ON STUDYING HÜCKEL PARAMETERS BY USE OF A

MAPPING CONCEPT

R. Brüggemann and J. Voitländer

Institut für Physikalische Chemie der Universität München,

8000 München 2, Sophienstr. 11

A new method to study HÜCKEL parameters is suggested, wherein the essential point is the number of solutions of a system of nonlinear equations. This system of equations is derived from a simple model effective HAMILTONIAN and is discussed for the case of $2p_\sigma$ and $2p_\pi$ SLATER type orbitals.

1. Introduction

Interpreting semiempirical theories one can construct effective HAMIL-
TONIANs and deduce properties of semiempirical parameters[1].
Here we want to describe a less involved method of studying HÜCKEL
parameters, which is entirely different from ab-initio oriented efforts.

To describe molecular properties resulting from a quantum mechanical
formalism one may assign fixed numerical values to the HÜCKEL para-
meters "y_i". On the other hand in general one has a system of non-
linear equations with the HÜCKEL parameters given in a functional
dependence of parameters "z_j":

$$y_1 = \varphi_1(z_1, z_2, z_3, ..)$$
$$y_2 = \varphi_2(z_1, z_2, z_3, ..) \qquad\qquad (1)$$
$$\cdots\cdots\cdots$$

The parameters z_j (for example: "effective nuclear charge", bond distance", etc) have some "a priori character", appearing in molecular electronic SCHRÖDINGER equations of different sophisticated levels (cf sec. 3). Collecting the parameters z_j to a tuple z and resolving the system (1) for given numerical values of y_i we are interested in whether we get one or more tuples z. We think that relating the HÜCKEL parameters to parameters of a SCHRÖDINGER equation by a one to one correspondece might be a qualitative foundation of HÜCKEL theory or HÜCKEL type theories. We do not want to treat explicitly interactions or "effects" which are not explicitly considered in HMO theory, such as electron interaction and correlation effects. Knowing that in HMO theory such interactions are built in appropriately[1], after discussing equations (1) we are looking for procedures by which one can extrapolate from approximate model equations given through "φ_i^M" to better equations, given through more exact φ_i, which also depends on correlation effects etc . This extrapolation is done by suggesting the inequality (13) (see below).

2. The method

We now redefine the following terms:

$z: = (z_1, z_2, ..),$ a tuple of "a priori parameters" which can be found in SCHRÖDINGER equations. They may be identified with (effective) nuclear charge and geometric informations like bond distances.

$y: = (y_1, y_2, ..),$ a tuple of HÜCKEL parameters, including heteroatomic parameters and in general consisting of different COULOMB and resonance integrals. The overlap integral, not being a typical HÜCKEL para-

meter is not accounted for in y.

$\varphi: = (\varphi_1, \varphi_2, ..)$, a tuple of functions, which are occuring in system (1). φ will be understood as a mapping

$$\varphi: \quad z \longmapsto \quad y \tag{2}$$

Within the concept of mapping the HÜCKEL parameters are not necessarily fixed numbers, but functions of z, which simplifies some arguments. Beyond this, (2) be a unique relation. This is not only because of mathematical reasons, but expresses also the physical premise, that for example we restrict ourselves in describing HÜCKEL parameters as fitted exclusively to ground state one electron properties. In this context FREED[1] has shown that in general one should expect an implicit dependency of semiempirical parameters on energy, causing at least partially the ambiguity of semiempirical parameters. R. McWEENY[2] too, stated somewhat more explicitly that the HMO parameters are related to more than one expression of a self consistent generalization of HÜCKEL theory. Therefore it is not trivial to require the relation φ being unique. φ being unique justifies to denote this relation as a mapping.

The mapping φ can be thought of as constructed or at least identified by

$$\alpha_i = \int \eta_i \, \hat{H} \, \eta_i \, d\tau$$

$$\beta_{ij} = \int \eta_i \, \hat{H} \, \eta_j \, d\tau \tag{3}$$

$$y = (..., \alpha_i, ..., \beta_{ij}, ...) = \varphi(z)$$

with η_i being normalized and usually orthogonalized atomic orbitals, $\hat{H}$ being an operator acting only on a single particle according to the HMO concept[3]. This means that the derivation of the equation system is restricted to the "literal" or "theoretical" part of the HMO-parameters[1]. The influence of the correlation part and other neglected interactions are taken into regard by the inequality (13).

In our opinion a good semiempirical theory should have at least the following properties:

a) After fixing the values of all y_i of the tuple y (for example by fitting to experimental data) there must be such tuples z that the fixed tuple y, say y^*, is an image of at least one preimage tuple z, say z^* through the mapping φ.

The above statement, holding at best for all tuples y that are reasonable for quantum chemical purposes, means that φ is a surjective mapping (here: a surjective mapping with respect to the set of y-tuples "D")[4]. Now let be $L(\varphi,R,y)$ the number of solutions of (1) for a given tuple y, with given φ and with a set R of the z tuples, then we require

$$L(\varphi,R,y) \geqslant 1 \quad \text{for all } y \in D \tag{4}$$

b) To a given fixed tuple y there should exist only <u>one</u> solution z of (1), or in other words: the resolution of the system (1) should not lead to a multivalued result. This condition originates from an entirely different point of view than that taken by FREED and others[1], that is we simulate practical HMO work (which requires surjectivity), and then look for an unique relation between $y \in D$ and $z \in R$. This means: Turning from HMO parameters to the a priori parameters by means of an <u>injectivity</u> (one to one correspondence) yields a well defined theoretical background.

Equivalent to both criteria a) and b) is the following one: The mapping φ must be <u>bijective</u> (injective and surjective), or expressed operationally: φ must be <u>invertible</u>, i. e.

$$L(\varphi, R,y) = 1 \quad \text{for all } y \in D \tag{5}$$

Inclusion of the interactions relevant for the correct description of HMO-parameters would give a "best" mapping "φ^B", which to handle, however, would be troublesome. The two properties a) and b) in mind as criteria one is only interested in $L(\varphi,R,y)$. Not knowing the explicit

analytical form of φ^B which also includes correlation effects we discuss a simpler model mapping $\varphi^M = (\varphi_1{}^M, \varphi_2{}^M, \ldots)$, evaluate $L(\varphi^M, R, y)$ associated with system (1) and try to relate $L(\varphi^M, R, y)$ to $L(\varphi^B, R, y)$. The restriction to the number of solutions rather than the exact numerical construction of the solutions (which could only be done by rather involved ab-initio methods) should justify our approximations. The correlation parts etc are given reasonable care by the relation between $L(\varphi^M, R, y)$ and $L(\varphi^B, R, y)$.

3. The model to construct φ^M

We restrict ourselves to diatomic molecules, the simplest systems with two types of HMO parameters α_i and β. To get some feeling how criterion (5) works on the basis of the HMO definitions (3) in zeroth approximation we take a molecular electronic SCHRÖDINGER equation:

$$\hat{H}^M \psi = E \psi$$

with ψ being a linear combination of two atomic orbitals centered at different nuclei and the model HAMILTONIAN $\hat{H}^M$ given by

$$\hat{H}^M = -(1/2)\Delta - (n_1 z_1 / r_1) - (n_2 z_2 / r_2) \quad (a.u.) \tag{6}$$

This operator will be used in the RAYLEIGH SCHRÖDINGER quotient.

r_i: distance between the electron and the nucleus i

n_i: the effective principal quantum number of the orbital associated with atom i.

The z_i's have the following meaning:

z_1 and z_2: qualitative measure of the attraction between shielded nuclei and the electron in η_i.

z_3: distance between the two nuclei (z_3 is involved in the evaluation of two center integrals).

Apart from the nuclear repulsion term a HAMILTONIAN of type (6) has also been used by R.L.FLURRY[5] for an analysis of charge transfer

complexes. Normalized $2p_\sigma$ or $2p_\pi$ SLATER type orbitals are used, with z_1 and z_2 as orbital exponents, respectively. At this level of approximation we do not care for any orthogonalization problems. Certainly it would be very interesting to use LÖWDIN-orbitals[6] at least in the case of π-systems.

The HAMILTONIAN (6) could be thought of as part of a more sophisticated effective HAMILTONIAN. Thus the resulting mapping φ^M is expected not to be a too bad starting point. The HAMILTONIAN (6) inserted in (3) leads to matrix elements H_{ij} which are functions of the parameters z_j. Our model does not yield any energy dependency of y_i as functions of the "a priori parameters" z through φ^M. The mathematical requirement φ^M to be single valued is obviously given. To treat either the $2p_\sigma$- or $2p_\pi$-case underlying (6) in (3) we take $2p_\sigma$- or $2p_\pi$-STOs and do not differentiate any more. This means among other things that in both cases we use pure HMO-formalism.

With the HAMILTONIAN (6) and with specified η_i one obtains the following mapping:

$$
\begin{aligned}
y_1 &:= H_{11} = \int \eta_1 \hat{H}^M \eta_1 d\tau = \varphi_1^M(z_1, z_2, z_3) \\
y_2 &:= H_{12} = \int \eta_1 \hat{H}^M \eta_2 d\tau = \varphi_2^M(z_1, z_2, z_3) \qquad (7)\\
y_3 &:= H_{22} = \int \eta_2 \hat{H}^M \eta_2 d\tau = \varphi_3^M(z_1, z_2, z_3) \\
\Longrightarrow y &= \varphi^M(z)
\end{aligned}
$$

As a domain of $\varphi^M = (\varphi_1^M, \varphi_2^M, \varphi_3^M)$ for doing calculations we used the following subset of R:

$$
N := \left\{ (z_1, z_2, z_3) : 0.5 \le z_j \le 5, \ (a.u.) \quad j = 1,2,3 \right\}
$$

Our mapping corresponds to a simple electrostatic model (apart from the kinetic energy term $-(1/2)\Delta$) in which a $2p_\sigma$ or a $2p_\pi$ electron roughly

sees the shielded nuclei or the shielded nuclei together with the field of σ-electrons.

4. Results

In a previous paper[7] we employed the functional determinant

$$\det J\,(z)\;=\;\det\;\frac{\partial\,(y_1,y_2,y_3)}{\partial\,(z_1,z_2,z_3)} \qquad\qquad (8)$$

as a function of z to find out whether the mapping is <u>local</u> injective. Studying the behaviour of (8) leads to the following results:

a) For the $2p_\sigma$ case as well as the $2p_\pi$ case the functional determinant does not identically vanish in N, which means that the three functions φ_i^M are functionally not dependent[8] as necessary for practising HMO theory. Otherwise two given numbers for, say H_{11} and H_{12}, would patently imply a numerical value of the third parameter H_{22}.

b) For $2p_\pi$-STOs, $\det J \neq 0$ for all $z \in$ N. This means that $\varphi^M(2p_\pi)$ is local injective everywhere and hence locally possesses an inverse. This, however, does not necessarily involve the existence of a <u>global</u> inverse[4].

c) For $2p_\sigma$-STOs, one finds

$$Q:\,=\;\bigl\{(z_1,z_2,z_3)\quad:\quad \det J\,(z) = 0\bigr\},$$

which dissects the set N into two regions. In Q the mapping is singular. At least a <u>linearization</u> of $\varphi^M(2p_\sigma)$

$$y^{lin}\;=\;c\,+\,\bigl[J\,(z_o)\bigr]\,\cdot\,z$$

with a tuple c of constant numbers and the functional matrix $\bigl[J\,(z_o)\bigr]$ calculated in points z_o <u>of Q</u> has no inverse. We, however, are primarily interested in global inverse mappings.

Simplifying matters, we introduce the restriction

$$z_1\;=\;z_2\;=\;\tilde{z},\;\text{whence now}\,:\quad z\;=\;(\tilde{z},z_3)$$

and obtain the simpler mapping

$$
\begin{pmatrix} \tilde{\alpha} \\ \\ \tilde{\beta} \end{pmatrix} = \begin{pmatrix} \tilde{\varphi}_1{}^M(\tilde{z}, z_3) \\ \\ \tilde{\varphi}_2{}^M(\tilde{z}, z_3) \end{pmatrix} \tag{9}
$$

$$
\tilde{\varphi}^M = (\tilde{\varphi}_1{}^M, \tilde{\varphi}_2{}^M)
$$

with

$$
\tilde{N} = \left\{ (\tilde{z}, z_3) \; : \; 0.5 \leqslant \tilde{z} \leqslant 5, \; 0.5 \leqslant z_3 \leqslant 5 \quad (\text{a.u.}) \right\} \quad .
$$

From the behaviour of the functional determinant

$$
\det \tilde{J}(z) = \det \frac{\partial (\tilde{\alpha}, \tilde{\beta})}{\partial (\tilde{z}, z_3)} \tag{10}
$$

as a function of the tuple z and from a graphical analysis yielding the number of solutions of (9) for fixed values of $\tilde{\alpha}$ and $\tilde{\beta}$ we could deduce the following results:

a) For the $2p_\pi$ case in $\tilde{N}$ there exist no zeros of the function $\det \tilde{J}(z)$ (Eq. (10)). Furthermore one can show by means of the graphical analysis that in $\tilde{N}$ the mapping $\tilde{\varphi}^M$ is invertible. So one gets a one to one correspondence for the tuples $(\tilde{\alpha}, \tilde{\beta})$ and $(\tilde{z}, z_3)$ implied by the model mapping $\tilde{\varphi}^M$. According to (5) we have:

$$
L(\tilde{\varphi}^M, \tilde{N}, (\tilde{\alpha}, \tilde{\beta})) = 1 \quad \text{for } 2p_\pi \tag{11a}
$$

b) The zeros of the function $\det \tilde{J}(z)$ (Eq. (10)) in the case of $2p_\sigma$ point to a global ambiguity of the inverse relation $\left(\tilde{\varphi}^M \right)^{-1}$. In fact from given numerical values of the parameters $\tilde{\alpha}$ and $\tilde{\beta}$ one obtains <u>two</u> tuples of z and therefore $\tilde{\alpha}$ and $\tilde{\beta}$ do not well define the a priori parameters. Now we have:

$$
L(\tilde{\varphi}^M, \tilde{N}, (\tilde{\alpha}, \tilde{\beta})) = 2 \quad \text{for } 2p_\sigma \tag{11b}
$$

c) In both cases the subset of negative values of the $\tilde{\alpha}$ and $\tilde{\beta}$ is

restricted, so that there exist pairs $(\widetilde{\alpha}, \widetilde{\beta})$ which cannot be images of tuples z due to $\widetilde{\varphi}^M$. Therefore $\widetilde{\varphi}^M$ is not surjective with respect to the whole set of negative values of $\widetilde{\alpha}$ and $\widetilde{\beta}$. One of the boundaries is given by

$$\widetilde{\varphi}^M(\left\{ (\widetilde{z}, z_3) \; : \; z_3 = 0 \right\}) \quad \text{for } 2p_\pi \tag{12a}$$

and

$$\widetilde{\varphi}^M(\left\{ (\widetilde{z}, z_3) \; : \; \det \widetilde{J}(z) = 0 \right\}) \quad \text{for } 2p_\sigma \tag{12b}$$

Therefore the equalities (11a) and (11b) are valid only in an appropriate subset "$\widetilde{D}$" (not to confuse with D) of pairs $(\widetilde{\alpha}, \widetilde{\beta})$ within the above boundaries.

5. Discussion

Extending these results, which depend on the functional form of the special $\widetilde{\varphi}^M$, to the "best" mapping φ^B (cf sec. 2), we have to take into consideration that $L(\varphi^B, R, y)$ might exceed $L(\varphi^M, R, y)$ (note that now we use φ^M as general notation for model mappings) in the following sense: Let be $L^*(\varphi, R)$ the maximum of $L(\varphi, R, y)$ with respect to all y and with fixed φ and R:

$$L^*(\varphi, R) : = \max_y \left\{ L(\varphi, R, y) \right\}$$

Then one has to expect:

$$L^*(\varphi^B, R) \geqslant L^*(\varphi^M, R) \tag{13}$$

The inequality (13) qualitatively corresponds to further "informations"[9] inherent in better mappings resulting from additional interactions (correlation effects etc) and the use of more subtle descriptions.

Until now not being able to prove the inequality (13), we expect that the following statements hold, which at least should be plausible arguments for (13):

a) A transformation of a φ to another similar and better φ' does not

cancel an already given global bifurcation of solutions of system (1).
For example we presume that a further term in $\widetilde{\varphi}^M$ (Eq. (9)) for
$2p_\sigma$ leading to φ^B would not remove completely the second solution
in the whole set R.

b) Furthermore, if there exists a (global) homeomorphism $\mathcal{H}$

$$\mathcal{H} \ : \ \varphi^M(z) \longrightarrow \varphi^B(z) \qquad (14)$$

then (13) is correct and even an equality. This would be an ideal
extrapolation procedure because the determination of $L(\varphi^M, R, y)$ is
not so difficult as that of ab initio like models, or in other words:
If inclusion of additional effects or interactions in a mapping φ leads
to φ' such that we have an analogy to (14) then it is sufficient to
examine L with respect to the simpler φ.

For example, if the image $\varphi^B(z)$ of the "best" mapping φ^B can be
thought of as generated through a bijective linear transformation (e.g.
the scaling of the $\varphi_i^M(z)$) of $\varphi^M(z)$ then we get such a (very simple)
homeomorphism.

c) When applying the degree of mapping $d(\varphi, R, y)^{10 \ +)}$ with

$$Id(\varphi, R, y)I \ \leqslant \ L(\varphi, R, y) \quad \text{for "almost all" } y \in \widetilde{D} \qquad (15)$$

and with

$$d(\varphi, R, y) \ = \ \text{const of homotopic}^{+)} \text{ connected mappings } \varphi \qquad (16)$$

$^{+)}$ degree of mapping: $d(\varphi, R, y) \ = \ \Sigma \ \text{sign (det J (z))}$, the sum of all

z which are solutions of (1) and elements of R.

$^{+)}$ homotopy: Given two continuous mappings φ and φ' and

$(z,t): = (z_1, z_2, z_3, \ldots, t) \ 0 \leqslant t \leqslant 1$ we

have a homotopy through $H(z,t)$, if $H(z,t)$ is

simultaneous continuous in z and t and:

$H(z,0) \ = \ \varphi(z)$

$H(z,1) \ = \ \varphi'(z)$

the important property (16) yields stability at least of the <u>lower</u> bound of L due to homotopic variations of φ [10].

Of course the relation (13) does not follow mathematically from (16) in general, but for mappings of similar form $\overset{*}{L}$ should be rather insensitive.

Applications of the important concept of the degree of mapping will be given in greater detail in a following paper. For example it can be shown that respecting some (rather important) premises [10] the number of solutions should be $\geqslant$ 1 due to homotopic variations of $\varphi^M(2p_\pi)$ and $\geqslant$ 2 in case of $\varphi^M(2p_\sigma)$.

d) Finally according to L.C.ALLEN[9] in Eq. (7) crudely we simulate destabilizing interactions by simply adding the terms

$$z_1 z_2 K/z_3 \qquad \text{for } H_{ii}$$

$$z_1 z_2 S K/z_3 \qquad \text{for } H_{ij} \quad i \neq j \,, \tag{17}$$

K being a constant and S the overlap integral. Adding (17) to (7) and considering only the first two components we get a much more complicated situation, especially in the case of $2p_\sigma$ -STOs, with the graphical method being at the limit of its efficiency. The inequality (13), however, seems to be confirmed. To gain a deeper insight

---

For example, through

$$H(z,t) = t\varphi'(z) + (1-t)\varphi(z)$$

we have (assuming the simultaneous continuity) a homotopy. φ and φ' are connected by a homotopy, or φ' is generated by a homotopic variation of φ and conversely.

For the sake of distinctness in (16) we have not mentioned the restriction

$$y \notin H(\delta N, t) \text{ for all } t \in [0,1]$$

δN: the boundary of N or $\tilde{N}$ (Eq. (12)).

into this complicated case one could try to use modern mathematical concepts like "catastrophe theory" where problems arising from "fold- and cusp-points" etc can be handled systematically[11].

Given (13) we arrive at the following conclusions:

i) For $2p_\pi$ -orbitals the mapping $\widetilde{\varphi}^M$ is invertible, i.e. there exists a uniquely defined inverse relation

$$y = (\widetilde{\alpha}, \widetilde{\beta}) \longrightarrow z = (\widetilde{z}, z_3) \text{ for all } y \in \widetilde{D}$$

and therefore a semiempirical theory built up with $\widetilde{\varphi}^M$ corresponding to our premises is a good theory . If extended to the HMO theory (given by φ^B) this conclusion can be but is not necessarily valid because we have

$$L^*(\varphi^B, R) = \max_y \left\{ L(\varphi^B(2p_\pi), R, y) \right\} \geq L^*(\varphi^M, R) = 1$$

due to (13)

ii) For $2p_\sigma$ orbitals HMO theory respectively HMO type theory is not a good semiempirical theory. The application of $2p_\sigma$ HMO seems to be at least questionable and requires additional informations with respect to $\widetilde{N}$.

iii) When we use φ^M respectively $\widetilde{\varphi}^M$ the different behaviour of the two types $2p_\sigma$ and $2p_\pi$ is governed by the derivatives

$$\partial H_{12} / \partial z_3 \text{ respectively } \partial \widetilde{\beta} / \partial z_3$$

and can be correlated with the functional dependence of the overlap integral as a function of z_3 (nuclear distance).
Our conclusions i) and ii) are not altered by so typical empirical assumptions as

H_{ii} being a monotone function of the effective nuclear charge of atom i alone,

and

H_{ij} being proportional to the overlap integral or given by a

LINDERBERG type formula[12], approximately assuming

$$S(2p_\pi) : = \exp(-z_3(z_1 + z_2))$$

$$S(2p_\sigma) : = z_3^2 z_1 z_2 \cdot S(2p_\pi) .$$

($S(2p_\sigma)$ is not correct for small z_3 values.)

6. Summary

The presented method is based upon the assumption that for a qualitative explanation of semiempirical theories in terms of physical significant parameters one has to establish a well defined solution of system (1). If

$$L(\varphi, R, y) = 1 \quad \text{for all} \quad y \in D$$

holds, then one can relate the semiempirical parameters y to better defined ones z without being calculated actually. It is primarily important that $L(\varphi, R, y)$ is enough insensitive to φ for extrapolation procedures like homeomorphism and homotopy to work well.

Vice versa additional interactions leading to variations of L can thus be classified through L. Taking this point of view we find that in a simple electrostatic model without electron interaction, correlations and nuclear repulsion the HMO theory (given with $2p_\pi$) is a good theory apart from boundaries (12a) in contrary to the $2p_\sigma$ HMO type theory.

We hope that it might be possible in future to give restrictions such that $L(\varphi, R, y) = 1$ for $2p_\pi$ orbitals. This means that we must examine L relative to interactions (L dependent upon φ), relative to physically relevant tuples z (L dependent upon R) and finally relative to appropriate regions within the set of negative values of y (L dependent upon y). Up to now it seems that L is dependent only upon geometric informations, i.e. on the overlap distribution along the internuclear axis.

We expect that studying semiempirical theories (not only of quantum chemical regime) by using a criterion like

$$L(\varphi, R, y) = 1 \quad \text{for all} \quad y \in D$$

and applying mathematical methods letting L invariant might give further special insight into the semiempirical parameters.

References

[1] K.F. Freed: Mod. Theor. Chem. 7, 201 (1977) and literature
     cited therein

[2] R. McWeeny in: Molecular Orbitals in Chemistry, Physics and
     Biology, ed. P.O. Löwdin, B. Pullman (Acad. Press, London,
     New York 1964)

[3] e.g. A. Streitwieser, Molecular Orbital Theory for Organic Chemists
     (J. Wiley Sons, Inc., New York 1961)

[4] H. Grauert and W. Fischer, Differential- und Integralrechnung II
     (Springer Verlag Berlin etc. 1968)

[5] R.L. Flurry, J. Phys. Chem. 73, 2111 (1969)

[6] P.O. Löwdin, J. Chem. Phys. 19, 1570 (1951)

[7] R. Brüggemann and J. Voitländer, Zeitschr. f. Phys. Chem. N.F.
     90, 205 (1974)

[8] G. Doetsch, Math. Annalen 99, 590 (1927)

[9] L.C. Allen in: Sigma Molecular Orbital Theory, ed. O. Sinanoglu,
     K. Wiberg (Yale University Press, New Haven, London 1970)

[10] K. Deimling, Nichtlineare Gleichungen und Abbildungsgrade
     (Springer Verlag Berlin etc. 1974)

[11] Th. Bröcker, Differentiable Germs and Catastrophes (Cambridge
     University Press 1975)
     Y.C. Lu, Singularity Theory and an Introduction to Catastrophe
     Theory (Springer Verlag Berlin etc. 1976)

[12] J. Linderberg, Chem. Phys. Lett. 1, 39 (1967)
     D. Kracht and W.C. Nieuwpoort, Chem. Phys. Lett. 9, 48 (1971)