

Construction of Some Molecular Orbitals to Be Approximately Invariant for Changes from One Molecule to Another

S. F. Boys

Department of Theoretical Chemistry, University of Cambridge, Cambridge, England

THE outstanding feature of the chemistry of nonconjugated covalent compounds is the high insensitivity of the structure of particular chemical groups to changes in the structure of distant parts of the molecule. No quantitative mathematical procedure which corresponds closely to this has hitherto been formulated. It has, however, been the opinion of the writer that, when quantitative wave functions for molecules with a few atoms become available, it should be possible to calculate directly orbitals with the smallest possible variation for distant changes. Such functions might be described as the best chemically invariant orbitals. At the present, functions for small polyatomic molecules, such as those for H_2CO described in following papers, are becoming available, and shortly, even more complicated cases will be calculated. It is thought that if such chemically invariant orbitals are calculated by imposing a mathematical constraint of maximum insensitivity to alterations in the distant nuclear charges, then the result will be orbitals which are localized around the chemical valency links and the atomic lone pairs. The possibility of transforming the orbitals in a wave function to satisfy an additional imposed requirement is very well known and is a consequence of the antisymmetry of the function.

Very recently another interesting aspect has been found. Some other functions which will be defined in the paper by Foster and Boys¹ have been found to show something of the localization which had been expected for the chemically invariant orbitals. These functions which are being described as exclusive orbitals have been introduced for the purpose of systematizing the correlation corrections in a complicated calculation, but to do so they are defined in a way which causes high localization. How closely the two types approximate to each other is a matter for quantitative tests on a number of molecular calculations. Later it can be decided whether it is worthwhile to perform the more laborious work of chemically invariant orbitals or to use the simpler exclusive orbitals. It is, however, important in any case to specify as precisely as possible this desirable property of chemical invariance, and to give a method for its calculation. The following analysis has been developed to provide solutions of these problems.

¹J. M. Foster and S. F. Boys, *Revs. Modern Phys.* **32**, 300 (1960), this issue.

It is difficult to specify mathematically to which types of molecule the direct calculation of invariant orbitals could be applied. In principle it could apply only to those atomic associations in which the orbitals were found to be highly invariant in the calculation itself. In practice, it is very probable that these are just those systems which are generally described by simple covalent bonds. These need not be stable molecules but could be any atomic association in a distorted molecule or in the course of a reaction provided the simple valency description is valid.

The most direct way to specify a chemically invariant orbital φ_i appears to be to define it to be associated with a particular chemical link or atomic lone pair. Then a set of atoms $K(i)$ which are to be regarded as distant from it must be defined. In this simple calculation these would be taken to be all atoms other than the first neighbors, but in more refined investigations a larger number of more distant and possibly fictitious centers might be included. The φ_i could be defined to be those linear combinations of molecular orbitals which show the least changes when the charges on the centers $K(i)$ are altered by $\delta\epsilon$. Let $\chi_{iI}\delta\epsilon$ be the change in φ_i due to an increase of $\delta\epsilon$ on the (I)th nucleus. It is desired to choose all the φ_i so that the sum of the squares of all the χ_{iI} over all space is a minimum. It is considered to be most practical to postulate that the φ_i must be orthonormal, and hence this constraint must be considered in their calculation.

Before the formulation of the compact mathematical definitions and analyses, let us consider a possible application to the case H_2CO . It is considered that the most practical method is to exclude the core electrons from the invariance analysis, and that φ_1 can be defined to be the orbital with lowest potential energy in the field of the C nucleus and φ_2 to have the lowest energy in the field of the O nucleus. The valence link CH can be considered to specify φ_3 which should be adjusted to have the minimum variations for $\delta\epsilon$ on the O nucleus and on the other H nucleus. φ_4 should be specified similarly for the other CH link. Let φ_5 be one of the CO double bonds with its variation for $\delta\epsilon$ on the H nuclei to be minimized. The other double bond could be similarly specified. The particular specification of how the double bond is made up of two bonds is entirely arbitrary, but these could be identified either as the σ and π types or as an equivalent pair above and below the plane of the molecule. One of the lone pairs

on the O atom could be taken as φ_7 and the other as φ_8 with a similar arbitrary choice with respect to linear combinations of them. These should both be made as insensitive as possible to variations in the charge on the C and H atoms.

In a practical calculation it would be necessary to start with some set of orbitals for which the Slater determinant constructed from them would have the lowest possible energy for any such single determinant. Such a set of orbitals is not at all uniquely specified and could, in fact, be solutions of Roothaan's equations, but there are very many other sets of orbitals which also satisfy this minimum requirement and which it might be more convenient to obtain. To perform the calculation one could guess some linear combinations of these given orbitals in the hope that they would correspond roughly to the chemically invariant orbitals. In actual fact it would not matter what linear combinations or what starting orbitals were used: they would all converge to the same chemically invariant orbitals in a shorter or longer time. The quantity

$$J(\varphi_1, \varphi_2, \dots) = \sum_i \sum_I^{K(i)} (\chi_{iI} | \chi_{iI}) \quad (1)$$

would be evaluated, and then the φ_i varied systematically to minimize J . A systematic way of varying the φ_i while preserving the orthonormality is described in the following. It has been found quite effective in the case of a similar calculation for the exclusive orbitals described later.

The main hope would be that the final orbitals would show the least possible change for any alterations in the distant part of the molecule. This might be so to such a degree that the orbital for a CH group in HCHO would be very nearly the same as for the CH in C₂H₄ or in a more complicated molecule. It might be hoped that the O and CO orbitals would be very similar to those obtained if calculations were made for other aldehydes or ketones. If this were found to be true in sufficient cases, then we have a possibility of tabulating descriptions of the orbitals for particular chemical groups and of writing down approximate wave functions for further molecules without direct calculation.

Before the formal mathematical statement, it can be noted that additional centers of charge variation can be introduced without any alteration to the scheme. The values of I in the insensitivity sums could be taken over specified fictitious nuclear positions. All that is necessary is that the sets of such positions $K(i)$ for each φ_i are to be chosen in the regions of space to which the φ_i are desired to be particularly insensitive. For example, there might be a case for defining the CO link in HCHO insensitive to some positions where additional nuclei would be introduced in the homologous series derived from this compound. Such considerations would have to be judged on their merits, but we can

let $K(i)$ be defined to be the appropriate specified set in each case.

The general statement of the method can be written:

(1) Define the $\varphi_1, \varphi_2, \dots$ to be calculated for a molecule by first specifying the core electrons to have minimum potential energy around their particular nuclei; and then define the chemical link and lone pair φ_i by stipulating for each a set of positions $K(i)$ to which it is to be least sensitive for a charge change.

(2) Let $\chi_{iI}\delta\epsilon$ be the variation in φ_i due to placing a charge $\delta\epsilon$ at the point I . Then the following theorem gives a practical method of calculation. For brevity, the proof of this is given in Appendix A, and in Appendix B an allied theorem is established which enables the χ_{iI} during the iteration process to be evaluated fairly simply from the first set.

THEOREM 1

Let $\eta_r(x, y, z)$ for $r=1, 2, \dots, n$ be a set of space functions in terms of which all the orbitals φ_r are to be expressed as linear combinations. Let

$$\Phi = \mathcal{A} \varphi_1 \alpha \varphi_1 \beta \varphi_2 \alpha \varphi_2 \beta \dots \varphi_N \beta \quad (2)$$

be a Slater determinant of the orthonormal functions φ_i , where $i=1, \dots, N$, for which $(\Phi | H | \Phi)$ has the minimum value possible for a single determinant. Let Φ' be the corresponding determinant giving the lowest value of $(\Phi' | H + \epsilon V | \Phi')$ for the perturbation ϵV . Let

$$\Phi' = \mathcal{A} \varphi_1' \alpha \varphi_1' \beta \varphi_2' \alpha \varphi_2' \beta, \quad (3)$$

where

$$\varphi_i' = \varphi_i + \epsilon \varphi_i^1 + \epsilon^2 \varphi_i^2, \quad (4)$$

and where the φ_i satisfy the condition $(\varphi_i | \varphi_j) = 0$, for which it is always possible to find a solution. Let $\Phi(a, i)$ denote sum of the determinants in which each φ_i is replaced by φ_a , where φ_a is one of a set of $(n-N)$ linear combinations of η_r , which are orthonormal to the φ_i and to each other. Let $\Phi(a, i, b, j)$ denote the sum of all terms given by replacing φ_j in $\Phi(a, i)$ by φ_b . Then it follows that $\varphi_i^1 = \sum_{a=N+1}^n \varphi_a Y_{ia}$, where Y_{ia} are the solutions of

$$\sum_{jb} [(\Phi(a, i, b, j) | H - W | \Phi) + (\Phi(a, i) | H - W | \Phi(b, j))] Y_{jb} + (\Phi(a, i) | V | \Phi) = 0. \quad (5)$$

If V is taken in turn to be equal to the V_I , which are unit charges on all the centers, then

$$\chi_{iI} = \sum_a Y_{ia} \varphi_a \quad (6)$$

by definition. All the χ_{iI} can thus be evaluated.

(3) For any trial set of molecular chemical orbitals φ_i , it is now possible to evaluate the sum of the squares of all the orbital perturbations for the specified distant

centers. This is defined to be

$$J = \sum_i \sum_I^{K(i)} (\chi_{iI} | \chi_{iI}). \quad (7)$$

Suppose we examine a transformation of a single pair of orbitals

$$\varphi_i' = \varphi_i \cos\theta + \varphi_j \sin\theta, \quad \varphi_j' = -\varphi_i \sin\theta + \varphi_j \cos\theta. \quad (8)$$

This preserves the orthogonal requirements and we can, by systematic trial and error, obtain the value of θ which gives the minimum J . This involves calculating J for many values of θ . In principle, this can be done by the complete reevaluation of J for each φ_i, φ_j , but in practice there is the much shorter method shown in Appendix B. Then take the new values of φ_i, φ_j for the minimum J and proceed to make a similar minimization for every pair of φ 's in turn. Then repeat the whole process again and again until the largest adjustment in a whole i, j range is less than some limit such as 0.001. This is the kind of method which is easily carried out on an automatic computing machine. The final set of φ_i are the chemically invariant orbitals.

It is still a conjecture that when invariant orbitals for two molecules both containing some particular chemical group are found, then the orbitals for these groups are nearly the same. However, this appears to be probable and it is obviously very important to make all possible quantitative tests on this as soon as possible.

If this constancy is found in sufficient cases that we can assume its general occurrence, then we can use it to write down approximate simple determinant wave functions for further molecules, without direct calculations, by taking the appropriate orbitals from previous calculations. In practice, this could be applied to make a theoretical chemical synthesis of large molecules from calculations on small molecules.

This would provide only a source of approximate wave functions. But it appears probable to the writer that high-accuracy calculations in the future will consist of three stages such as the calculations of

(a) an approximate wave function of a few Slater determinants;

(b) correction functions to the orbitals by a linear expansion theory;

(c) correction functions for the omitted part of the electronic correlation by a linear expansion theory.

Considerable investigation may be required for (b) and (c), but these are very desirable for all calculations. For large molecules the synthesis by chemically invariant orbitals provides a method for (a). A reasonable long range estimate of future development might be that numerous chemically invariant orbitals would be tabulated for common radicals and in addition sufficient adjustment coefficients to enable those to be corrected to their new environments with respect to (b) and (c) by relatively short calculations.

It is a very interesting question whether the exclusive orbitals proposed in the following papers may provide a working substitute for chemically invariant orbitals. This appears to be merely a matter of sufficient quantitative tests. The exclusive orbitals are not calculated by a method which demands invariance from molecule to molecule. However, in view of our knowledge of the molecular characteristics of covalent systems, it appears as if they will be similarly localized. The first results on HCHO suggest this very strongly. If this were found to be generally true, it would not alter the conceptual picture of chemically invariant orbitals, but it would provide an economical substitute. The exclusive orbitals have been defined to provide a basis for a systematic method of dealing with correlation corrections, and if one set of orbitals can fulfil this requirement and that of chemical invariance, then both computation and the theoretical picture would be simplified. This issue must be left to be decided by future quantitative examination.

In conclusion, it may be noted that although this concept of insensitivity to distant change must have played some part in the formulations of early systems of localized orbitals such as the hybrid orbitals of Pauling, Slater, etc., or the equivalent orbitals of Lennard-Jones, Pople, and Hall, this was not introduced explicitly. Here it has been suggested that a more powerful approach may be achieved by stipulating this as the essential condition. It both makes a criterion suitable for automatic computation and provides a concept directly related to the comparison of one molecule with another.

APPENDIX A

Proof of Theorem I

First let us show that any perturbation solution of the form

$$\Phi' = A \bar{\varphi}_1 \alpha \bar{\varphi}_1 \beta \bar{\varphi}_2 \alpha \bar{\varphi}_2 \beta \cdots, \quad (9)$$

where

$$\bar{\varphi}_i = \varphi_i + \epsilon \sum_{j=1}^n Y_{ij} \varphi_j + \epsilon^2 q_i \quad (10)$$

and

$$(\bar{\varphi}_i | \bar{\varphi}_j) = \delta_{ij}$$

can be put in a form with $Y_{ij} = 0$ for $j \leq N$. The notation qt is used to denote higher terms which need not be specified.

We note that Φ is unchanged if φ_1 and φ_2 are replaced by

$$\begin{aligned} \varphi_1'' &= (\bar{\varphi}_1 - \epsilon Y_{12} \bar{\varphi}_2) (1 + \epsilon^2 Y_{12} Y_{12}^*)^{-\frac{1}{2}}, \\ \varphi_2'' &= (\bar{\varphi}_2 + \epsilon Y_{12} \bar{\varphi}_1) (1 + \epsilon^2 Y_{12} Y_{12}^*)^{-\frac{1}{2}}, \end{aligned} \quad (11)$$

since these can be verified to be orthonormal combinations of φ_1 and φ_2 . In this case it follows that

$$\begin{aligned} \varphi_1'' &= \varphi_1 + \epsilon \sum_{j=3}^n Y_{ij} \varphi_j + \epsilon^2 q t, \\ \varphi_2'' &= \varphi_2 + \epsilon \sum_{j=3}^n Y_{2j} \varphi_j + \epsilon^2 q t. \end{aligned} \quad (12)$$

The fact that $Y_{11}=Y_{22}=0$ follows, since otherwise there is a term of order ϵ in $(\varphi_1|\varphi_1)$ and $(\varphi_2|\varphi_2)$. The term $Y_{12}=0$ follows by the direct substitution; $Y_{21}=0$ follows since otherwise there would be a term of order ϵ in $(\varphi_1|\varphi_2)$. Hence we have removed the Y_{12} , Y_{21} terms and left the other Y_{ij} unchanged. Similarly, all other Y_{ij} for $i, j \leq N$ can be removed.

We have now proved the first assertion of the theorem, since the perturbed determinant of minimum energy can be written with orbitals

$$\varphi_i' = \varphi_i + \epsilon \varphi_i^1 + \epsilon^2 q t, \quad (13)$$

where

$$\varphi_i^1 = \sum_{j=N+1}^n Y_{ij} \varphi_j.$$

Consider the general condition that $(\Phi|H|\Phi)$, where Φ is a detor, has a minimum value. It follows from the minimum condition that $(\Phi|H-W|\Phi)$ has a stationary value for any variation which does not affect orbital orthonormality, and it follows from its form that this is stationary for any orbital variation which is a linear combination of the other orbitals. Hence $(\Phi|H-W|\Phi)$ is stationary for all orbital variations, and a trivial examination of the complex quantities concerned shows that

$$(\Phi(a, i)|H-W|\Phi) = 0. \quad (14)$$

If we consider each φ_i in this equation to be replaced by its expansion of the preceding form (13) and H to be replaced by $H + \epsilon V$ for a perturbed problem, then the main equation of the theorem follows by equating the terms of order ϵ to zero.

APPENDIX B

Short Method for the Evaluation of Successive χ_{iI}

It is assumed that the χ_{iI} have been first evaluated for all of a set φ_i ($i \leq N$) and for all of the possible I 's.

These are somewhat more in number than the χ_{iI} which were required for a single step in the preceding calculation, since there, only the I in $K(i)$ were required for each φ_i . It is shown in the following theorem that the χ_{iI} for a particular value of I transform by just the same matrix as the φ_i to which they correspond. Hence for any new set $\bar{\varphi}_j = \sum_{j'} S_{ij'} \varphi_{j'}$, we have $\bar{\chi}_{iI} = \sum_{j'} S_{ij'} \chi_{iI}$. This reduces the calculations after the first evaluation to a very simple process. It is sufficient to prove this for a single perturbation φ_i^1 , since the χ_{iI} correspond to this for the different charges.

THEOREM 2

Let

$$\bar{\varphi}_i = \varphi_i + \epsilon \varphi_i^1 + \epsilon^2 q t \quad (15)$$

be the perturbed set of orbitals with $(\varphi_i|\varphi_j^1) = 0$, which give a minimum of

$$W = (\mathcal{A} \bar{\varphi}_1 \alpha \bar{\varphi}_1 \beta \bar{\varphi}_2 \alpha \cdots | H + \epsilon V | \mathcal{A} \bar{\varphi}_1 \alpha \bar{\varphi}_1 \beta \bar{\varphi}_2 \alpha \cdots). \quad (16)$$

If S_{ij} is a unitary transformation, and $\mu_i = \sum_j S_{ij} \varphi_j$, then the perturbed form of μ_i with $(\mu_i|\mu_j^1) = 0$ is

$$\bar{\mu}_i = \mu_i + \epsilon \mu_i^1 + \cdots \quad (17)$$

where

$$\mu_i^1 = \sum_j S_{ij} \varphi_j^1. \quad (18)$$

Proof

Since W is a minimum for variations in the φ_i with respect to its adjustable parameters previously written as Y_{ia} , then

$$(\mathcal{A} \bar{\mu}_1 \alpha \bar{\mu}_1 \beta \cdots | H + \epsilon V | \mathcal{A} \bar{\mu}_1 \alpha \bar{\mu}_1 \beta \bar{\mu}_2 \alpha \cdots) \quad (19)$$

must also be a minimum with respect to these parameters, since it is the identical integral merely written with different internal combinations. It must also be a minimum with respect to the combinations $\bar{Y}_{ia} = \sum_j S_{ij} Y_{ja}$, and hence $\bar{\mu}_i$ satisfy just the minimization conditions of a perturbed solution with zero-order solutions μ_i .

In general perturbation theory, it follows that there can be only one solution for nondegenerate zero-order solutions, and hence the $\bar{\mu}_i$ is the only solution and the theorem is established.