Original Investigations

Extended Hückel Theory: A New Population Analysis and Its Application to Forecast Chemical Bond Lengths

C. R. Guerillot, Roland Lissillour and A. Le Beuze

Université de Rennes, U.E.R. "Structures et Propriétés de la Matière",
Laboratoire de Chimie Théorique, F-35042 Rennes Cedex, France

Although widely used, the population analysis proposed by Mulliken has been contested by several authors. A new analysis, very easily computed on the orthogonal basis, is here proposed and applied to the EHT wave function.

Under its usual presentation, the EHT method is unable to directly evaluate bond lengths through an energy minimum condition. However, it is possible to settle an empirical quadratic relation between the bond length R_{rs} and a quantity called p_{rs}, similar to a bond population. Such relations are given for bonds of the CC, CN, CO, CS, CF, CCl, CBr, CH, NO and OH types.

The examination of the variation of the bond population under a variation of the bond length has enabled us to prove that this semi-empirical relation was usable in an iterative process: starting from bond lengths taken from any systematic table, it is possible, for a given molecule, to evaluate the bond length consistent with experimental values within an accuracy of 0.03 Å. Some examples, concerning cyclic or acyclic molecules and various kinds of bonds, are given.

Key words: EHT population analysis

The extension of the Hückel method to the whole set of valence electrons (EHT) was first formulated by R. Hoffmann [1] in 1963. Since then, many authors have proposed some modifications or corrections. For instance, P. Schuster gave arguments to introduce an electrostatic correction [2], and Engelke and Beckel [3] made the suggestion that the variation of the EHT molecular energy with interatomic distance could be improved if their expression of the K constant was used instead of the classic Wolfsberg-Helmholtz approximation.

In fact, even if the genuine K^0 method proposed by Hoffmann or if the more

elaborate $K(R)$ proposed by Engelke and Beckel are used, with or without taking in account Schuster's correction, any attempt to forecast chemical bond lengths from an energy minimum condition leads to very erratic and unrealistic results for polyatomic species.[1] The same failure is encountered for valency angles.

The aim of this paper is to show that a semi-empirical procedure may be formulated in order to forecast, within an acceptable accuracy, chemical bond lengths.

1. EHT Formalism and Mulliken's Population Analysis

The EHT procedure may be summarized as follows:

First step: a subset of real Slater's functions ϕ_{ri} is associated with each center in the following way

i	1	2	3	4	$\cdots$
$\phi_{ri} =$	$\phi_{r;ns}$	$\phi_{r;np_x}$	$\phi_{r;np_y}$	$\phi_{r;np_z}$	$\cdots$

the whole set of ϕ_{ri} functions associated with the different centers being a normed but non-orthogonal basis $|\phi)$ called, in this paper, *Slater's basis*.

The nature of an atom r is specified by:

Q_r^0, the number of its valence electrons;
ζ_{ri}, the orbital parameters of its Slater's functions;
I_{ri}, its valence state ionization potential.

To the Slater basis, an overlap matrix $[S^\phi]$ is now associated defined by

$$S^\phi = (\phi \mid \phi). \tag{1}$$

A matrix $[h^\phi]$ is then built up in order to represent a certain effective Hamiltonian on the Slater's basis by setting

$$h^\phi_{ri,ri} = -I_{ri}; \qquad h^\phi_{ri,sj} = -\tfrac{1}{2}K_{risj}S^\phi_{risj}[I_{ri} + I_{sj}], \tag{2}$$

the choice of K_{risj} depending on the used version.

Second step: by use of the Löwdin symmetric orthogonalization scheme [4] an orthogonal basis set $|\lambda)$, called in this paper *Löwdin's basis*, is built on the *Slater* basis from

$$|\lambda) = [\phi][S^\phi]^{-1/2}. \tag{3}$$

The Hamiltonian matrix becomes

$$[h^\lambda] = [S^\phi]^{-1/2}[h^\phi][S^\phi]^{-1/2} \tag{4}$$

and the equation

$$[h^\lambda][C^\lambda] - [C^\lambda][\epsilon] = 0 \tag{5}$$

---

[1]   Of course, the $K(R)$ version leads to results a bit less inaccurate!

is solved, $[\varepsilon]$ being a diagonal matrix. The molecular orbitals Φ_a are then obtained and they are expressed, in the Löwdin basis, by

$$|\Phi] = |\lambda)[C^\lambda]. \tag{6}$$

Third step: one comes back to the Slater basis through the deorthogonalization procedure:

$$|\Phi) = |\phi)[C^\phi]$$
$$[C^\phi] = [S^\phi]^{-1/2}[C^\lambda] \tag{7}$$

and, then, the usual Mulliken population analysis is done [5]. From

$$\Phi_a = \sum_{ri} C_a^{\phi,ri} \phi_{ri} \tag{8}$$

the electronic density at point ρ is given by

$$P_1(\rho) = \sum_{ri} \sum_{sj}' P_{risj}^\phi \phi_{ri}(\rho)\phi_{sj}(\rho) \tag{9}$$

which, *through a whole space integration,* leads to

$$\int P_1(\rho)\, dv = N = \sum_r \sum_{i\in r} P_{riri}^\phi + \sum_{r,s}' \sum_{i\in r} \sum_{j\in s} P_{risj}^\phi S_{risj}^\phi, \tag{10}$$

where N is the total number of valence electrons.

Mulliken identifies the orbital charge Q_{ri}^ϕ to

$$Q_{ri}^\phi = P_{riri}^\phi + \sum_{s\neq r} \sum_{j\in s} P_{risj}^\phi S_{risj}^\phi \tag{11}$$

and, then, goes to a global atomic charge Q_r^ϕ by

$$Q_r^\phi = \sum_{i\in r} Q_{ri}^\phi \tag{12}$$

to an atomic net charge by

$$q_r^\phi = Q_r^0 - Q_r^\phi \tag{13}$$

and, last, to an overlap orbital population $p_{ri;sj}^\phi$ through

$$p_{risj}^\phi = P_{risj}^\phi S_{risj}^\phi \tag{14}$$

leading to an overlap diatomic population $P_{r;s}^\phi$ by

$$p_{r,s}^\phi = \sum_{i\in r} \sum_{j\in s}' p_{ri,sj}^\phi. \tag{15}$$

This analysis, quite exact from the mathematical point of view, has been widely used and is popular. However, many authors have contested it [6–10]. For instance, it has been pointed out that Mulliken's analysis

1. may be unfit for heteropolar bonds [9];
2. leads to atomic central charges, which is clearly unrealistic for lone pairs [9];
3. depends, to a great extent, on the used basis [10].

It must be quoted too that the atomic net charges arising from Mulliken's analysis seem to many authors to be exaggerated. Politzer [8] proposes a partition of the molecular volume into regions associated with each atom, in order to integrate the electronic density. Hirshfeld [9] obtains very similar results from the introduction of a sharing function allowing to assign univocally a part of the electronic density to a specified atom. It must be emphasized that these criticisms were not connected with the use of a particular method of obtainment of the wave function.

If we turn to the bond populations, it seems clear that the Mulliken overlap population, bound to a *whole space integration*, cannot be retained. To remain consistent with Politzer's charge definition, a given bond ought to be associated with a given "bond region" in which the electronic density will be integrated.

2. Population Analysis on the Löwdin Basis

Implicitly, the π-electron restricted Hückel theory implies the use of a basis both *localized* and *orthonormalized*. The Slater basis is of course localized but *is not orthonormalized*. In the second step, the EHT procedure uses another basis, Löwdin's basis, which is, clearly, *orthonormalized* but *not localized*. But Löwdin's basis, as already shown [4b], is transformed, *under a unitary operation*, such as a symmetry transformation, *exactly in the same way as Slater's basis and approaches this basis in the closest way for an orthogonal basis*.

Therefore, it is questionable whether this second step, far from being a simple matter of convenience, could not be the more important step, *Löwdin's basis being the best approximation to Hückel's basis*. Thereafter, the population analysis ought to be done with *Löwdin's basis*.

If we do so, from the molecular orbitals

$$\Phi_a = \sum_{ri} C_a^{\lambda,ri} \lambda_{ri} \tag{8'}$$

and the electronic density at point

$$P_1(\rho) = \sum_{ri} \sum_{sj} P_{risj}^{\lambda} \lambda_{ri}(\rho) \lambda_{sj}(\rho) \tag{9'}$$

a whole space integration gives

$$\int P_1(\rho) \, dv = N = \sum_{ri} P_{riri}^{\lambda}, \tag{10'}$$

and, then, in the usual way,

$$Q_{ri}^{\lambda} = P_{riri}^{\lambda}, \tag{11'}$$

$$Q_r^{\lambda} = \sum_{i \in r} Q_{ri}^{\lambda}, \tag{12'}$$

$$q_r^{\lambda} = Q_r^0 - Q_r^{\lambda}. \tag{13'}$$

As the $|\lambda)$ basis is orthonormalized, the whole space integration gives a vanishing contribution for the $p_{ri;sj}$ cross-terms.

Thereby, the population analysis leads to charges Q_{ri}^λ, Q_r^λ, and q_r^λ that could be exposed to the same criticisms as those obtained from the usual Mulliken analysis.

In fact, the wanted quantities would be computed from an integration on some little volumes associated some with atoms and others with bonds, instead of a whole space integration. Now, if we introduce the $[\mathscr{S}]$ matrix, defined as

$$[\mathscr{S}] = [S^\phi] - [1], \tag{16}$$

a given λ_{ri} function may be developed as a series on the ϕ functions as

$$\lambda_{ri} = \phi_{ri} - \tfrac{1}{2}\sum_{tk} \mathscr{S}_{ri,tk}\phi_{tk} + \tfrac{3}{8}\sum_{tk}\sum_{ul} \mathscr{S}_{tknl}\mathscr{S}_{riul}\phi_{tk} + \cdots. \tag{17}$$

Then, *a charge distribution* $\lambda_{ri}\lambda_{sj}$ will be

$$\lambda_{ri}\lambda_{sj} = \phi_{ri}\phi_{sj} - \tfrac{1}{2}\sum_{tk} \mathscr{S}_{ritk}\phi_{tk}\phi_{sj} - \tfrac{1}{2}\sum_{ul} \mathscr{S}_{sjul}\phi_{nl}\phi_{ri} + \cdots \tag{18}$$

and the electronic density has the expression

$$P_1(\rho) = \sum_{ri}\sum_{sj} p_{risj}^\lambda \{\phi_{ri}(\rho)\phi_{sj}(\rho)$$
$$- \tfrac{1}{2}\sum_{tk} [\mathscr{S}_{ritk}\phi_{tk}(\rho)\phi_{sj}(\rho) + \cdots + \mathscr{S}_{sjtk}\phi_{tk}(\rho)\phi_{ri}(\rho)] + \cdots\}. \tag{19}$$

If we integrate this expression *within a finite ellipsoidal volume* V_{rs}, the centers r and s being at the foci, we get

$$\int_{V_{rs}} P_1(\rho)\, dv \simeq \sum_{i\in r}\sum_{j\in s} p_{ri\,sj}^\lambda \int_{V_{rs}} \phi_{ri}(\rho)\phi_{sj}(\rho)\, dv + \cdots, \tag{20}$$

this first term of the series being the more especially important as the ratio χ_{rs} of the principal axis of the ellipsoid to the interatomic distance R_{rs} is nearer to 1.

The quantity $\int_{V_{rs}} \phi_{ri}(\rho)\phi_{sj}(\rho)\, dv$ is a monotonous function of χ_{rs}, going to $S_{ri;sj}^\phi$ when χ_{rs} goes to infinity. Therefore,

$$\int_{V_{rs}} \phi_{ri}(\rho)\phi_{sj}(\rho)\, dv = A_{rs}(\chi_{rs})S_{risj}^\phi; \quad 0 \leqslant A_{rs}(\chi_{rs}) \leqslant 1. \tag{21}$$

Hence, the integration of the electronic density, *within a bond volume* equivalent to a finite ellipsoid with the two centers as foci, leads to *a bond population* defined as

$$p_{r,s} = A_{rs}\sum_{i\in r}\sum_{j\in s} p_{risj}^\lambda S_{risj}^\phi \tag{22}$$

and, then, proportional to the quantities

$$p_{ri,sj}^\lambda = P_{risj}^\lambda S_{risj}^\phi \tag{14'}$$

and

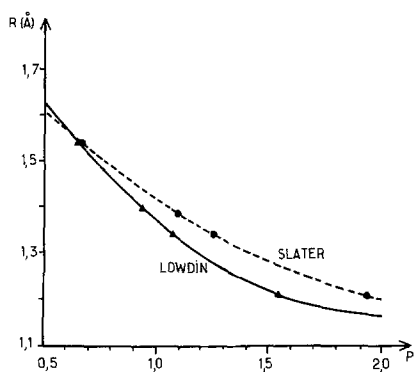$$p_{r,s}^\lambda = \sum_{i\in r}\sum_{j\in s} p_{risj}^\lambda \tag{15'}$$

Fig. 1. Standard curves, C—C bond, K^0 version

where the term $p^\phi_{ri\,;sj}$ comes from the orthogonal analysis and the term $S^\phi_{ri\,;sj}$ is just an overlap integral on Slater's functions.

Of course, the charges may be studied, in the very same way.

3. Proving the Existence of an Empirical Relationship

The idea that an empirical relationship may occur between the length and the order of a given bond is clearly not a new one. As early as 1939, C. A. Coulson [10] has given such a relationship within the frame of the Hückel method. Since then, H. C. Longuet-Higgins and L. Salem [11] have proved that such a relationship may exist when both bond energy and bond order are univocal functions of interatomic distances.

Then, as the quantities p^λ_{rs} and p^ϕ_{rs} may be regarded as generalizations of the bond order concept (with a scaling factor, of course), we can expect to find a relationship such as

$$R_{rs} = F(p_{rs}). \tag{23}$$

We can look for several forms. For instance, Coulson has used a continued fraction. The examination of Figs. 1 and 2, representing the standard curves for the C—C bond for the K^0 and $K(R)$ versions, leads us to a simple quadratic form:

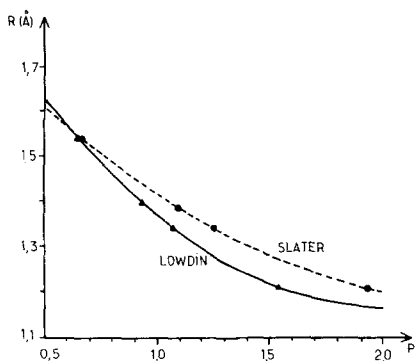$$R_{rs} = a_0 + a_1 p_{rs} + a_2 p^2_{rs}. \tag{24}$$

Fig. 2. Standard curves, C—C bond, $K(R)$ version

Table 1. Coefficients in the semi-empirical quadratic expression for the C–C bond

Variant	Basis	a_0	a_1	a_2
K^0	Löwdin	1.98679	−0.82794	0.20946
	Slater	1.85123	−0.54204	0.10790
$K(R)$	Löwdin	1.99515	−0.84810	0.21887
	Slater	1.87006	−0.57534	0.11986

Of course, the actual values of the coefficients depend on the choice of the version and of the population analysis used. Our results for the C—C bond are summarized in Table 1.

As we are intent on building up an empirical relationship, it is worth proceeding to a statistical examination of the results obtained from the four possibilities for a large set of molecules the geometries of which are known from microwave spectrometry [12]. These results, and their statistical analysis, are collected in Table 2.

Table 2. Comparison between R_{exp} and R_{cal} for the C–C bond

| Molecule	R_{exp} (Å)	K^0 variant		$K(R)$ variant	
		Löwdin	Slater	Löwdin	Slater
C_2H_6	1.536	1.535	1.536	1.536	1.537
C_2H_4	1.338	1.336	1.338	1.337	1.340
C_2H_2	1.206	1.206	1.206	1.206	1.206
C_6H_6	1.393	1.396	1.393	1.395	1.390
CH_3—$C{\equiv}CH$	1.459	1.456	1.471	1.459	1.457
	1.206	1.211	1.209	1.212	1.210
$FC{\equiv}CH$	1.198	1.209	1.210	1.209	1.210
$ClC{\equiv}CH$	1.204	1.212	1.214	1.214	1.213
CH_3—CN	1.458	1.443	1.454	1.436	1.440
CF_3—CH_3	1.492	1.493	1.478	1.494	1.483
$CH{=}CH$—CH_2	1.515	1.550	1.577	1.546	1.570
	1.300	1.321	1.321	1.322	1.327
$CH_2{=}CH$—CN	1.339	1.347	1.354	1.349	1.358
	1.426	1.419	1.429	1.412	1.476
C_5H_5N	1.395	1.394	1.388	1.393	1.385
	1.394	1.398	1.398	1.397	1.395
CH_3CHO	1.501	1.485	1.483	1.482	1.476
$CH_2{=}C{=}O$	1.314	1.316	1.328	1.315	1.325
CF_3—$C{\equiv}CH$	1.460	1.468	1.473	1.467	1.471
	1.201	1.206	1.207	1.206	1.207
(R_{exp}/R_{cal})		0.9975	0.9953	0.9980	0.9949
Probable error in %		0.5	0.7	0.6	0.9
Linear regression					
a		1.0089	0.9831	1.0164	0.9814
b		−0.0155	0.0166	−0.0250	0.0185
r		0.9954	0.9903	0.9947	0.9860
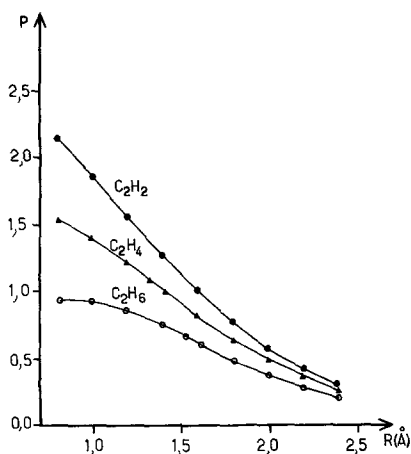
Fig. 3. Bond population as a function of interatomic distance $K(R)$ version, Löwdin's basis

Although the four possibilities lead to good results, a slight advantage must be proclaimed in favor of the analysis using Löwdin's basis. The two versions appear to be nearly equivalent. As the $K(R)$ version seems to us a little more theoretically founded, we have chosen to carry on our work in this last version.

We must recall now that the use of a semi-empirical relationship implies an univocal correspondence between the bond population p_{rs} and the interatomic distance R_{rs}. In Figs. 3 and 4 the variations of p_{rs} versus R_{rs} are shown, both for the Löwdin and Slater bases. In the first case, it can be seen that the three functions are monotonous for distances longer than 0.8 Å (below, this is, of course, of no practical importance).

On the other hand, for Slater's basis, the functions associated with ethane and ethylene go through a maximum, and, therefore, there is no biunivocal correspondence between p_{rs}^{ϕ} and R_{rs}.[2] Then, we choose to carry on this work with the population analysis using Löwdin's basis only.

4. Extension to Other Bonds

In order to evaluate the coefficients of the quadratic expression associated to a given bond, we shall proceed as follows:

first, we select four molecules, the known microwave geometries of which give, within an accuracy better than 0.01 Å, four bond lengths distributed in a convenient way between the shorter and the longer distances associated with the studied bond;

second, we compute, in the selected version, the corresponding bond populations;

third, we compute the least square coefficients of the quadratic relationship;

fourth, we check the relationship fit on a set of more than ten different bond lengths coming themselves from microwave data.

---

[2]   We must notice that the same results, even more accentuated, are exhibited when K^0 version is used.
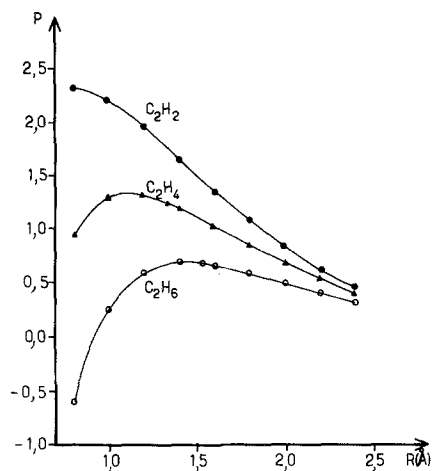
Fig. 4. Bond population as a function of interatomic distance $K(R)$ version, Slater's basis

The computed coefficients are collected in Table 3. They are associated with $K(R)$ version and Löwdin's basis population analysis.

The goodness of the fit is evaluated through the linear regression

$$R_{exp} = a R_{cal} + b \tag{25}$$

and the corresponding figures are given in Table 4.

We give also some results obtained with K^0 version, always for Löwdin's basis population analysis (Table 5).

It can be seen in Table 6 that the fit remains good.

In fact, the results obtained both with $K(R)$ and K^0 versions are close enough so that it is possible to use the coefficients collected in Table 2, even if the actual computations were done in the K^0 version. Some examples of this cross calculation are given in Table 7.

Table 3. Coefficients for empirical relationships ($K(R)$ version, Löwdin's basis)

Bond	a_0	a_1	a_2
C—C	1.99515	−0.84800	0.21887
C—N	1.90064	−0.82957	0.20812
C—O	1.87316	−1.01108	0.32704
C—S	2.31938	−0.99880	0.29237
C—F	1.92763	−1.63290	0.95066
C—Cl	2.45379	−1.63913	0.78718
C—Br	2.41404	−0.89824	0.17074
N—O	1.45012	−0.30358	−0.00415
C—H	2.0800	−2.2343	1.1526
O—H	1.51346	−1.30910	0.49688

Table 4. Analysis of fit ($K(R)$ version, Löwdin's basis)

Bond	a	b	r	n
C—C	1.0336	−0.0461	0.9946	31
C—N	0.9985	0.0225	0.9985	13
C—O	1.0207	−0.0293	0.9965	14
C—S	1.0003	−0.0005	0.9939	10
C—F	0.9898	0.0134	0.9961	14
C—Cl	1.0366	−0.0619	0.9957	10
C—Br	0.9999	0.0003	0.9979	10
N—O	0.98607	0.03164	0.9733	10
C—H	0.9980	0.0013	0.9614	10

Table 5. Coefficients for empirical relationships (K^0 version, Löwdin's basis)

Bond	a_0	a_1	a_2
C—C	1.98679	−0.82794	0.20946
C—N	1.89800	−0.83129	0.21089
C—O	1.82313	−0.92101	0.28341
C—F	1.88682	−1.61586	0.98719
C—Cl	2.48741	−1.99342	0.91051

Table 6. Analysis of fit (K^0 version, Löwdin's basis)

Bond	a	b	r	n
C—C	1.0311	−0.0433	0.9953	31
C—N	0.9807	0.0285	0.9988	13
C—O	1.0062	−0.0112	0.9922	14
C—F	1.0020	−0.0338	0.9992	14
C—Cl	1.0467	−0.0779	0.9912	7

Table 7. Bond length evaluation from K^0 version calculated bond populations and empirical coefficients listed in Table 2

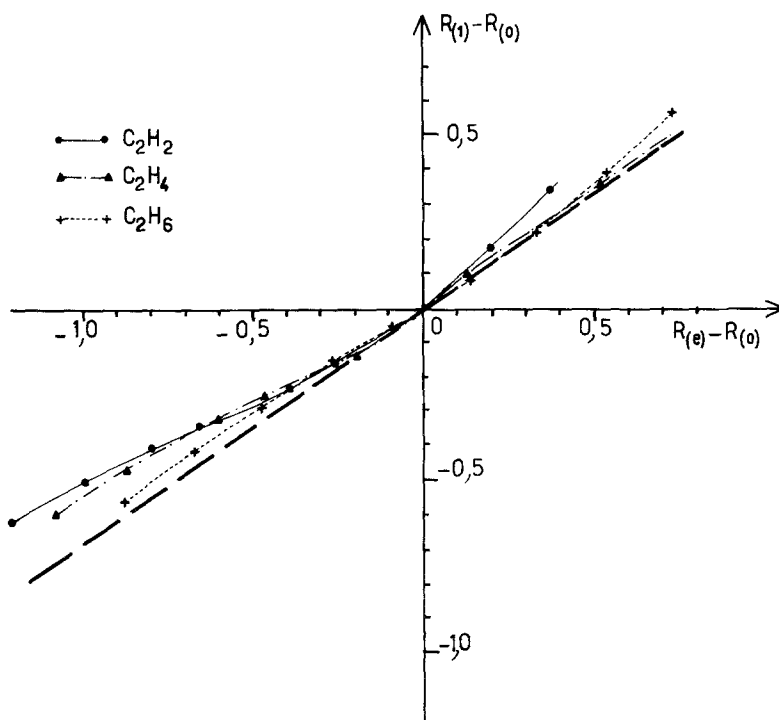| Molecule	Bond	p_{rs}^{λ}	R_{cal} (Å)	R_{exp} (Å)
CH_3—CH_2F	C—C	0.6781	1.521	1.533
CH_3—C≡CH	C—C	0.8058	1.454	1.459
	C≡C	1.5279	1.210	1.206
C_6H_6	C—C	0.9355	1.393	1.393
C_5H_5N	C—N	0.8482	1.347	1.341
CH_3NH_2	C—N	0.6008	1.477	1.474
HCN	C≡N	1.4131	1.144	1.153
CH_3OH	C—O	0.5069	1.442	1.434
CH_3CHO	C=O	0.9192	1.212	1.216

Fig. 5. Relation between $R^{(1)} - R^{(0)}$ and $R^{\text{exp}} - R^{(0)}$ C—C bond, $K(R)$ version, Löwdin's basis

5. Feasibility of an Iterative Procedure

All the preceding results were obtained through EHT computations with assigned known geometries. Our next problem is: are we now able to forecast a given bond length in an unknown-geometry molecule?

We have already seen that the bond population p_{rs}^λ varies as a function of the length R_{rs}. Then, if we take an initial length $R_{rs}^{(0)}$, we can compute, through the EHT procedure, a bond population $p_{rs}^{\lambda(0)}$. Then, we can deduce a length $R_{rs}^{(1)}$.

If R_{rs}^{exp} is the true distance between atoms r and s, an iterative procedure is feasible if:

1. the bond population p_{rs}^λ is a monotonous function of the length R_{rs}, at least in the region of practical physical importance;

2. the difference $R_{rs}^{(1)} - R_{rs}^{(0)}$ is nearly a linear function of the difference $R_{rs}^{\text{exp}} - R_{rs}^{(0)}$, these two functions going simultaneously to zero;

3. in the neighborhood of $R_{rs}^{\text{exp}} - R_{rs}^{(0)} = 0$, the figurative points, for a given bond, lie on the same straight line whatever the studied molecule may be.

We had already seen (Fig. 3) that, for lengths longer than 0.8 Å, the bond population p_{rs}^λ meets the first condition. Then, we had studied the linear regression

$$R_{rs}^{(1)} - R_{rs}^{(0)} = a(R_{rs}^{\text{exp}} - R_{rs}^{(0)}) + b \qquad (26)$$
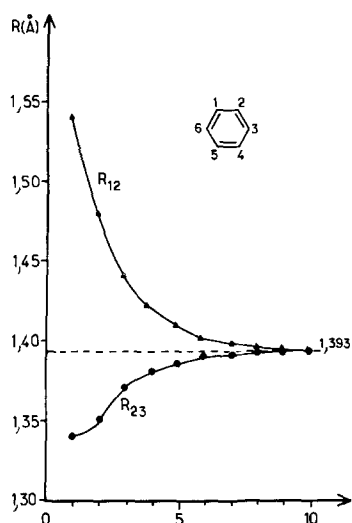
Fig. 6. Iterative routine on benzene

for the three molecules ethane, ethylene, and acetylene, the starting length $R_{rs}^{(0)}$ going from 0.8 to 2.4 Å. Results are given in Table 8. In this table, the last line corresponds to the whole set of points belonging to the three molecules, the values of $R_{rs}^{\exp} - R_{rs}^{(0)}$ standing between -0.3 and $+0.3$ Å.

The examination of Table 8 and of the corresponding Fig. 5 allows us to assert that the two last conditions are also met.[3] More, it is easy to show that, if the a coefficient in the linear regression (26) is included between 0 and 1, then the successive differences $D^{(n)} = R^{(n)} - R^{(n-1)}$ form a convergent sequence for Cauchy's ratio, $D^{(n+1)}/D^{(n)}$, is equal to $(1 - a)$. Therefore, the sequence will always converge to $R^{\exp}$, within an accuracy of the magnitude of the product ab. For the C—C bond, for instance, this accuracy would be about 0.0066 Å.[4]

In order to check the feasibility of such an iterative routine, we have studied the benzene molecule. As a starting point, we have taken a Kekulé formula, the C—C

Table 8. Linear relationship between $R_{rs}^{(1)} - R_{rs}^{(0)}$ and $R_{rs}^{\exp} - R_{rs}^{(0)}$

Molecule	a	b	r
C_2H_6	0.6968	0.0303	0.9973
C_2H_4	0.5994	0.0295	0.9961
C_2H_2	0.5875	0.0411	0.9880
C—C	0.6802	0.0097	0.9910

[3]   As Longuet-Higgins and Salem [11] have already foreseen, for a large difference $R_{rs}^{\exp} - R_{rs}^{(0)}$, figurative points lie on slightly different lines in account of hybridization.

[4]   Computational hazards may perhaps introduce oscillations when a will be very near to 1. We have not found this trouble in our present computations. It could be eliminated by taking as a new starting point, the mean of the two oscillating values.
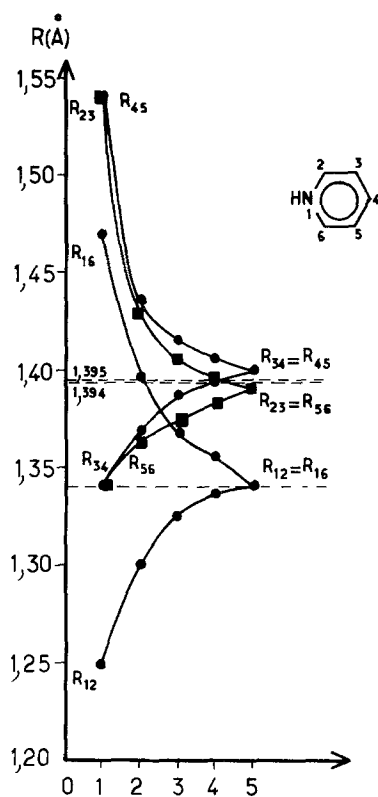
Fig. 7. Iterative routine on pyridine

lengths being, alternately, 1.34 and 1.54 Å. For each step, we have computed the bond populations and deduced new C—C lengths. Our results, shown in Fig. 6, are also collected in Table 9.

Therefore, it is clear that an iterative routine is feasible. For evident reasons we have modified the procedure used for benzene, in order to include an accelerating factor and a stop test.

Table 9. Iterative routine on benzene

Step	R_a^0	p_z	R_a^1	R_b^0	p_b	R_b^1
1	1.54	0.7556	1.479	1.34	1.0462	1.347
2	1.48	0.8263	1.444	1.35	0.9947	1.368
3	1.44	0.8732	1.421	1.37	0.9782	1.375
4	1.42	0.8959	1.411	1.38	0.9602	1.383
5	1.41	0.943	1.400	1.385	0.9468	1.388
6	1.40	0.9259	1.398	1.39	0.9438	1.390
7	1.398	0.9309	1.395	1.390	0.9380	1.392
8	1.395	0.9350	1.394	1.392	0.9372	1.393
9	1.394	0.9354	1.393	1.393	0.9356	1.393
10	1.393	0.9352	1.393	1.393	0.9353	1.393

We propose to perform as follows:

First step: take an initial set $\{R_{rs}^{(0)}\}$ of bond lengths deduced either from a limit formula and a localized bond length system (for instance Pauling and Huggins [13] or Schomaker, Stevenson and Gordy [14] or from the well-known Pople's standard lengths table [15]);

Second step: compute the bond populations $\{p_{rs}^{\lambda(0)}\}$ corresponding to this initial geometry, get a set of distances $\{R_{rs}'\}$ from the quadratic relationship, and then compute the bond lengths for the next iteration with

$$R_{rs}^{(1)} = R_{rs}^{(0)} + 1.5(R_{rs}'^{(0)} - R_{rs}^{(0)}) \tag{27}$$

where 1.5 is the accelerating factor.

Iterate step 2 till the stop condition

$$[R_{rs}'^{(n)} - R_{rs}^{(n)}]_{max} \leqslant \varepsilon \tag{28}$$

will be met.

In our present computation, we have taken $\varepsilon = 0.006$ Å.

6. Some Examples

This routine was applied to molecules of pyridine, furan, butadiene, formic acid, and formamide. When dealing with cyclic molecules, we have evaluated, at each step, the valency angles through a least square method, admitting the equality of all the angular force constants. This is clearly a rough approximation.

More, it is clear that bond populations depend on valency angles. For instance, in water molecules, when the HOH angle increases from 80 to 140 degrees, the O—H bond population increases also, from 0.4942 to 0.5662, letting out a decrease of the O—H length from 0.99 to 0.95 Å.

But such large variations are linked with large angular variations. In our furan calculation, the larger angular variation was less than 5 degrees. So the error due to the dependence of bond population on angular variation is quite less than those arising from the supposed equality of the angular force constants.

However, in more accurate computations, it is necessary to take this dependence into account. Then the iterative routine must be fulfilled for each angular value. Unfortunately the EHT method, in its usual formulation, is unable to give valency angle. Therefore, no energy criterion can be used. Angular force constants must be taken from literature tables. . . . But they are scarcely accurate for a given molecule, specially an unknown one.

Therefore, we think that it is sufficient to iterate simultaneously both the lengths and the angles.

Thereafter our results are as shown in Tables 10–14.

Table 10. Geometry of the pyridine molecule

Step	R_{12}	R_{23}	R_{34}	R_{45}	R_{56}	R_{61}
Pyridine						
0	1.25	1.54	1.34	1.54	1.34	1.47
1	1.300	1.429	1.368	1.433	1.363	1.396
2	1.325	1.405	1.388	1.414	1.373	1.368
3	1.337	1.396	1.394	1.405	1.384	1.355
4	1.340	1.390	1.400	1.400	1.390	1.340
exp [16]	1.340	1.395	1.394	1.394	1.395	1.340

Table 11. Geometry of the furan molecule

Furan

Step	R_{12}	R_{23}	R_{34}
0	1.40	1.34	1.54
1	1.369	1.364	1.458
2	1.344	1.363	1.445
3	1.342	1.371	1.439
exp [17]	1.362	1.361	1.431
[18]	1.371	1.354	1.440

Table 12. Geometry of the butadiene molecule

Butadiene. $H_2C_{(1)}$—$C_{(2)}H$—$C_{(3)}H$—$C_{(4)}H_2$

Step	R_{12}	R_{23}
0	1.34	1.54
1	1.343	1.458
2	1.348	1.448
3	1.349	1.450
exp [19]	1.337	1.483

Table 13. Geometry of the formic acid molecule

Formic acid. H—C with O_1 (double bond) and O_2—H
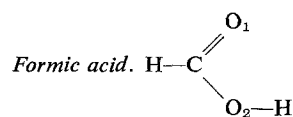
Step	$C{=}O_1$	$C{-}O_2$
0	1.20	1.40
1	1.205	1.360
2	1.212	1.339
3	1.217	1.328
4	1.220	1.322
exp [20]	1.202	1.343
[21]	1.245	1.312

Formamide. $H-C\overset{\displaystyle O}{\underset{\displaystyle NH_2}{\diagdown}}$

Table 14. Geometry of the formamide molecule

Step	C=O	C—N
0	1.20	1.47
1	1.216	1.392
2	1.231	1.369
3	1.240	1.357
4	1.243	1.353
exp [22]	1.193	1.376
exp [23]	1.243	1.343
exp [24]	1.243	1.319
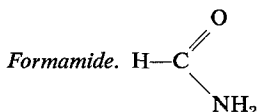
For every example, our results are fairly accurate, the relative error remaining near 1%, except for the C—N bond in formamide, when compared with the data from Ottersen Tor [24]. But this author has worked on the crystal phase and he quoted that hydrogen bonding plays an important part and that there is a large shortening of the C—N bond, precisely due to the hydrogen bonding. It can be seen that, on an average, our results lie in an intermediary position between micro-wave and radiocristallographical data.

7. Conclusion

Either in its genuine K^0 version or in its more elaborated $K(R)$ one, the EHT method looks like a topological algorithm unable to forecast bond lengths from an energy minimum condition.

The Löwdin basis, up to now regarded just as a matter of convenience, probably plays a more important part. The population analysis, when done upon this basis, enables us to determine a *bond population* which can be used in the same way as Hückel's bond order.

Specially, an empirical relationship between the length of a bond and the value of the corresponding bond population does exist.

Standard curves may be laid down. An iterative routine may be formulated which allows to forecast, within an accuracy near 1%, a given bond length in a given unknown molecule.

References

1. Hoffman, R.: J. Chem. Phys. **39**, 1397, 2745 (1963); **40**, 2472, 2480 (1964)
2. Schuster, P.: Monatsh. Chem. **100**, 1033 (1969)
3. Engelke, Z. Beckel, C. L.: Intern. J. of Quantum Chem. **58**, 209 (1974).
4. a) Löwdin, P. O.: J. Chem. Phys. **18**, 365 (1950); b) Advan. Quantum Chem. **5**, 185 (1970)

5. Mulliken, R. S.: J. Chem. Phys. **23**, 1833, 2343 (1955)
6. Roby, K. R.: Theoret. Chim. Acta (Berl.) **33**, 105 (1974)
7. Rinaldi, D., Rivail, J. L., Barriol, J.: Theoret. Chim. Acta (Berl.) **22**, 291 (1971)
8. Politzer, R., Harris, R. R.: J. Am. Chem. Soc. **92**, 6451 (1970); Politzer, P., Mulliken, R. S.: J. Chem. Phys. **55**, 5135 (1971); Politzer, P.: Theoret. Chim. Acta (Berl.) **23**, 203 (1971); Pollitzer, P., Reggio, P. H.: J. Am. Chem. Soc. **94**, 8308 (1972)
9. Hirshfeld, F. L.: Theoret. Chim. Acta (Berl.) **44**, 129 (1977)
10. Coulson, C. A.: Proc. Roy. Soc. (London) **A169**, 413 (1939)
11. Longuet-Higgins, H. C., Salem, L.: Proc. Roy. Soc. (London) **A251**, 172 (1959); see also Salem, L.: The molecular orbital theory of conjugated systems, p. 136. New York: W. A. Benjamin (1960)
12. Gordy, W., Cook, R. L.: "Microwave molecular spectra", in: Technic of organic chemistry, A. Weissberger, Ed., vol. **55**, part II, p. 675–710. New York: Interscience 1970
13. Pauling, L., Huggins, M. L.: Z. Krist. **A87**, 205 (1934)
14. Schomaker, V., Stevenson, D. P.: J. Am. Chem. Soc. **63**, 37 (1941); Gordy, W.: J. Chem. Phys. **15**, 81 (1947)
15. Pople, J. A., Beveridge, D. L.: Approximate molecular orbital theory, p. 111. New York: McGraw-Hill 1970
16. Bak, B., Hansen-Nygaard, L., Rastrup-Andersen, J.: J. Mol. Spectry. **22**, 361 (1958)
17. Bak, B., Cristensen, D., Dixon, W. B., Hansen-Nygaard, L., Rastrup-Andersen, J., Schottlander, M.: J. Mol. Spectry. **9**, 124 (1962
18. Bak, B., Hansen, L., Rastrup-Andersen, J.: Disc. Faracay Soc. **19**, 30 (1955)
19. Almennigen, A., Bastiansen, O., Traetteberg, M.: Acta Chem. Scand. **12**, 1221 (1958)
20. Scharpen, L. H., Laurriee, V. M.: J. Chem. Phys. **43**, 2765 (1965)
21. Lerner, R. G., Dailey, B. P., Friend, J. P.: J. Chem. Phys. **26**, 680 (1957)
22. Wollrab, J. E., Scharpen, L. H., Ames, D. P.: J. Chem. Phys. **49**, 2405 (1965)
23. Kurland, R. J., Wilson, E. B.: J. Chem. Phys. **27**, 585 (1957)
24. Ottersen Tor: Acta Chem. Scand. **A29**(10), 939 (1975)