbol{v_R} = (4, -2, 0) = c_1(1, -1, 1) + c_2(-1, 0, 1),$$

solving the system we get $c_1 = 2$ and $c_2 = -2$.
Hence, the studied general process corresponds to two conversions $(2a)$ and two conversions $(2d)$.

Geometric Representation of Conversions of States of Atomic Vectors

In the preceding section every process running simultaneously over both atoms of an atomic vector can be expressed by the linear combination of two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. All isoelectronic atomic vectors (*e.g.* $X\!\equiv\!Y$, $\overline{X}\!=\!Y$, $\overline{X}\!-\!Y$, *etc.*) are geometrically visualized at the same plane of 3-dimensional Euclidean space E^3. The vectors

$u = (1, -1, 1)$ and $v = (-1, 0, 1)$ are orthogonal, $uv = 0$ (the scalar product is defined in the standard manner), we enlarge this system of two orthogonal vectors by a third vector w which is orthogonal to previous ones (*e.g.* by Schmidt's orthogonalization method), we get

$$u = (1, -1, 1),$$
$$v = (-1, 0, 1), \qquad (4)$$
$$w = (1/3, 2/3, 1/3).$$

These three orthogonal vectors will serve as a new coordinate system at the 3-dimensional Euclidean space E^3. An arbitrary vector $a = (a_1, a_2, a_3) \in E^3$ is expressed by

$$a = x_1 u + x_2 v + x_3 w, \qquad (5)$$

where the coeficients x_1, x_2, x_3 are called the coordinates of a in the system of orthogonal vectors in Eq. (4). Introducing Eq. (4) into Eq. (5) we get

$$x_1 = (a_1 - a_2 + a_3)/3, \qquad (6a)$$

$$x_2 = (a_3 - a_1)/2, \qquad (6b)$$

$$x_3 = (a_1 + 2a_2 + a_3)/2. \qquad (6c)$$

If a state of atomic vector is determined by the vector $a = (a_1, a_2, a_3)$, then its number of valence electrons is equal to $a_1 + 2a_2 + a_3 = 2x_3$ (*cf. Example 1*). It means that all isoelectronic atomic vectors have the same coordinate x_3, they are situated at the plane which is parallel with a plane containing the axes u and v and distance between them is x_3 (*cf.* Fig. 1). In this new coordinate system the elementary processes described by Eqs (2) have the coordinates x_1, x_2, x_3 determined as follows:

Elementary process	Coordinates x_1, x_2, x_3	
(2a)	1, 0, 0	(7a)
(2b)	−1, 0, 0	(7b)
(2c)	0, 1, 0	(7c)
(2d)	0, −1, 0	(7d)

We have considered only those processes that are running over both atoms from a given atomic vector, for our forthcoming considerations it will be worthwhile to study also processes which involve only one atom from the atomic vector.

Example 3: Next step of the mechanism studied in example 1 may look as follows

$$\text{>}C^{(+)}\!\!-\!\overline{O}|^{(-)} \quad \rightarrow \quad \text{>}C^{(+)}\!\!-\!\overline{O}\!\!-\!H \, .$$
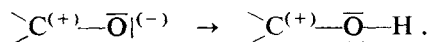
From the standpoint of atomic vector (C, O) this process is
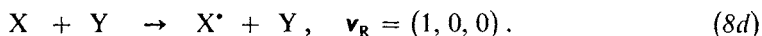
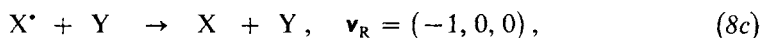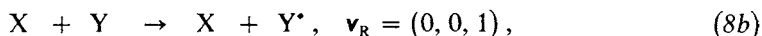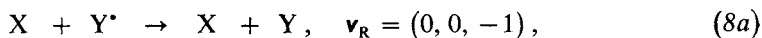$$C\!\!-\!\overline{O}| \;\rightarrow\; C\!\!-\!\overline{O} \, .$$
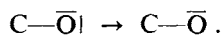
In the vector formalism we get

$$(0, 1, 6) + (0, 0, -2) = (0, 1, 4) \, ,$$

where the **R** vector is $\mathbf{v_R} = (0, 0, -2)$.

The processes that are similar to the above example are expressible *via* the following four elementary steps:

$$X \; + \; Y^{\bullet} \;\rightarrow\; X \; + \; Y \, , \quad \mathbf{v_R} = (0, 0, -1) \, , \tag{8a}$$

$$X \; + \; Y \;\rightarrow\; X \; + \; Y^{\bullet} \, , \quad \mathbf{v_R} = (0, 0, 1) \, , \tag{8b}$$

$$X^{\bullet} \; + \; Y \;\rightarrow\; X \; + \; Y \, , \quad \mathbf{v_R} = (-1, 0, 0) \, , \tag{8c}$$

$$X \; + \; Y \;\rightarrow\; X^{\bullet} \; + \; Y \, , \quad \mathbf{v_R} = (1, 0, 0) \, . \tag{8d}$$

We note that the processes described by Eqs $(8a-b)$ and, similarly, the processes $(8c-d)$ are mutually inverse. It means that **R** vector assigned to an arbitrary process running over only one atom of a given atomic vector may be determined by the linear combination of **R** vectors corresponding to elementary processes $(8a)$ and $(8c)$. The coordinates x_1, x_2, x_3 of Eqs (8) in the orthogonal system $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are:

Elementary process	Coordinates x_1, x_2, x_3
$(8a)$	$-1/3, \; -1/2, \; -1/2$
$(8b)$	$1/3, \; 1/2, \; 1/2$
$(8c)$	$-1/3, \; 1/2, \; -1/2$
$(8d)$	$1/3, \; -1/2, \; 1/2$.

The obtained results allow us to visualize geometrically all possible atomic vectors with arbitrary number of valence electrons and multiplicity of bond between atoms from atomic vectors, and also, their mutual interconversions, see Figs 1*a* and 1*b*.
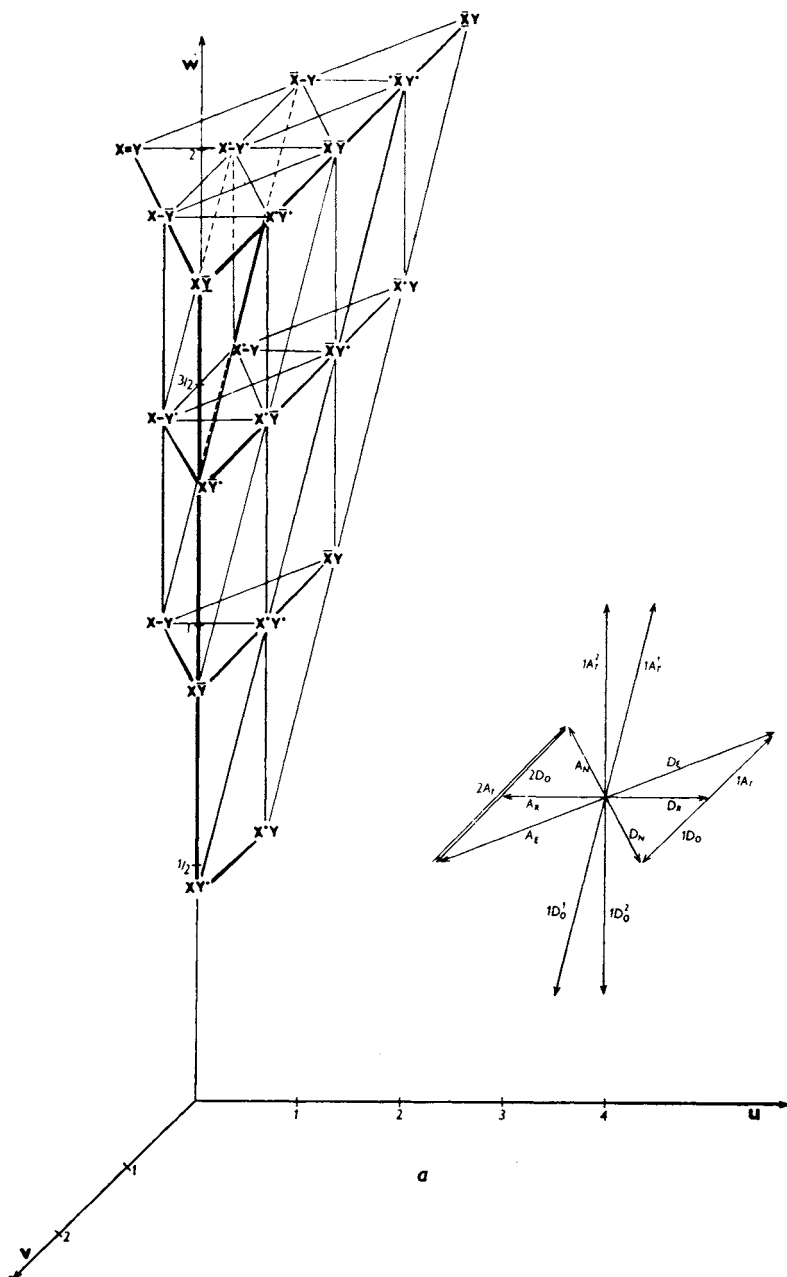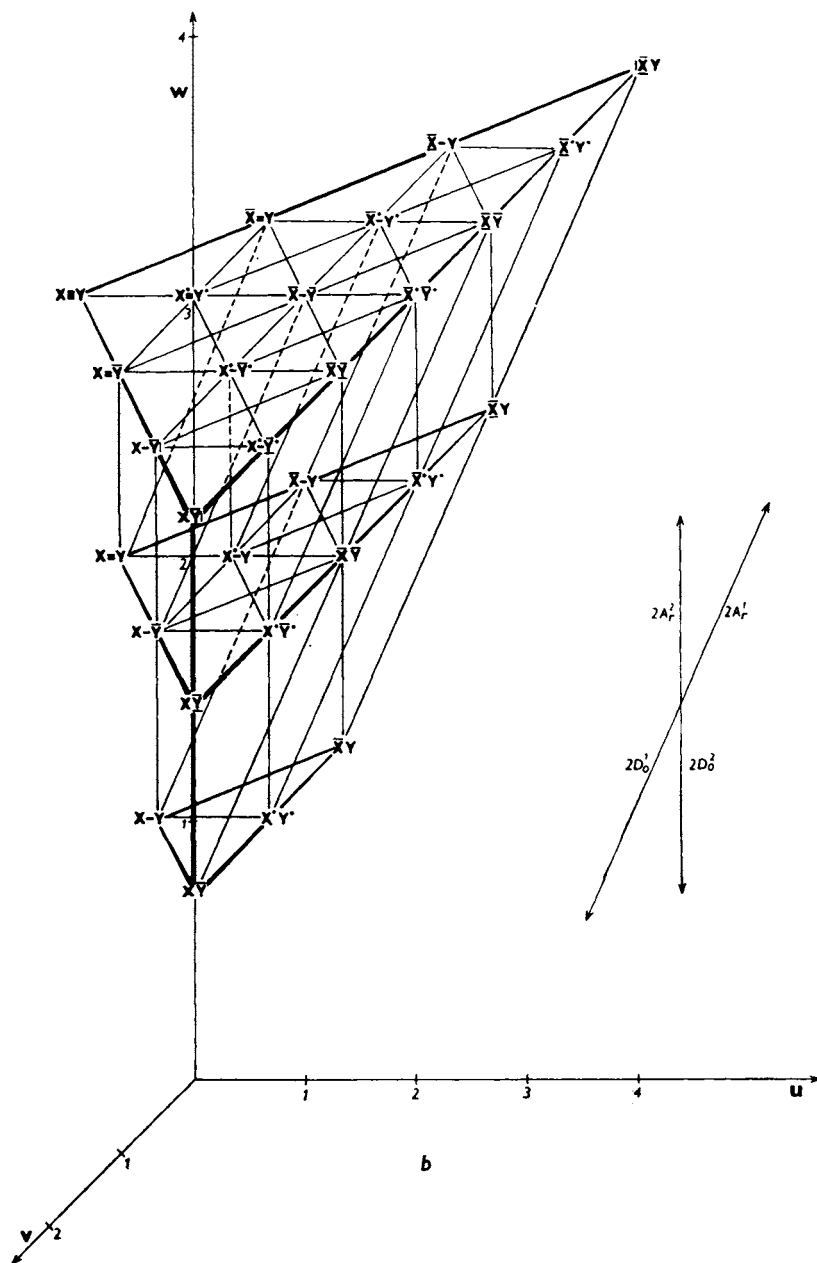
FIG. 1

A subsection of over-all geometrical representation of the states of atomic vectors and their mutual conversions. *a* the states of atomic vectors involving even or odd number of electrons; *b* the states of atomic vectors involving only even number of electrons. The specification of con-

versions is identical with the previous one[1]. If a conversion runs over a pair of atoms, then the used specification is controlled by its first atom. For a conversion involving only one atom the used specification is controlled by this atom (*e.g.* the symbol $1D_0^1$ describes a one electron oxidative dissociation running on the first atom of a given atomic vector, for example, $\overset{\bullet}{X}$—Y $\rightarrow$ X—Y)

ALGEBRA FOR STATES OF ATOMIC VECTORS IN E^9

The elaborated algebraic theory is suggested for a description of atomic vectors from the standpoint of valence states of both their atoms.

Formal Description of States of Atomic Vectors

In our previous communication[1] the valence states of atoms were described by a 4-dimensional vector. Since the state of atomic vector is fully characterized by the valence states of its constituent atoms and the multiplicity of a bond between them, it should be represented by a 9-dimensional vector

$$\mathbf{v} = (v_1, v_2, \ldots, v_9), \tag{9}$$

where v_1, v_6 are the numbers of lone valence electrons on the first and second atom, respectively; v_2, v_7 are the numbers of single bonds starting at the first and second atom, respectively; v_3, v_8 are the numbers of double bonds starting at the first and second atom, respectively; v_4, v_9 are the numbers of triple bonds starting at the first and second atom, respectively; v_5 is the multiplicity of bond between atoms forming the atomic vector.

Example 4: The state of atomic vector assigned to the carbonyl group $\diagdown\mathrm{C}{=}\overline{\mathrm{O}}$ is
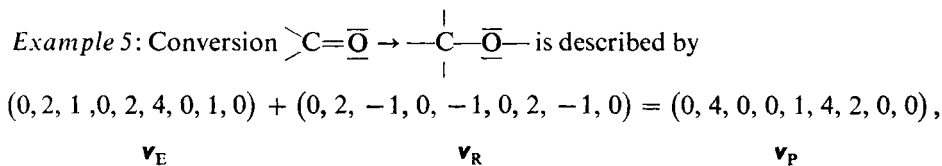
$$\mathbf{v} = (0, 2, 1, 0, 2, 4, 0, 1, 0).$$

The total number of all possible states of atomic vectors may be roughly enumerated by the following simple considerations: We start from 139 valence states of the octet chemistry[1], we can construct from them $139.139 = 19\,321$ pairs. Let us assume that an atom in fixed valence state may form at most two distinct pairs with another atom at a fixed valence state (*e.g.* for valence states —X= and —Y≡ we get a pair of bonded atoms =X—Y≡ and a pair of isolated atoms =X Y≡). This means that we can estimate, roughly, 40 000 distinct states of atomic vectors.

Conversions of States of Atomic Vectors

In formal description of conversions of states of atomic vectors we assume that the elementary conversions of states are running over either simultaneously both atoms or a single atom. For all these cases the conversions is accompanied by a transformation of valence states of either both atoms or a single atom; formally, it can be described by the algebraic approach suggested for conversions of valence atomic states[1] (containing 22 elementary conversions). Since, as was mentioned above, a conversion of a state of a given atomic vector can be distinguished by three different ways (valence states are changed for both atoms, first atom, and second atom, respectively),

the total amount of different elementary conversions of states of atomic vectors is composed of $3.22 = 66$ different types, see Table I. A transition from the initial to final state is described, formally, by Eq. (1), where the used vectors are 9-dimensional.
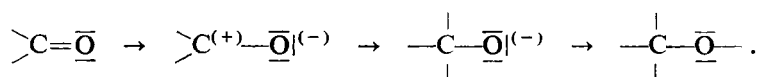
Example 5: Conversion $\diagdown C{=}\overline{O} \rightarrow -\overset{|}{\underset{|}{C}}-\overline{O}-$ is described by

$$(0, 2, 1, 0, 2, 4, 0, 1, 0) + (0, 2, -1, 0, -1, 0, 2, -1, 0) = (0, 4, 0, 0, 1, 4, 2, 0, 0),$$

$$\mathbf{v_E} \qquad\qquad\qquad \mathbf{v_R} \qquad\qquad\qquad \mathbf{v_P}$$

where the used **R** vector $\mathbf{v_R}$ does not belong to the set of elementary conversions, see Table I. This fact means that the studied conversion should be characterized as "multi-step".

Distance Between Two States of Atomic Vectors

Let us assume that all states of atomic vectors form the so-called vertices of a graph G. Two vertices are connected by an edge if both states (assigned to the considered vertices) are mutually related by an elementary conversion. The edges of G are evaluated by alphanumerical symbols of elementary conversions (see Table I). The distance between two states of an atomic vector is defined as the graph-theory distance between the corresponding pair of vertices[1], *i.e.* by the smallest number of edges that form a path connecting the vertices. This corresponds to the smallest number of elementary conversions that transforms the initial (educt) state to the final (product) state.

Example 6: The distance between states $\diagdown C{=}\overline{O}$ and $-\overset{|}{\underset{|}{C}}-\overline{O}-$ is 3, one of the possible paths in the graph G should be of the form

$$\diagdown C{=}\overline{O} \;\rightarrow\; \diagdown C^{(+)}{-}\overline{O}|^{(-)} \;\rightarrow\; -\overset{|}{\underset{|}{C}}-\overline{O}|^{(-)} \;\rightarrow\; -\overset{|}{\underset{|}{C}}-\overline{O}-\,.$$

Now, the task standing before us is to formulate an algorithm for the calculation of distance between two states of an atomic vector. For this purpose we have derived a function d_1 which expresses the distance between two states with the same multiplicity of bond connecting both atoms of the atomic vector (*e.g.* for $\equiv X{-}\overline{Y}$ and $-\overline{X}{-}Y-$). The function d_1 is used in the forthcoming algorithm LENGHT which calculates the distance d between two states of an atomic vector with arbitrary multiplicity of bonds connecting both atoms in the two states.
The function d_1 is specified in Appendix.

Algorithm LENGHT

Step 1 (initialization): Let $\mathbf{a} = (a_1, a_2, ..., a_9)$ and $\mathbf{b} = (b_1, b_2, ..., b_9)$ be two vectors of the evaluated states of an atomic vector, they are choosen in such a way that $a_5 \geqq b_5$ (where

Elementary conversions of states of atomic vectors (symbols from the left column are specified in ref.[1] and by Fig. 1).

Notation	Reaction scheme (example)		Vector $\underset{\sim}{v}_R$
	Heterolytic reactions		
$1D_E$	$X{-}\!\{Y \longrightarrow$	$\overline{X} + Y$	$(2,-1,0,0,-1,0,-1,0,0)$
$2D_E$	$X{=}\!\{Y \longrightarrow$	$\overline{X} - Y$	$(2,1,-1,0,-1,0,1,-1,0)$
$3D_E$	$X{\equiv}\!\{Y \longrightarrow$	$\overline{X} = Y$	$(2,0,1,-1,-1,0,0,1,-1)$
$1D_E^1$	$-X-Y \longrightarrow$	$\overline{X} - Y$	$(2,-1,0,0,0,0,0,0,0)$
$2D_E^1$	$=X-Y \longrightarrow$	$-\overline{X} - Y$	$(2,1,-1,0,0,0,0,0,0)$
$3D_E^1$	$\equiv X-Y \longrightarrow$	$=\overline{X} - Y$	$(2,0,1,-1,0,0,0,0,0)$
$1D_E^2$	$X - Y- \longrightarrow$	$X - \overline{Y}$	$(0,0,0,0,0,2,-1,0,0)$
$2D_E^2$	$X - Y= \longrightarrow$	$X - \overline{Y}-$	$(0,0,0,0,0,2,1,-1,0)$
$3D_E^2$	$X - Y\equiv \longrightarrow$	$X - \overline{Y}=$	$(0,0,0,0,0,2,0,1,-1)$
$1D_N$	$X\{-Y \longrightarrow$	$X + \overline{Y}$	$(0,-1,0,0,-1,2,-1,0,0)$
$2D_N$	$X\{=Y \longrightarrow$	$X - \overline{Y}$	$(0,1,-1,0,-1,2,1,-1,0)$
$3D_N$	$X\{\equiv Y \longrightarrow$	$X = \overline{Y}$	$(0,0,1,-1,-1,2,0,1,-1)$
$1D_N^1$	$-\{X-Y \longrightarrow$	$X - Y$	$(0,-1,0,0,0,0,0,0,0)$
$2D_N^1$	$=\!\{X-Y \longrightarrow$	$-X-Y$	$(0,1,-1,0,0,0,0,0,0)$
$3D_N^1$	$\equiv\!\{X-Y \longrightarrow$	$=X-Y$	$(0,0,1,-1,0,0,0,0,0)$
$1D_N^2$	$X-Y\{- \longrightarrow$	$X - Y$	$(0,0,0,0,0,0,-1,0,0)$
$2D_N^2$	$X-Y\{= \longrightarrow$	$X - Y-$	$(0,0,0,0,0,0,1,-1,0)$
$3D_N^2$	$X-Y\{\equiv \longrightarrow$	$X - Y=$	$(0,0,0,0,0,0,0,1,-1)$

TABLE I

(*Continued*)

$1A_E$	$\widetilde{X}+ Y \longrightarrow$	$X-Y$	$(-2,1,0,0,1,0,1,0,0)$
$2A_E$	$\widetilde{X}-Y \longrightarrow$	$X=Y$	$(-2,-1,1,0,1,0,-1,1,0)$
$3A_E$	$\widetilde{X}=Y \longrightarrow$	$X\equiv Y$	$(-2,0,-1,1,1,0,0,-1,1)$
$1A_E^1$	$\widetilde{X}-Y \longrightarrow$	$-X-Y$	$(-2,1,0,0,0,0,0,0,0)$
$2A_E^1$	$-\widetilde{X}-Y \longrightarrow$	$=X-Y$	$(-2,-1,1,0,0,0,0,0,0)$
$3A_E^1$	$=\widetilde{X}-Y \longrightarrow$	$\equiv X-Y$	$(-2,0,-1,1,0,0,0,0,0)$
$1A_E^2$	$X-\widetilde{Y} \longrightarrow$	$X-Y-$	$(0,0,0,0,0,-2,1,0,0)$
$2A_E^2$	$X-\widetilde{Y}- \longrightarrow$	$X-Y=$	$(0,0,0,0,0,-2,-1,1,0)$
$3A_E^2$	$X-\widetilde{Y}= \longrightarrow$	$X-Y\equiv$	$(0,0,0,0,0,-2,0,-1,1)$
$1A_N$	$X+\widetilde{Y} \longrightarrow$	$X-Y$	$(0,1,0,0,1,-2,1,0,0)$
$2A_N$	$X-\widetilde{Y} \longrightarrow$	$X=Y$	$(0,-1,1,0,1,-2,-1,1,0)$
$3A_N$	$X=\widetilde{Y} \longrightarrow$	$X\equiv Y$	$(0,0,-1,1,1,-2,0,-1,1)$
$1A_N^1$	$X-Y \longrightarrow$	$-X-Y$	$(0,1,0,0,0,0,0,0,0)$
$2A_N^1$	$-X-Y \longrightarrow$	$=X-Y$	$(0,-1,1,0,0,0,0,0,0)$
$3A_N^1$	$=X-Y \longrightarrow$	$\equiv X-Y$	$(0,0,-1,1,0,0,0,0,0)$
$1A_N^2$	$X-Y \longrightarrow$	$X-Y-$	$(0,0,0,0,0,0,1,0,0)$
$2A_N^2$	$X-Y- \longrightarrow$	$X-Y=$	$(0,0,0,0,0,0,-1,1,0)$
$3A_N^2$	$X-Y= \longrightarrow$	$X-Y\equiv$	$(0,0,0,0,0,0,0,-1,1)$
	Homolytic reactions		
$1D_R$	$X+Y \longrightarrow$	$\dot{X} \ \dot{Y}$	$(1,-1,0,0,-1,1,-1,0,0)$
$2D_R$	$X\pm Y \longrightarrow$	$\dot{X}-\dot{Y}$	$(1,1,-1,0,-1,1,1,-1,0)$
$3D_R$	$X\doteq Y \longrightarrow$	$\dot{X}=\dot{Y}$	$(1,0,1,-1,-1,1,0,1,-1)$

TABLE I
(*Continued*)

$1D_R^1$	$+X-Y \longrightarrow \dot{X}-Y$	$(1,-1,0,0,0,0,0,0,0)$
$2D_R^1$	$\doteq X-Y \longrightarrow -\dot{X}-Y$	$(1,1,-1,0,0,0,0,0,0)$
$3D_R^1$	$\triangleq X-Y \longrightarrow =\dot{X}-Y$	$(1,0,1,-1,0,0,0,0,0)$
$1D_R^2$	$X-Y+ \longrightarrow X-\dot{Y}$	$(0,0,0,0,0,1,-1,0,0)$
$2D_R^2$	$X-Y\doteq \longrightarrow X-\dot{Y}-$	$(0,0,0,0,0,1,1,-1,0)$
$3D_R^2$	$X-Y\triangleq \longrightarrow X-\dot{Y}=$	$(0,0,0,0,0,1,0,1,-1)$
$1A_R$	$\overset{\curvearrowright\ \curvearrowright}{X \ + \ Y} \longrightarrow X-Y$	$(-1,1,0,0,1,-1,1,0,0)$
$2A_R$	$\overset{\curvearrowright\ \curvearrowright}{X-Y} \longrightarrow X=Y$	$(-1,-1,1,0,1,-1,-1,1,0)$
$3A_R$	$\overset{\curvearrowright\ \curvearrowright}{X=Y} \longrightarrow X\equiv Y$	$(-1,0,-1,1,1,-1,0,-1,1)$
$1A_R^1$	$\overset{\curvearrowright}{X}-Y \longrightarrow -X-Y$	$(-1,1,0,0,0,0,0,0,0)$
$2A_R^1$	$\overset{\curvearrowright}{-X}-Y \longrightarrow =X-Y$	$(-1,-1,1,0,0,0,0,0,0)$
$3A_R^1$	$\overset{\curvearrowright}{=X}-Y \longrightarrow \equiv X-Y$	$(-1,0,-1,1,0,0,0,0,0)$
$1A_R^2$	$X-\overset{\curvearrowright}{Y} \longrightarrow X-Y-$	$(0,0,0,0,0,-1,1,0,0)$
$2A_R^2$	$X-\overset{\curvearrowright}{Y-} \longrightarrow X-Y=$	$(0,0,0,0,0,-1,-1,1,0)$
$3A_R^2$	$X-\overset{\curvearrowright}{Y=} \longrightarrow X-Y\equiv$	$(0,0,0,0,0,-1,0,-1,1)$

Redox reactions

$1D_o$	$\dot{X} + Y \longrightarrow X + \dot{Y}$	$(-1,0,0,0,0,1,0,0,0)$
$2D_o$	$\bar{X} + Y \longrightarrow X + \bar{Y}$	$(-2,0,0,0,0,2,0,0,0)$
$1D_o^1$	$\dot{X} + Y \longrightarrow X + Y$	$(-1,0,0,0,0,0,0,0,0)$
$2D_o^1$	$\bar{X} + Y \longrightarrow X + Y$	$(-2,0,0,0,0,0,0,0,0)$
$1D_o^2$	$X + \dot{Y} \longrightarrow X + Y$	$(0,0,0,0,0,-1,0,0,0)$
$2D_o^2$	$X + \bar{Y} \longrightarrow X + Y$	$(0,0,0,0,0,-2,0,0,0)$

TABLE I
(*Continued*)

$1A_r$	$X + \overset{\bullet}{Y} \longrightarrow \overset{\bullet}{X} + Y$	$(1,0,0,0,0,-1,0,0,0)$
$2A_r$	$X + \overline{Y} \longrightarrow \overline{X} + Y$	$(2,0,0,0,0,-2,0,0,0)$
$1A_r^1$	$X + Y \longrightarrow \overset{\bullet}{X} + Y$	$(1,0,0,0,0,0,0,0,0)$
$2A_r^1$	$X + Y \longrightarrow \overline{X} + Y$	$(2,0,0,0,0,0,0,0,0)$
$1A_r^2$	$X + Y \longrightarrow X + \overset{\bullet}{Y}$	$(0,0,0,0,0,1,0,0,0)$
$2A_r^2$	$X + Y \longrightarrow X + \overline{Y}$	$(0,0,0,0,0,2,0,0,0)$

the components a_5, b_5 are the multiplicities of bonds between atoms in the first and second state, resp.).

Step 2: If $a_5 = b_5$ then $d(a, b): = d_1(a, b)$ and go to step 5.

Step 3: $N: = a_5 - b_5$. Realize a break of N bonds between atoms from the first state a by all possible permissible ways, the result of these breaks forms a set M. For example, for $N = 1$ and if the first state is X=Y, the set $M = \{\overset{\bullet}{X}-\overset{\bullet}{Y}, \overline{X}-Y, X-\overline{Y}\}$; for $N = 2$ and the same state we get $M = \{|\overline{X}Y, |\overset{\bullet}{X}\overset{\bullet}{Y}, \overline{X}\overline{Y}, \overset{\bullet}{X}\overset{\bullet}{Y}|, X\underline{Y}\}$, and for $N = 3$ and X=Y we get $M = \{|\underline{\overline{X}}Y,$ $\overline{X}{}^{\bullet}Y^{\bullet}, \overline{X}\overline{Y}, \overline{X}{}^{\bullet}\overline{Y}{}^{\bullet}, \overline{X}\underline{\overline{Y}}, \overset{\bullet}{X}\underline{Y}|\}$. The set M is composed of those states that have the same multiplicity of the bond as the second state.

Step 4: $d(a, b): = \min\{d_1(k, b); k \in M\} + N$.

Step 5: End of the algorithm LENGTH.

Example 7: For states —C≡$\overline{N}$ and =C=$\overline{N}$— we obtain (by using the algorithm LENGTH) the distance $d = 3$. One of the possible paths corresponding to this distance is

$$-C\equiv\overline{N} \quad \rightarrow \quad -\overline{C}=\overline{N} \quad \rightarrow \quad =C=\overline{N} \quad \rightarrow \quad =C=\overline{N}-\ .$$
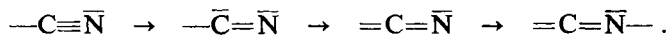
Paths in Graph of States of an Atomic Vector

The aim of this section is to look for shortest paths in the graph G assigned to a given atomic vector, we know that its vertices and edges represent the states of atomic vector and their mutual interconversions, respectively. The know algorithms in graph theory[4,5] are not very useful for our purposes, they usually need to keep explicitly, at least, some subgraph (*e.g.* containing for (C, C) atomic vector approximately 2 000 vertices) of the graph G. We suggest the algorithm PATH which involves as an input the initial (E-educt) and final (P-product) vertices (states of the atomic vector)

and the distance N calculated for these two vertices by the algorithm LENGTH. The algorithm PATH generates only a subgraph of G which is composed of shortest paths from E to P vertices, whereas their distance is determined by the positive integer N. Its complexity does not depend on the total number of vertices of a parent subgraph of G (assigned to the given atomic vector), its time consuming is mainly determined by the distance N, which for realistic applications is ranged from 2 to 6.

Algorithm PATH

Step 1 (initialization): The vertices E and P are algebraically determined by 9-dimensional vectors **e** and **p**, resp., their distance is equal to N, i.e. $N := d(\mathbf{e}, \mathbf{p})$. The set $\mathscr{B}$ is composed of all possible 66 elementary conversions listed in Table I.

Step 2: $i := 0$ and the vertex E is included at the zeroth level.

Step 3: $j := 1$.

Step 4: Take j-th vertex from the i-th level, if it does not exist then go step 6. Apply to this vertex all the elementary conversions from the set $\mathscr{B}$ and retain only those ones that are chemically permissible and their distance d (calculated by the algorithm LENGTH) from the vertex P is equal to $N - i + 1$. Include these vertices at the $(i + 1)$-th level and connect them in a new generated graph by edges with the j-th vertex from the i-th level.

Step 5: $j := j + 1$ and go to step 4.

Step 6: $i := i + 1$, if $i < N - 1$ then go to step 3 else connect by edges all vertices from the $(N - 1)$-th level with the vertex P and go to step 7.

Step 7: You have generated a graph whose vertices and edges form shortest paths between vertices E and P with the length determined by N, this graph forms a subgraph of G assigned to the studied atomic vector.
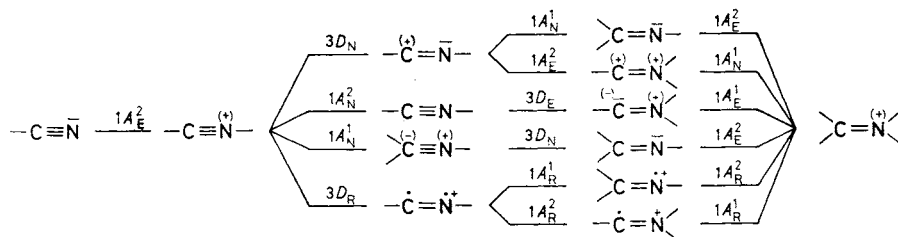
The algorithms LENGTH and PATH were implemented in FORTRAN for the computer EC 1033, they form main part of the program PEACE (Program for Elementary conversions of Atomic Couples Elaboration). This program is able to generate also paths with length longer by a prescribed constant.

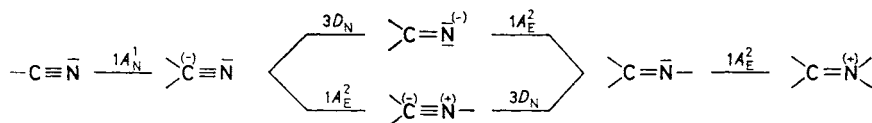**ILLUSTRATIVE CHEMICAL EXAMPLE**

For better understanding of the present theory we give an illustrative example in a detailed form; in particular, we shall consider the conversion of the atomic vector (C, N) from the state $-C \equiv N|$ to the state $\gtrdot C = N^{(+)} \lessdot$. The educt and product states are described by 9-dimensional vectors $(0, 1, 0, 1, 3, 2, 0, 0, 1)$ and $(0, 2, 1, 0, 2, 0, 2, 1, 0)$, respectively. The distance between these input states is $N = 4$; hence, there exist four elementary conversions, sequences of which converse the educt state on to the product state. All the paths produced by the algorithm PATH will be classified by the alphanumerical symbol of first elementary conversion.

The first conversion step $1A_E^2$ produces six paths, where one can recognize among them typical remarks of mechanisms of the Pinner reaction, Stephen reduction of

nitriles or its side realization $\left(\text{with product } \begin{array}{c} R \\ RO \end{array}\!\!\!\diagup\!\!\! C{=}N^{(+)}\!\!\!\diagdown\!\!\!\begin{array}{c} H \\ CH_2R \end{array}\right)$, addition of thioamides or sulfoamides to nitriles, cycloaddition of substituted alkylchlorides to thiocyanates catalyzed by $SnCl_4$, etc. Some paths do not correspond to known reactions since they are either involving rather "exotic" intermediates (e.g. $-C^{(+)}{=}{=}N^{(+)}{\diagdown}$) or still not yet realized.

$$-C{\equiv}\bar{N}\ \xrightarrow{1A_E^2}\ -C{\equiv}\overset{(+)}{N}-\ \Big\langle \begin{array}{l} 3D_N \\ 1A_N^2,\ 1A_N^1 \\ 3D_R \end{array} \cdots \Big\rangle\ {\diagup}C{=}\overset{(+)}{N}{\diagdown}$$

If the first step is $1A_N^1$ (an association of the HOMO of a nucleophile with the LUMO of a carbon), then the program produces two alternative paths.

$$-C{\equiv}\bar{N}\ \xrightarrow{1A_N^1}\ {\diagup}\overset{(-)}{C}{\equiv}\bar{N}\ \Big\langle \cdots \Big\rangle\ {\diagup}C{=}\bar{N}-\ \xrightarrow{1A_E^2}\ {\diagup}C{=}\overset{(+)}{N}{\diagdown}$$

These are typical, for example, for mechanisms of a formation of cyclic imino-carbonates, reactions of nitriles with conjugated bases with C-acides, additions of Grinard reagents to nitriles, alcalic hydrolyses of nitriles, additions of the dimethyl-sulfoxide to nitriles catalyzed by C-electrophiles (tritylperchlorate), etc.

Third alternative possibility determined by the first step $1A_R^2$ represents an association of radical with nitrogen. In these paths one can see an intermediate $-\dot{C}{=}N-$ resulting to nitrile–isonitrile isomerisation, steps of radical addition, etc.

$$-C{\equiv}\bar{N}\ \xrightarrow{1A_R^2}\ -C{\equiv}\dot{N}-\ \Big\langle \cdots \Big\rangle\ {\diagup}C{=}\overset{(+)}{N}{\diagdown}$$
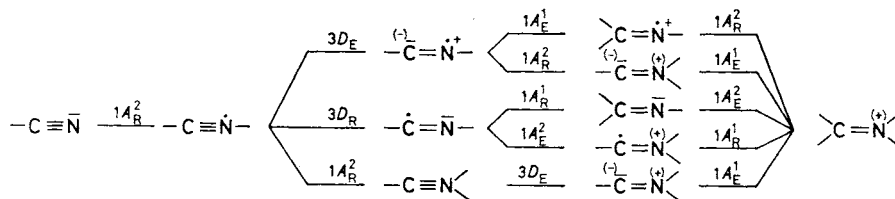
Next three possibilities are characterized in the first step by a dissociation (hetero-lytic or homolytic) of triple bond $C{\equiv}N$. If this dissociation is nucleofugal (from the

standpoint of the carbon) $3D_N$, then we get three paths.

$$-C\equiv\bar{N} \xrightarrow{3D_N} \overset{(+)}{-C}\equiv\underline{\bar{N}}^{(-)}$$
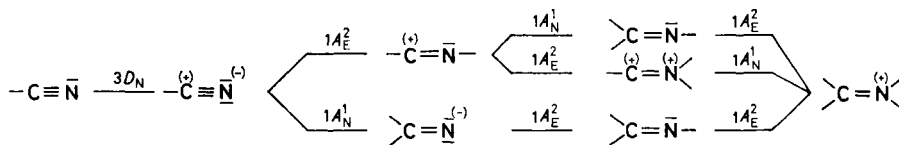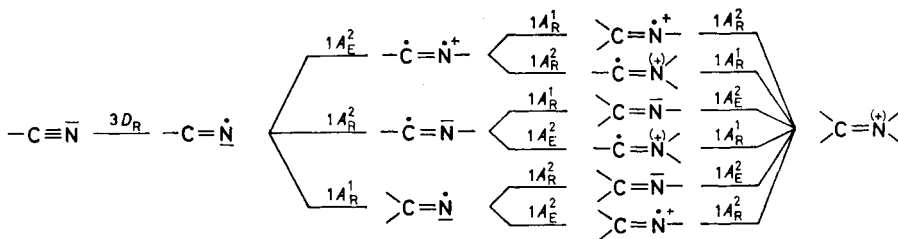
They correspond to intermediate steps of formation of imidoylhalides, O-alkylimido salts, additions of the benzylamine to nitriles catalyzed by $LiAlH_4$, *etc.*

Models of reactions initiated formally by the heterolytic dissociation $3D_E$ of the bond $C\equiv\bar{N}$,

$$-C\equiv\bar{N} \xrightarrow{3D_E} \overset{(-)}{-C}=\underline{\bar{N}}^{(+)}$$
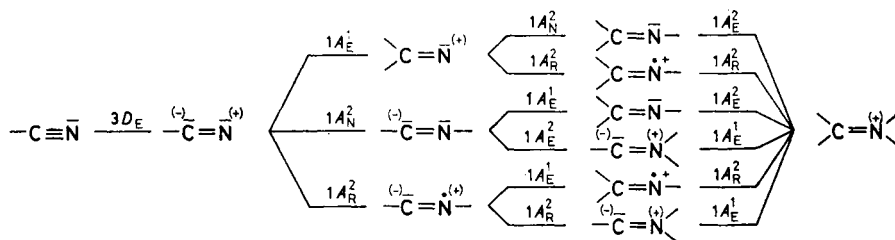
can be applied to additions initiated either by strong electrophiles (with respect to carbon) or strong nucleophiles (with respect to nitrogen). Mixed ion–radical sequences appear rather unusual.

Paths with the first elementary step $3D_R$ may be assigned to models of mechanisms of catalytic hydrogenation of nitriles, the reaction of F_3C—CN with AgF_2

$$\left(\rightarrow \overset{F_3C}{\underset{F}{>}}C=N^{(+)}\overset{F}{\underset{X}{<}}\right), etc.$$

$$-C\equiv\bar{N} \xrightarrow{3D_R} -C=\dot{\underline{N}}$$

Some paths with mixed ion–radical elementary reactions are not still corresponding to known reactions.

If the list of states of considered atomic vector is complete (*e.g.* for atomic vector (C, N) this list contains 156 items, where 50 items are stable), then we may expect that the program produces all possibilities for conversions of an educt state to a product state. Hence, taking these states of a given atomic vector as "subsynthons", we get very powerful algorithmic method for the modelling of syntheses and retrosyntheses, respectively.

In our recent communication[6] we have suggested two alternative methods for the description of reaction mechanisms. The first one is based on the notion of reaction mechanism graph, it can be simply related to the graph G produced by the present theory. The second one, more appropriate for computer implementation, uses the so-called ANCOD (AlphaNumerical CODes) strings. Both these approaches represent very valuable and fruitful methods for formal modelling of reaction mechanisms, their interconnections with the present theory will be studied in forthcoming communication of this series.

APPENDIX

Let $a = (a_1, a_2, ..., a_9)$ and $b = (b_1, b_2, ..., b_9)$ be two vectors describing states of a given atomic vector, and let us assume that they are restricted by $a_5 = b_5$, *i.e.* in both states the multiplicity of bond between atoms is the same. A function F_2 maps a valence state onto another valence state in which the common bond is not counted,

$$F_2(x) = F_2(x_1, x_2, ..., x_9) = \begin{cases} x' = (x_i, x_2, ..., x_9) & \text{for} \quad x_5 = 0 \\ x = (x_i, ..., x_{i+1} - 1, ..., x_{i+5} - 1, ..., x_9) \\ \qquad\qquad\qquad\qquad \text{for} \quad x_5 = i \end{cases}$$

New state vectors corresponding to a and b are denoted by a' and b',

$$a' = F_2(a) = (a'_1, a'_2, ..., a'_9),$$
$$b' = F_2(b) = (b'_1, b'_2, ..., b'_9).$$

Let v be a vector defined by

$$v = (\Delta v_1, \Delta v_2, ..., \Delta v_9) = b' - a'.$$

The entities $I_1, ..., I_6$ are defined as follows

$$I_1 = \Delta v_2 + \Delta v_3 + \Delta v_4, \quad I_2 = \Delta v_3 + \Delta v_4, \quad I_3 = \Delta v_4,$$

$$I_4 = \Delta v_7 + \Delta v_8 + \Delta v_9 \,, \quad I_5 = \Delta v_8 + \Delta v_9 \,, \quad I_6 = \Delta v_9 \,.$$

Let us define the constants B, A_1, A_2, C_1, C_2, D_1, D_2, E_1, and E_2 by $B = \sum_{i=1}^{6} |I_i|$,

$$D_1 = F_1(\Delta v_1, I_1, I_2, I_3), \, D_2 = F_1(\Delta v_6, I_4, I_5, I_6) \,,$$

$$A_1 = F(D_1) \div 2, \, A_2 = F(D_2) \div 2 \,,$$

$$C_1 = \mathrm{mod}_2 (\Delta v_1) \, G(D_1), \, C_2 = \mathrm{mod}_2 (\Delta v_6) \, G(D_2)$$

$$E_1 = H(\Delta v_1 \, \Delta v_6, \min \{\min \{A_1, |\Delta v_1| \div 2\}, \min \{A_2, |\Delta v_6| \div 2\}\})$$

$$E_2 = H(\Delta v_1 \, \Delta v_6, \min \{\min \{C_1, \mathrm{mod}_2 (\Delta v_1)\}, \min \{C_2, \mathrm{mod}_2 (\Delta v_6)\}\}) \,,$$

where $\div$ denotes the integer divide, and $\mathrm{mod}_2 (\,)$ is the modulo-two function, and functions F, G, F_1, and H are defined as follows:

$$F(x) = \begin{cases} x & \text{for} \quad x > 0 \,, \\ 0 & \text{for} \quad x \leqq 0 \,, \end{cases}$$

$$G(x) = \begin{cases} 1 & \text{for} \quad x > 0 \,, \\ 0 & \text{for} \quad x \leqq 0 \,, \end{cases}$$

$$H(x, y) = \begin{cases} 0 & \text{for} \quad x \cdot y \geqq 0 \,, \\ |y| & \text{for} \quad x \cdot y < 0 \,, \end{cases}$$

$$F_1(x_1, x_2, x_3, x_4) = |x_1| - 2 \sum_{i=2}^{4} H(x_1, x_i) \,.$$

Finally, the distance d_1 between states a and b is determined by

$$d_1(a, b) = B + A_1 + A_2 + C_1 + C_2 - E_1 - E_2 \,.$$

Example: We shall calculate the distance between states $=\mathrm{C}=\overline{\mathrm{N}}-$ and $>\mathrm{C}\!=\!\mathrm{N}^{(+)}\!\!<$ we get

$$a = (0, 0, 2, 0, 2, 2, 1, 1, 0) \quad b = (0, 2, 1, 0, 2, 0, 2, 1, 0)$$

$$a' = (0, 0, 1, 0, 2, 2, 1, 0, 0) \quad b' = (0, 2, 0, 0, 2, 0, 2, 0, 0)$$
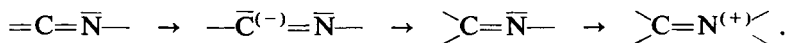
$$v = (0, 2, -1, 0, 0, -2, 1, 0, 0)$$

$$I_1 = -1 \quad I_2 = 0 \quad I_3 = 0 \quad I_4 = -1 \quad I_5 = 0 \quad I_6 = 0$$

$$B = 3 \quad D_1 = 0 \quad D_2 = 1 \quad A_1 = 0 \quad A_2 = 0 \quad C_1 = 0$$

$$C_2 = 0 \quad E_1 = 0 \quad E_2 = 0$$

$$d_1(a, b) = 3 + 0 + 0 + 0 + 0 - 0 - 0 = 3$$

A path in the graph G between the vertices assigned to states *a* and *b* may look as follows

$$=C=\overline{N}\!\!- \quad \rightarrow \quad -\overline{C}^{(-)}\!\!=\!\overline{N}\!\!- \quad \rightarrow \quad {>}C=\overline{N}\!\!- \quad \rightarrow \quad {>}C=N^{(+)}{<} \; .$$

We emphasize that the above theory is applicable also for the calculation of distances between valence states of single atoms[1]. For example, the distance between valence states $\equiv\overline{X}$ and ${>}\overline{X}{-}{-}$ is equal to the distance between states $\equiv\overline{X}Y$ and ${>}\overline{X}Y$ of a formal atomic vector (X, Y).

REFERENCES

1. Koča J., Kratochvíl M., Kunz M., Kvasnička V.: This Journal *49*, 1247 (1984).
2. Dugundji J., Ugi I.: Fortschr. Chem. Forsch. *39*, 19 (1973).
3. Schubert W.: Match *6*, 213 (1979).
4. Dreyfus S. E.: *An Appraisal of Some Shortest Path Algorithms*. Oper. Res. *17*, 395 (1969)
5. Kučera L.: *Kombinatorické algoritmy*. Published by SNTL, Praha 1983.
6. Koča J., Kratochvíl M., Kvasnička V.: This Journal *50*, 1433 (1985).