

# ELECTRONIC STRUCTURE IN AROMATIC COMPOUNDS

ENRICO CLEMENTI

IBM Research Laboratory, San Jose, California 95114

Received January 26, 1968

## CONTENTS

|   |     |
|---|-----|
| I. Introduction.....  | 341 |
| II. Molecular Orbital Method and Its Direct Extensions.....     | 342 |
| A. Introduction.....  | 342 |
| B. Multi-Configuration Self-Consistent Field Theory.....        | 345 |
| C. Electron Population Analysis.....                            | 349 |
| D. Bond-Energy Analysis.....                                    | 351 |
| III. Additional Comments Concerning the Correlation Energy..... | 352 |
| IV. Contracted Basis Set of Gaussian Functions.....             | 358 |
| V. Pyrrole.....   | 361 |
| VI. Pyridine Ground State.....                                  | 363 |
| VII. Pyridine Ion, ${}^2A_2$ .....                              | 365 |
| VIII. Pyrazine Ground State.....                                | 371 |
| IX. Conclusions.....  | 373 |

## I. INTRODUCTION

Aromatic compounds have been the subject of quantum-mechanical studies for the last 40 years. A main emphasis in these studies was to provide classification of the electronic spectra. At the very beginning of the quantum theory of molecules we find Hückel's<sup>1</sup> application to the benzene ring, Mulliken's very similar treatment,<sup>2</sup> and later Lennard-Jones<sup>3</sup> and Coulson's work<sup>4</sup> and that of Wheland<sup>5</sup> and others.<sup>6</sup> The early treatments are characterized by complete neglect of the  $\sigma$  electrons and by neglect of explicit computation, accurate or approximated, of the electron-electron interaction. Mayer and Sklar<sup>7</sup> set about a much more refined work whereby the electron-electron interaction was explicitly computed for the  $\pi$  electrons. The work

of Mayer and Sklar was a significant improvement and had enormous impact on the field of quantum-mechanical computation of molecules. It was the first attempt to obtain quantitative information from quantitative computations of electron-electron integrals in a large molecule. The validity of the assumption of a rather rigid separation between the  $\sigma$  and the  $\pi$  electrons was discussed by Altmann,<sup>8</sup> but after that many workers essentially ignored any further study of the problem. A semiempirical version of Mayer and Sklar's work was proposed by Pariser and Parr<sup>9</sup> in what is now known as the Pariser and Parr approximation. The latter made feasible a large number of computations for molecules even more complicated than benzene. The main emphasis of the Pariser-Parr and Pople theory (or P-P-P theory as later it was baptized in view of parallel work by Pople<sup>10</sup>) was to elucidate the electronic spectra. Since the  $\sigma$  electrons were approximated and since part of the two-electron integrals for the  $\pi$  electrons were neglected or approximated, the P-P-P technique had to recur to semiempirical parameters obtained from spectra data.

The theory works formally within the one-electron approximation and had to explain spectral transition, a problem inherently connected with the correlation energy. At that time quantum chemists had not developed a quantitative appreciation for the magnitude of the correlation energy and, therefore, the problem of spectral transition prediction was remarkably difficult, since the same set of data was capable of fitting a few

(1) See, for example, J. D. Roberts, "Notes on Molecular Orbital Calculations," W. A. Benjamin, Inc., New York, N. Y., 1961; A. Streitwieser, Jr., "Molecular Orbital Theory for Organic Chemists," John Wiley and Sons, Inc., New York, N. Y., 1961. The original work was by E. Hückel, *Z. Physik*, **70**, 204 (1931); **72**, 310 (1931); **76**, 628 (1932).

(2) R. S. Mulliken, C. Rieke, and S. Brown, *J. Am. Chem. Soc.*, **63**, 41 (1941); R. S. Mulliken, *J. Chim. Phys.*, **46**, 497, 695 (1949).

(3) J. E. Lennard-Jones, *Proc. Roy. Soc. (London)*, **A158**, 280 (1937).

(4) (a) C. A. Coulson, *ibid.*, **A169**, 413 (1939); C. A. Coulson and H. C. Longuet-Higgins, *ibid.*, **A191**, 39(1947); **A193**, 16 (1947); **A193**, 447, 456 (1948); **A195**, 188 (1948); (b) H. C. Longuet-Higgins, "Advances in Chemical Physics," Vol. I, I. Prigogine, Ed., Interscience Publishers, New York, N. Y., 1958.

(5) G. W. Wheland, *J. Am. Chem. Soc.*, **63**, 2025 (1941).

(6) R. Daudel, chapter in ref 4b; E. Cartmell and G. W. A. Fowles, "Valency and Molecular Structure," Butterworth and Co., Ltd., London, 1956; C. A. Coulson, "Valence," Clarendon Press, Oxford, 1959; W. Kautzmann, "Quantum Chemistry," Academic Press Inc., New York, N. Y., 1957; J. A. A. Ketelaar, "Chemical Constitution," Elsevier Publishing Co., Amsterdam, 1958; R. G. Parr, "Quantum Theory of Molecular Electronic Structure," W. A. Benjamin, Inc., New York, N. Y., 1963; M. J. S. Dewar, *Rev. Mod. Phys.*, **35**, 586 (1963).

(7) M. G. Mayer and A. L. Sklar, *J. Chem. Phys.*, **6**, 645 (1938).

(8) S. L. Altmann, *Proc. Roy. Soc. (London)*, **A210**, 327, 343 (1951).

(9) R. Pariser and R. G. Parr, *J. Chem. Phys.*, **21**, 466 (1953); **23**, 711 (1955).

(10) J. A. Pople, *Proc. Phys. Soc. (London)*, **A68**, 81 (1954).

states well, but was much less satisfactory for other states within a given molecular system. In addition, the problem of transferring empirical parameters from molecule to molecule was considerably difficult. By its own nature such type of research can only be extremely time consuming since, at least in principle, each set of new empirical parameters had to be as consistent as possible with previous parameters. This careful analysis was speeded up by experimentation with computers, and a very large number of molecules were analyzed. To some degree this type of work is continued today. Personally, this work is a monument to the ingenuity of the users, who often were able to obtain spectral assignments with a method which inherently should be rather incapable of doing so. It is also worthwhile to note that more often than not the goal of the work was to predict new spectra or to provide theoretical help into the *interpretation* and *classification* of the spectra. The goal of *explaining* and of *understanding* was less prominent. A series of papers which summarize and systematize the  $\pi$ -electron approach was presented by Rudenberg.<sup>11</sup> Rudenberg, in collaboration with Platt,<sup>12</sup> attempted to use a very simple model for explaining aromatic spectra; the model is the one useful for free electrons in a given potential.

Much of this work, however, is of historical importance and represents the extent of quantum chemistry hope in the existence of a simple shortcut whereby simple calculations and far-reaching simplifying assumptions could lead to permanent or so solution of the very complicated theory of the electronic interaction in a molecular field. However, it might very well turn out that the simplest and quickest way to understand aromatic molecules lies in extending to molecules those techniques we know are successful in atoms. This approach is much less glamorous among colleagues in chemistry; it sounds hard and unyielding. Yet, either we seek a quantum-mechanical answer or not. If yes, then we had better use quantum mechanics, however expensive and tedious it might be. If not, *i.e.*, if we think we presently understand the basic theory of chemistry, then why bother at all with quantum chemistry. And if assumptions are to be made and if semiempirical theory is to be used, then their full implications should be understood. (Parenthetically, I would like to note that I have no doubt of the importance in developing new assumptions and new semiempirical attempts.)

In the following section the present status of molecular computations is summarized. No attempt is made to present a full survey; emphasis on only one ap-

proach is made, the one in which I think we can make significant progress.

## II. MOLECULAR ORBITAL METHOD AND ITS DIRECT EXTENSIONS

### A. INTRODUCTION

In the last decades two approaches have been prominent in the study of the electronic structure of molecules, namely the valence-bond approximation and the molecular orbital approximation. Other techniques have been proposed. Of special interest are those computational techniques, which can be applied to molecular systems with few electrons. For example, Kolos and Wolniewicz<sup>13</sup> have obtained very accurate wave functions for a number of electronic states in the  $H_2$  molecule at many internuclear distances. Kolos' computer program can definitely not only supplement but can compete with most sophisticated spectrographs. The resolution of his program is within fractions of a wavenumber, and the use of the program presents so little difficulty that it can be correctly handled by any "technician." The results by Kolos and Wolniewicz provide a concrete, although partial, example of what one wishes to obtain from theoretical chemistry. However, most molecules have more than two electrons and exact wave functions are presently not easily obtainable. Nevertheless, a number of important steps have been made and there is good reason for optimism in the future. Some of the steps which have been taken, their accomplishments, and limitations will be outlined.

The valence-bond approximation is progressively losing its impact, particularly in theoretical chemistry, and somewhat more slowly in chemistry in general. It is becoming more and more apparent that any "all-electron" treatment of a moderately complex molecule is unfeasible even with modern high-speed computers. In addition, a full valence-bond treatment which considers all the electrons of a molecule introduces unreasonable highly positive and negative ions of dubious physical meaning. (For example, a valence-bond treatment of benzene will use a large number of structures, including  $C^+$ ,  $C^{2+}$ ,  $C^{3+}$ ,  $C^-$ ,  $C^{2-}$ ,  $C^{3-}$ , etc.) Its appeal remains in its basic simplicity and in having brought about the concept of resonating structures which remains a basic concept in theoretical interpretations of chemistry.

The molecular orbital theory has the advantage of being conceptually based on atomic theory, with techniques that can be tested for atoms. An electronic theory of molecules should in principle and in practice be applicable to the limiting case of a single atom. Electrons do not change nature from atoms to molecules and the same should hold for any model which describes

(11) K. Rudenberg, *J. Chem. Phys.*, **34**, 391 (1961).

(12) J. R. Platt, *ibid.*, **17**, 484 (1949).

(13) S. Kolos and L. Wolniewicz, *ibid.*, **41**, 3663 (1964); **43**, 2429 (1965).

the electronic structure either of atoms or of molecules.

A pleasant characteristic of the molecular orbital theory is that each progressive improvement or step has a natural physical explanation. Rather arbitrarily we shall present the molecular orbital theory as a five-step evolution.

The first step is the LCAO-MO approximation. There are actually two approximations in the above step: the first is the MO approximation; the second is the LCAO approximation of an MO. As known, the short notation LCAO-MO stands for "linear combination of atomic orbitals-molecular orbitals."

The MO is a one-electron function which is factored into a spatial component and a spin component. The expression "one-electron function" means that only the coordinates of one electron are explicitly used in a given MO. This factorization into spatial and spin components is permissible since generally one uses a Hamiltonian which does not explicitly contain spin-dependent terms. The MO's are the exact analog of the atomic orbitals, which describe the electrons in an atom to a first approximation. Indeed, one can read several chapters of the classical work of Condon and Shortly,<sup>14</sup> replace the word "AO" with the word "MO," and read a book on molecular physics instead of atomic physics.

This situation has some important consequences; namely, a large amount of testing and development for molecular wave-function techniques can be done with atoms. For this reason atomic and molecular examples are freely mixed throughout this review.

If the molecule contains  $2n$  electrons (let us consider a closed-shell case for simplicity), the MO approximation will distribute the electrons in  $2n$  molecular orbitals  $\varphi_1, \varphi_2, \dots, \varphi_{2n}$ . Since there are two possible spin orientations ( $\alpha$  and  $\beta$  spins), a space distribution function has either spin  $\alpha$  or  $\beta$  and, therefore, the  $2n$  electron system is described by  $n$  space functions and  $2n$  spin orbitals. Thus  $\varphi_1$  and  $\varphi_2$  will have the same space distribution (will depend on the coordinates of one electron alone), but, in accordance with the Pauli exclusion principle, will have different spin functions. It is stressed that the one-electron model is justified only because it simplifies the treatment. Indeed, since the very beginning of quantum theory, Hylleraas introduced a wave function for the He atom in which one orbital is described in terms of the coordinates of *both* electrons.

The total wave function  $\Psi$  of the  $2n$  electron system is then

$$\Psi = \frac{1}{\sqrt{(2n)!}} \begin{vmatrix} \varphi_1(1) & \dots & \varphi_1(2n) \\ \varphi_2(1) & \dots & \varphi_2(2n) \\ \vdots & & \vdots \\ \varphi_{2n}(1) & \dots & \varphi_{2n}(2n) \end{vmatrix} \quad (1)$$

where the number in parentheses indicates a given electron. This determinant wave function guarantees that any interchange of two electrons ( $i$  and  $j$ ) brings about a sign change in the wave function. This is the Pauli principle constraint for fermions. The energy for such a system is given by the relation

$$E = \langle \Psi^* | H | \Psi \rangle$$

where the Hamiltonian  $H$  is

$$H = -\sum_{ia} \frac{1}{2} \Delta_i^2 - \sum_{ia} \frac{Z_a}{r_{ia}} + \sum_{ij} \frac{1}{r_{ij}} - \sum_{ab} \frac{Z_a Z_b}{R_{ab}} \quad (2)$$

The first term is the kinetic operator for the  $i$ th electron, the second term is the potential operator between the  $i$ th electron and the  $a$ th nucleus (with charge  $Z_a$ ), the third term is the electron-electron potential between the  $i$ th and the  $j$ th electrons, and finally the last term is the nucleus-nucleus potential with  $R_{ab}$  the distance between the  $a$ th and  $b$ th nucleus of respective charges  $Z_a$  and  $Z_b$ .

The first and second terms are subsequently referred to as the one-electron Hamiltonian and will be indicated as  $h_0$ . The total energy for such a determinant was given by J. C. Slater, and it is

$$E = 2 \sum_i h_i + \sum_{ij} (2J_{ij} - K_{ij}) + E_{NN} \quad (3)$$

where

$$h_i = \langle \varphi_i^* | h_0 | \varphi_i \rangle \quad (4)$$

$$J_{ij} = \langle \varphi_i(1)^* \varphi_j(2)^* | r_{12}^{-1} | \varphi_i(1) \varphi_j(2) \rangle \quad (5)$$

$$K_{ij} = \langle \varphi_i(1)^* \varphi_j(2)^* | r_{12}^{-1} | \varphi_i(2) \varphi_j(1) \rangle \quad (6)$$

$$E_{NN} = \sum_{ab} (Z_a Z_b / R_{ab}) \quad (7)$$

As known,  $J$  and  $K$  are usually referred to as Coulomb and exchange terms, respectively. Equation 7 can be rewritten as

$$E = \sum_i (2h_i + J_{ii}) + \sum_{i>j} (2J_{ij} - K_{ij}) + E_{NN} \quad (8)$$

$$\equiv \sum_i \tau_i + \sum_{i>j} \rho_{ij} + E_{NN} \quad (9)$$

where  $\tau_i = h_i + J_{ii}$  and  $\rho_{ij} = (2J_{ij} - K_{ij})$ .

What form should the MO have? Clearly, the molecular orbitals are subjected to symmetry constraints (as in the case of atomic orbitals) and any molecular orbital will transform as an irreducible representation of the molecular symmetry group. This statement, however, is not a sufficient one; indeed it tells us mainly how the molecular orbital should *not* be. In principle we could insist on the analogy between atomic one-electron functions and molecular one-electron functions and "tabulate" the MO in a way analo-

(14) E. V. Condon and G. H. Shortley. "The Theory of Atomic Spectra." University Press, Cambridge, 1957.

gous to the method of Hartree and Fock in the 1930's. This would ensure that we have the best possible molecular orbitals. It is noted that numerical Hartree-Fock functions for diatomic molecules are a somewhat tempting possibility; this, however, has not seriously been explored at the present time.

Nevertheless, chemistry is concerned with more than only diatomic molecules. An answer is provided by the LCAO approximation, in which the MO's are built up as linear combinations of atomic functions. We refer to R. S. Mulliken's classical series of papers for the early development and application of the LCAO-MO approximation.

The *second step* in the evolution of quantum theory is the introduction of self-consistency. Again, the physical model is provided by atomic physics, namely by the Hartree-Fock model. The LCAO approximation to the MO requires the best possible linear combination: this is what one intends for self-consistency. A good review paper on this subject is the one by Roothaan.<sup>15</sup> There the self-consistent field technique in the LCAO-MO approximation (SCF-LCAO-MO) is systematically exposed for the closed-shell case.

Up to now we are strictly in the one-electron approximation. The electrons interact among themselves only *via* the average field and the MO has no explicit electron-electron parameters. Fortunately, the Pauli principle keeps electrons with parallel spin (in different MO's) away from each other, but it has nothing to offer to electrons with antiparallel spin in the same MO. The full catastrophe might be appreciated by recalling that, in the SCF-LCAO-MO approximation, two fluorine atoms are incapable of giving molecular bonding when brought together; *i.e.*, the SCF-LCAO-MO does not recognize the existence of the F<sub>2</sub> molecule.<sup>16</sup> Of course, it does not require a computation of F<sub>2</sub> to realize this point. For example, when the Roothaan work appeared (1950), another less familiar paper was written by Fock<sup>17</sup> to a large degree solving the problem and introducing the concept of two-electron molecular functions or "geminals," as they are called today. At the same time Lennard-Jones and collaborators<sup>18</sup> put forward a classical series of papers in which part of the correlation problem was tentatively solved, but at the expense of drastic orthogonality restrictions. For a variety of reasons, neither of the two avenues was numerically explored and in the meantime a third possibility slowly emerged.

Hylleraas,<sup>19</sup> and later Boys,<sup>20</sup> proposed the possibility of using not only one determinant, but as many as

needed. This technique is known as the configuration interaction or superposition of configuration technique; since the first designation is more common, it shall be adopted hereafter abbreviated as CI.

Let us consider, for example, the case of the beryllium atom in its ground state. Considering atoms instead of molecules is appropriate, since we are really considering electrons. The electronic configuration is 1s<sup>2</sup>2s<sup>2</sup> and, therefore, the Hartree-Fock function  $\Psi_0$  is

$$\Psi_0 = \frac{1}{4!} \begin{vmatrix} 1s(1) & 1s(2) & 1s(3) & 1s(4) \\ \bar{1}s(1) & \bar{1}s(2) & \bar{1}s(3) & \bar{1}s(4) \\ 2s(1) & 2s(2) & 2s(3) & 2s(4) \\ \bar{2}s(1) & \bar{2}s(2) & \bar{2}s(3) & \bar{2}s(4) \end{vmatrix} \equiv \{1s(1)\bar{1}s(2)2s(3)\bar{2}s(4)\}$$

where the bar designates  $\beta$  spin.

Let us consider the following functions,  $\Psi_1 = \{1s(1)-\bar{1}s(2)2p(3)2s(4)\}$ ,  $\Psi_2 = \{2s(1)\bar{2}s(2)2p(3)\bar{2}p(4)\}$ , and  $\Psi_3 = \{2p(1)\bar{2}p(2)2p(3)\bar{2}p(4)\}$ , and, after having ensured that each  $\Psi$  has <sup>1</sup>S symmetry, let us build the following functions

$$\Psi = a_0\Psi_0 + a_1\Psi_1 + a_2\Psi_2 + \dots$$

By optimizing the orbitals in each function and by variationally selecting the CI coefficients  $a_0, a_1, a_2, \dots$ , we shall have a solution necessarily as good as or better than  $\Psi_0$ , and, if the series of the above equations is sufficiently long, we shall reach an exact solution. The only trouble is that the necessary series is *too* long. The slow convergence of the series is due to the fact that in most cases one insists on using a 1s orthogonal to the 2s and to the 3s, a 2p orthogonal to the 3p, etc., with the 2p, 3p, and 3s functions overlapping the 1s and 2s functions very little. If the added functions overlap very little, they will interact very little and correlate equivalently.

However, let us assume that when we construct  $\Psi_0$  we construct  $\Psi_1, \Psi_2$ , etc. *at the same time*, and we do not insist on the best possible  $\Psi_0$ , but on the best possible  $\Psi$ ; then the variational principle, used simultaneously on both the  $a$ 's (the CI coefficients) and the  $\varphi$ 's (the atomic orbitals), will ensure that the  $\Psi_i$  will overlap as much as possible. This is accomplished in the *multi-configuration SCF-LCAO-MO* technique (MC-SCF-LCAO-MO), *the third step*. Before entering into the details of the MC-SCF-LCAO-MO theory, let us briefly mention a fourth and fifth step.

A *fourth step* in the molecular orbital theory is the inclusion of relativistic effects. There is little work done in this area at present (and this is not only true for molecular functions, but for atomic functions as well). Recent advances in metalloorganic chemistry, with heavy metals as constituents, demand a relativistic interpretation of the electronic structure. Even in molecules containing low  $Z$  atoms, the importance

(15) C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).

(16) A. C. Wahl, *J. Chem. Phys.*, **41**, 2600 (1964).

(17) V. Fock, *Izv. Akad. Nauk SSSR, Ser. Fiz.*, **18**, 161 (1954).

(18) A. C. Hurley, J. E. Lennard-Jones, and J. A. Pople, *Proc. Roy. Soc. (London)*, **A220**, 446 (1953).

(19) E. Hylleraas, *Z. Physik.*, **54**, 347 (1929); **65**, 759 (1930).

(20) F. S. Boys and G. B. Cook, *Rev. Mod. Phys.*, **32**, 285 (1960).

of spin-orbit effects in transition intensities is demanding more studies and computations in this area. It is gratifying to note that a simple perturbation treatment on atoms (Hartmann and Clementi<sup>21</sup> and Clementi<sup>22</sup>) gives energies as good as a full relativistic Hartree-Fock treatment.<sup>23</sup>

Finally, the "electronic structure" of molecules should always be considered a limiting case of the vibronic structure of molecules. Real molecules vibrate (and rotate and translate, too) and, therefore, the question of how much we can rely on the Born-Oppenheimer approximation should not be ignored. This is a *fifth step* and shall be referred to the work of Kolos and Wolniewicz for more details.<sup>24</sup>

#### B. MULTI-CONFIGURATION SELF-CONSISTENT FIELD THEORY

The MC-SCF theory seems to have been first proposed by Frenkel<sup>25</sup> (1934), Hartree, Hartree, and Swirles<sup>26</sup> (1939), and Yutsis<sup>27</sup> (1952). Recently, it has been reanalyzed and applied by Yutsis, Vizbaraite, Strockite, and Bandzaitis<sup>28</sup> (1962), Veillard<sup>29</sup> (1966), Veillard and Clementi<sup>30</sup> (1966), Clementi<sup>31</sup> (1967), and Das and Wahl<sup>32</sup> (1966). We shall first consider the simpler case of two configurations and expand it later to many configurations.<sup>30</sup>

Let us consider a configuration of the type  $1s^2 2s^2 2p^n$  (called configuration A) and a configuration of the type  $1s^2 2s^0 2p^{n+2}$  (called configuration B). States of like symmetry from A and B will interact and the resultant function will be (we are interested in its lowest eigenvalue)

$$\Psi = A\Psi_A + B\Psi_B$$

with energy

$$E = \langle \Psi^* | H | \Psi \rangle = A^2 E_A + B^2 E_B + ABE_{AB}$$

The SCF theory can be used in solving first  $\Psi_A$  and then  $\Psi_B$ , and a secular equation can be solved for  $\Psi = A\Psi_A + B\Psi_B$ . This is standard configuration interaction. However, the problem can be solved in one step, *i.e.*, an optimal  $\Psi_A$  and  $\Psi_B$  can be found so that, when the

two interact, an optimal  $\Psi$  is given. In other words, for a given basis set in  $\Psi_A$  and  $\Psi_B$  an optimal two-determinant combination can be obtained. The standard SCF guarantees an optimal  $\Psi_A$  or an optimal  $\Psi_B$ ; the MC-SCF guarantees an optimal  $\Psi$ , *but not an optimal  $\Psi_A$  or  $\Psi_B$* . For the specific case in consideration the  $E_A$ ,  $E_B$ , and  $E_{AB}$  are standard energy expressions, namely

$$\begin{aligned} E_A &= \tau_{1s} + \tau_{2s} + 2\rho_{1s,2s} + f_1 2 \sum_m h_m + \\ &\quad f_1 \sum_{mn} (2a_1 J_{mn} - b_1 K_{mn}) + 2 \sum_m \rho_{1s,m} + \sum_m \rho_{2s,m} \\ E_B &= \tau_{1s} + f_2 \sum_m h_m + f_2 \sum_{mn} (2a_2 J_{mn} - b_2 K_{mn}) + \\ &\quad 2 \sum_m \rho_{1s,m} \\ E_{AB} &= c \sum_m K_{2s,m} \end{aligned}$$

where  $\tau$ ,  $\rho$ ,  $h$ ,  $J$ , and  $K$  have been previously defined (see eq 8 and 9),  $f_1$  and  $f_2$  are occupation numbers ( $f_1 = n/6$ ,  $f_2 = (n+2)/6$ ,  $n$  is the number of electrons in the 2p shell), and  $a_1$ ,  $b_1$ ,  $a_2$ ,  $b_2$ , and  $c$  are numerical constants which ensure proper bookkeeping in the energy expression (these are called vector coupling coefficients). The indices  $m$  and  $n$  refer to the 2p orbitals.

The SCF technique is then applied; namely, an infinitesimal variation is applied on each orbital in  $\Psi_A$  and  $\Psi_B$ , which brings about a variation  $\delta E$  in the energy. The optimal solutions are those for which  $\delta E = 0$ . The orbitals are constrained to be orthonormal, and the mixing coefficients  $A$  and  $B$  are subjected to the relation  $A^2 + B^2 = 1$ . The constraints are sufficient in number as to ensure that a unique solution for the problem is found. The MC-SCF technique in this respect parallels the traditional SCF technique.

Let us analyze the results for Be(<sup>1</sup>S), B(<sup>2</sup>P), and C(<sup>3</sup>P) with electronic configurations  $1s^2 2s^2$ ,  $1s^2 2s^2 2p$ , and  $1s^2 2s^2 2p^2$ . The MC-SCF functions are found to be

$$\begin{aligned} \text{Be } \Psi(^1\text{S}) &= (1s^2)[0.9484(2s)^2 - 0.317(2p)^2] \\ \text{B } \Psi(^2\text{P}) &= (1s^2)[0.9728(2s)^2 - 0.2316(2p)^3] \\ \text{C } \Psi(^3\text{P}) &= (1s^2)[0.9888(2s)^2 - 0.1490(2p)^4] \end{aligned}$$

The experimental energies are  $-14.6685$ ,  $-24.6580$ , and  $-37.8557$  au, respectively. The relativistic corrections are computed to be  $-0.0022$ ,  $-0.0061$ , and  $-0.0138$  au. The single configuration energies (Hartree-Fock) are  $-14.5730$ ,  $-24.5290$ , and  $-37.6886$  au, respectively. When this energy is mass corrected, we have  $-14.5721$ ,  $-24.5278$ , and  $-37.6869$  au, respectively.

The correlation energy is defined as the difference between the experimental energy and the sum of the Hartree-Fock and relativistic energies. In the following equations we shall give the correlation energy ( $E_C$ ) for the three atoms in consideration.

- (21) H. Hartmann and E. Clementi, *Phys. Rev.*, **133**, A1295 (1964).  
 (22) E. Clementi, *J. Mol. Spectry.*, **12**, 18 (1964).  
 (23) Y.-K. Kim, *Phys. Rev.*, **154**, 17 (1967).  
 (24) W. Kolos and L. Wolniewicz, *J. Chem. Phys.*, **41**, 3674 (1964).  
 (25) J. Frenkel, "Wave Mechanics, Advanced General Theory," Clarendon Press, Oxford, 1934.  
 (26) D. R. Hartree, W. Hartree, and B. Swirles, *Phil. Trans. Roy. Soc. (London)*, **A238**, 223 (1939).  
 (27) A. P. Yutsis, *Zh. Eksperim. i. Teoret. Fiz.*, **23**, 129 (1952); **24**, 425 (1954); A. P. Yutsis, *Soviet Phys.-JETP*, **2**, 481 (1956); see, in addition, T. L. Gilbert, *J. Chem. Phys.*, **43**, S248 (1965).  
 (28) A. P. Yutsis, Ya. I. Vizbaraite, T. D. Strockite, and A. A. Bandzaitis, *Opt. Spectry.* (USSR), **12**, 83 (1962).  
 (29) A. Veillard, *Theoret. Chim. Acta*, **4**, 22 (1966).  
 (30) A. Veillard and E. Clementi, *ibid.*, **7**, 133 (1967).  
 (31) E. Clementi, *J. Chem. Phys.*, **46**, 3842 (1967); IBM Technical Report RJ 413, Feb 1967.  
 (32) G. Das and A. C. Wahl, *J. Chem. Phys.*, **44**, 87 (1966).

|                     |                    |                      |                  |                            |
|---------------------|--------------------|----------------------|------------------|----------------------------|
|                     | $E_{\text{exptl}}$ | $E_{\text{HF}}$      | $E_{\text{rel}}$ | $E_{\text{C}}$             |
| Be( <sup>1</sup> S) | -14.6685           | -(-14.5721 - 0.0022) | =                | -0.00942 au =<br>-2.563 eV |
| B( <sup>2</sup> P)  | -24.6580           | -(-24.5278 - 0.0061) | =                | -0.1241 au =<br>-3.378 eV  |
| C( <sup>3</sup> P)  | -37.8557           | -(-37.6869 - 0.0138) | =                | -0.1550 au =<br>-4.217 eV  |

Introduction of the second configuration lowers the correlation energy error by 0.0424, 0.0311, and 0.0173 au, respectively. The remaining error is partly due to the 1s<sup>2</sup> electrons, and these are about 0.0443, 0.0447, and 0.0451 au, respectively (these values are taken from the two-electron isoelectronic series). In Table I we give

TABLE I  
ENERGY CONTRIBUTIONS TO THE TOTAL ENERGY (%)

|                     | HF      | HF + R  | MC-SCF +<br>R | MC-SCF +<br>R + E <sub>C</sub> (1s) | Remainder |
|---------------------|---------|---------|---------------|-------------------------------------|-----------|
| Be( <sup>1</sup> S) | 99.3428 | 99.3578 | 99.6474       | 99.9495                             | 0.0505    |
| B( <sup>2</sup> P)  | 99.4720 | 99.4967 | 99.6228       | 99.8041                             | 0.1959    |
| C( <sup>3</sup> P)  | 99.5541 | 99.5906 | 99.6362       | 99.7554                             | 0.2446    |

the per cent error of the Hartree-Fock energy, the Hartree-Fock plus relativistic correction (*R*), the two-configuration SCF calculation plus relativistic correction, and the two-configuration calculation plus relativistic correction and the 1s<sup>2</sup> correlation energy contribution. The remaining error is 0.0074 au for Be(<sup>1</sup>S), 0.0483 au for B(<sup>2</sup>P), and 0.1126 au for C(<sup>3</sup>P). It is noted that there are two electrons with parallel spin in addition to the 1s<sup>2</sup>2s<sup>2</sup> electrons in C(<sup>3</sup>P), one unpaired electron in B(<sup>2</sup>P). The following algebra is quite tempting: 2 × (0.0483) + 0.0074 = 0.1040 au to be compared with 0.1126 au above reported. More accurately we should not use the value of 0.0074 which was obtained for the beryllium atom, but rather the value 0.0093 which is derived for C<sup>2+</sup>(<sup>1</sup>S). Therefore, we have not explained the correlation effect by the amount of 0.1126 - 0.1059 = 0.0067 au. This error is a cumulation of small contributions like the neglect of the relativistic energy difference between the 1s<sup>2</sup>(a2s<sup>2</sup>2p<sup>2</sup> + b2p<sup>4</sup>) and the 1s<sup>2</sup>2s<sup>2</sup>2p<sup>2</sup> configurations, oversimplifications in the estimate of the p-p correlation correction, the use of the correlation energy for the 1s<sup>2</sup>(C<sup>4+</sup>) in C(<sup>3</sup>P), and other small errors.

How many configurations should be added in order to obtain an accurate Be(<sup>1</sup>S) ground-state energy is not certain, without a numerical check. One could expect, however, that two configurations are sufficient to give 90% of the 1s<sup>2</sup> correlation (one for radial correlation, ns<sup>2</sup>2s<sup>2</sup>, and one for angular correction, np<sup>2</sup>2s<sup>2</sup>). If we are correct, the configurations 1s<sup>2</sup>2s<sup>2</sup>, ns<sup>2</sup>2s<sup>2</sup>, np<sup>2</sup>2s<sup>2</sup>, and 1s<sup>2</sup>2p<sup>2</sup> should improve the Hartree-Fock energy by about 0.08 or 0.085 (to be compared with 0.0942 au, the total correlation correction). A standard configuration interaction treatment would require over 20 configurations to reach this energy.

We shall now extend the MC-SCF theory to the case of *n* configurations for a closed-shell ground state. We assume that the 2*n* electrons of a given closed-shell system are distributed in *n* doubly occupied orbitals  $\varphi_1 \dots \varphi_n$ , and we shall refer to this set as the "(*n*)" set. A second set of orbitals  $\varphi_{(n+1)} \dots \varphi_{\omega}$  is used, and this will be referred to as the "( $\omega - n$ )" set. We consider all the possible excitations from the (*n*) set to the ( $\omega - n$ ) set; i.e., we consider *n*( $\omega - n$ ) configurations. A given excitation from the (*n*) set to the ( $\omega - n$ ) set will be indicated as  $t \rightarrow u$ , where *t* is a number from 1 to *n* and *u* is a number from *n* + 1 to  $\omega$ .

We shall designate as the complete multiconfiguration-self-consistent field (CMC-SCF) technique the one where a given orbital of the (*n*) set is excited to all orbitals of the ( $\omega - n$ ) set; if an orbital of the (*n*) set is excited to one or more, but not all orbitals of the ( $\omega - n$ ) set, then we shall describe the technique as *incomplete MC-SCF* (IMC-SCF).

In the following, the MCM formalism is described following the analysis of Veillard and Clementi.<sup>30</sup> A program for atoms and molecules of general geometry is in the coding process.

The wave function of the system is

$$\Psi = a_{\infty}\psi_{\infty} + \sum_{t=1}^n \sum_{u=1}^{\omega-n} a_{tu}\psi_{tu} \quad (10)$$

If one wishes to exclude a number of occupied orbitals from the excitations, i.e., if a number of orbitals are left uncorrelated, then this requires simply starting the summation over the index *t* at some value of *t* larger than 1. We shall use *t* or *t'* as indices for the (*n*) set and *u* or *u'* as indices for the ( $\omega - n$ ) set. The energy corresponding to  $\Psi$  is

$$E = \langle \psi_{\infty}^* | H | \psi_{\infty} \rangle a_{\infty}^2 + \sum_{t=1}^n \sum_{u=1}^{\omega-n} a_{tu}^2 \langle \psi_{tu}^* | H | \psi_{tu} \rangle + \sum_{t=1}^n \sum_{u=1}^{\omega-n} a_{t'u'} \sum_{t'=1}^n \sum_{u'=1}^{\omega-n} a_{t'u'} \langle \psi_{t'u'}^* | H | \psi_{t'u'} \rangle + 2a_{\infty} \sum_{t=1}^n \sum_{u=1}^{\omega-n} a_{tu} \langle \psi_{\infty}^* | H | \psi_{tu} \rangle \quad (11)$$

$$E = a_{\infty}^2 E_{\infty} + \sum_t \sum_u a_{tu}^2 E_{tu} + \sum_t \sum_u a_{tu} \sum_{t'} \sum_{u'} a_{t'u'} I_{t't'} (1 - \delta_{t't'}) + \sum_t \sum_u a_{tu} \sum_{u'} \sum_{u''} a_{t'u''} I_{u'u''} (1 - \delta_{u'u''}) + 2a_{\infty} \sum_t \sum_u a_{tu} I_{tu} \quad (12)$$

where

$$E_{\infty} = \sum_{t=1}^n 2h_t + \sum_{t=1}^n \sum_{t'=1}^n P_{t't'} \quad (13)$$

$$E_{tu} = E_{\infty} - 2h_t + 2h_u - \sum_{t'=1}^n 2P_{t't'} + \sum_{t'=1}^n 2P_{t't'} - 2P_{tu} + P_{uu} + P_{tt} \quad (14)$$

$$I_{ij} = K_{ij} \quad (15)$$

$$P_{ij} = J_{ij} - \frac{1}{2} K_{ij} \quad (16)$$

By simple algebraic manipulations the energy expression can be rewritten as

$$E = 2 \sum_{i=1}^n h_i + \sum_{i'=1}^n P_{ii'} - A_i h_i + 2 \sum_{i'=1}^n P_{ii'} - P_{ii} + \sum_{u=1}^{\omega-n} 2B_u h_u + P_{uu} + 2 \sum_{i=1}^n P_{iu} + 2 \sum_{i=1}^n \sum_{u=1}^{\omega-n} a_{iu} (a_{oo} K_{iu} - 2a_{iu} P_{iu}) + \sum_{i=1}^n \sum_{i'=1}^n A_{ii'} K_{ii'} (1 - \delta_{ii'}) + \sum_{u=1}^{\omega-n} \sum_{u'=1}^{\omega-n} B_{uu'} K_{uu'} (1 - \delta_{uu'}) \quad (17)$$

where

$$1 = a_{oo}^2 + \sum_{i=1}^n \sum_{u=1}^{\omega-n} a_{iu}^2 \quad (18)$$

$$A_{ii'} = \sum_{u=1}^{\omega-n} a_{iu} a_{i'u} \quad (19a)$$

$$A_i \equiv A_{ii} = \sum_{u=1}^{\omega-n} a_{iu}^2 \quad (19b)$$

$$B_{uu'} = \sum_{i=1}^n a_{iu} a_{iu'} \quad (20a)$$

$$B_u \equiv B_{uu} = \sum_{i=1}^n a_{iu}^2 \quad (20b)$$

The coefficients  $a_{iu}$ ,  $A_u$ , and  $B_u$  are related by the following equations

$$1 = a_{oo}^2 + \sum_{i=1}^n A_i = a_{oo}^2 + \sum_{u=1}^{\omega-n} B_u \quad (21)$$

or

$$\sum_{i=1}^n A_i = \sum_{u=1}^{\omega-n} B_u \quad (22)$$

The coefficient  $A_i$  represents the "fraction of an electron" which is excited from the  $\varphi_i$  orbital of the  $(n)$  set to the  $\varphi_u$  orbitals of the entire  $(\omega - n)$  set. The coefficient  $B_u$  represents the "fraction of an electron" in the  $\varphi_u$  orbital of the  $(\omega - n)$  set as a result of the excitation from the entire  $(n)$  set. It is therefore tempting to reexamine the configuration structure of a  $2n$  electron system. The standard electronic configuration for the  $2n$  electrons is a product of  $n$  orbitals. For example,  $\varphi_{oo}$  has the configuration

$$\varphi_1^2 \varphi_2^2 \cdots \varphi_n^2$$

Let us call such a configuration a "zero-order electronic configuration." The MC-SCF-LCAO-MO function will be a set of  $(\omega n - n^2)$  zero-order configurations with appropriate coefficients,  $a_{iu}$ . It is rather difficult to visualize in a simple way the effect of such a rather long

expansion. However, we can make use of the  $A_i$  and  $B_u$  coefficients and write the following configuration

$$\varphi_1^{2(1-A_1)} \varphi_2^{2(1-A_2)} \cdots \varphi_n^{2(1-A_n)}$$

$(n)$  set

$$\varphi_{n+1}^{2B_1} \varphi_{n+2}^{2B_2} \cdots \varphi_{\omega}^{2B(\omega-n)}$$

$(\omega - n)$  set

which we shall refer to as the "complete electronic configuration." The set of  $(n)$  orbitals has a fractional occupation equal to  $(1 - A_i)$  for the orbital  $\varphi_i$ , whereas the remaining orbitals [the  $\varphi_u$ 's of the  $(\omega - n)$  set] will have, in general, relatively small fractional occupation values,  $B_u$ . Clearly, the sum of the fractions of electrons annihilated from the  $(n)$  set is equal to the sum of the fractions created in the  $(\omega - n)$  set, since  $\sum_i A_i = \sum_u B_u$  (eq 22).

The energy  $E_{oo}$  defined in eq 13 is formally the SCF-MO closed-shell energy expansion; however, the  $\varphi_i$  in the CMC-SCF formalism are not equal to the  $\varphi_i$  of the Hartree-Fock formalism. If we indicate with  $E_{HF}$  the usual Hartree-Fock energy, we can state that  $E_{HF}$  is somewhat lower than  $E_{oo}$ , by an amount which is almost proportional to the correlation error of  $E_{HF}$ , as can be seen by analysis of Clementi and Veillard's ICM results for first-row atoms. We now define a quantity  $E_c = E - E_{oo}$  which is larger than the correlation energy by the amount the  $E_{oo}$  is larger than  $E_{HF}$ . It is noted that the correlation energy is commonly defined as  $E - E_{HF}$ . Therefore the CMC-SCF formalism differs from most many-body techniques presented to date insofar as we do not assume the Hartree-Fock energy to be the zero-order energy.

We shall briefly analyze the energy expression 14 in terms of  $E$ ,  $E_c$ , and  $E_{oo}$ . For this purpose we introduce the following definitions.

$$E_c(t) = -2h_t + 2 \sum_{i'=1}^n 2P_{ti'} - 2P_{tt} \quad (23)$$

$$E_c(u) = 2h_u + 2P_{uu} + 2 \sum_i 2P_{iu} \quad (24)$$

$$E_c(tu) = 2a_{oo} K_{tu} - 4a_{tu} P_{tu} \quad (25)$$

$$E_c(tt1) = K_{ii'} (1 - \delta_{ii'}) \quad (26)$$

$$E_c(uu') = K_{uu'} (1 - \delta_{uu'}) \quad (27)$$

We can now write

$$E = E_{oo} + \sum_t A_t E_c(t) = \sum_u B_u E_c(u) + \sum_{tu} a_{iu} E_c(tu) + \sum_{ii'} A_{ii'} E_c(tt') + \sum_{uu'} B_{uu'} E_c(uu') \quad (28)$$

The first term is the contribution to  $E$  given by the one-electron model. The second term is a correction to  $E_{oo}$  obtained by annihilation of electrons in the  $(n)$  set. The third term is the energy of the electrons created in the  $(\omega - n)$  set. The fourth term is interaction of

created and annihilated electrons. The fifth term is the interaction energy resulting from any pair of electrons in a  $\varphi_i$  orbital interacting with any pair of electrons in a  $\varphi_{i'}$  orbital. Therefore, it is the pair-pair interaction in the  $(n)$  set. The last term is the pair-pair interaction in the  $(\omega - n)$  set.

Inspection of the energy expression (eq 28) reveals the reason for the often-found poor agreement between computed orbital energies  $\epsilon_i$  and ionization potentials or excitation potentials in the standard SCF computations, where  $E_\infty = \sum_i(\epsilon_i + h_i)$ . As known, one reason is that the orbitals in the excited configuration or in the ionized molecule often differ sufficiently from the ground-state orbitals even in the SCF-LCAO-MO approximation. The second reason is clearly obvious by inspection of eq 28, namely that the numerical values of the  $A_i$  and  $B_u$  coefficients will, in general, vary from the ground state to the excited states of a neutral molecule or from the ground state of the neutral molecule to the ground state of the ionized molecule.

It is tempting to consider the possibility of a semi-empirical scheme whereby the correct ionization potential or the correct excitation energies are obtained by empirically determining the  $A_i$  and  $B_u$  fractional occupation values. It is noted that the justification of the use of empirical parameters in the Pariser-Parr technique lies exactly in the fact that the one-electron approximation assumes  $A_i = B_u = 0$ , whereas in an exact theory  $A_i$  and  $B_u$  are different from zero.

Let us now continue with the development of complete MC-SCF-LCAO-MO theory. We wish to obtain the best  $\varphi_i$ 's and  $\varphi_u$ 's, making use of the variational principle, *i.e.*, by requiring that  $(\partial E/\partial \varphi_i) = 0$  and  $(\partial E/\partial \varphi_u) = 0$ . In addition we have to satisfy the equations  $(\partial E/\partial a_{\infty}) = 0$  and  $(\partial E/\partial a_{iu}) = 0$  in order to obtain the best multiconfiguration expansion. We shall make use of the lagrangian multiplier technique for determining  $\varphi_i$  and  $\varphi_u$ , and of the solution of the secular equation for determining the  $a_{iu}$  coefficients.

Let us define the following operators

$$F_i = (1 - A_i)h + 2(1 - A_i - A_{i'})P_i + 2A_iP_i + 2\sum_u B_u P_u + \sum_u (a_{\infty} a_{iu} K_u - 2a_{iu}^2 P_u) + \sum_{u i'} A_{i' u} K_{i'} (1 - \delta_{i i'}) \quad (29a)$$

and

$$F_u = B_u(h + 2P_u + \sum_i 2P_i) + \sum_i (a_{iu} a_{\infty} K_i - 2a_{iu}^2 P_{iu}) + \sum_{u u'} B_{u u'} K_{u'} (1 - \delta_{u u'}) \quad (29b)$$

where  $P_{ij} = \langle \varphi_i^* | P_j | \varphi_i \rangle$  and  $K_{ij} = \langle \varphi_j | K_i | \varphi_j \rangle$ . Differentiation of  $E$  with respect to the variational parameters  $\varphi_i$ ,  $\varphi_u$ ,  $a_{\infty}$ ,  $A_u$ ,  $B_u$  brings about the following relation.

$$\begin{aligned} \delta E = & 2\langle \delta \varphi_i | F | \varphi_i \rangle + 2\langle \varphi_i | F_i | \delta \varphi_i \rangle + 2\langle \delta \varphi_u | F_u | \varphi_u \rangle + \\ & 2\langle \varphi_u | F_u | \delta \varphi_u \rangle + \sum_i \delta A_i (-2h_i - 4\sum_{i'} P_{i i'}) + \\ & \sum_u \delta B_u (2h_u + P_{uu} + \sum_i 4P_{iu}) + \\ & 2\delta a_{\infty} \sum_i \sum_u a_{iu} K_{iu} + 2\sum_i \sum_u \delta a_{iu} [2a_{\infty} K_{iu} - 8a_{iu} P_{iu} + \\ & 2\sum_{i'} a_{i' u} K_{i' u} (1 - \delta_{i i'}) + 2\sum_{u'} a_{u u'} K_{u u'} (1 - \delta_{u u'})] \quad (30) \end{aligned}$$

The variational principle is satisfied for  $\varphi_i$  and  $\varphi_u$  if  $(\partial E/\partial \varphi_i) = 0$  and  $(\partial E/\partial \varphi_u) = 0$ . However, the variation in the  $\varphi$ 's is constrained by imposition of the orthogonality relations

$$\langle \varphi_i | \varphi_j \rangle = \delta_{ij} \quad (31)$$

where the indices  $i$  and  $j$  run over the full  $(n)$  and  $(\omega - n)$  sets. By setting eq 30 to zero, then by differentiation of the above equation, and finally by joining the resulting equations, we obtain the relation which defines  $\varphi_i$  and  $\varphi_u$

$$|F_i - \sum_{i'}^{(i)} |\varphi_{i'}\rangle \langle \varphi_{i'}| F_i | \varphi_i\rangle - \sum_u |\varphi_u\rangle \langle \varphi_u| F_u | \varphi_i\rangle | \varphi_i\rangle = | \varphi_i \rangle_{ii} \quad (32a)$$

$$|F_u - \sum_{u'}^{(u)} |\varphi_{u'}\rangle \langle \varphi_{u'}| F_u | \varphi_u\rangle - \sum_i |\varphi_i\rangle \langle \varphi_i| F_i | \varphi_u\rangle | \varphi_u\rangle = | \varphi_u \rangle_{uu} \quad (32b)$$

which can be written as

$$|F_i - T_i - T_u | \varphi_u\rangle = | \varphi_i \rangle_{ii} \quad (33a)$$

$$|F_u + U_u - U_i | \varphi_u\rangle = | \varphi_u \rangle_{uu} \quad (33b)$$

where  $T_i$  and  $T_u$  are the second and third operators in eq 33a and  $U_u$  and  $U_i$  are the second and third operators in eq 33b.

In the past, use has been made of the "virtual orbitals" in the configuration interaction technique. It is noted in this regard that the  $\varphi_u$ 's of the  $(\omega - n)$  set are quite different from the virtual orbitals of a standard SCF-LCAO-MO computation. The reason is that virtual orbitals have very little physical meaning: they are obtained from diagonalization of the Fock equation and are orthogonal to the occupied orbitals, but the variational principle cannot act on them since they do not contribute to the total energy. In general the virtual orbitals have very little overlap with the occupied orbitals, and therefore are of little use in correlating the electrons of the occupied orbitals. A discussion on this point can be found in Yutsis' review paper<sup>27</sup> as well as in A. C. Weiss' work.

It is noted that a given  $\varphi_u$  will mainly be used to correlate one, or at most two or three  $\varphi_i$ 's, and therefore the remaining  $(n - 1)$  or  $(n - 2)$  or  $(n - 3)$   $\varphi_i$ 's which are promoted to that given  $\varphi_u$  will add little to the



correlation correction. However, by including the  $(n - 1)$  or  $(n - 2)$  or  $(n - 3)$  remaining set of  $\varphi_u$ 's, we will include part of the pair-pair correlation in the total energy at no extra cost. In addition, the inclusion of the additional excitation allows us to make use of the equality  $\sum B_u = \sum A_i$  with the simple physical meaning for each  $B_u$  and  $A_i$  as previously explained. Therefore, for a given  $A_i$  and a given  $B_u$  there are one or at most very few leading terms in the  $\sum_u a_{iu}^2$  or in the  $\sum_i a_{iu}^2$  summation, respectively. The IMC-SCF treatments consider only the leading terms in  $A_i$  or  $B_u$ , and this requires a more accurate optimization of the basis set for the  $\varphi_i$  and  $\varphi_u$  which is very time consuming in the computation.

Recently, the ICM technique was applied to the first-<sup>30</sup> and second-row<sup>33</sup> atoms for two configurations. (The theory was applied to open- and closed-shell atoms.) For closed shells the ICM-SCF-LCAO-MO theory was developed by Das and Wahl for the  $H_2$ ,  $Li_2$ , and  $F_2$  molecules. Since only one  $\varphi_i$  orbital was excited, owing to program limitations, the results are quite good for  $H_2$ , good for  $Li_2$  (from a molecular viewpoint), and rather poor for  $F_2$  (as expected).

It is finally noted that Nesbet<sup>34</sup> has applied to atoms the Bethe-Goldstone formalism which is to some extent analogous to the CMC-SCF-LCAO-MO formalism (but which does not fully employ self consistency, and therefore has to work with larger numbers of configurations).

It seems that most of the earlier literature on the subject was not noticed by those groups which were rather involved in machine computations. This is somewhat unfortunate, because the CMC-SCF technique does not require any large increase in computational effort. The likely reason for the retarded explosion of CMC-SCF computations is that an undue amount of expectation was placed on the Hartree-Fock technique, despite quite extended theoretical proof to the contrary available in the last 15-20 years. In this respect the work of Nesbet indicates full awareness of the problem.<sup>35</sup> The same can be stated for the work of Löwdin<sup>36</sup> where many of the pitfalls of the Hartree-Fock technique have been predicted, and where much of the CMC-SCF theory had been developed along different lines, in Löwdin's "natural orbitals."

### C. ELECTRON POPULATION ANALYSIS

The SCF wave functions can be analyzed indirectly via a study of the physical properties of the molecule under consideration (like moments, polarizabilities, vibrational analysis, etc.) or directly by what is known as "electron population analysis." In the following, we shall briefly expose the method of Mulliken,<sup>37</sup> here somewhat modified and extended.

From the previous exposition of the SCF approximation, a molecular orbital is written as

$$\varphi_{\lambda i} = \sum_p c_{\lambda i p} \chi_{\lambda p}$$

where  $\lambda$ ,  $i$ , and  $p$  are indices which refer to symmetry representation, a specific orbital, and a specific basis set, respectively. The basis set is in general a symmetry-adapted function (SAF); *i.e.*, it transforms as  $\lambda$ . In the LCAO approximation the  $\chi_{\lambda p}$  is a linear combination of functions, designated by  $\gamma_q$ , centered on the atoms. The linear combination coefficients of the SAF are determined on the basis of symmetry alone, and we can write

$$\varphi_{\lambda i} = \sum_p c_{\lambda i p} \sum_q d_{\lambda p q} \gamma_q \quad (34)$$

By combining the  $c$ 's and the  $d$ 's into a new coefficient  $w$ , we have

$$\varphi_{\lambda i} = \sum_{ms} w_{\lambda i m s} \gamma_{m s} \quad (35)$$

where for each  $\lambda$  and  $i$  the index  $m$  refers to a given atom and the index  $s$  refers to a given  $\gamma_q$  on the  $m$  atom.

For real functions, the electronic density of  $\varphi_{\lambda i}$  is

$$(\varphi_{\lambda i})^2 = \sum_{ms} \sum_{m's'} w_{\lambda i m s} w_{\lambda i m' s'} \langle \gamma_{m s} | \gamma_{m' s'} \rangle \quad (36)$$

This relation is the base of Mulliken's analysis. The above sum can be written as

$$(\varphi_{\lambda i})^2 = \sum_{mm'} \sum_{ss'} w_{\lambda i m s} w_{\lambda i m' s'} \langle \gamma_s | \gamma_{s'} \rangle_{mm'} = \sum_{mm'} \sum_{ss'} [mm' ss']_{\lambda i} \quad (37)$$

where for  $m = m'$  and  $s = s'$  the overlap  $\langle \gamma_s | \gamma_s \rangle$  is unity since the SAF as well as the  $\gamma$  are normalized. For each MO, a table such as Table II can be constructed. There is one such table for each MO,  $\varphi_{\lambda i}$ , and each table is by construction symmetrical. We shall call "quadrant" the matrix of numbers with a given  $m$  and  $m'$  and indicate this as  $\{mm'\}_{\lambda i}$ . The sum of its terms is indicated as  $S\{mm'\}_{\lambda i}$ . The diagonal elements of a quadrant are indicated as  $\{mm'\}_{\Delta \lambda i}$  and its sum as  $S\{mm'\}_{\Delta \lambda i}$ .

For  $m = m'$ , the quadrant  $\{mm\}_{\lambda i}$  contains quantities which are specific to the atomic set for the atom  $m$ ; for  $m \neq m'$  the quadrant  $\{mm'\}_{\lambda i}$  contains quantities

(33) A. Veillard and E. Clementi, IBM Technical Report RJ-447, June 1967.

(34) R. K. Nesbet, *Phys. Rev.*, **155**, 51, 56 (1967).

(35) For a comprehensive documentation, see R. K. Nesbet, "Quantum Theory of Atoms, Molecules and Solid State," P. O. Löwdin, Ed., Academic Press Inc., New York, N. Y., 1966, pp 157-165.

(36) See, for example, P.-O. Löwdin, Quantum Chemistry Group, Uppsala University, Uppsala, Sweden, Technical Note No. 2, 1957; Technical Note No. 48, 1960; Quantum Theory Project for Research in Atomic, Molecular and Solid State Chemistry and Physics, University of Florida, Gainesville, Fla., Preprint No. 53, 1964; Preprint No. 65, 1964.

(37) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833, 1841, 2338, 2343 (1955).

TABLE II  
GRAPHICAL REPRESENTATION OF THE POPULATION ANALYSIS

| $m'$ | $s'$    | $m = 1$      |              | ... | $m = 2$       |               | ... | ... |
|------|---------|--------------|--------------|-----|---------------|---------------|-----|-----|
|      |         | $s = 1$      | $s = 2$      |     | $s = (i + 1)$ | $s = (i + 2)$ |     |     |
| 1    | 1       | [1111]       | [1121]       | ... | [21(i + 1)1]  | [21(i + 2)1]  | ... | ... |
| 1    | 2       | [1112]       | [1122]       | ... | [21(i + 1)2]  | [21(i + 2)2]  | ... | ... |
| 1    | .       | .            | .            | ... | .             | .             | ... | ... |
| 2    | $i + 1$ | [121(i + 1)] | [122(i + 1)] | ... | [             | ]             | ... | ... |
| 2    | $i + 2$ | [121(i + 2)] | [122(i + 2)] | ... | [             | ]             | ... | ... |
| 2    | .       | .            | .            | ... | .             | .             | ... | ... |
| .    | .       | .            | .            | ... | .             | .             | ... | ... |
| .    | .       | .            | .            | ... | .             | .             | ... | ... |
| .    | .       | .            | .            | ... | .             | .             | ... | ... |
| .    | .       | .            | .            | ... | .             | .             | ... | ... |

which are specific to atomic sets of the atoms  $m$  and  $m'$ .

For the atom  $m$  the following definitions, borrowed from Mulliken, are given.

Net atomic population

$$P_m = \sum_{\lambda} \sum_i S\{mm\}_{\lambda i} \quad (38)$$

Overlap population with atom  $m'$

$$P_{mm'} = \sum_{\lambda} \sum_i S\{mm'\}_{\lambda i} \quad (39)$$

Gross atomic population

$$G_m = P_m + \sum_{m'} P_{mm'} \quad (40)$$

Let us now focus our attention on the quadrant  $\{mm\}$ . The atomic set ( $\gamma$ 's) of any such quadrant will be of s, p, d, etc. type; therefore, within the  $\{mm\}$  quadrant we can have subquadrants of the type  $\{m_s m_s\}$ ,  $\{m_s m_p\}$ ,  $\{m_p m_p\}$ , etc. designated in general as  $\{m_l m_{l'}\}$  where  $l$  and  $l'$  are the angular quantum numbers for the  $\gamma$ 's. In full analogy to the previous definitions for the quadrants  $\{mm'\}$  we can define  $S\{m_l m_{l'}\}_{\lambda i}$  for the subquadrants  $\{m_l m_{l'}\}$ . With this in mind we can introduce the following definitions.

Nonhybrid net atomic

$$P_{m^i} = \sum_{\lambda} \sum_i S\{m_i m_i\}_{\lambda i} \quad (41)$$

Hybrid net atomic

$$P_{m^{ii'}} = \sum_{\lambda} \sum_i S\{m_i m_{i'}\}_{\lambda i} \quad (42)$$

Nonhybrid overlap

$$P_{m^i m'^i} = \sum_{\lambda} \sum_i S\{m_i m_{i'}\}_{\lambda i} \quad (43)$$

Hybrid overlap

$$P_{m^i m'^{ii'}} = \sum_i S\{m_i m_{i'}\}_{\lambda i} \quad (44)$$

Nonhybrid gross atomic

$$G_{m^i} = P_{m^i} + \sum_{\lambda} \sum_i \sum_{m'} P_{m^i m'^i} \quad (45)$$

Hybrid gross atomic

$$G_{m^{ii'}} = P_{m^{ii'}} + \sum_{\lambda} \sum_i \sum_{m'} P_{m^{ii'} m'^{ii'}} \quad (46)$$

Hybridization is a very familiar concept in theoretical chemistry. Its meaning, however, is often used in an exceedingly restrictive sense, usually when we have more than one atom. However, hybridization is no more than polarization, and therefore we can talk of hybridization between two atoms or between two electrons on the same atom. As a consequence we have *internal hybridization* (within a given atom and due to the electrons of that atom) as well as external hybridization (within a given atom and due to a field originated outside the atom). External hybridization is the familiar one. An example of internal hybridization is the beryllium ground-state atom previously discussed. As we know the 2s orbital is strongly hybridized (internally) with the 2p orbitals. Therefore, the correlation problem in atoms can be viewed as a problem of describing in the best possible way the internal hybridization, and the correlation problem in molecules can be viewed as a problem of describing in the best possible way the internal polarization of the component atoms *plus* the external hybridization. Personally I would not mind avoiding the word "hybridization" in theoretical chemistry, since the term "polarization" seems to be more accurate. However, I shall continue to use it in deference to previous workers in the field.

It is stressed that the above definitions have meaning only for a given basis set. Therefore, they provide quantitative data of qualitative character. However, it is exactly this *type* of data which we like to analyze in order to obtain some correlation between molecular structures. An exact wave function for a molecule provides a tool for obtaining exact expectation values. These can be obtained, as an alternative, from experimental data. However, taken alone, neither an exact list of expectation values nor an accurate list of experimental data constitutes understanding of the electronic structure of molecules.

Let us consider the population analysis in the CMC-SCF-LCAO-MO formalism. Since the  $\varphi_i$ 's and the  $\varphi_u$ 's are orthogonal, we shall have that  $\langle \psi_{i'u} | \psi_{i'u} \rangle = \langle \psi_{i'u} | \psi_{i'u} \rangle = 0$ , and recalling eq 10 we shall have

$$\langle \Psi | \Psi \rangle = a_{00}^2 \langle \psi_{00} | \psi_{00} \rangle + \sum_i \sum_u a_{iu}^2 \langle \psi_{iu} | \psi_{iu} \rangle$$

For each of the determinants  $\psi_{00}$  or  $\psi_{iu}$ , the previous definitions for the electronic population are valid. Therefore, we shall have

Net atomic population

$$\bar{P}_m = a_{00}^2 P_m^{00} + \sum_i \sum_u a_{iu}^2 p_m^{iu} \quad (47)$$

Overlap population

$$\bar{P}_m = a_{00}^2 P_{mm}^{00} + \sum_i \sum_u a_{iu}^2 P_{mm}^{iu} \quad (48)$$

Gross atomic population

$$\bar{G}_m = \bar{P}_m + \sum_{m'} \bar{P}_{mm'} \quad (49)$$

and equivalent expressions for the hybrid and nonhybrid populations.

#### D. BOND-ENERGY ANALYSIS

It is customary in the literature to report wave functions, expectation values, orbital energies, and total energies. The molecular total energy is then compared with the total energy of the separated atoms, and deductions about bond energies are made.

In theoretical chemistry the orbital energies are used mainly in connection with the Koopman theorem<sup>38</sup> but cannot be identified with bond energies.

In this work we shall introduce the definitions of the "Bond-Energy Analysis."<sup>31</sup> We shall make use of a number of bond-energy classifications which are derived at first from the usual SCF-LCAO-MO energy expression for the total energy (see eq 13)

$$E = 2 \sum_i h_i + \sum_{ij} P_{ij} + E_{NN}$$

The above energy expression can be written as

$$E = 2 \sum_i \sum_A h_{iA} + \sum_{AB} h_{iAB} + \sum_{ABC} h_{iABC} + \sum_i \sum_j \sum_A P_{ijA} + \sum_{AB} P_{ijAB} + \sum_{ABC} P_{ijABC} + \sum_{ABCD} P_{ijABCD} + \sum_{AB} E_{NA.NB} \quad (50)$$

where  $A, B, C, D$  are indices running over the atoms and where  $h_{iA}, h_{iAB}, h_{iABC}$  are the one-, two-, and three-center components of  $h_i$ ,  $E_{NA.NB}$  is the two-center component of  $E_{NN}$ , and  $P_{ijA}, P_{ijAB}, P_{ijABC}, P_{ijABCD}$  are the one-, two-, three-, and four-center components of the electron-electron interaction energy.

We shall then denote as the zero-order energy diagram

$$E_0 = \sum_A E_A = 2 \sum_A \sum_i j_{iA} + 2 \sum_{ij} P_{ijA} \quad (51)$$

as the first-order energy diagram ( $A \neq B$ )

$$E_1 = E_{AB} = \sum_{AB} \sum_i h_{iAB} - E_{NA.NB} - P_{ijAB} \quad (52)$$

as the second-order diagram ( $A \neq B \neq C$ )

$$E_2 = E_{ABC} = \sum_{ABC} \sum_i h_{iABC} - \sum_{ij} P_{ijABC} \quad (53)$$

and finally as the third-order diagram ( $A \neq B \neq C \neq D$ )

$$E_3 = \sum_{ABCD} E_{ABCD} = \sum_{ABCD} \sum_{ij} P_{ijABCD} \quad (54)$$

The  $E_0$  should be compared with the sum of the energy of the separated atoms. The correlation correction within  $E_0$  can be estimated directly from atomic energy computation and can be taken as equal to the electronic population within "a given atom," not summing up the full contribution to a given  $m$  column, but only the first quadrant.

$E_1$ , the first-order diagram energy, is the first "bond" energy and links any two atoms in a molecule, two at a time. Clearly, the *classical chemical formulas are a representation of the first-order diagram*—an incomplete one, however, since in them only some of the nearest neighbor atoms are connected. So we usually do not write a bond between neighbor hydrogen atoms in benzene, although there is clearly an interaction between them. The first-order energy differs from the other bond energies in that it includes the nuclear-nuclear repulsion, clearly present only in  $E_1$ . The quantity is numerically quite significant; it is part of the potential energy. Therefore, in view of the virial theorem, one can expect that  $E_1$  will be the dominant part of the binding energy. A large number of other evidences, both theoretical and experimental, are known to support this point. As is the case for  $E_0$ ,  $E_1$  is composed of a number of terms which will satisfy molecular symmetry consideration. In other words, the equivalent atom in  $E_0$  will have equal energy, so the equivalent bonds in  $E_1$  will be equal in energy. The correlation energy correction associated with  $E_1$  for the atoms A and B will clearly depend on the electronic density between atoms A and B. But here we should be careful in the use of atomic correlation energy computations, since no bond analogy can be uniquely drawn between the electrons in an atom and those in a molecule.

$E_2$ , the second-order diagram energy, is the three-atom interaction less the direct atom-atom pair interaction. This term does not contain a nuclear-nuclear repulsion term, and the one-electron energy (kinetic and potential) is relatively small.

$E_3$ , the third-order diagram energy, is the only term which does not include one-electron terms; therefore, it is totally a potential-energy (Coulomb and exchange) term. This term includes, by definition, mainly long-range interaction.  $E_3$  as well as  $E_2$  are in effect neglected in many semiquantitative computations (like the Pariser-Par approximation).

The population analysis, as previously described, has an energy analog in  $E_0$  and  $E_1$ , not in  $E_2$  and  $E_3$ . Classi-

(38) See, for example, J. C. Lorquett, *Rev. Mod. Phys.*, **32**, 312 (1960).

cal chemistry formulas have a partial analog in  $E_0$  and  $E_1$ .

The present breakdown of the total energy should represent a natural frame for transferability of bond energy which has a large body of thermodynamic evidence within families of compounds. It provides a framework which will not change by introducing more and more configurations in the MC-SCF-LCAO-MO formalism. It seems to offer advantages to *ab initio* vibrational analysis and to offer easy interpretation of the over-all constancy of group frequency in different molecules. Finally, it maintains the basic ideas of the molecular orbital theory, but it associates it intimately with our basic intuitive approach in chemistry, *i.e.*, that there are atoms in molecules.

This breakdown does not follow traditional ideas on the number of bonds one would like to see associated with an atom. So we could have in a molecule of  $N$  atoms a hydrogen with  $(N - 1)$  first-order bonds. However, this number will be reduced (a) by consideration of the nearest neighbors and (b) by quantitative consideration of the energy associated with each bond.

It is noted that there are at most four atomic bonds in our analysis: this is an effect of having chosen a basis set centered at the atoms. Alternatively, one could use localized orbitals and this would alter the number of atoms involved in bonds.

The above discussion brings about the following conclusions: (1) a chemical bond is an arbitrary concept and can be defined in as many ways as one wishes; (2) of the many possible representations, some are more useful than others; and (3) the "chemical bond" concept is to a certain extent simply a bookkeeping device, though of great importance.

One might prefer to have an exact correspondence between electron-population analysis and bond-energy analysis. However, the starting point of the electron population is  $(\varphi_{\lambda i})^2$ , and this can lead only to a subdivision involving one and two centers, whereas the bond energy analysis, in view of our definitions, involves one, two, three, and four centers.

Let us briefly digress to CMC-SCF-LCAO-MO bond-energy diagrams and compare these with SCF-LCAO-MO bond-energy diagrams. Let us start with the first-order diagrams and the first-order energy  $E_0 = \sum_A E_A^0$ . The main correction to these diagrams can be obtained from atomic computations, more explicitly from associating to the  $E_A^0$  a correction proportional to the  $\{mm\}$  and the atomic correlation less degeneracy effects (see Veillard and Clementi<sup>30,33</sup> for numerical values). This is a very simple correction to introduce. Alternatively, the correction to  $E_A^0$  can be computed using the device of correcting the  $P_{ij}$  integrals with a pseudo-potential (for example, see Clementi's work<sup>39</sup>). This alternative is not too different from the previous one, since the pseudo-potential uses the overlap between

orbitals as a parameter. A formal equation for  $E_0$  can be most easily obtained from eq 17.

The first-order diagrams and their energy  $E_1 = \sum_{A>B} E_{AB}$  are the most important in general for molecular correlation corrections. The correction to these diagrams should be proportional for a given atom to  $\sum_{m' \neq m} \{mm'\}$ . Alternatively, one could use a pseudo-potential computed correction in a manner exactly equivalent to our work on atoms. Again the formal expression of  $E_1$  can be obtained from eq 17.

The second- and third-order diagrams cannot be corrected by the use of population analysis parameters. However, it can be done by the use of the pseudo-potential technique or by the CMC-SCF-LCAO-MO method. The total energy will, therefore, be subdivided as

$$E = \sum_A (E_A + \eta_A) + \sum_{AB} (E_{AB} + \eta_{AB}) + \sum_{ABC} (E_{ABC} + \eta_{ABC}) + \sum_{ABCD} (E_{ABCD} + \eta_{ABCD})$$

where the  $E$ 's are the energies computed in the SCF-LCAO-MO approximation and the  $\eta$ 's are the energies obtained either from the CMC-SCF-LCAO-MO theory or from empirical correction to the  $E$ , or a proper mixture of both.

One of the expectations in proposing the bond-energy diagrams partitioning of  $E$  is that, by a systematic comparison of a number of molecules, a simple correlation will emerge which will put "transferability of bond energies" on a sound basis. It is noted that in the early attempts at studying molecular kinetics, the approach of using Morse-type potentials between any pair of atoms was often adopted. The present analysis in some respects does exactly that, if one considers only zero- and first-order diagrams. In addition it does much more. Therefore it is expected that this type of analysis will be of help in the formulation of a theory for reaction mechanism.

Other problems, like vibrational analysis, the study of the barrier to internal rotation, and charge transfer can be explained quite naturally in this framework of analysis. We shall return later to these points.

### III. ADDITIONAL COMMENTS CONCERNING THE CORRELATION ENERGY

The CMC-SCF-LCAO-MO technique we have previously discussed has one rather strong drawback; namely, we have not allowed for single and higher than double excitations. This can be remedied rather simply, however, by noting that once we have the orbitals of the  $(\omega - n)$  set, we can construct with these single and higher than double excitations, which are not included in the CMC theory, and use the standard CI

(39) E. Clementi, "Tables of Atomic Functions," Special IBM Technical Report, IBM Research Laboratory, San Jose, Calif., 1965.

technique. Therefore, we envision a full solution of the problem via the CI-CMC-SCF-LCAO-MO technique, the main correlation correction being computed via the CMC technique, the remainder via the standard CI technique.

In this section we wish to add some more quantitative information to the numeric value of the atomic correlation energy. The reason is that by so doing we shall develop confidence in using simple SCF molecular wave functions and be able to estimate rather accurately the correlation corrections.

The correlation energy is commonly defined as the difference between the exact nonrelativistic energy and the Hartree-Fock energy.<sup>40</sup>

It is worthwhile to note that there are several Hartree-Fock schemes,<sup>41</sup> each leading to a somewhat different energy and, consequently, to different values of the correlation energy. For this reason we state from the beginning that in the following, when we refer to the Hartree-Fock energy, we refer to the best energy one can obtain by the analytical self-consistent field method as put forward by Roothaan.<sup>42</sup> The reason for this choice is simply that by now this method has been used to obtain many atomic functions and energies and a large number of molecular functions and energies.

From a conceptual point of view one might prefer to define the correlation energy as the difference between the exact nonrelativistic energy and the Hartree-Fock energy, since the Hartree-Fock method presents an unbalanced situation when we look at the way in which electrons with like spins and those with different spins are considered.<sup>43</sup> The Hartree-Fock method partially correlates electrons with the same spins. This correlation present in the Hartree-Fock method will be hereafter referred to as *precorrelation*, where we define the precorrelation energy as the difference between the Hartree-Fock energy and the Hartree energy. This energy difference is a correlation energy, but, in view of the accepted definition of correlation, we might say that it is a correlation energy *ante literam*.

We note that the emphasis on the nonrelativistic exact energy in the definition of the correlation energy has mainly a practical value. The relativistic energy itself can be partitioned into a correlated and an uncorrelated relativistic energy.

Since electrons with parallel spins are somewhat correlated in the Hartree-Fock method and since parallel spins occur to a varying extent in the low-energy states

of atoms, one can expect that the correlation energy in the ground states of neutral atoms is not a linear function of the number of electrons. The Hartree-Fock method uses antisymmetrized wave functions; this is done to satisfy the Pauli principle and brings about the exchange energy which is the origin of the precorrelation energy. Electrons with the *same* spin find themselves encircled by a Fermi hole which prevents electrons with the same spin from approaching each other.

We can expect a large correlation energy for pairs of electrons of the same shell (intrashell correlation), a smaller correlation energy between electrons of different shells (intershell correlation), and a quasi-constancy for the correlation of given types of pairs of electrons with opposite spin.

With the Fröman<sup>44</sup> and Linderberg-Shull<sup>45</sup> work in mind, one will predict that the correlation energy of the ground-state first-row atoms will behave as follows. There is a given correlation for the pair of electrons in the He atom. For the Be atom, the correlation is about twice that of helium. Lithium will have an intermediate correlation energy between He and Be. Since the extra electron (compared with He) is a 2s electron, which has a maximum radial probability far from the 1s electrons, its correlation with the 1s electrons is certainly small. In fact, from the Linderberg and Shull<sup>45</sup> values, we know it to be very small (the intershell correlation for 1s-2s is *much smaller* than for the 1s (or 2s) intrashell correlation). The correlation energy of B, C, and N in their ground states can be estimated by keeping in mind that the 2p electrons have all parallel spins and consequently the precorrelation existing in the Hartree-Fock energy will take care of most of the correlation for the 2p electrons. There will certainly be some intershell correlation of 1s-2p type and 2s-2p type. Since the 2s electrons are in the same spatial neighborhood as the 2p electrons, one is tempted to assume that (1s-2p) intershell correlation  $\ll$  (2s-2p) intershell correlation.

The correlation energy for O, F, and Ne should increase sharply. With those atoms we build one, two, and three pairs of unparallel spin electrons, respectively, in the same shell (the 2p shell). The sharp increase is due to the lack of precorrelation for those newly added electrons.

It is fairly simple to be more quantitative about all the above reasoning. Accurate Hartree-Fock energies are available.<sup>46</sup> The necessary relativistic energies were

(40) This follows P.-O. Löwdin's definition of correlation energy.

(41) See, for example, E. Clementi, *J. Chem. Phys.*, **38**, 2248 (1963).

(42) C. C. J. Roothaan, *Rev. Mod. Phys.*, **32**, 179 (1960); C. C. J. Roothaan and P. Bagus, "Methods in Computational Physics," Vol. 2, Academic Press Inc., New York, N. Y., 1963.

(43) See, for example, the review paper by P.-O. Löwdin, "Advances in Chemical Physics," Vol. II, I. Prigogine, Ed., Interscience Publishers, New York, N. Y., 1959.

(44) A. Fröman, *Phys. Rev.*, **112**, 870 (1958); *Rev. Mod. Phys.*, **32**, 317 (1960).

(45) J. Linderberg and H. Shull, *J. Mol. Spectry.*, **5**, 1 (1960).

(46) E. Clementi, *J. Chem. Phys.*, **40**, 1944 (1964); **38**, 996, 1001 (1963); **41**, 303 (1964); E. Clementi, C. C. J. Roothaan, and M. Yoshimine, *Phys. Rev.*, **127**, 1618 (1962); E. Clementi and A. D. McLean, *ibid.*, **133**, A419 (1964); E. Clementi, A. D. McLean, D. L. Raimondi, and M. Yoshimine, *ibid.*, **133**, A1274 (1964); E. Clementi, *ibid.*, **135**, A980 (1964).

TABLE IIIA  
 EXPERIMENTAL, HARTREE-FOCK, AND RELATIVISTIC DATA FOR FIRST AND SECOND ROW (IN ATOMIC UNITS)<sup>a</sup>

|    |    | Exptl    | Hartree-Fock <sup>b</sup> | HF. mass cor | $E$ (rel) <sup>c,d</sup> | Lamb cor <sup>e</sup> |
|----|----|----------|---------------------------|--------------|--------------------------|-----------------------|
| He | 1S | -2.9038  | -2.8616799                | -2.86129     | -0.00007                 | 0.000022              |
| Li | 2S | -7.4780  | -7.4327257                | -7.43214     | -0.00055                 | 0.000106              |
| Be | 1S | -14.6685 | -14.573020                | -14.57229    | -0.00220                 | 0.000323              |
| B  | 2P | -24.6579 | -24.529052                | -24.52782    | -0.00603                 | 0.000740              |
| C  | 3P | -37.8558 | -37.688611                | -37.68690    | -0.01381                 | 0.001439              |
| C  | 1D | -37.8093 | -37.631317                | -37.62961    | -0.01377                 | 0.001439              |
| C  | 1S | -37.7572 | -37.549535                | -37.54783    | -0.01379                 | 0.001439              |
| N  | 4S | -54.6122 | -54.400911                | -54.39879    | -0.02732                 | 0.002500              |
| N  | 2D | -54.5246 | -54.296152                | -54.29404    | -0.02736                 | 0.002500              |
| N  | 2P | -54.4808 | -54.228087                | -54.22598    | -0.02739                 | 0.002500              |
| O  | 3P | -75.1101 | -74.809369                | -74.80683    | -0.04940                 | 0.004000              |
| O  | 1D | -75.0378 | -74.729213                | -74.72667    | -0.04940                 | 0.004000              |
| O  | 1S | -74.9562 | -74.610955                | -74.60842    | -0.04935                 | 0.004000              |
| F  | 2P | -99.8053 | -99.40928                 | -99.40644    | -0.08289                 | 0.006015              |
| Ne | 1S | -129.056 | -128.54701                | -128.54355   | -0.13121                 | 0.008614              |
| Na | 2S | -162.441 | -161.85889                | -161.85506   | -0.20021                 | 0.011856              |
| Mg | 1S | -200.333 | -199.61458                | -199.61011   | -0.292505                | 0.015791              |
| Al | 2P | -242.752 | -241.87665                | -241.87177   | -0.42062                 | 0.020460              |
| Si | 3P | -289.927 | -288.85426                | -288.84867   | -0.58351                 | 0.025887              |
| Si | 1D | -289.898 | -288.81500                | -288.80941   | -0.58382                 | 0.025887              |
| Si | 1S | -289.857 | -288.75845                | -288.75286   | -0.58414                 | 0.025887              |
| P  | 4S | -342.025 | -340.71866                | -340.71268   | -0.79111                 | 0.032085              |
| P  | 2D | -341.973 | -340.64872                | -340.64274   | -0.79110                 | 0.032085              |
| P  | 2P | -341.940 | -340.60316                | -340.59718   | -0.79126                 | 0.032085              |
| S  | 3P | -399.144 | -397.50475                | -397.49801   | -1.05076                 | 0.039051              |
| S  | 1D | -399.102 | -397.45210                | -397.44536   | -1.05084                 | 0.039051              |
| S  | 1S | -399.043 | -397.37444                | -397.36770   | -1.05090                 | 0.039051              |
| Cl | 2P | -461.514 | -459.48187                | -459.47482   | -1.37168                 | 0.046765              |
| Ar | 1S | -529.303 | -526.81734                | -526.81017   | -1.76094                 | 0.055190              |

<sup>a</sup> The first and second columns are case identifications; the following columns are the computed Hartree-Fock energy, the same energy mass corrected, the computed relativistic energy (with the spin-spin and spin-orbit energy estimated), and the Lamb shift correction. <sup>b</sup> E. Clementi, "Tables of Atomic Functions," Supplement to *IBM J. Res. Develop.*, 9, 2 (1965). <sup>c</sup> H. Hartman and E. Clementi, *Phys. Rev.*, 133, A1295 (1964). <sup>d</sup> E. Clementi, *J. Mol. Spectry.*, 12, 18 (1964).

available from our work,<sup>21,22</sup> and the total energy can be obtained experimentally by adding the ionization potentials from Moore.<sup>47</sup> Then the correlation energy is simply the total energy minus the Hartree-Fock energy minus the relativistic energy.

With these data we obtain an accurate estimate of the correlation energy for 2 to 18 electrons in atoms.

In Table IIIA we report the first- and second-row atomic symbols and states (columns one and two), the total experimental energy (column three), the Hartree-Fock energy (column four), the Hartree-Fock energy, mass corrected, (column five), the relativistic energy (previously reported for the second period<sup>11</sup> and here given for both first and second period; column six), and the lowest order Lamb shift correction (column seven). All quantities are in atomic units.

In Table IIIB we report first- and second-row atomic symbols and states (columns one and two) and the correlation energy  $E_C(1)$  (column three) obtained by subtracting from the experimental energy the Hartree-Fock energy (mass corrected) and the relativistic energy. In column four we report the average value of the correlation energy for electron, *i.e.*,  $E_C(1)/Z$ . In column five we report  $E_C(1)$  minus the lowest order

Lamb shift. (It is noted that these quantities have a number of theoretical uncertainties, as explained elsewhere,<sup>21</sup> and we shall designate them as  $E_C(2)$ .) In column six we report  $E_C(2)/Z$ . It is noted that in the two-electron problem the electronic cloud is more contracted around the nucleus than in the neutral atoms. Therefore, the Lamb shift for the neutral atoms is expected to be smaller in absolute value than the Lamb shift for the two-electron isoelectronic series. As a consequence  $E_C(3)$  has only indicative value.

Finally, in column seven we report  $E_C(3)$ , the correlation energy  $E_C(1)$  minus the degeneracy effect. In column eight we report  $E_C(3)/Z$ .

The data of the first and second rows reveal a strong dependency of the correlation energy on the number of direct pairs in the system. Indeed we can "estimate" the correlation energy by simply adding a somewhat constant correlation energy value per pair. One should be careful in not identifying this constant amount with the standard designation of electron pair as given by the electronic configuration. For example, the two electrons of the ground state of beryllium atoms are commonly referred to as 2s electrons. From the work of Linderberg and Shull<sup>48</sup> we know that these are really part 2s and part 2p. Indeed most of the correlation associated with this pair is due to the 2p electrons in the more complete electronic configuration written as

(47) C. Moore, "Atomic Energy Levels," National Bureau of Standards, Circular 467, U. S. Government Printing Office, Washington 25, D. C., 1949.

TABLE IIIB

VARIOUS ESTIMATES OF CORRELATION ENERGY BASED ON TABLE IIIA AND ON THE TWO CONFIGURATION SELF-CONSISTENT FIELD COMPUTATIONS (IN ATOMIC UNITS) BY VEILLARD AND CLEMENTI

| Case |       | $E_C(1)^a$ | $E_C(1)/Z$ | $E_C(2)^b$ | $E_C(2)/Z$ | $E_C(3)^c$ | $E_C(3)/Z$ |
|------|-------|------------|------------|------------|------------|------------|------------|
| He   | $^1S$ | -0.0420    | -0.0210    | ...        | ...        | ...        | ...        |
| Li   | $^2S$ | -0.0454    | -0.0151    | -0.0453    | -0.0151    | ...        | ...        |
| Be   | $^1S$ | -0.0940    | -0.0235    | -0.0937    | -0.0234    | -0.0465    | -0.0116    |
| B    | $^2P$ | -0.1240    | -0.0248    | -0.1233    | -0.0247    | -0.0929    | -0.0186    |
| C    | $^3P$ | -0.1551    | -0.0258    | -0.1537    | -0.0256    | -0.1378    | -0.0230    |
| C    | $^1D$ | -0.1659    | -0.0276    | -0.1645    | -0.0274    | ...        | ...        |
| C    | $^1S$ | -0.1956    | -0.0326    | -0.1942    | -0.0324    | ...        | ...        |
| N    | $^4S$ | -0.1861    | -0.0266    | -0.1836    | -0.0262    | ...        | ...        |
| N    | $^2D$ | -0.2032    | -0.0290    | -0.2007    | -0.0287    | ...        | ...        |
| N    | $^2P$ | -0.2274    | -0.0325    | -0.2249    | -0.0321    | -0.1932    | -0.0276    |
| O    | $^3P$ | -0.2539    | -0.0317    | -0.2499    | -0.0312    | ...        | ...        |
| O    | $^1D$ | -0.2617    | -0.0327    | -0.2577    | -0.0322    | ...        | ...        |
| O    | $^1S$ | -0.2985    | -0.0373    | -0.2945    | -0.0368    | -0.2449    | -0.0306    |
| F    | $^2P$ | -0.3160    | -0.0351    | -0.3100    | -0.0344    | ...        | ...        |
| Ne   | $^1S$ | -0.381     | -0.0381    | -0.372     | -0.0372    | ...        | ...        |
| Na   | $^2S$ | -0.386     | -0.0351    | -0.374     | -0.034     | ...        | ...        |
| Mg   | $^1S$ | -0.428     | -0.0357    | -0.412     | -0.034     | -0.3984    | -0.0332    |
| Al   | $^2P$ | -0.459     | -0.0353    | -0.439     | -0.034     | -0.4412    | -0.0339    |
| Si   | $^3P$ | -0.494     | -0.0353    | -0.468     | -0.033     | -0.4837    | -0.0345    |
| Si   | $^1D$ | -0.505     | -0.0361    | -0.479     | -0.034     | ...        | ...        |
| Si   | $^1S$ | -0.520     | -0.0371    | -0.494     | -0.035     | ...        | ...        |
| P    | $^4S$ | -0.521     | -0.0347    | -0.489     | -0.033     | ...        | ...        |
| P    | $^2D$ | -0.539     | -0.0359    | -0.507     | -0.034     | ...        | ...        |
| P    | $^2P$ | -0.552     | -0.0368    | -0.520     | -0.035     | -0.5306    | -0.0354    |
| S    | $^3P$ | -0.595     | -0.0372    | -0.556     | -0.035     | -0.595     | -0.0372    |
| S    | $^1D$ | -0.606     | -0.0379    | -0.567     | -0.035     | ...        | ...        |
| S    | $^1S$ | -0.624     | -0.0390    | -0.585     | -0.037     | -0.5902    | -0.0369    |
| Cl   | $^2P$ | -0.667     | -0.0392    | -0.620     | -0.036     | ...        | ...        |
| Ar   | $^1S$ | -0.732     | -0.0407    | -0.677     | -0.038     | ...        | ...        |

<sup>a</sup> Correlation energy obtained by subtracting from the total experimental energy of the Hartree-Fock (mass corrected) and relativistic energy (excluding Lamb shifts). <sup>b</sup> Correlation energy obtained by subtracting from  $E_C(1)$  the estimate of the Lamb-shift correction computed for the two-electron isoelectronic series. <sup>c</sup> Correlation energy obtained by subtracting from  $E_C(1)$  the near-degeneracy effects as reported in this and a previous paper of this series.

$s(a2s + b2p)$ . This is known from the quoted work of Linderberg and Shull,<sup>45</sup> Watson,<sup>48</sup> Sinanoglu,<sup>49</sup> and our systematic study.<sup>30,33</sup>

We can comment on the excited-state correlation energies. The correlation energy for the multiplet components of a given term is approximately the same. The difference in correlation energy, for example, between  $B(^2P_{1/2})$  and  $B(^2P_{3/2})$  is very small and within the error of the estimate. For this reason no such data are reported.

For different states of the same electronic configuration, the correlation energy has the following characteristics. First, the lowest correlation energy is for the state of highest spin multiplicity. For example, in the  $^3P$ ,  $^1D$ , and  $^1S$  series beginning at C (or O),  $E_C(^3P) < E_C(^1D)$  and  $E_C(^3P) < E_C(^1S)$ . This is, as mentioned previously at length, a consequence of the spin pre-correlation in the Hartree-Fock method. Second, for states with the same spin multiplicity the correlation energy is smaller for the states of highest angular momentum. For example,  $E_C(^1D) < E_C(^1S)$  for the carbon and the oxygen series, and  $E_C(^2D) < E_C(^2P)$  for the nitrogen series. Since states with the same spin multiplicity but dif-

ferent angular momenta do not have the same correlation energy (for given  $Z$  and number of electrons), one concludes that in the Hartree-Fock method we have not only spin-related pre-correlation but also angular pre-correlation, the angular pre-correlation being in the sense that the higher the angular momentum (total angular momentum) the higher the angular pre-correlation. *This is quite interesting because it tells us that we cannot obtain excitation energies of the correct magnitude with the Hartree-Fock method even for states of the same multiplicity.* A simple explanation of the differences of the correlation energies between states of the same multiplicity but different total angular momentum is that the larger the angular momentum, the more "preferential" is the electron's motion about the nucleus.

Up to now the results we have obtained for the correlation energy in the first<sup>30,50</sup> and second<sup>33,51</sup> rows have indicated a remarkably simple picture, where one can simply divide the correlation energy into "strong" and "weak" pairs, the former for intrashell electron pairs, the latter for intershell electron pairs. In addition, for the first and second row, the "weak pairs" have much smaller correlation where the electrons in the pair have different principal quantum numbers; for example, the

(48) R. E. Watson, *Ann. Phys.*, **13**, 250 (1961).

(49) O. Sinanoglu, "Advances in Chemical Physics," Vol. 6, I. Prigogine, Ed., Interscience Publishers, New York, N. Y., 1963.

(50) E. Clementi, *J. Chem. Phys.*, **38**, 2248 (1963).

(51) E. Clementi, *ibid.*, **39**, 175 (1963).

1s-2s pair correlation is smaller than the 2s-2p correlation.

There seems to be evidence<sup>52</sup> that the above simple pairing model<sup>53</sup> is only the limit of a more complex situation, which we have called the complex pairing model. For this we mean that the simple division between "strong" and "weak" pairs is inadequate, that the "weak" interactions increase not only in number (simply because there are more electrons in the systems) but also in strength. In other words, the new situation is that the correlation energy to a first approximation is not the simple sum of the "strong" pairs correlation, but one should add to this the contribution of the intershell correlation; this contribution is comparable in value to that of the "strong" pairs. One reason for this behavior is that the  $n$ ,  $l$ ,  $m$ , and  $s$  quantum numbers ( $L$ - $S$  coupling) do not describe adequately the atomic system. If the atomic system under examination is partially described by  $j$ - $j$  coupling, then this prevents assigning strong and weak pairs to the valence electron configuration. This point can be simply stated in the following way. "If a system is, for example, not a pure singlet, but a mixture of singlet, triplet, and quintet states, then why consider the correlation as due entirely to its singlet component? Further, if the system does not possess a well-defined total orbital or spin angular momentum, what is the meaning of pairs based on the assumption of a well-defined total orbital or spin angular momentum?" A second reason is that the number of subshells is more important; for example, in the third group we have 4s, 4p, 4d, and 4f degeneracy as compared with only 2s and 2p, the case for the first period. These two reasons affect the correlation energy picture in the same way; namely, they emphasize the role of the "weak pairs" of the "simple pair model." We might say that the larger the number of electrons, the more linear the correlation behavior becomes with respect to the number of all possible strong and weak pairs. This is tantamount to saying that we see the emerging of a statistical picture which is very likely the final limit of the complex pairing model.

Let us examine, for example, the scandium atom<sup>51</sup> and consider the correlation energy for the  $\text{Sc}(^2D)$ ,  $\text{Sc}^+(^3F)$ ,  $\text{Sc}^{2+}(^2D)$ , and  $\text{Sc}^{3+}(^1S)$ , with corresponding configuration:  $4s^23d^1$ ,  $4s^03d^2$ ,  $4s^03d^1$ , and  $4s^03d^0$ . The correlation energy difference from Sc to  $\text{Sc}^+$ , from  $\text{Sc}^+$  to  $\text{Sc}^{2+}$  and from  $\text{Sc}^{2+}$  to  $\text{Sc}^{3+}$  are 0.037, 0.034, and 0.031 au, respectively. In the first step, Sc to  $\text{Sc}^+$ , a

"strong pair" is destroyed and the 4s electron promoted to the 3d shell can bring about only "weak pairs." In the second and third steps, we leave unaltered the number of "strong pairs" and we vary only the number of "weak pairs." But the correlation energy is very insensitive to such distinction of "weak" and "strong" pairs and behaves as if the 4s and 3d electrons do not depend on the  $n$ ,  $l$ ,  $m$ , and  $s$  quantum numbers at all. (The remarkable linearity of the computed values, 0.037, 0.034, 0.031 au, should not be taken too literally, because of the angular momenta, near-degeneracy, uncertainty in ionization potentials, lack of accuracy in the computation of the relativistic effects, etc.) Unfortunately, the lack of reliable data for higher ionization potentials for many cases prevents a final conclusion. It is noted that the highest ionization potentials available in the literature are likely to be in error, because of the heavy reliance on extrapolation and analogy which characterizes the determination of the high-order ionization potentials for the third group.

The net outcome of the analysis of these data is that we see the emerging of the "complex pairing model" and the collapse of the "simple pairing model."

The above considerations on the two-particle model are obtained by analyzing the correlation-energy data. It should be noted that the two-particle method has been proposed and analyzed by Hurley, Lennard-Jones, and Pople,<sup>18</sup> and later by others.<sup>54</sup> Their analysis is not within the self-consistent framework. Huzinaga<sup>55</sup> has derived a set of coupled Hartree-Fock-type equations to determine the two-electron geminals.

Presently there are no numerical computations to prove that the two-particle model functions (called geminals) represent a general answer to the correlation problem. However, the numerical results on the correlation energy which we have reported seem to indicate that this model will work well for the Be atom, less well for the Ne atom, and poorly for atoms with more than 20 electrons. The reason is the large amount of correlation energy due to the intershell correlation. In addition, we note that the model at present makes no provision for those cases where one should work in terms of  $j$ - $j$  coupling. On the other hand, it is expected that the two-particle model will give a satisfactory answer for saturated molecules with strongly localized bonds. The Hartree-Fock (HF) models assume that each electron experience the average field of all the remaining electrons and that the total wave function can be expressed as an antisymmetrized product of one-electron orbitals. Thus, the exact function

(52) The simple pairing model is exposed in detail by L. C. Allen, E. Clementi, and H. Gladney, *Rev. Mod. Phys.*, **35**, 465 (1963). The model fails to sufficiently recognize the importance of  $Z$  dependency<sup>50</sup> as well as the implication contained in the comments about excited states as reported in this work and elsewhere.

(53) E. Clementi, "Comprehensive Analyses of the First Three Periods of the Atomic System," presented at the Symposium Molecular Structure and Spectra, Columbus, Ohio, June 1963; *Ab Initio Computations in Atoms and Molecules*, *IBM J. Res. Develop.*, **9**, 2 (1965).

(54) R. McWeeny and B. T. Sutcliffe, *Proc. Roy. Soc. (London)*, **A273**, 103 (1963); T. L. Allen and H. Shull, *J. Chem. Phys.*, **35**, 1644 (1961); *J. Phys. Chem.*, **66**, 2281 (1962); W. Brenig, *Nucl. Phys.*, **4**, 363 (1957); F. Coester and H. Kümmel, *ibid.*, **17**, 477 (1960).

(55) S. Huzinaga, "SCF Method in the Paired-Electron Approximation," IBM Technical Report, available on request.



TABLE IIIC  
 CORRELATION ENERGY FROM THE CHF METHOD (IN ATOMIC UNITS)

| System                              | Hartree-Fock | CHF        | $E_C(\text{caled})^a$ | $E_C(\text{exptl})$ |
|-------------------------------------|--------------|------------|-----------------------|---------------------|
| He( <sup>1</sup> S)                 | -2.86166801  | -2.9037222 | -0.0420421            | -0.0421             |
| Li( <sup>2</sup> S)                 | -7.4327257   | -7.4850509 | -0.0523252            | -0.0453             |
| Be( <sup>1</sup> S)                 | -14.649920   | -14.573070 | -0.076860             | -0.0944             |
| B( <sup>2</sup> P)                  | -24.529052   | -24.632040 | -0.102988             | -0.125              |
| C( <sup>2</sup> P)                  | -37.688611   | -37.829531 | -0.140920             | -0.158              |
| M( <sup>4</sup> S)                  | -54.400911   | -54.590641 | -0.189730             | -0.188              |
| O( <sup>3</sup> P)                  | -74.809359   | -75.055357 | -0.245998             | -0.258              |
| F( <sup>2</sup> P)                  | -99.409284   | -99.725809 | -0.316521             | -0.324              |
| Ne( <sup>1</sup> S)                 | -128.54636   | -128.94431 | -0.39795              | -0.393              |
| Na( <sup>2</sup> S)                 | -161.85734   | -162.26045 | -0.40311              | -0.403              |
| Mg( <sup>1</sup> S)                 | -199.61430   | -200.05139 | -0.43709              | -0.451              |
| Al( <sup>2</sup> P)                 | -241.87625   | -242.35842 | -0.48217              | -0.482              |
| Si( <sup>2</sup> P)                 | -288.85109   | -288.38888 | -0.53779              | -0.522              |
| P( <sup>4</sup> S)                  | -340.71846   | -341.30388 | -0.58538              | -0.561              |
| S( <sup>2</sup> P)                  | -397.50460   | -398.14259 | -0.63799              | -0.60               |
| Cl( <sup>2</sup> P)                 | -459.48169   | -460.20544 | -0.72375              | -0.71               |
| Ar( <sup>1</sup> S)                 | -526.81703   | -527.64943 | -0.83240              | -0.79               |
| Be <sup>2+</sup> ( <sup>1</sup> S)  | -13.611256   | -13.654058 | -0.042802             | -0.0443             |
| C <sup>4+</sup> ( <sup>1</sup> S)   | -32.361154   | -32.404169 | -0.043015             | -0.0451             |
| O <sup>7+</sup> ( <sup>1</sup> S)   | -59.111119   | -59.154215 | -0.043096             | -0.0455             |
| Ne <sup>8+</sup> ( <sup>1</sup> S)  | -93.861103   | -93.904409 | -0.043306             | -0.0457             |
| Kr <sup>34+</sup> ( <sup>1</sup> S) | -1273.6110   | -1273.6543 | -0.0433               | -0.047              |

<sup>a</sup> Difference of  $E(\text{CHF}) - E(\text{HF})$ .

is replaced by a single determinant of one-electron orbitals (at least for closed-shell systems), and the  $1/r_{ij}$  operator of the exact Hamiltonian is replaced by Coulomb and exchange operators, representing the average field interaction.

Formally, the Hartree-Fock model can be equated to an "unperturbed system" and the difference between the exact and the average electron-electron interaction will be a "perturbation potential." This much is well known from the Möller and Plesset<sup>56</sup> analysis of the correlation problem. As a consequence, one can attempt to obtain the exact function by taking the HF function as a zero-order function and then add some correction *via* perturbation and/or variation techniques.

Physically, one can equate the Hartree-Fock model to a system where the "Coulomb hole" for electron pairs with antiparallel spin is not accounted for. We refer to Wigner's work on this point.<sup>57</sup>

We could attempt to introduce in the Hartree-Fock potential an additional term which directly represents the "Coulomb hole." Since we are interested in the quantum chemistry of molecular systems, we are concerned with not increasing the mathematical complexity of the problem beyond the Hartree-Fock formalism.

The "Coulomb hole" is introduced directly as a modification of the Coulomb integrals  $J_{\lambda pq, \mu rs}$ . This modification consists in replacing the integration range of the first electron from zero to  $r$  and from  $r$  to infinity (the usual limits of the  $J_{\lambda pq, \mu rs}$  elements) with the integration range from zero to  $(r - \delta)$  and from  $(r + \delta)$

to infinity. Since at the integration limit  $r$  the two electrons of the Coulomb element occupy the same radial position, the effect of replacing  $r$  by  $(r - \delta)$  and  $(r + \delta)$  introduces a discontinuity in the potential. Thus, we have a "Coulomb hole." In our method there are as many  $\delta$  as  $J$  integrals, thus  $\delta$  is designated as  $\delta_{\lambda pq, \mu rs}$ .

It is not difficult to obtain an expression for the  $\delta_{\lambda pq, \mu rs}$ . In our work<sup>53</sup> we have made use of two empirical parameters, one for the case of  $\delta$  with  $\lambda = \mu$  and the second for the case  $\delta$  with  $\lambda \neq \mu$ . The first parameter has been obtained by fitting the He atom (<sup>1</sup>S state) and the second by fitting the Ne atom (<sup>1</sup>S state).

With these two fittings we have analytically computed the  $J'_{\lambda pq, \mu rs}$  elements which differ from the standard  $J_{\lambda pq, \mu rs}$  because of the discontinuity in the integration range.

In summary, the technique of computation is as follows: (a) compute a Hartree-Fock function, (b) compute the  $\delta_{\lambda pq, \mu rs}$  and then the  $J'_{\lambda pq, \mu rs}$  matrix elements, and (c) compute again the self-consistent field function, but with the newly obtained  $J'$  matrix elements.

The resultant energies (in atomic units) for the first and second period are given in Table IIIC. The first column gives the Hartree-Fock energy for the functions we have used as a starting point; the second column gives the Coulomb-Hartree-Fock (CHF) energies (we shall call this the Coulomb-Hartree-Fock method, CHF). The third column gives the difference between the CHF and HF energies (*i.e.*, the correlation energy computed by the CHF method). The results are in fair agreement with the "experimental correlation energies" presented in Table IIIB.

The main results of this work seem to indicate that (a) the HF method can be improved within the spirit of

(56) C. Möller and M. S. Plesset, *Phys. Rev.*, **46**, 618 (1934).

(57) E. P. Wigner, *ibid.*, **46**, 1002 (1934); *Trans. Faraday Soc.*, **34**, 678 (1938).

the one-electron approximation (notice that the HF model is a direct extension of the Hartree model, *via* introduction of the Fermi hole; in an analogous way the CHF model is an extension of the HF model, *via* introduction of the Coulomb hole); (b) the CHF semi-empirical method proposed and tested here gives correlation energies in rough agreement with the experimental values; and (c) the empirical CHF functions seem to be as good as the HF functions, but this point must be studied further.

At present we are expanding this analysis and we are attempting to obtain the  $\delta$ 's directly from some physical model without making use of empirical parameters. It is noted that the free-electron gas correlation energy expression can be a useful starting point in this direction.<sup>58</sup> Simultaneously, we are attempting to extend the CHF method to molecular systems.

We note that the CHF method could be reformulated by referring to the  $F$  integrals (introduced by Slater) in place of the  $J$  integrals (as defined by Roothaan).

From the previous discussion one could be led to the hurried conclusion that the "simple pairing model" should hold well for those molecules with component atoms of low  $Z$  value (say,  $Z < 15$ ). Indeed, for such molecules the spin-orbit effect (at least at the equilibrium distances) is small and the complications due to the near-degeneracy in the atoms are removed, because of symmetry requirements in the molecule. It is noted that the above conclusion might be in error because in multiple-bonded molecules the intershell effect can be substantial.

The availability of the correlation energies for the first three periods of the atomic system proved that we can use atomic correlation data for predicting the correlation energy in molecules. This was done first in our works on HF and LiF and CH<sub>4</sub>,<sup>59</sup> then in McLean's work on LiF,<sup>60</sup> and subsequently in the CH<sub>4</sub> computations by Carlson and Skancke,<sup>61</sup> the N<sub>2</sub>, CO, and BF analyses by Nesbet,<sup>62</sup> and the BeO analysis of Yoshimine.<sup>63</sup>

It seems worthwhile to define a few quantities somewhat more critically than was previously done.<sup>64</sup> First we shall partition the correlation energy per orbital. If we have  $n$  orbitals and  $2n$  electrons, we assume, on the basis of our empirical knowledge on atomic data and from the CHF work that the total molecular correlation energy  $E_c(m)$  for a molecule  $m$  is given by

$$E_c(m) = \sum_i \eta_i + \sum_i \sum_j \eta_{ij}$$

$\eta_{ij}$  is the pair-pair correlation (weak pairing), and  $\eta_i$  is the pair correlation (strong pairing). If the orbitals are not the Hartree-Fock orbitals, but localized orbitals, then one would expect that the quantities  $\eta_{ij}$  be quite small. Of course the same is true for Hartree-Fock orbitals which do not overlap.

The atomic correlation energy  $E_c(A)$  for the component atoms of the molecule  $m$  is given by

$$E_c(A) = \sum_i \eta_i^A + \sum_{ij} \eta_{ij}^A$$

We define a quantity, designated as the *molecular extra correlation energy*,  $\Delta_c(m)$  as follows

$$\Delta_c(m) = \sum_c(m) - \sum_A E_c(A)$$

where  $A$  is now an index running over the set of atoms of  $m$ . Clearly for the inner-shell electrons, which retain nearly the same electronic distribution in the molecule as in the separated atoms, we can expect the same contribution to  $\eta$ . In other words, the correlation energy of the inner-shell electrons of the atoms in a molecule is nearly equal to the correlation energy of the same electrons in the separated atoms.

If we use closed-shell atoms in forming a molecule, or closed-shell molecules in forming a more complex molecule, then the number of pairs in the new molecule is equal to the sum of the number of pairs of the separated species (atomic or molecular). In these cases one would expect a relatively small value for  $\Delta_c(m)$  due mainly to the facts that (a) there are many more  $\eta_{ij}$  in  $E_c(m)$  than in  $E_c(A)$ , and (b) the valency electrons have different density in the molecule than in the atom.

If we use open-shell atoms in forming a molecule, and if the molecule pairs the previously unpaired spins, then  $\Delta_c(m)$  is a rather large quantity. From the atomic computations it seems that about 1 to 2 eV are added for each new pair.

In Table IV we report accurate computations for a number of diatomic and other linear molecules and we list both  $\Delta_c(m)$  and the computed Hartree-Fock dissociation energy as well as the experimental dissociation energy. This table takes some unpublished results presented and distributed at the 1964 Gordon Conference on Quantum Chemistry. A similar analysis has later been published by Hollister and Sinanoglu<sup>65</sup> and Cade.<sup>66</sup>

#### IV. CONTRACTED BASIS SET OF GAUSSIAN FUNCTIONS

In an SCF computation, the number of two-electron integrals (one center or many centers) is proportional to

(58) References on the free-electron gas correlation energy are available from ref 33.

(59) E. Clementi, *J. Chem. Phys.*, **36**, 33 (1962); **38**, 2780 (1963); **39**, 487 (1963).

(60) A. D. McLean, *ibid.*, **39**, 2653 (1963).

(61) K. D. Carlson and P. N. Skancke, *ibid.*, **40**, 613 (1964).

(62) R. K. Nesbet, *ibid.*, **40**, 3619 (1964).

(63) M. Yoshimine, *ibid.*, **40**, 2970 (1964).

(64) This analysis follows the results presented at the 1964 Gordon Conference on Quantum Mechanics. The data in Table IV are part of the work distributed at that conference.

(65) C. Hollister and O. Sinanoglu, *J. Am. Chem. Soc.*, **88**, 13 (1966).

(66) P. E. Cade and W. M. Huo, "The Electronic Structure of Diatomic Molecules," VI. A. "Hartree-Fock Wavefunctions and Energy Quantities for the Ground States of the First Row Hydrides," Technical Report, Laboratory of Molecular Structure and Spectra, University of Chicago, Chicago, Ill., 1966, p 203; VII. A. "Hartree-Fock Wavefunctions and Energy Quantities for the Ground States of the Second Row Hydrides," p 313.

TABLE IV  
MOLECULAR EXTRA CORRELATION ENERGY (MECE) AND  
DISSOCIATION ENERGY ( $D_e$ ) FOR DIATOMIC HETERONUCLEAR  
MOLECULES (IN ELECTRON VOLTS)<sup>a</sup>

| Molecule        | MECE                 | $D_e$ (HF) | $D_e$ (exptl) | New pairs |
|-----------------|----------------------|------------|---------------|-----------|
| CO              | $^1\Sigma^+$ 3.406   | 7.836      | 11.242        | 2         |
| BF              | $^1\Sigma^+$ 2.397   | 6.183      | ~8.58         | 1         |
| LiH             | $^1\Sigma^+$ 1.040   | 1.476      | 2.516         | 1         |
| FH              | $^1\Sigma^+$ 1.682   | 4.378      | ~6.06         | 1         |
| CH <sub>4</sub> | $^1A_{1g}$ ~5.20     | ~13.0      | ~18.20        | 4         |
| N <sub>2</sub>  | $^1\Sigma_g^+$ 4.631 | 5.271      | 9.902         | 3         |
| Li <sub>2</sub> | $^1\Sigma_g^+$ 0.884 | 0.169      | 1.050         | 1         |
| C <sub>2</sub>  | $^1\Sigma_g^+$ 5.469 | 0.781      | ~6.25         | 2         |
| O <sub>2</sub>  | $^3\Sigma_g^-$ 4.910 | 1.227      | 5.178         | 1         |
| O <sub>2</sub>  | $^1\Delta_g$ 4.171   | -0.520     | 4.171         | 2         |
| O <sub>2</sub>  | $^1\Sigma_g^+$ 3.951 | -1.392     | 3.518         | 2         |
| F <sub>2</sub>  | $^1\Sigma_g^+$ 3.047 | -1.374     | 1.679         | 1         |

<sup>a</sup> By inspection we see that the new pairs in the molecules bring about a correlation energy contribution of the same order as found for the pairs in the atoms. We can be more specific and first of all note that MECE here is given by using the data of  $E_C(1)$  from Table IIIB. If we correct these results by using the data from the correlation energy obtained from the two configuration computations  $E_C(2)$ , the values are somewhat different. For example, MECE for BF is not 2.397 eV, but 3.234 eV. Note that in BF we have formed an extra pair, mainly of  $\sigma$  type and in F<sub>2</sub> we have formed an extra pair, mainly of  $\sigma$  type, and the corresponding energy is 3.047 eV. However, in HF we have formed an extra pair, mainly of  $\sigma$  type, and the corresponding energy is 1.682 eV. In previous work we have attempted to give some general rule on the pair contribution for molecules. This work was subsequently continued by Nesbet and Sinanoglu. However, the rules presently available are quite successful for a few cases and fail remarkably in others. (Of course, one can avoid the problem by simply presenting only the case which fits the rules well.) For example, if we assume that a pair of p electrons is associated to a correlation energy of about 0.059 au (or 1.605 eV) as suggested by our atomic data, then the following algebra is tempting: for CO,  $2 \times 1.605 = 3.210$  (to be compared with 3.87); for HF,  $1 \times 1.605 = 1.605$  (to be compared with 1.682); for N<sub>2</sub>,  $3 \times 1.605 = 4.815$  (to be compared with 4.631); for O<sub>2</sub>( $^1\Sigma_g^+$ ),  $2 \times 1.605 = 3.210$  (to be compared with 3.951). Clearly, the agreement improves by using somewhat larger values than 1.605 eV. We can compare LiH with Be atom, HF with Ne atom, and CH<sub>4</sub> with Ne atom and obtain the corresponding values of MECE in agreement with the data of our table. The work by Sinanoglu<sup>49</sup> is a step in the correct direction, but fails to recognize the dependency of the correlation energy from the density gradient and uses the population analysis data with excessive confidence. It is noted that, in the population analysis, the distribution of charges between atoms A and B is assumed to be governed by the overlap of the charge distribution of A and B. However, such assumption is exceedingly crude whenever A and B are different atoms. Our present inclination is that we can transfer more easily data on correlation energy from small molecules (now available in large volume) to large molecules, than from atoms to large molecules. In addition, it seems likely that correlation energy from atoms to molecules can be transferred more easily if we compare correlation energies of atoms in the valency states rather than in the spectroscopic states.

the fourth power of the numbers of atomic functions. To adequately describe an atom of the first period with s and p electrons, one needs nine to twelve s-type Gaussian functions and five to six p-type Gaussian functions. Therefore such an atom, which in a molecule, will be described by about 30 Gaussian functions (ten of s type, five of p<sub>x</sub> type, five of p<sub>y</sub> type, five of p<sub>z</sub> type, plus some d and f types). The corresponding

number of Slater-type functions is about 20 (four to five of s type, three of p<sub>x</sub> type, three of p<sub>y</sub> type, three of p<sub>z</sub> type, plus some d and f types).

A relatively simple molecule like C<sub>2</sub>H<sub>6</sub> requires about 90 Gaussian functions, and this brings about the need to compute about 7 million integrals over Gaussian functions. These integrals are then transformed into integrals over symmetry-adapted functions; no matter what transformation technique is used, the computational time required for the transformation is proportional to the size of the integral list. Availability of large core memory certainly ameliorates the situation but does not eliminate the complexity of the transformation. Alternatively, one could compute directly the integrals over the symmetry-adapted functions, therefore eliminating the need for the transformations. However, in this case one will either compute a redundant number of integrals over the Gaussian function or carry a very long "integral request" list.

These difficulties can be overcome by replacing the individual Gaussians with some appropriate linear combination of Gaussian such as to reduce and therefore "contract" the number of stored integrals. This has been suggested for large molecular computations some time ago.<sup>39</sup> Now we have implemented this suggestion, and from our preliminary results it appears that one can finally compute large molecular systems rather accurately.

Two possible schemes are available for contracting the original basis set. On one hand, one could use as the contracted set the atomic orbitals of the separated atom. In this case one would start with as many contracted functions as the orbitals of the component atoms in their ground state. Computationally, one would construct the integrals over the atomic orbitals, making use of the atomic expansion coefficients. The drawback of this scheme is that the atomic orbitals are in general poor representations for molecular functions, except for the inner shells.

A second possibility is to analyze the Gaussian functions of the atomic orbitals and make appropriate linear combinations of the atomic functions. For an illustration of this technique, let us consider the Ne( $^1S$ ) atom. We shall compare the result of a standard, but optimal, basis set obtained by Huzinaga<sup>67</sup> with the results ob-

(67) The correct SCF total energy expression for a  $^2A_2$  ion (the case of the pyridine ion, analyzed in this work) is

$$E_{HF}(^2A_2) = \sum_{i'} (H_{i'}^{i'} - \epsilon_{i'}^{i'}) - \frac{1}{2} H_{i'}^{i'} - \frac{1}{2} \epsilon_{i'}^{i'} - \sum_{i'} K_{i'}^{i'}$$

where the superscript  $i$  indicates the above quantities are obtained from the open-shell wave function for the ion, and where the subscript  $t$  refers to the  $2a_2$  MO's singly occupied. In this paper the quantities with superscript + refer to the ion quantities obtained from the closed-shell ground function. A large part of the confusion existing in the literature is due to the identification of quantities here distinguished by the upper index + and by the upper index  $i$ . In the above equation the summation over  $t'$  excludes the terms where  $t' = t$ .

TABLE V

NE ATOM. UNCONTRACTED GAUSSIAN SET, ORBITAL ENERGIES, AND EXPANSION COEFFICIENTS (TOTAL ENERGY -128.5447 au)

|            | Orbital exponent | 1s orbital | 2s orbital |            | Orbital exponent | 2p orbital |
|------------|------------------|------------|------------|------------|------------------|------------|
| $X_1$      | 47870.2          | 0.00021    | -0.00005   | $X_{13}$   | 129.802          | 0.00426    |
| $X_2$      | 7385.83          | 0.00162    | -0.00038   | $X_{13}$   | 30.4192          | 0.03061    |
| $X_3$      | 1660.18          | 0.00863    | -0.00206   | $X_{14}$   | 9.62151          | 0.11927    |
| $X_4$      | 460.539          | 0.03617    | -0.00856   | $X_{15}$   | 3.54645          | 0.26912    |
| $X_5$      | 146.038          | 0.12134    | -0.03097   | $X_{16}$   | 1.41435          | 0.35733    |
| $X_6$      | 50.4137          | 0.30702    | -0.08388   | $X_{17}$   | 0.578893         | 0.33183    |
| $X_7$      | 18.7165          | 0.43944    | -0.17194   | $X_{18}$   | 0.216044         | 0.16084    |
| $X_8$      | 7.39702          | 0.22518    | -0.10947   | $\epsilon$ |                  | -0.84999   |
| $X_9$      | 2.6768           | 0.01554    | 0.37643    |            |                  |            |
| $X_{10}$   | 0.775195         | -0.00230   | 0.57102    |            |                  |            |
| $X_{11}$   | 0.29176          | 0.00095    | 0.20449    |            |                  |            |
| $\epsilon$ |                  | -32.772    | -1.9300    |            |                  |            |

tained with a "contracted set." Huzinaga's set consists of 11 Gaussian functions of s type, and seven Gaussian functions of p type. We shall designate the s function as  $X_1 \cdots X_{11}$ , and the seven p functions as  $X_{12} \cdots X_{18}$ . The orbital exponents and expansion coefficients of Huzinaga's computation are not reported.

The total energy is -128.5447 au. By simply inspecting the orbital exponents and the expansion coefficients, it is clear that (a) a number of Gaussians are needed only in order to represent the 1s cusp (the very high orbital exponents), and (b) several Gaussians belong only to the 2s orbital, and the small coefficient in the 1s is present for orthogonality requirements. Therefore, the set of 18 Gaussians can be contracted to a smaller set, a "contracted" set. For example, we could use the following set, and re-perform an SCF computation with it.

$$\begin{aligned}
 X_1' &= 0.00021X_1 + 0.00162X_2 \\
 X_2' &= 0.00863X_3 + 0.03617X_4 \\
 X_3' &= 0.12194X_5 + 0.30702X_6 \\
 X_4' &= 0.43944X_7 + 0.22518X_8 \\
 X_5' &= X_9 \\
 X_6' &= 0.57102X_{10} + 0.20449X_{11} \\
 X_7' &= 0.00426X_{12} + 0.03061X_{13} \\
 X_8' &= 0.11927X_{14} + 0.26912X_{15} \\
 X_9' &= 0.35733X_{16} + 0.33183X_{17} \\
 X_{10}' &= X_{18}
 \end{aligned}$$

Table V shows the results for orbital energies, total energies, and the new expansion coefficients.

The total energy for the contracted set,  $X'$ , is -128.5440. Let us contract even more and use the following contracted set for a new SCF computation.

$$\begin{aligned}
 X_1'' &= 0.00021X_1 + 0.00162X_2 + 0.00863X_3 \\
 X_2'' &= 0.03617X_4 + 0.12194X_5 + 0.30702X_6 \\
 X_3'' &= 0.43944X_7 + 0.22518X_8 \\
 X_4'' &= 0.37643X_9 + 0.57102X_{10} + 0.20449X_{11} \\
 X_5'' &= 0.00426X_{12} + 0.03061X_{13} + 0.11927X_{14} \\
 X_6'' &= 0.26912X_{15} + 0.35733X_{16} \\
 X_7'' &= 0.16084X_{17} + 0.33183X_{18}
 \end{aligned}$$

The recomputed orbital energies (au) are -32.76031 for 1s, -1.92977 for 2s, and -0.848104 for 2p; and the total energy is -128.5412 au.

Table VI summarizes the above analysis and gives

TABLE VI  
CONTRACTED SET FOR NE ATOM

|            | 1s orbital | 2s orbital |            | 2p orbital |
|------------|------------|------------|------------|------------|
| $X_1'$     | 0.001747   | -0.000416  | $X_7'$     | 0.033118   |
| $X_2'$     | 0.043015   | -0.010125  | $X_8$      | 0.366407   |
| $X_3'$     | 0.41259    | -0.107915  | $X_9$      | 0.651301   |
| $X_4'$     | 0.647292   | -0.267860  | $X_{10}$   | 0.160667   |
| $X_5'$     | 0.0145941  | 0.363564   | $\epsilon$ | -0.85091   |
| $X_6'$     | -0.000715  | 0.757529   |            |            |
| $\epsilon$ | -32.7711   | -1.93016   | $E =$      | -128.5440  |

the total energies, the orbital energies, the number of two-electron integrals for each type of set, and the equivalent number of two-electron integrals (the 2p orbitals were considered to be subdivided in  $2p_x$ ,  $2p_y$ , and  $2p_z$ , as they are in most molecules.

In Table VII we have reported in the last column the number of elements in the  $\mathcal{O}$  matrix. It is noted that the number of integrals is much larger, since (a) we compute all the distinct and possible integrals which can be derived from the basis set; (b) the  $\mathcal{O}$  supermatrix contains both Coulomb and exchange integrals.

A final example of the usefulness of the contracted set is reported for the  $N_2$  molecule. Again we started with Huzinaga's N(<sup>3</sup>P) atomic computation [24] with eleven s-type Gaussian and seven p-type Gaussians. This set was contracted to four functions of s type and two of p type. The computed total energy is -108.81163 au. This contracted set of four contracted s-type Gaussians and two contracted Gaussians of 2p type is equal in number to a double- $\zeta$  Slater-type set. However, a Slater double- $\zeta$ -type set gives an energy of -108.79508 au, or 0.01655 au higher than the Gaussian contracted set.

Before concluding our remarks on the contraction, we wish to point out that one can optimize the contracted set by performing, for a given selected contracted set, a series of atomic computations whereby the contraction coefficients are optimized (instead of

TABLE VII  
 COMPARISON OF CONTRACTED AND UNCONTRACTED SETS

| Type    | Total energy, au | $\epsilon(1s)$ | $\epsilon(2s)$ | $\epsilon(2p)$ | No. of s basis | No. of p basis | No. of $\mathcal{P}$ elements |
|---------|------------------|----------------|----------------|----------------|----------------|----------------|-------------------------------|
| X set   | -128.5447        | -32.772        | -1.9300        | -0.84999       | 11             | 7, 7, 7        | 11325                         |
| X' set  | -128.54398       | -32.771        | -1.9301        | -0.85091       | 6              | 4, 4, 4        | 861                           |
| X'' set | -128.54114       | -32.700        | -1.9298        | -0.84810       | 4              | 3, 3, 3        | 406                           |

 TABLE VIII  
 BASIS SET OF CONTRACTED GAUSSIANS FOR C, N, AND H

| Atom | Contracted basis   |
|------|--|
| N    | $\gamma(1s) = 0.01817\beta_1 + 0.10776\beta_2 + 0.32320\beta_3 + 0.47673\beta_4 + 0.22046\beta_5$                |
| N    | $\gamma(2s) = 0.50027\beta_6 + 0.63917\beta_7$   |
| N    | $\gamma(2p) = 0.13843\beta_8 + 0.49760\beta_9 + 0.57505\beta_{10}$   |
| C    | $\gamma(1s) = 0.02220\beta_{11} + 0.13285\beta_{12} + 0.38435\beta_{13} + 0.45798\beta_{14} + 0.15441\beta_{15}$ |
| C    | $\gamma(2s) = 0.56673\beta_{16} + 0.55692\beta_{17}$   |
| C    | $\gamma(2p) = 0.10845\beta_{18} + 0.46116\beta_{19} + 0.63043\beta_{20}$   |
| H    | $\gamma(1s) = 0.64767\beta_{21} + 0.40789\beta_{22} + 0.07048\beta_{23}$   |

 TABLE IX  
 UNCONTRACTED GAUSSIAN SET FOR N, C, AND H

| No.       | Type | Orb. exp | No.          | Type | Orb. exp | No.          | Type | Orb. exp |
|-----------|------|----------|--------------|------|----------|--------------|------|----------|
| $\beta_1$ | 1s   | 636.101  | $\beta_9$    | 2p   | 1.10716  | $\beta_{17}$ | 1s   | 0.155139 |
| $\beta_2$ | 1s   | 105.386  | $\beta_{10}$ | 2p   | 0.261750 | $\beta_{18}$ | 2p   | 4.31613  |
| $\beta_3$ | 1s   | 27.5167  | $\beta_{11}$ | 1s   | 391.445  | $\beta_{19}$ | 2p   | 0.873682 |
| $\beta_4$ | 1s   | 9.02708  | $\beta_{12}$ | 1s   | 64.7358  | $\beta_{20}$ | 2p   | 0.20286  |
| $\beta_5$ | 1s   | 3.33086  | $\beta_{13}$ | 1s   | 16.2247  | $\beta_{21}$ | 1s   | 0.151374 |
| $\beta_6$ | 1s   | 0.828625 | $\beta_{14}$ | 1s   | 5.33460  | $\beta_{22}$ | 1s   | 0.681277 |
| $\beta_7$ | 1s   | 0.243109 | $\beta_{15}$ | 1s   | 2.00995  | $\beta_{23}$ | 1s   | 4.50037  |
| $\beta_8$ | 2p   | 5.19829  | $\beta_{16}$ | 1s   | 0.502323 |              |      |          |

 TABLE X  
 MOLECULAR GEOMETRY FOR PYRROLE MOLECULE<sup>a</sup>

| Center | N        | C(1)     | C(2)     | C(3)     | C(4)     | H(1)     | H(2)     | H(3)     | H(4)     | H(5)     |
|--------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| y      | 0.0      | 2.128939 | 2.128939 | 1.360632 | 1.360632 | 4.093449 | 4.093449 | 2.571379 | 2.571379 | 0.0      |
| z      | 1.633593 | 0.0      | 0.0      | 2.433109 | 2.433109 | 0.619408 | 0.619408 | 4.099559 | 4.099559 | 3.598950 |

<sup>a</sup> Distances are given in atomic units; the value of the  $x$  coordinate is 0, since the molecule is in the  $yz$  plane.

the orbital exponents as usual). This first requires an optimal uncontracted basis set and then new optimization on the contraction coefficients. Work is in progress for the programming of this problem.

From the above results on the Ne atom and  $N_2$  molecule where direct comparison with Slater-type functions can be made, and from computations on  $H_2O$ ,  $C_2H_6$ , and  $H_6$  that we have performed, we tentatively conclude that the basis set with more than 150-200 functions must necessarily resort to some "contraction" technique; otherwise the handling of integrals becomes a very expensive process in terms of computational time.

## V. PYRROLE

The basis set of contracted Gaussians we shall use for the C, N, and H atoms are given in Table VIII. These contracted Gaussians are built up from single (or uncontracted) Gaussian functions given in Table IX. The geometry of the pyrrole molecule is given in Table X. The molecules of pyrrole have  $C_{2v}$  symmetry and the following electronic configuration.

|               |  |
|---------------|--|
| $\sigma$ type | $1a_1^2 2a_1^2 3a_1^2 4a_1^2 5a_1^2 6a_1^2 7a_1^2 8a_1^2 9a_1^2$ |
| $\sigma$ type | $1b_2^2 2b_2^2 3b_2^2 4b_2^2 5b_2^2$                             |
| $\pi$ type    | $1b_1^2 2b_1^2 1a_2^2$   |

 TABLE XI  
 TOTAL ENERGY AND ORBITAL ENERGIES FOR PYRROLE  
 (IN ATOMIC UNITS)

| $A_1(\sigma)$              | $B_1(\sigma)$ | $B_1(\pi)$ | $A_2(\pi)$ |
|----------------------------|---------------|------------|------------|
| -15.71000                  | -11.42526     | -0.63133   | -0.38794   |
| -11.42520                  | -11.37850     | -0.42529   |            |
| -11.37931                  | -1.03448      |            |            |
| -1.32387                   | -0.79702      |            |            |
| -1.09548                   | -0.62429      |            |            |
| -0.82508                   | -0.60219      |            |            |
| -0.77787                   |               |            |            |
| -0.64759                   |               |            |            |
| -0.57659                   |               |            |            |
| Total energy -207.93135 au |               |            |            |

The total energy and the orbital energies of our computations are given in Table XI.

Let us analyze the MO's. This is done by making use of the gross population analyses given in Table XII (for the  $a_1$  orbitals), Table XIII (for the  $b_2$  orbitals), and Table XIV (for the  $b_1$  and  $a_2$  orbitals).

The  $1a_1$ ,  $2a_1$ , and  $3a_1$  MO's are clearly the 1s atomic orbitals for the nitrogen and carbon atoms, respectively. The remaining two 1s electrons for the carbon atoms are the  $1b_2$  and  $2b_2$  MO's. This can be seen from the population analysis results. The  $1a_1$ ,  $2a_1$ ,  $3a_1$ ,  $1b_2$ , and  $2b_2$  are the inner-shell MO and to a large extent they can

TABLE XII

| PYRROLE. GROSS POPULATION ANALYSIS (A <sub>1</sub> ) |                 |                 |                 |                 |                 |                 |                 |                 |                 |          |
|--|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------|
|  | 1a <sub>1</sub> | 2a <sub>1</sub> | 3a <sub>1</sub> | 4a <sub>1</sub> | 5a <sub>1</sub> | 6a <sub>1</sub> | 7a <sub>1</sub> | 8a <sub>1</sub> | 9a <sub>1</sub> | Total    |
| N(1s)  | 1.99356         | 0.0             | 0.0             | 0.00236         | 0.00132         | 0.00020         | 0.00002         | 0.00002         | 0.00004         | 1.99752  |
| N(2s)  | 0.00718         | -0.00002        | 0.0             | 0.85534         | 0.42728         | 0.07166         | 0.00094         | 0.00150         | 0.00702         | 1.37090  |
| C <sub>1</sub> (1s) + C <sub>2</sub> (1s)            | 0.0             | 1.99700         | 0.00042         | 0.00096         | 0.00008         | 0.00050         | 0.00024         | 0.0             | 0.0             | 1.99920  |
| C <sub>1</sub> (2s) + C <sub>2</sub> (2s)            | -0.00020        | 0.00266         | -0.00008        | 0.51234         | 0.03734         | 0.22734         | 0.23380         | 0.00140         | 0.00752         | 1.02212  |
| C <sub>3</sub> (1s) + C <sub>4</sub> (1s)            | 0.0             | 0.00040         | 1.99654         | 0.00040         | 0.00166         | 0.00040         | 0.0             | 0.0             | 0.0             | 1.99940  |
| C <sub>3</sub> (2s) + C <sub>4</sub> (2s)            | 0.0             | -0.00002        | 0.00316         | 0.22062         | 0.85246         | 0.26344         | 0.01096         | 0.00900         | -0.00076        | 1.35886  |
| C <sub>1</sub> (2p <sub>y</sub> )                    |                 |                 |                 |                 |                 |                 |                 |                 |                 |          |
| - C <sub>2</sub> (2p <sub>y</sub> )                  | -0.00020        | 0.00002         | -0.00002        | 0.17066         | 0.00174         | 0.00280         | 0.69752         | -0.00228        | 0.13268         | 1.00292  |
| C <sub>3</sub> (2p <sub>y</sub> )                    |                 |                 |                 |                 |                 |                 |                 |                 |                 |          |
| - C <sub>4</sub> (2p <sub>y</sub> )                  | 0.0             | 0.0             | 0.00004         | 0.01364         | 0.08014         | 0.11560         | 0.28388         | 0.03796         | 1.16986         | 1.70112  |
| N(2p <sub>z</sub> )                                  | 0.0             | 0.0             | 0.0             | 0.05148         | 0.03748         | 0.48572         | 0.24706         | 0.41460         | 0.00008         | 1.23642  |
| C <sub>1</sub> (2p <sub>z</sub> )                    |                 |                 |                 |                 |                 |                 |                 |                 |                 |          |
| + C <sub>2</sub> (2p <sub>z</sub> )                  | -0.00012        | 0.0             | -0.00002        | 0.00240         | 0.36376         | -0.00002        | -0.00034        | 0.29072         | 0.27066         | 0.92704  |
| C <sub>3</sub> (2p <sub>z</sub> )                    |                 |                 |                 |                 |                 |                 |                 |                 |                 |          |
| + C <sub>4</sub> (2p <sub>z</sub> )                  | 0.0             | 0.0             | 0.0             | 0.07786         | 0.01662         | 0.35292         | 0.00888         | 0.72772         | 0.00552         | 1.18952  |
| H <sub>1</sub> (1s) + H <sub>2</sub> (1s)            | 0.0             | -0.0002         | 0.0             | 0.02566         | 0.00180         | 0.08692         | 0.40844         | 0.01750         | 0.23456         | 0.77486  |
| H <sub>3</sub> (1s) + H <sub>4</sub> (1s)            | 0.0             | 0.0             | -0.00002        | 0.00888         | 0.08144         | 0.14746         | 0.04660         | 0.30424         | 0.17042         | 0.75902  |
| H <sub>5</sub> (1s)                                  | -0.0002         | 0.0             | 0.0             | 0.05740         | 0.09678         | 0.24510         | 0.06200         | 0.19748         | 0.00244         | 0.66098  |
| Total  | 2.00000         | 2.00002         | 2.00002         | 2.00000         | 1.99990         | 2.00004         | 2.00000         | 1.99986         | 2.00004         | 17.99988 |

TABLE XIII

| PYRROLE. GROSS POPULATION ANALYSIS (B <sub>2</sub> )                  |                 |                 |                 |                 |                 |                 |          |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------|
|   | 1b <sub>2</sub> | 2b <sub>2</sub> | 3b <sub>2</sub> | 4b <sub>2</sub> | 5b <sub>2</sub> | 6b <sub>2</sub> | Total    |
| N(2p <sub>z</sub> )   | 0.0             | 0.0             | 0.21860         | 0.47370         | 0.26210         | 0.18958         | 1.14398  |
| C <sub>1</sub> (1s) - C <sub>2</sub> (1s)                             | 1.99676         | 0.00044         | 0.00178         | 0.00016         | 0.0             | 0.00008         | 1.99922  |
| C <sub>1</sub> (2s) - C <sub>2</sub> (2s)                             | 0.00276         | -0.00004        | 0.96008         | 0.08428         | 0.00078         | 0.02378         | 1.07164  |
| C <sub>3</sub> (1s) - C <sub>4</sub> (1s)                             | 0.00044         | 1.99682         | 0.00072         | 0.00070         | 0.00006         | 0.00004         | 1.99878  |
| C <sub>3</sub> (2s) - C <sub>4</sub> (2s)                             | 0.0             | 0.00282         | 0.34680         | 0.33742         | 0.01026         | 0.05590         | 0.75320  |
| C <sub>1</sub> (2p <sub>y</sub> ) + C <sub>2</sub> (2p <sub>y</sub> ) | -0.00002        | 0.0             | -0.01558        | -0.00004        | 1.05212         | 0.00396         | 1.04044  |
| C <sub>3</sub> (2p <sub>y</sub> ) + C <sub>4</sub> (2p <sub>y</sub> ) | 0.0             | -0.00004        | 0.12384         | 0.17368         | 0.00246         | 0.06090         | 0.36084  |
| C <sub>1</sub> (2p <sub>z</sub> ) - C <sub>2</sub> (2p <sub>z</sub> ) | 0.0             | 0.00002         | 0.04066         | 0.61552         | 0.02064         | 0.32090         | 0.99774  |
| C <sub>3</sub> (2p <sub>z</sub> ) - C <sub>4</sub> (2p <sub>z</sub> ) | 0.00006         | 0.0             | 0.13082         | 0.00570         | 0.06042         | 0.76172         | 0.95872  |
| H <sub>1</sub> (1s) - H <sub>2</sub> (1s)                             | 0.0             | 0.0             | 0.14564         | 0.08342         | 0.58696         | 0.00150         | 0.81752  |
| H <sub>3</sub> (1s) - H <sub>4</sub> (1s)                             | 0.0             | 0.0             | 0.04656         | 0.22544         | 0.00420         | 0.58166         | 0.85786  |
| Total   | 2.00000         | 2.00002         | 1.99992         | 1.99998         | 2.00000         | 2.00002         | 11.99994 |

TABLE XIV

| PYRROLE. GROSS POPULATION ANALYSIS (π) A <sub>2</sub> AND B <sub>1</sub> |                 |                 |                 |
|--|-----------------|-----------------|-----------------|
|  | 1b <sub>1</sub> | 2b <sub>1</sub> | 1a <sub>2</sub> |
| N(2p <sub>z</sub> )  | 0.89596         | 0.76294         |                 |
| C <sub>1</sub> (2p <sub>z</sub> ) ± C <sub>2</sub> (2p <sub>z</sub> )    | 0.71820         | 0.02808         | 1.40416         |
| C <sub>3</sub> (2p <sub>z</sub> ) ± C <sub>4</sub> (2p <sub>z</sub> )    | 0.38580         | 1.20896         | 0.59580         |
| Total  | 1.99996         | 1.99998         | 1.99996         |

be considered undistorted atomic orbitals of the separated atoms in the molecular field.

The second group of  $\sigma$  electrons are responsible for the C-C, C-N, C-H, and N-H bond formation. The 4a<sub>1</sub> is mainly constructed from the 2s atomic orbital on C and N. This orbital flows over the entire molecular skeleton with maximum density at the nitrogen atom [the gross charge on 2s(N) is 0.85], lesser density at the C<sub>1</sub> and C<sub>2</sub> positions (gross charges on C<sub>1</sub> and C<sub>2</sub> are 0.34 partially polarized), and again lesser density at the C<sub>3</sub> and C<sub>4</sub> positions (gross charges on C<sub>3</sub> and C<sub>4</sub> are 0.12, again partially polarized). This MO, therefore, envelops the molecule, has maximum density at nitrogen, and is polarized. It has an analog in the lowest  $\pi$  MO, the 1b<sub>1</sub>, which has very similar charge distribution. This makes us question the long-standing idea in chemistry that the  $\pi$  electrons are much more delocalized than the  $\sigma$  electrons. As far as we can see, the 4a<sub>1</sub> MO

and the 1b<sub>1</sub> are very similar in character (of course, the former is much more bound than the latter).

The 5a<sub>1</sub> MO has maximum charge at the C<sub>3</sub> and C<sub>4</sub> positions, lesser charge at the C<sub>1</sub> and C<sub>2</sub> positions, and intermediate charges at the N position. This orbital in part tends to reverse the charge distribution given by the 4a<sub>1</sub> by concentrating charges on C<sub>3</sub> and C<sub>4</sub>. In addition, whereas the 4a<sub>1</sub> orbital is mainly 2p<sub>y</sub> polarized, the 5a<sub>1</sub> is mainly 2p<sub>z</sub> polarized. Note that the H<sub>5</sub> contributes to both the 4a<sub>1</sub> and the 5a<sub>1</sub>, by 0.06 and 0.10 fractional electron, respectively.

The remaining MO's of a<sub>1</sub> symmetry have the charge distribution on all the ten atoms of the molecule. The 5a<sub>1</sub>, the 6a<sub>1</sub>, and the 8a<sub>1</sub> are responsible for the N-H bond. The set 6a<sub>1</sub> to 9a<sub>1</sub> as well as the set 3b<sub>2</sub> to 6b<sub>2</sub> are responsible for the C<sub>1</sub>-H<sub>1</sub>, C<sub>2</sub>-H<sub>2</sub>, C<sub>3</sub>-H<sub>3</sub>, and C<sub>4</sub>-H<sub>4</sub> bonds. There is no single MO which can be identified with a given bond; this is the nature of the MO theory.

The  $\pi$ -occupied MO's are the 1b<sub>1</sub>, 2b<sub>1</sub>, and 1a<sub>2</sub>. The 1b<sub>1</sub> flows over the full molecular skeleton, with a density maximum at the nitrogen atom. The 2b<sub>1</sub> has a node between the nitrogen atom and the rest of the molecule; it has a density maximum both at the C<sub>3</sub> and the C<sub>4</sub> positions and at the nitrogen atom; a density minimum is present at the C<sub>1</sub> and C<sub>2</sub> positions

(the  $C_3$ ,  $C_4$  maximum is higher than the nitrogen maximum). The  $a_2$  has no charge on the nitrogen (by symmetry); it has high density at the  $C_1$  and  $C_2$  positions, and low density at the  $C_3$  and  $C_4$  positions; for symmetry considerations, it has a node in the  $C_2$  symmetry axis and, being the lowest of that symmetry, has the same phase on  $C_1$  and  $C_2$  and on  $C_3$  and  $C_4$ . Therefore, there is extended conjugation. Pyrrole is an aromatic compound, because of the six  $\pi$  electrons; however, there is one  $\pi$  MO deep in the  $\sigma$  MO's.

The charge distribution in pyrrole is summarized in Table XV. The nitrogen atom has the following

TABLE XV  
PYRROLE. GROSS CHARGES—SUMMARY

| N                                   |        |                 |        |    |        |                  |         |
|-------------------------------------|--------|-----------------|--------|----|--------|------------------|---------|
| 1s                                  | 1.9975 | 1s              | 1.9975 | 1s | 1.9975 | $\delta(\sigma)$ | -0.7488 |
| 2s                                  | 1.3709 | 2s              | 1.3709 | 2s | 1.3709 | $\delta(\pi)$    | 0.3411  |
| 2p <sub>x</sub>                     | 1.2364 | 2p <sub>y</sub> | 2.3804 | 2p | 4.0393 | $\delta$         | -0.4077 |
| 2p <sub>y</sub>                     | 1.1440 | 2p <sub>x</sub> | 1.6589 |    |        |                  |         |
| 2p <sub>z</sub>                     | 1.6589 |                 |        |    |        |                  |         |
| C <sub>1</sub> (or C <sub>2</sub> ) |        |                 |        |    |        |                  |         |
| 1s                                  | 1.9992 | 1s              | 1.9992 | 1s | 1.9992 | $\delta(\sigma)$ | -0.0302 |
| 2s                                  | 1.0469 | 2s              | 1.0469 | 2s | 1.0469 | $\delta(\pi)$    | -0.0752 |
| 2p <sub>x</sub>                     | 0.9624 | 2p <sub>y</sub> | 1.9841 | 2p | 3.0593 | $\delta$         | -0.1054 |
| 2p <sub>y</sub>                     | 1.0217 | 2p <sub>x</sub> | 1.0752 |    |        |                  |         |
| 2p <sub>z</sub>                     | 1.0752 |                 |        |    |        |                  |         |
| C <sub>3</sub> (or C <sub>4</sub> ) |        |                 |        |    |        |                  |         |
| 1s                                  | 1.9991 | 1s              | 1.9991 | 1s | 1.9991 | $\delta(\sigma)$ | -0.1602 |
| 2s                                  | 1.0560 | 2s              | 1.0560 | 2s | 1.0560 | $\delta(\pi)$    | -0.0953 |
| 2p <sub>x</sub>                     | 1.0741 | 2p <sub>y</sub> | 2.1051 | 2p | 3.2004 | $\delta$         | -0.2555 |
| 2p <sub>y</sub>                     | 1.0310 | 2p <sub>x</sub> | 1.0953 |    |        |                  |         |
| 2p <sub>z</sub>                     | 1.0953 |                 |        |    |        |                  |         |
| H <sub>5</sub>                      |        |                 |        |    |        |                  |         |
| 1s                                  | 0.6610 | 1s              | 0.6610 | 1s | 0.6610 | $\delta(\sigma)$ | 0.3390  |
| H <sub>1</sub> (or H <sub>2</sub> ) |        |                 |        |    |        |                  |         |
| 1s                                  | 0.7962 | 1s              | 0.7962 | 1s | 0.7962 | $\delta(\sigma)$ | 0.2038  |
| H <sub>3</sub> (or H <sub>4</sub> ) |        |                 |        |    |        |                  |         |
| 1s                                  | 0.8084 | 1s              | 0.8084 | 1s | 0.8084 | $\delta(\sigma)$ | 0.1916  |

charges,  $1s^{2.2}s^{1.37}2p_{\sigma}^{2.38}2p_{\pi}^{1.66}$ , which could be compared with the original (separated atom) distribution,  $1s^{2.2}s^{2.2}2p_{\sigma}^2 2p_{\pi}^2$ . Therefore, the nitrogen has gained 0.41 electron. This gain is the sum of two effects: a gain of 0.75 electron from the  $\sigma$  orbital, and a loss of 0.34 electron from the  $\pi$  orbitals. The charge transfer acts two ways: *the nitrogen is a  $\pi$  donor and  $\sigma$  acceptor*, with the net result of a gain of 0.41 electron. This two-way charge transfer brings about the problem of how reasonable are the charge distributions with the  $\pi$ -electron approximation, where one assumes in general an undistorted core ( $\sigma$  electrons).

For the carbon atoms and the hydrogen atoms we have only one-way charge transfer. The carbon atoms are both  $\sigma$  and  $\pi$  acceptors, whereas the hydrogens are  $\sigma$  donors.

Finally, the hybridization of the nitrogen atom is  $s^{1.37}p^{2.38}$ ; the hybridization of the  $C_1$  and  $C_2$  is  $s^{1.05}p^{1.98}$ ; and the hybridization of  $C_3$  and  $C_4$  is  $s^{1.06}p^{2.10}$ . These

values are not too different from the  $s^{1p^2}$  of a trigonal hybrid, except for the nitrogen atom.

## VI. PYRIDINE GROUND STATE

The geometry of the pyridine molecule is reported in Table XVI. The  $z$  axis is taken perpendicular to the

TABLE XVI  
MOLECULAR GEOMETRY FOR THE PYRIDINE MOLECULE\*

|                | $x$       | $y$       |                | $x$       | $y$       |
|----------------|-----------|-----------|----------------|-----------|-----------|
| N              | 0.0       | 1.328750  | C <sub>1</sub> | -2.185210 | 0.0       |
| H <sub>1</sub> | -3.943160 | 1.029325  | C <sub>2</sub> | 2.185210  | 0.0       |
| H <sub>2</sub> | 3.943160  | 1.029325  | C <sub>3</sub> | -2.268960 | -2.609880 |
| H <sub>3</sub> | -4.061760 | -3.581670 | C <sub>4</sub> | 2.268960  | -2.609880 |
| H <sub>4</sub> | 4.061760  | -3.581670 | C <sub>5</sub> | 0.0       | -3.970520 |
| H <sub>5</sub> | 0.0       | -6.007500 |                |           |           |

\* Distances are given in atomic units; the value of the  $z$  coordinate is 0.

molecular plane; the  $y$  axis is taken as the principal axis of symmetry. The basis set of Gaussian functions is equal to the basis set used for the pyrrole function.

The electronic configuration for pyridine is as follows.

|                    |  |
|--------------------|--|
| $\sigma$ electrons | $1a_1^2 2a_1^2 3a_1^2 4a_1^2 5a_1^2 6a_1^2 7a_1^2 8a_1^2 9a_1^2 10a_1^2 11a_1^2$ |
| $\sigma$ electrons | $1b_2^2 2b_2^2 3b_2^2 4b_2^2 5b_2^2 6b_2^2 7b_2^2$                               |
| $\pi$ electrons    | $1b_1^2 2b_1^2$  |
| $\pi$ electrons    | $1a_2^2$   |

The computed orbital energies and total energy are given in atomic units in Table XVII. In the following

TABLE XVII  
TOTAL ENERGY AND ORBITAL ENERGIES FOR PYRIDINE  
(IN ATOMIC UNITS)

| A <sub>1</sub> ( $\sigma$ ) | B <sub>2</sub> ( $\sigma$ ) | B <sub>1</sub> ( $\pi$ ) | A <sub>2</sub> ( $\pi$ ) |
|-----------------------------|-----------------------------|--------------------------|--------------------------|
| -15.67755                   | -11.46115                   | -0.62227                 | -0.44725                 |
| -11.46112                   | -11.43440                   | -0.45856                 |                          |
| -11.44301                   | -1.11025                    |                          |                          |
| -11.43432                   | -0.90435                    |                          |                          |
| -1.32774                    | -0.72600                    |                          |                          |
| -1.15774                    | -0.67003                    |                          |                          |
| -0.92182                    | -0.57950                    |                          |                          |
| -0.77915                    |                             |                          |                          |
| -0.70116                    |                             |                          |                          |
| -0.63938                    |                             |                          |                          |
| -0.46543                    |                             |                          |                          |

Total energy -245.62194 au

we shall analyze the results by comparing the orbital energies and the gross electronic charge distribution obtained from the population analysis (Tables XVIII, XIX, and XX).

The first group of MO's are the inner shells, *i.e.*, the 1s on the N atom and the 1s on the five carbon atoms. These are represented by the following MO's:  $1a_1$  (the 1s for N),  $2a_1$  (the 1s for  $C_1$  and  $C_2$ ),  $3a_1$  (the 1s for  $C_3$  and  $C_4$ ),  $4a_1$  (the 1s for  $C_5$  and  $C_6$ ),  $1b_2$  (the 1s for  $C_1$  and  $C_2$ ), and  $2b_2$  (the 1s for  $C_3$  and  $C_4$ ). The orbital energies of these MO's are nearly identical with the 1s orbital

TABLE XVIII  
 PYRIDINE. GROSS POPULATION ANALYSIS (A<sub>1</sub>)

| Symmetry function   | 1a <sub>1</sub> | 2a <sub>1</sub> | 3a <sub>1</sub> | 4a <sub>1</sub> | 5a <sub>1</sub> | 6a <sub>1</sub> | 7a <sub>1</sub> | 8a <sub>1</sub> | 9a <sub>1</sub> | 10a <sub>1</sub> | 11a <sub>1</sub> | Total    |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|----------|
| N(1s)   | 1.99366         | 0.0             | 0.0             | 0.0             | 0.00226         | 0.00082         | 0.00044         | 0.00002         | 0.00016         | 0.00004          | 0.00018          | 1.99758  |
| N(2s)   | 0.00720         | 0.0             | 0.0             | 0.0             | 0.77424         | 0.25442         | 0.12372         | 0.02660         | 0.02928         | 0.05102          | 0.24248          | 1.50896  |
| C <sub>1</sub> (1s) + C <sub>2</sub> (1s)                             | 0.0             | 1.99666         | 0.0             | 0.00080         | 0.00108         | 0.0             | 0.00042         | 0.00004         | 0.00018         | 0.00004          | 0.00002          | 1.99924  |
| C <sub>1</sub> (2s) + C <sub>2</sub> (2s)                             | -0.00034        | 0.00258         | 0.0             | -0.00004        | 0.58370         | -0.00036        | 0.24870         | 0.05190         | 0.13222         | 0.00428          | 0.01246          | 1.03510  |
| C <sub>3</sub> (1s) + C <sub>4</sub> (1s)                             | 0.0             | 0.00078         | 0.01374         | 1.98270         | 0.00030         | 0.00112         | 0.00018         | 0.0             | 0.00022         | 0.00004          | 0.00004          | 1.99912  |
| C <sub>3</sub> (1s) + C <sub>4</sub> (2s)                             | 0.0             | -0.00002        | 0.0             | 0.00286         | 0.19752         | 0.60648         | 0.08166         | 0.01004         | 0.15610         | 0.02214          | 0.00798          | 1.08476  |
| C <sub>5</sub> (1s)   | 0.0             | 0.0             | 1.98348         | 0.01378         | 0.00010         | 0.00108         | 0.00060         | 0.00002         | 0.00014         | 0.0              | 0.0              | 1.99920  |
| C <sub>5</sub> (2s)   | 0.0             | 0.0             | 0.00280         | -0.00006        | 0.06172         | 0.60108         | 0.30768         | 0.03600         | 0.05362         | 0.00934          | 0.00004          | 1.07222  |
| H <sub>1</sub> (1s) + H <sub>2</sub> (1s)                             | 0.0             | -0.00002        | 0.0             | 0.0             | 0.07678         | 0.00050         | 0.08218         | 0.19996         | 0.28018         | 0.05212          | 0.12616          | 0.76786  |
| H <sub>3</sub> (1s) + H <sub>4</sub> (1s)                             | 0.0             | 0.0             | 0.0             | -0.00002        | 0.06420         | 0.05050         | 0.03034         | 0.13242         | 0.48700         | 0.01534          | 0.03690          | 0.75890  |
| H <sub>5</sub> (1s)   | 0.0             | 0.0             | -0.00002        | -0.00002        | 0.00192         | 0.05052         | 0.10802         | 0.14528         | 0.02880         | 0.41928          | 0.02590          | 0.77968  |
| C <sub>1</sub> (2p <sub>z</sub> ) - C <sub>2</sub> (2p <sub>z</sub> ) | -0.00044        | 0.0             | 0.0             | 0.0             | 0.16054         | 0.05096         | 0.20438         | 0.29440         | 0.21742         | 0.07050          | -0.00264         | 0.99512  |
| C <sub>3</sub> (2p <sub>z</sub> ) - C <sub>4</sub> (2p <sub>z</sub> ) | 0.00000         | 0.0             | 0.00002         | 0.0             | 0.01642         | 0.11210         | 0.19246         | 0.31520         | 0.45192         | 0.03146          | 0.00080          | 1.12038  |
| N(2p <sub>y</sub> )   | 0.00002         | 0.00004         | 0.0             | 0.0             | 0.11498         | 0.00056         | -0.00194        | 0.11322         | 0.00608         | 0.25404          | 1.17376          | 1.66076  |
| C <sub>1</sub> (2p <sub>y</sub> ) + C <sub>2</sub> (2p <sub>y</sub> ) | -0.00008        | 0.0             | 0.0             | 0.0             | 0.00148         | 0.20116         | 0.20458         | 0.21698         | 0.04148         | 0.20728          | 0.13986          | 1.01274  |
| C <sub>3</sub> (2p <sub>y</sub> ) + C <sub>4</sub> (2p <sub>y</sub> ) | 0.00000         | 0.00002         | 0.00002         | 0.0             | 0.03636         | 0.02802         | 0.39294         | 0.12804         | 0.11496         | 0.14904          | 0.21450          | 1.06390  |
| C <sub>5</sub> (2p <sub>y</sub> )                                     | 0.0             | 0.0             | 0.0             | 0.0             | 0.01418         | 0.04098         | 0.02362         | 0.32976         | 0.00028         | 0.71408          | 0.02154          | 1.14444  |
| Total   | 2.00002         | 2.00004         | 2.00004         | 2.00000         | 2.00000         | 1.99994         | 1.99998         | 1.99988         | 2.00004         | 2.00004          | 1.99998          | 21.99996 |

 TABLE XIX  
 PYRIDINE. GROSS POPULATION ANALYSIS (B<sub>2</sub>)

| Symmetry function   | 1b <sub>2</sub> | 2b <sub>2</sub> | 3b <sub>2</sub> | 4b <sub>2</sub> | 5b <sub>2</sub> | 6b <sub>2</sub> | 7b <sub>2</sub> | Total    |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------|
| C <sub>1</sub> (1s) - C <sub>2</sub> (1s)                             | 1.99636         | 0.00090         | 0.00138         | 0.00062         | 0.0             | 0.0             | 0.0             | 1.99926  |
| C <sub>1</sub> (2s) - C <sub>2</sub> (2s)                             | 0.00270         | -0.00004        | 0.74362         | 0.30820         | 0.00146         | -0.00390        | 0.00722         | 1.05926  |
| C <sub>3</sub> (1s) - C <sub>4</sub> (1s)                             | 0.00088         | 1.99624         | 0.00128         | 0.00074         | 0.0             | 0.00002         | 0.0             | 1.99916  |
| C <sub>3</sub> (2s) - C <sub>4</sub> (2s)                             | -0.00002        | 0.00290         | 0.66958         | 0.36620         | 0.00302         | -0.00360        | 0.00864         | 1.04672  |
| H <sub>1</sub> (1s) - H <sub>2</sub> (1s)                             | 0.0             | 0.0             | 0.08190         | 0.11650         | 0.00934         | 0.27574         | 0.30518         | 0.78866  |
| H <sub>3</sub> (1s) - H <sub>4</sub> (1s)                             | 0.0             | -0.00002        | 0.07406         | 0.14388         | 0.00252         | 0.28314         | 0.30322         | 0.80680  |
| N(2p <sub>z</sub> )   | 0.00006         | 0.0             | 0.14316         | 0.37044         | 0.36822         | 0.00516         | 0.16160         | 1.04864  |
| C <sub>1</sub> (2p <sub>z</sub> ) + C <sub>2</sub> (2p <sub>z</sub> ) | -0.00002        | 0.0             | -0.00518        | 0.00040         | 0.25710         | 0.42002         | 0.41966         | 1.09198  |
| C <sub>3</sub> (2p <sub>z</sub> ) + C <sub>4</sub> (2p <sub>z</sub> ) | 0.0             | -0.00002        | -0.00476        | -0.00262        | 0.12432         | 0.46272         | 0.52766         | 1.10730  |
| C <sub>5</sub> (2p <sub>z</sub> )                                     | 0.0             | 0.00002         | 0.12302         | 0.29356         | 0.33594         | 0.00726         | 0.25124         | 1.01104  |
| C <sub>1</sub> (2p <sub>y</sub> ) - C <sub>2</sub> (2p <sub>y</sub> ) | 0.0             | 0.00002         | 0.07822         | 0.24162         | 0.38656         | 0.29742         | 0.01321         | 1.01706  |
| C <sub>3</sub> (2p <sub>y</sub> ) - C <sub>4</sub> (2p <sub>y</sub> ) | 0.00004         | 0.0             | 0.09372         | 0.16052         | 0.51160         | 0.25602         | 0.00232         | 1.02422  |
| Total   | 2.00000         | 2.00000         | 2.00000         | 2.00006         | 2.00008         | 2.00000         | 1.99996         | 14.00010 |

 TABLE XX  
 PYRIDINE. GROSS POPULATION ANALYSIS (π) A<sub>2</sub> AND B<sub>1</sub>

|   | 1b <sub>1</sub> | 2b <sub>1</sub> | 1a <sub>2</sub> |
|---|-----------------|-----------------|-----------------|
| N(2p <sub>z</sub> )   | 0.41500         | 0.59524         |                 |
| C <sub>1</sub> (2p <sub>z</sub> ) ± C <sub>2</sub> (2p <sub>z</sub> ) | 0.76454         | 0.24148         | 1.00354         |
| C <sub>3</sub> (2p <sub>z</sub> ) ± C <sub>4</sub> (2p <sub>z</sub> ) | 0.56764         | 0.44064         | 0.99648         |
| C <sub>5</sub> (2p <sub>z</sub> )                                     | 0.25284         | 0.72264         |                 |
| Total   | 2.00002         | 2.00000         | 2.00002         |

energies of the corresponding free atoms (<sup>3</sup>P for carbon and <sup>4</sup>S for nitrogen).

The second group of MO's are responsible for the bonds of the pyridine molecule. The 5a<sub>1</sub> MO has the charge concentrated mainly on the N atom, less on the C<sub>1</sub> and C<sub>2</sub> atoms, and even less on the C<sub>3</sub> and C<sub>4</sub> atoms. The charge is built up mainly of 2s functions polarized parallel to the principal symmetry axis (including 2p<sub>y</sub>). The 6a<sub>1</sub> MO has the charge distributed approximately equally on the six atoms of the ring. The 7a<sub>1</sub> MO is the first MO of a<sub>1</sub> symmetry where there is nonnegligible (0.22) electronic charge on the five hydrogen atoms. The 11a<sub>1</sub> MO is the main contributor to the lone-pair electrons on nitrogen. The 2p<sub>y</sub> population is 1.17 electrons, and the 2s population on nitrogen is 0.24 electron,

giving a net contribution of 1.41 electrons to the lone pair. However, the remaining 0.59 electron is delocalized (mainly as 2p<sub>y</sub>) on C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, and C<sub>4</sub> and somewhat on H<sub>1</sub> and H<sub>2</sub>. This computation therefore does not support the supposition that the lone-pair electrons on pyridine are about equally distributed on the N atom and on the rest of the molecule. The above values seem to indicate that our traditional ideas of a localized lone pair are not too much in error. The 4b<sub>2</sub>, the 6b<sub>2</sub>, and the 7b<sub>2</sub> account for 1.59 of the 5 electrons in the hydrogen atoms. The a<sub>1</sub> MO's (from 7a<sub>1</sub> to 11a<sub>1</sub>) account for 2.3 electrons in the hydrogen atoms. Therefore, the 5 electrons of the hydrogen atoms are present in the 14 molecular orbitals. This gives an idea of the extent of the delocalization in the σ frames.

The two π electrons of 1b<sub>1</sub> are concentrated on the N atom, less at the C<sub>1</sub> and C<sub>2</sub> positions, less at the C<sub>3</sub> and C<sub>4</sub> positions, and even less at the C<sub>5</sub> position. The trend is partially reversed in the 2b<sub>1</sub> MO. The 1a<sub>2</sub> MO has about equal distribution on the C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, and C<sub>4</sub> atoms.

Table XXI summarizes the charge migrations in the pyridine molecule. The N, C<sub>1</sub> (or C<sub>2</sub>), C<sub>3</sub> (or C<sub>4</sub>), and C<sub>5</sub> atoms have the following charge distribution.



TABLE XXI  
PYRIDINE. GROSS CHARGES—SUMMARY

| N                                   |         |                 |  |    |         |                  |          |
|-------------------------------------|---------|-----------------|--|----|---------|------------------|----------|
| 1s                                  | 1.99758 | 1s              | 1.99758  | 1s | 1.99758 | $\delta(\sigma)$ | -0.21594 |
| 2s                                  | 1.50896 | 2s              | 1.50896  | 2s | 1.50896 | $\delta(\pi)$    | -0.01024 |
| 2p <sub>x</sub>                     | 1.04864 | 2p <sub>y</sub> | 2.70940  | 2p | 3.71964 | $\delta$         | -0.22618 |
| 2p <sub>y</sub>                     | 1.66076 | 2p <sub>z</sub> | 1.01024  |    |         |                  |          |
| 2p <sub>z</sub>                     | 1.01024 |                 |  |    |         |                  |          |
| C <sub>1</sub> (or C <sub>2</sub> ) |         |                 |  |    |         |                  |          |
| 1s                                  | 1.99925 | 1s              | 1.99925  | 1s | 1.99925 | $\delta(\sigma)$ | -0.10488 |
| 2s                                  | 1.04718 | 2s              | 1.04718  | 2s | 1.04718 | $\delta(\pi)$    | -0.00478 |
| 2p <sub>x</sub>                     | 1.04355 | 2p <sub>y</sub> | 2.05845  | 2p | 3.06323 | $\delta$         | -0.10966 |
| 2p <sub>y</sub>                     | 1.01490 | 2p <sub>z</sub> | 1.00478  |    |         |                  |          |
| 2p <sub>z</sub>                     | 1.00478 |                 |  |    |         |                  |          |
| C <sub>3</sub> (or C <sub>4</sub> ) |         |                 |  |    |         |                  |          |
| 1s                                  | 1.99914 | 1s              | 1.99914  | 1s | 1.99914 | $\delta(\sigma)$ | -0.22278 |
| 2s                                  | 1.06574 | 2s              | 1.06574  | 2s | 1.06574 | $\delta(\pi)$    | -0.00238 |
| 2p <sub>x</sub>                     | 1.11384 | 2p <sub>y</sub> | 2.15790  | 2p | 3.16028 | $\delta$         | -0.22516 |
| 2p <sub>y</sub>                     | 1.04406 | 2p <sub>z</sub> | 1.00238  |    |         |                  |          |
| 2p <sub>z</sub>                     | 1.00238 |                 |  |    |         |                  |          |
| C <sub>5</sub>                      |         |                 |  |    |         |                  |          |
| 1s                                  | 1.99920 | 1s              | 1.99920  | 1s | 1.99920 | $\delta(\sigma)$ | -0.22690 |
| 2s                                  | 1.07222 | 2s              | 1.07222  | 2s | 1.07222 | $\delta(\pi)$    | +0.02452 |
| 2p <sub>x</sub>                     | 1.01104 | 2p <sub>y</sub> | 2.15548  | 2p | 3.13096 | $\delta$         | -0.20238 |
| 2p <sub>y</sub>                     | 1.14444 | 2p <sub>z</sub> | 0.97548  |    |         |                  |          |
| 2p <sub>z</sub>                     | 0.97548 |                 |  |    |         |                  |          |
| H <sub>1</sub> (or H <sub>2</sub> ) |         |                 |  |    |         |                  |          |
| 1s                                  | 0.77826 | 1s              | 0.77826  | 1s | 0.77826 | $\delta(\sigma)$ | 0.22174  |
| H <sub>3</sub> (or H <sub>4</sub> ) |         |                 |  |    |         |                  |          |
| 1s                                  | 0.78285 | 1s              | 0.78285  | 1s | 0.78285 | $\delta(\sigma)$ | 0.21715  |
| H <sub>5</sub>                      |         |                 |  |    |         |                  |          |
| 1s                                  | 0.77968 | 1s              | 0.77968  | 1s | 0.77968 | $\delta(\sigma)$ | 0.22032  |
|                                     |         | N               | 1s <sup>2</sup> 2s <sup>1.51</sup> 2p <sub>x</sub> <sup>2.71</sup> 2p <sub>y</sub> <sup>1.01</sup> |    |         |                  |          |
|                                     |         | C <sub>1</sub>  | 1s <sup>2</sup> 2s <sup>1.05</sup> 2p <sub>x</sub> <sup>2.06</sup> 2p <sub>y</sub> <sup>1.00</sup> |    |         |                  |          |
|                                     |         | C <sub>3</sub>  | 1s <sup>2</sup> 2s <sup>1.06</sup> 2p <sub>x</sub> <sup>2.16</sup> 2p <sub>y</sub> <sup>1.00</sup> |    |         |                  |          |
|                                     |         | C <sub>5</sub>  | 1s <sup>2</sup> 2s <sup>1.07</sup> 2p <sub>x</sub> <sup>2.16</sup> 2p <sub>y</sub> <sup>0.97</sup> |    |         |                  |          |

The hybridization of the carbon atoms is about exactly the s<sup>1</sup>p<sup>2</sup> trigonal hybridization. There is no two-way charge transfer on the nitrogen atom, as encountered in pyrrole.

The nitrogen atom is negatively charged by 0.22 electron, C<sub>1</sub> by 0.11 electron, C<sub>3</sub> by 0.22 electron, and C<sub>5</sub> by 0.20 electron. Each of the five hydrogen atoms has lost 0.22 electron.

Clearly the nitrogen atoms in pyrrole and pyridine are in a different situation, one being bound to two carbon and one hydrogen atoms (pyrrole), the other only to two carbon atoms (pyridine). Therefore, the bond is N-H in the first case and a lone pair in the second case. In addition, in pyrrole the N atom brings two electrons to the  $\pi$  system, whereas in pyridine it is only one. Therefore, in pyrrole the N atom is a  $\pi$  donor, whereas in the pyridine molecule the N atom retains all its charge (actually it is very slightly a  $\pi$  acceptor). The N atom is a  $\sigma$  acceptor in both molecules, but the N-H bond in pyrrole (which is partly N<sup>+</sup>H<sup>-</sup>) makes the N atom a stronger acceptor than in pyridine. All this can be obtained quantitatively by simple inspection of the population analyses of pyrrole<sup>1</sup> and pyridine.

In the following we shall attempt to correlate the MO's of pyrrole with the MO's of pyridine. This can be done rather simply by noticing that to a very gross approximation pyridine can be transformed into pyrrole by removal of C<sub>5</sub> and H<sub>5</sub> and by adding a hydrogen atom to the nitrogen. Therefore, by looking at the gross charge tables we should see which are the most affected orbitals in this transformation. On symmetry arguments it is clear that the pyridine electronic configuration transforms itself into the pyrrole configuration by subtracting two MO's of a<sub>1</sub> symmetry and one of b<sub>2</sub> symmetry. The first MO of a<sub>1</sub> symmetry to be subtracted is the 3a<sub>1</sub>. The removal of C<sub>5</sub> will affect mostly the 6a<sub>1</sub> and the 7a<sub>1</sub> MO's; since there are four electrons in these two MO's and we wish to remove only two, we can think that the 6a<sub>1</sub> and the 7a<sub>1</sub> MO's will coalesce into one with an orbital energy intermediate between the orbital energies of 6a<sub>1</sub> and 7a<sub>1</sub>. This lowering in orbital energy will bring about a lowering of the full remaining set of a<sub>1</sub> orbitals. In particular the 11a<sub>1</sub> MO will be lowered quite drastically because the lone pair is now replaced by a  $\sigma$  bond. The most affected orbitals in the b<sub>2</sub> symmetry upon removal of C<sub>5</sub> are the 4b<sub>2</sub> and 5b<sub>2</sub>. A removal of one MO from this symmetry will bring the two into one as in the case of the a<sub>1</sub> symmetry. It is therefore possible to correlate these two molecules rather simply. For the  $\pi$  electrons there is no difficulty since there are six  $\pi$  electrons in both molecules: four in the b<sub>1</sub> symmetry and two in the a<sub>2</sub> symmetry. This type of correlation suggests the possibility of transferability of bond charges between molecules.

## VII. PYRIDINE ION, <sup>2</sup>A<sub>2</sub>

Here we shall consider the wave function for one of the positive ions of the pyridine molecule. The ion is obtained by the process a<sub>2</sub>(2) → a<sub>2</sub>(1), whereby one of the two  $\pi$  electrons in the a<sub>2</sub> molecular orbital is ionized. The resulting function is therefore a doublet, <sup>2</sup>A<sub>2</sub>. The geometry of the pyridine ion used in the computation is given in Table XVI.

Before analyzing the results of the computation, it seems worthwhile to state the limitations of this work and to clarify some confusion existing in the current literature concerning the use of computed SCF ground-state functions for prediction of excited states and ionization potentials.

The basis set we have used is quite small and insufficient to obtain a Hartree-Fock limit (*i.e.*, the best possible single determinant function with a basis set of such nature that any more extended or different basis set will not produce any difference in the computed density distribution). Our program can handle much larger basis sets than the one adopted. However, we have used consistently this size basis set for pyrrole, pyridine, pyrazine, and pyridazine, and we are now

working on substantially larger molecules. The choice of the size was so as to ensure standardization in a relatively large number of molecules. Whereas for the molecules presented up to now in this series of papers, we are far below the computer program limit; the same will not hold in future papers of this series.

It is a rather standard practice in quantum chemistry to attempt to extract from a given computation as much information as feasible. For example, the orbital energies obtained from self-consistent-field ground-state functions are compared with ionization potentials, and the orbital energies of virtual orbitals are often used for prediction of excitation energies. These attempts are correct as long as it is fully realized that we are attempting to obtain semiquantitative information from quantitative computations. To put it differently, the attempts quoted above are questionable on a theoretical ground since the ionization potential as well as the excitation energies involve necessarily quantities which cannot be extracted from the ground-state wave functions alone.

Let us analyze the rather old assumptions which were made when one equates an orbital energy with an ionization potential. One standard argument was that the ground-state wave function (let us take the case of a closed-shell ground state),  ${}^1\psi_0$ , and an ionic state wave function, indicated as  ${}^2\psi_i$  (where the index  $i$  designates the  $i$ th MO doubly filled in  ${}^1\psi_0$  and singly occupied in the double state  ${}^2\psi_i$ ) differs formally by  $-\epsilon_i$ , the orbital energy of the  $i$ th MO of  ${}^1\psi_0$ . Formally, one can write that the ground-state energy,  $E({}^1\psi_0)$ , and the energy of the ion,  $E({}^2\psi_i)$ , are related by the equation

$$E({}^2\psi_i) - E({}^1\psi_0) = -\epsilon_i = \text{IP}_i$$

where  $\text{IP}_i$  is the  $i$ th ionization potential. The standard argument is that if  ${}^2\psi_i$  and  ${}^1\psi_0$  are *reasonable approximations* to the exact wave functions of the corresponding electronic states,  $-\epsilon_i$  should give a reasonable approximation to the experimental ionization potential. However, we know that the Hartree-Fock technique often gives unreasonable results. It seems, therefore, useful to study this matter somewhat more rigorously.

In the following we shall restrict ourselves to vertical ionization potentials; namely, we shall consider the ion when it has the same geometry as the ground-state molecule. We shall use the following symbols: the orbitals 1 to  $n$  are doubly occupied in the ground state (assumed to be a singlet state for simplicity) with wave function  ${}^1\psi_0$ ; an index  $t$  refers to any orbital in the  $n$  set; an index  $u$  will refer to an empty orbital (virtual orbital). Let us consider, in addition to  ${}^1\psi_0$ , a doublet positive ion,  ${}^2\psi_{i1}$ , where the subscript  $t1$  indicates that in the  $n$  set of orbital 1, the  $t$ th one has only one electron.

By definition, the exact energy of these states is the sum of the Hartree-Fock, relativistic, and correlation energy, or

$$E({}^1\psi_0) = E_{\text{HF}}({}^1\psi_0) + E_{\text{R}}({}^1\psi_0) + E_{\text{c}}({}^1\psi_0) \quad (55)$$

$$E({}^2\psi_{i1}) = E_{\text{HF}}({}^2\psi_{i1}) + E_{\text{R}}({}^2\psi_{i1}) + E_{\text{c}}({}^2\psi_{i1}) \quad (56)$$

The exact (or experimental) ionization potential (IP) is

$$\text{IP}_i = E({}^2\psi_{i1}) - E({}^1\psi_0) \quad (57)$$

Equation 3 requires knowledge of the relativistic correction. From our analysis of the relativistic correction in atoms, we can conclude that, if the ionization or the excitation of the electrons in the ground state refers to an *inner-shell* electron, the relativistic correction cannot be ignored. Quantitatively, removal of the 1s electrons from Be, Ne, Mg, and Ar brings about a variation in the relativistic energy of the order of 0.001, 0.003, 0.228, and 1.217 au (or 0.027, 9.24, 6.20, and 33.11 eV). One can interpolate between the above values; however, one should remember that removal of inner-shell electrons causes a nonnegligible rearrangement of the remaining inner-shell electrons and, by consequence, of the valency electrons. Again, from our analysis of the relativistic correction in atoms one learns that the inner core electrons contribute about equally in the neutral atom or ion. Quantitatively the 1s contributions to the relativistic corrections for  $\text{Ar}^{16+}$ ,  $\text{Ar}^{14+}$ ,  $\text{Ar}^{8+}$ ,  $\text{Ar}^{6+}$ , and neutral Ar are  $-1.255$ ,  $-1.248$ ,  $-1.229$ ,  $-1.221$ , and  $-1.271$  au, respectively.

For the valency electron the relativistic energy correction is clearly much less important. Quantitatively the relativistic corrections for the one 2s electron in Be, one 2p electron in Ne, one 3s electron in Mg, and one 3p electron in Ar are of the order of  $-0.00008$ ,  $-0.00167$ ,  $-0.00068$ , and  $-0.00374$  au, respectively.

From this analysis one can estimate the relativistic correction in molecules, case by case. It is not too bad an approximation to use the gross population analysis data relative to a given molecule and then transfer atomic data for the relativistic correction. For example, if a molecular computation yields a wave function with the following gross population for a carbon atom, 1s(2)2s(1)2p(2)2p(1), then we can assume the following relativistic correction:  $2 \times 0.00609 + 1 \times 0.00066 + 3 \times 0.00014 = 0.01326$  au, where the data are obtained from computation of the  ${}^3\text{P}$  state of the carbon atom. Had we used data from the  ${}^1\text{S}$  carbon atom, we would have  $2 \times 0.00607 + 1 \times 0.00067 + 3 \times 0.00014 = 0.01323$  au. It is noted that the 2s and 2p orbitals in the  ${}^1\text{S}$  excited state are quite different (more extended away from the nucleus) from those of the  ${}^3\text{P}$  state. Therefore, the above difference in the estimate for the carbon atom in the  $\text{sp}^2$  trigonal hybrid is indicative of the type of error one would accept in using atomic data in order to estimate molecular relativistic correction for low  $Z$  atoms.

Finally, we would like to point out that the approximation here introduced fails totally for internuclear distances much shorter than the equilibrium distances.

In addition, if the population analysis indicates the existence of ionic character, then the use of neutral atomic relativistic correction data is questionable to a degree proportional to the extent of the ionic character.

In addition to knowledge of the relativistic correction, eq 57 requires knowledge of the correlation-energy correction. The following argument can be presented. First, in the ground-state function  ${}^1\psi_0$ , there are  $n$  pairs (since there are  $n$  orbitals doubly filled) and  $n(n-1)/2$  pair-pair interactions. In general, the correlation energy of a pair is larger than the correlation energy associated with a pair-pair interaction. However, the sum of the pair-pair interaction might very well be larger than the individual pair correlation energy (this depending on case by case and on the size of  $n$ ). Let us indicate with  $\eta_t$  the pair correlation energy of electron pair in the  $t$ th orbital and with  $2\eta_{tt'}$  the pair-pair interaction between the pair of electrons in the orbital  $t$  and the pair of electrons in the orbital  $t'$ . Assuming, but not too seriously, that the  $\eta_t$  and  $\eta_{tt'}$  are constant in the various states, and assuming for simplicity that we are dealing with a case where the difference in the relativistic correction between neutral state and excited or ionic states is negligible, then we have

$$\text{IP}_{\text{a}} = \Delta E_{\text{HF}}({}^2\psi_{\text{a}}; {}^1\psi_0) - (\eta_t + \sum_{t'} \eta_{tt'}) \quad (58)$$

since

$$E_c({}^1\psi_0) = \sum_{t'} \eta_{t'} + \sum_{t'} \sum_{t''} \eta_{t't''} (1 - \delta_{t't''}) \quad (59)$$

$$E_c({}^2\psi_{\text{a}}) = \sum_{t'} \eta_{t'} + \sum_{t'} \sum_{t''} \eta_{t't''} (1 - \delta_{t't''}) - \eta_t - \sum_{t'} \eta_{tt'} \quad (60)$$

It is noted that  $\eta_{t'} = 2\eta_{tt'}$  and that the notation  $\Delta E_{\text{HF}}(a; b)$  stands for  $E_{\text{HF}}(b) - E_{\text{HF}}(a)$ . In eq 59 and 60, higher order correlation than the pair-pair are ignored. In addition, the angular dependency of the correlation energy is also ignored, in what concerns differences between  ${}^1\psi_0$  and  ${}^2\psi_{\text{a}}$ .

The correlation energy per pair varies between somewhat more than 1 to about 2 eV. Therefore, it is clear that in comparing Hartree-Fock computations for neutral ground-state molecules with positive ions or excited states, one should not expect agreement with experimental values. *If and when agreement exists between orbital energies and ionization potentials, the reason is not that the Hartree-Fock function is an adequate approximation of the exact functions but results from a cancellation of errors.*

In the remainder of this section we shall indicate where and when such cancellation is to be expected. This brings about a brief analysis of the Hartree-Fock energy contribution to the ionization potential.

The total energy for the ground state of the neutral

closed-shell molecule in the Hartree-Fock approximation is given by

$$E_{\text{HF}}({}^1\psi_0) = \sum_{t'} (H_{t'} + \epsilon_{t'}) \quad (61)$$

and the ground state for the positive ion is given by

$$E_{\text{HF}}({}^2\psi_{t'}) = \sum_{t''} (H_{t''}^+ + \epsilon_{t''}^+) - \epsilon_{t'}^+ \quad (62)$$

where the superscript  $+$  refers to quantities obtained from a Hartree-Fock solution for the neutral molecule.<sup>67</sup> Therefore, the ionization potential (excluding relativistic correction), assuming that the pair correlation as well as the pair-pair correlation are constant from neutral molecule to the ion, and assuming correlation interaction can be ignored, is

$$\begin{aligned} \text{IP}_{\text{a}} = & \sum_{t'} (H_{t'}^+ + \epsilon_{t'}^+ - \epsilon_t^+ - \sum_{t''} (H_{t''} + \epsilon_{t''}) - \\ & \eta_t - \sum_{t''} \eta_{tt''}) = \sum_{t'} (H_{t'}^+ + \epsilon_{t'}^+ - H_{t'} - \epsilon_{t'}) + \\ & H_t^+ + \epsilon_t - \epsilon_t - H_t - \epsilon_t - \eta_t - \sum_{t''} \eta_{tt''} \quad (63) \end{aligned}$$

where the summation  $\sum_{t''}^t$  indicates that  $t''$  can assume any value from 1 to  $n$  excluding  $t$ . The assumption that  $\text{IP}_{\text{a}}$  can be represented by  $-\epsilon_t$  is equivalent to the assumption, derived from eq 56, that a quantity  $\delta$ , here defined, is zero.

$$\begin{aligned} \theta = & \sum_{t'} (H_{t'}^+ + \epsilon_{t'}^+ - H_{t'} - \epsilon_{t'}) + \\ & H_t^+ - H_t - \eta_t - \sum_{t''} \eta_{tt''} = 0 \quad (64) \end{aligned}$$

Since eq 64 has been derived using the ground-state function both for the neutral molecule and the ion, clearly only the correlation term remains. In the following we analyze the term  $\delta$  assuming that the wave function for the ion and the neutral molecule are exactly identical for all the orbitals with exception of the one where ionization takes place. This assumption is, therefore, equivalent to stating

$$\sum_{t''}^t (H_{t''}^+ + \epsilon_{t''}^+ - H_{t''} - \epsilon_{t''}) = 0 \quad (65)$$

However, a further assumption that the electronic distribution of the orbital from which the electron is ejected is unchanged, whether it is singly or doubly occupied, is rather difficult to accept in general; therefore, one can hardly accept that  $H_t^+ - H_t = 0$ . In addition, it is noted that  $(-\eta_t)$  is, in general, a positive quantity (and so is  $-\sum_{t''} \eta_{tt''}$ ). The integral  $H_t^+$  and  $H_t$  are the sum of the kinetic and electronic-nuclear attraction energy for the orbital  $t$ . The positive ion tends to shrink the electronic cloud (since the Coulomb field of the nuclear charge is unbalanced and has one positive charge in excess): this is equivalent to increasing the kinetic energy and the electron-nuclear attraction. Therefore,  $H_t^+$  can be larger in absolute value

TABLE XXII  
 PYRIDINE ION. GROSS POPULATION ANALYSIS (A<sub>1</sub>)

| Symmetry function   | 1a <sub>1</sub> | 2a <sub>1</sub> | 3a <sub>1</sub> | 4a <sub>1</sub> | 5a <sub>1</sub> | 6a <sub>1</sub> | 7a <sub>1</sub> | 8a <sub>1</sub> | 9a <sub>1</sub> | 10a <sub>1</sub> | 11a <sub>1</sub> | Total    |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|----------|
| N(1s)   | 1.99366         | 0.0             | 0.0             | 0.0             | 0.00211         | 0.00091         | 0.00041         | 0.00005         | 0.00012         | 0.00005          | 0.00021          | 1.99752  |
| N(2s)   | 0.00717         | -0.00001        | 0.0             | 0.0             | 0.72094         | 0.27771         | 0.11380         | 0.04062         | 0.01701         | 0.04929          | 0.23704          | 1.51357  |
| C <sub>1</sub> (1s) + C <sub>2</sub> (1s)                             | 0.0             | 1.99708         | 0.00052         | 0.0             | 0.00106         | 0.00001         | 0.00042         | 0.00002         | 0.00016         | 0.00005          | 0.00004          | 1.99936  |
| C <sub>1</sub> (2s) + C <sub>2</sub> (2s)                             | -0.00034        | 0.00242         | -0.00006        | 0.0             | 0.59702         | 0.00017         | 0.25920         | 0.05324         | 0.14802         | 0.00451          | 0.02017          | 1.08435  |
| C <sub>3</sub> (1s) + C <sub>4</sub> (1s)                             | 0.0             | 0.00052         | 1.99621         | 0.00062         | 0.00033         | 0.00111         | 0.00013         | 0.0             | 0.00020         | 0.00007          | 0.00003          | 1.99922  |
| C <sub>3</sub> (2s) + C <sub>4</sub> (2s)                             | 0.0             | -0.00002        | 0.00276         | -0.00008        | 0.21120         | 0.62562         | 0.05873         | 0.02822         | 0.16693         | 0.03481          | 0.00507          | 1.13324  |
| C <sub>5</sub> (1s)   | 0.0             | 0.0             | 0.00081         | 1.99663         | 1.00009         | 0.00099         | 0.00063         | 0.00004         | 0.00014         | 0.00001          | 0.0              | 1.99914  |
| C <sub>5</sub> (2s)   | 0.0             | 0.0             | -0.00004        | 0.00281         | 0.06018         | 0.53747         | 0.32182         | 0.05694         | 0.05103         | 0.04480          | 0.00055          | 1.07556  |
| H <sub>1</sub> (1s) + H <sub>2</sub> (1s)                             | 0.0             | 0.0001          | 0.0             | 0.0             | 0.02264         | 0.00051         | 0.06650         | 0.15453         | 0.27337         | 0.02594          | 0.11214          | 0.65564  |
| H <sub>3</sub> (1s) + H <sub>4</sub> (1s)                             | 0.0             | 0.0             | 0.00001         | -0.00001        | 0.00575         | 0.04190         | 0.01974         | 0.13265         | 0.39209         | 0.02689          | 0.03193          | 0.65095  |
| H <sub>5</sub> (1s)   | 0.0             | 0.0             | -0.00001        | 0.00002         | 0.00151         | 0.03492         | 0.07214         | 0.10031         | 0.01596         | 0.44465          | 0.02336          | 0.69286  |
| C <sub>1</sub> (2p <sub>x</sub> ) - C <sub>2</sub> (2p <sub>x</sub> ) | -0.00046        | 0.00001         | 0.0             | 0.0             | 0.16937         | 0.06128         | 0.22066         | 0.29520         | 0.27277         | 0.04621          | -0.00346         | 1.06158  |
| C <sub>3</sub> (2p <sub>x</sub> ) - C <sub>4</sub> (2p <sub>x</sub> ) | 0.0             | 0.0             | 0.00001         | -0.00002        | 0.02085         | 0.11957         | 0.23236         | 0.32776         | 0.48072         | 0.01381          | 0.00149          | 1.19655  |
| N(2p <sub>y</sub> )   | 0.00002         | 0.00003         | 0.0             | 0.0             | 0.12714         | 0.00343         | -0.00331        | 0.10377         | 0.01807         | 0.18788          | 1.19766          | 1.63469  |
| C <sub>1</sub> (2p <sub>y</sub> ) + C <sub>2</sub> (2p <sub>y</sub> ) | -0.00006        | 0.0             | -0.00002        | 0.0             | -0.00008        | 0.22224         | 0.20412         | 0.25322         | 0.07797         | 0.16619          | 0.11047          | 1.03405  |
| C <sub>3</sub> (2p <sub>y</sub> ) + C <sub>4</sub> (2p <sub>y</sub> ) | 0.0             | 0.0             | 0.0             | 0.00002         | 0.04342         | 0.01812         | 0.43130         | 0.19136         | 0.08451         | 0.11839          | 0.18937          | 1.07649  |
| C <sub>5</sub> (2p <sub>y</sub> )                                     | 0.0             | -0.00001        | -0.00001        | 0.0             | 0.01644         | 0.05402         | 0.00134         | 0.26205         | 0.00092         | 0.83647          | 0.02392          | 1.19514  |
| Total   | 1.99999         | 2.00003         | 1.99998         | 1.99999         | 1.99997         | 1.99998         | 1.99999         | 1.99998         | 1.99999         | 2.00002          | 1.99999          | 21.99991 |

 TABLE XXIII  
 PYRIDINE ION. GROSS POPULATION ANALYSIS (B<sub>2</sub>)

| Symmetry function   | 1b <sub>2</sub> | 2b <sub>2</sub> | 3b <sub>2</sub> | 4b <sub>2</sub> | 5b <sub>2</sub> | 6b <sub>2</sub> | 7b <sub>2</sub> | Total    |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------|
| C <sub>1</sub> (1s) - C <sub>2</sub> (1s)                             | 1.99686         | 0.00057         | 0.00133         | 0.00061         | 0.0             | 0.0             | 0.0             | 1.99937  |
| C <sub>1</sub> (2s) - C <sub>2</sub> (2s)                             | 0.00249         | -0.00005        | 0.75512         | 0.33377         | 0.00025         | 0.00202         | 0.01249         | 1.10609  |
| C <sub>3</sub> (1s) - C <sub>4</sub> (1s)                             | 0.00057         | 1.99672         | 0.00125         | 0.00072         | 0.00001         | 0.00001         | 0.0             | 1.99928  |
| C <sub>3</sub> (2s) - C <sub>4</sub> (2s)                             | -0.00004        | 0.00271         | 0.67537         | 0.39492         | 0.00992         | 0.00132         | 0.01193         | 1.09613  |
| H <sub>1</sub> (1s) - H <sub>2</sub> (1s)                             | 0.00005         | 0.0             | 0.06292         | 0.09729         | 0.00621         | 0.23498         | 0.26678         | 0.66823  |
| H <sub>3</sub> (1s) - H <sub>4</sub> (1s)                             | 0.0             | 0.00003         | 0.05660         | 0.11887         | 0.00427         | 0.23796         | 0.27254         | 0.69027  |
| N(2p <sub>x</sub> )   | 0.00008         | 0.0             | 0.13062         | 0.37491         | 0.35506         | 0.00194         | 0.14664         | 1.00925  |
| C <sub>1</sub> (2p <sub>x</sub> ) + C <sub>2</sub> (2p <sub>x</sub> ) | -0.00002        | 0.0             | -0.00676        | 0.00218         | 0.26870         | 0.47298         | 0.45314         | 1.19022  |
| C <sub>3</sub> (2p <sub>x</sub> ) + C <sub>4</sub> (2p <sub>x</sub> ) | 0.0             | -0.00001        | -0.00587        | -0.00304        | 0.13037         | 0.49994         | 0.58769         | 1.20908  |
| C <sub>5</sub> (2p <sub>x</sub> )                                     | 0.0             | 0.00005         | 0.10998         | 0.27895         | 0.33627         | 0.00180         | 0.23123         | 0.95828  |
| C <sub>1</sub> (2p <sub>y</sub> ) - C <sub>2</sub> (2p <sub>y</sub> ) | 0.0             | 0.0             | 0.10169         | 0.24899         | 0.37266         | 0.29503         | 0.01683         | 1.03520  |
| C <sub>3</sub> (2p <sub>y</sub> ) - C <sub>4</sub> (2p <sub>y</sub> ) | 0.00002         | 0.0             | 0.11776         | 0.15184         | 0.51627         | 0.25200         | 0.00073         | 1.03860  |
| Total   | 2.00001         | 2.00002         | 2.00001         | 2.00001         | 1.99999         | 1.99998         | 2.00000         | 14.00001 |

than  $H_i$ . Since  $H$  is a negative quantity we have from eq 64 and 65 that

$$\theta = (H_i^+ - H_i) - \eta_i + \sum_j (\eta_{ij}) \quad (66)$$

This is essentially the sum of a negative quantity (the first parentheses) and a positive quantity (the second parentheses). Whenever the  $\epsilon_i$  happens to nearly coincide with the experimental ionization potential, it simply indicates that the conditions of eq 65 and 66 are valid or that the terms of eq 66 are cancelled by the terms of eq 65. Clearly, the fulfillment of these conditions has very little to do with the ability of the Hartree-Fock approximation to reasonably represent our exact wave function. In addition, this analysis holds both for Hartree-Fock functions or for less accurate self-consistent field functions. As a corollary to this analysis, we note that since  $H_i^+$  and  $H_i$  are quantities very simple to obtain, once a Hartree-Fock function is obtained (these are part of the standard output in our molecular program), then eq 66 is a useful starting point for obtaining quantitative information on the pair and pair-pair correlation energy, provided that the ionization potentials are known.

As a second corollary we might add that it is questionable to expect agreement between computed orbital energies and exact ionization potentials for the cases where the electron is ejected from an inner orbital. In such cases eq 66 can seldom be justified.

The comparison of the wave functions for the neutral ground state of pyridine and the positive ion, obtained by ionization of one  $\pi$  electron, is performed by analyzing the electronic population data. In Tables XXII, XXIII, and XXIV we report the gross charge population for the MO's of A<sub>1</sub> and B<sub>2</sub> and for the  $\pi$  MO's, respectively. These tables are now compared with the corresponding one previously reported for the pyridine neutral molecule in its ground state (see Tables XVIII-XXI).

Comparison of the population analysis of the 11 occupied MO's of A—a symmetry indicates that there is little difference between the ion and the ground-state molecule, if we compare a single MO at a time. The 1a, 2a, and 3a are practically identical in the neutral molecule and in the ion; the 3a<sub>1</sub> and 4a<sub>1</sub> switch themselves between ion and ground state; namely, the 3a<sub>1</sub> of the ion corresponds to the 4a<sub>1</sub> of the neutral molecule and the 4a<sub>1</sub> of the ion corresponds to the 3a<sub>1</sub> of

TABLE XXIV

PYRIDINE ION. GROSS POPULATION ANALYSIS FOR  $\pi$  ELECTRONS<sup>a</sup>

| Symmetry function   |                | 1b <sub>1</sub> | 2b <sub>1</sub> | 1a <sub>2</sub> |
|---|----------------|-----------------|-----------------|-----------------|
| N(2p <sub>z</sub> )   | P              | 0.41500         | 0.59524         | ...             |
|   | P <sup>+</sup> | 0.47160         | 0.56375         | ...             |
| C <sub>1</sub> (2p <sub>z</sub> ) $\pm$ C <sub>2</sub> (2p <sub>z</sub> ) | P              | 0.76454         | 0.24148         | 1.00354         |
|   | P <sup>+</sup> | 0.76580         | 0.21812         | 0.50482         |
| C <sub>3</sub> (2p <sub>z</sub> ) $\pm$ C <sub>4</sub> (2p <sub>z</sub> ) | P              | 0.56764         | 0.44064         | 0.99648         |
|   | P <sup>+</sup> | 0.51778         | 0.49018         | 0.49518         |
| C <sub>5</sub> (2p <sub>z</sub> )   | P              | 0.25284         | 0.72264         | ...             |
|   | P <sup>+</sup> | 0.24480         | 0.72795         | ...             |

<sup>a</sup> P refers to the neutral pyridine molecule ground-state electronic gross charges; P<sup>+</sup>, to the positive ion.

tron. Therefore, the apparent 1:1 correspondence in the individual MO's of the ion and the molecule is misleading, because there is a large over-all charge rearrangement. It is noted that we are studying the effect of the removal of one electron from a  $\pi$  orbital and we find over one-half electron displacement into the  $\sigma$  electrons of A<sub>1</sub> symmetry.

Let us now analyze the B<sub>2</sub> symmetry MO's, 1b<sub>2</sub> to 7b<sub>2</sub>. Again on the basis of a 1:1 comparison, one would conclude that the variation in population distribution between the pyridine ion and the neutral molecule is rather minimal. Let us compare the 14 electrons of B<sub>2</sub> symmetry in the ion with the 14 electrons of B<sub>2</sub> symmetry in the molecule. Again we find a substantial

TABLE XXV

GROSS CHARGE REARRANGEMENT OF THE  $\sigma$  ELECTRONS FOLLOWING  $\pi$  IONIZATIONS<sup>a</sup>

| $\chi$ 's   | A <sub>1</sub> symmetry MO's |                |          | $\chi$ 's   | B <sub>2</sub> symmetry MO's |                |          |
|-------------|------------------------------|----------------|----------|-------------|------------------------------|----------------|----------|
|             | P                            | P <sup>+</sup> | $\Delta$ |             | P                            | P <sup>+</sup> | $\Delta$ |
| $\chi_1$    | 1.99758                      | 1.99752        | +0.00006 | $\chi_{18}$ | 1.99926                      | 1.99937        | +0.00011 |
| $\chi_2$    | 1.50896                      | 1.51357        | -0.00461 | $\chi_{19}$ | 1.05926                      | 1.10609        | -0.04683 |
| $\chi_3$    | 1.99924                      | 1.99936        | -0.00008 | $\chi_{20}$ | 1.99916                      | 1.99928        | +0.00012 |
| $\chi_4$    | 1.03510                      | 1.08435        | -0.04925 | $\chi_{21}$ | 1.04672                      | 1.09613        | -0.04941 |
| $\chi_5$    | 1.99912                      | 1.99922        | -0.00010 | $\chi_{22}$ | 0.78866                      | 0.66823        | +0.12043 |
| $\chi_6$    | 1.08476                      | 1.13324        | +0.04848 | $\chi_{23}$ | 0.80680                      | 0.69027        | +0.11653 |
| $\chi_7$    | 1.99920                      | 1.99914        | +0.00006 | $\chi_{24}$ | 1.04864                      | 1.00925        | +0.03939 |
| $\chi_8$    | 1.07222                      | 1.07556        | -0.00334 | $\chi_{25}$ | 1.09198                      | 1.19022        | -0.09894 |
| $\chi_9$    | 0.76786                      | 0.65564        | +0.11222 | $\chi_{26}$ | 1.10730                      | 1.20908        | -0.10178 |
| $\chi_{10}$ | 0.75890                      | 0.65095        | +0.10795 | $\chi_{27}$ | 1.01104                      | 0.95828        | +0.05276 |
| $\chi_{11}$ | 0.77968                      | 0.69286        | +0.08682 | $\chi_{28}$ | 1.01706                      | 1.03520        | -0.01814 |
| $\chi_{12}$ | 0.99512                      | 1.06158        | -0.06646 | $\chi_{29}$ | 1.02422                      | 1.03860        | -0.01438 |
| $\chi_{13}$ | 1.12038                      | 1.19655        | -0.07617 | Total       | 14.00010                     | 14.00001       | 0.65882  |
| $\chi_{14}$ | 0.66076                      | 1.63469        | +0.02607 |             |                              |                |          |
| $\chi_{15}$ | 1.01274                      | 1.03405        | -0.02131 |             |                              |                |          |
| $\chi_{16}$ | 1.06390                      | 1.07649        | -0.01259 |             |                              |                |          |
| $\chi_{17}$ | 1.14444                      | 1.19514        | -0.05070 |             |                              |                |          |
| Total       | 21.99996                     | 21.99991       | 0.66627  |             |                              |                |          |

<sup>a</sup> The first column reports the symmetry function; the second and third columns report the total electronic charge for the 11 MO's of A<sub>1</sub> symmetry for the symmetry function of the first column. The charges for the pyridine neutral molecule are in the column designated by P; those for the ion are in the column designated as P<sup>+</sup>. The fourth column is the difference in charge between P and P<sup>+</sup>. The remaining four columns of this table report equivalent data for the seven MO's of B<sub>2</sub> symmetry. The  $\chi_1$  to  $\chi_{17}$  refer to the symmetry-adapted function (in the same order as given in Table XXII) for the a<sub>1</sub> MO's. The  $\chi_{18}$  to  $\chi_{29}$  refer to the symmetry adapted function (in the same order as given in Table XXIII) for the b<sub>2</sub> MO's.

the neutral molecule. However, this switching is probably of little significance since the two MO's are about degenerate in energy both in the neutral molecule and in the ion. Again, for the 5a<sub>1</sub> (to the 11a<sub>11</sub>) in the ion there is little difference with the 5a<sub>1</sub> (to the 11a<sub>11</sub>) of the neutral molecule. Each MO has a nearly exact correspondent in both ion and neutral pyridine. Perhaps we should analyze not the individual MO's but the composite group of 22 electrons of A<sub>1</sub> symmetry. This is done in columns 1 and 2 of Table XXV, where the number of figures reported is more than what is needed for the discussion of the physical phenomena, but the excessive number of figures is given for numerical accuracy reasons. Clearly, both ion and neutral molecule have 22 electrons, and therefore the difference on the total charge in the 11 MO's is zero. Of interest is the sum of the absolute deviations. This is given in the third columns and amounts to over one-half of an elec-

charge rearrangement involving over one-half of an electron.

The  $\pi$ -electron system has some variations in the B<sub>1</sub> MO's. Namely, the 1b has about 0.06 electronic charge shifted from the carbons to the nitrogen by comparing the neutral molecule with the ion. The 2b<sub>1</sub> continues the process of 1b<sub>1</sub> and shifts some more charge ( $\sim$ 0.03) from the nitrogen to the carbons. It is interesting to note that the four  $\pi$  electrons of B<sub>1</sub> symmetry shift around about 0.22 electron. The 22 electrons in A<sub>1</sub> shift around about 0.66 electron, and 0.66 for the 14 electrons in B<sub>2</sub>. Therefore, there is about 1.3 electrons out of 36  $\sigma$  electrons which are rearranged in the  $\sigma$  group and 0.22 in the four  $\pi$  electrons of B<sub>1</sub> symmetry. However, since there are 12 inner-core electrons in the  $\sigma$  group and since some of the low-lying  $\sigma$  orbitals undergo no rearrangement in going from neutral molecule to the ion, we conclude that the rearrangements in the  $\sigma$  elec-

TABLE XXVI  
 PYRIDINE ION. GROSS CHARGE SUMMARY

|                 |        |                                   |        | N                                   |        |                  | Neutral |                     |         |
|-----------------|--------|-----------------------------------|--------|-------------------------------------|--------|------------------|---------|---------------------|---------|
| 1s              | 1.9975 | 1s                                | 1.9975 | 1s                                  | 1.9975 | $\delta(\sigma)$ | -0.1550 | $P; \delta(\sigma)$ | -0.2159 |
| 2s              | 1.5136 | 2s                                | 1.5136 | 2s                                  | 1.5136 | $\delta(\pi)$    | -0.0353 | $P; \delta(\pi)$    | -0.0102 |
| 2p <sub>x</sub> | 1.0092 | 2p <sub><math>\sigma</math></sub> | 2.6439 | 2p                                  | 3.6792 | $\delta$         | -0.1903 | $P; \delta$         | -0.2262 |
| 2p <sub>y</sub> | 1.6347 | 2p <sub><math>\pi</math></sub>    | 1.0353 |                                     |        |                  |         |                     |         |
| 2p <sub>z</sub> | 1.0353 |                                   |        |                                     |        |                  |         |                     |         |
|                 |        |                                   |        | C <sub>1</sub> (or C <sub>2</sub> ) |        |                  |         |                     |         |
| 1s              | 1.9994 | 1s                                | 1.9994 | 1s                                  | 1.9994 | $\delta(\sigma)$ | -0.2551 | $P; \delta(\sigma)$ | -0.1049 |
| 2s              | 1.0952 | 2s                                | 1.0952 | 2s                                  | 1.0952 | $\delta(\pi)$    | +0.2257 | $P; \delta(\pi)$    | -0.0048 |
| 2p <sub>x</sub> | 1.1259 | 2p <sub><math>\sigma</math></sub> | 2.1605 | 2p                                  | 2.9348 | $\delta$         | -0.0294 | $P; \delta$         | -0.1097 |
| 2p <sub>y</sub> | 1.0346 | 2p <sub><math>\pi</math></sub>    | 0.7743 |                                     |        |                  |         |                     |         |
| 2p <sub>z</sub> | 0.7743 |                                   |        |                                     |        |                  |         |                     |         |
|                 |        |                                   |        | C <sub>3</sub> (or C <sub>4</sub> ) |        |                  |         |                     |         |
| 1s              | 1.9992 | 1s                                | 1.9992 | 1s                                  | 1.9992 | $\delta(\sigma)$ | -0.3742 | $P; \delta(\sigma)$ | -0.2228 |
| 2s              | 1.1147 | 2s                                | 1.1147 | 2s                                  | 1.1147 | $\delta(\pi)$    | +0.2484 | $P; \delta(\pi)$    | -0.0024 |
| 2p <sub>x</sub> | 1.2028 | 2p <sub><math>\sigma</math></sub> | 2.2603 | 2p                                  | 3.0119 | $\delta$         | -0.1258 | $P; \delta$         | -0.2252 |
| 2p <sub>y</sub> | 1.0575 | 2p <sub><math>\pi</math></sub>    | 0.7516 |                                     |        |                  |         |                     |         |
| 2p <sub>z</sub> | 0.7516 |                                   |        |                                     |        |                  |         |                     |         |
|                 |        |                                   |        | C <sub>5</sub>                      |        |                  |         |                     |         |
| 1s              | 1.9991 | 1s                                | 1.9991 | 1s                                  | 1.9991 | $\delta(\sigma)$ | -0.2081 | $P; \delta(\sigma)$ | -0.2269 |
| 2s              | 1.0756 | 2s                                | 1.0756 | 2s                                  | 1.0756 | $\delta(\pi)$    | +0.0273 | $P; \delta(\pi)$    | +0.0245 |
| 2p <sub>x</sub> | 0.9583 | 2p <sub><math>\sigma</math></sub> | 2.1534 | 2p                                  | 3.1261 | $\delta$         | -0.1808 | $P; \delta$         | -0.2024 |
| 2p <sub>y</sub> | 1.1951 | 2p <sub><math>\pi</math></sub>    | 0.9727 |                                     |        |                  |         |                     |         |
| 2p <sub>z</sub> | 0.9727 |                                   |        |                                     |        |                  |         |                     |         |
|                 |        |                                   |        | H <sub>1</sub> (or H <sub>2</sub> ) |        |                  |         |                     |         |
| 1s              | 0.6619 | 1s                                | 0.6619 | 1s                                  | 0.6619 | $\delta$         | 0.3381  | $P; \delta$         | 0.2217  |
|                 |        |                                   |        | H <sub>3</sub> (or H <sub>4</sub> ) |        |                  |         |                     |         |
| 1s              | 0.6706 | 1s                                | 0.6706 | 1s                                  | 0.6706 | $\delta$         | 0.3334  | $P; \delta$         | 0.2171  |
|                 |        |                                   |        | H <sub>5</sub>                      |        |                  |         |                     |         |
| s               | 0.6929 | 1s                                | 0.6929 | 1s                                  | 0.6929 | $\delta$         | 0.3071  | $P; \delta$         | 0.2203  |

trons are, in absolute value, larger than the rearrangement in the  $\pi$  electrons. If we add to the 0.22  $\pi$ -electron rearrangement one charge (the one which is ionized), then the  $\pi$ -charge rearrangement is nearly equal to the  $\sigma$  rearrangement.

This conclusion should caution against the use of the  $\pi$ -electron approximation, where one would like to assume equal  $\sigma$  core and consider only the  $\pi$  electrons.

Two comments should be added. The first concerns those quantum chemists who consider the population analysis a most questionable tool for understanding what the electrons are doing in molecules. Clearly there are drawbacks in the population analysis formulation, especially if one assumes a rigorous (or in this case, a puritanical) attitude in the use of quantum theory for understanding molecules. However, the rearrangements here described are not part of any expectation value as such, but do influence any expectation value. Therefore, it seems definitely worthwhile to present this type of data, even if not too rigorous. This comment is particularly true when we shall deal with larger and larger molecules.

The second comment is a qualification on the values we have reported. As noted previously, this computation does not represent the best possible SCF function, since we are working with a rather limited basis set. Therefore, the numerical conclusions about the  $\sigma$  rearrangement should be considered somewhat provisional.

The total charges on the atoms in the pyridine ion and in the pyridine neutral molecule are given in Table XXVI where the rearrangement is given, not in terms of groups of electrons of a given symmetry, but in terms of atomic sites. In the next to the last column we report the charge transfer,  $\delta$ , for the ion. In the last column we report the charge transfer,  $\delta$ , for the neutral molecule. The nitrogen atom has accepted 0.19 electron in the ion (0.15 of  $\sigma$  type, 0.04 of  $\pi$  type). In the neutral molecule the nitrogen accepts about the same change, 0.23 electron (0.21 of  $\sigma$  type and 0.02 of  $\pi$  type). The carbon atom C<sub>1</sub> (or C<sub>2</sub>) is the nearest neighbor to the nitrogen. Here the behavior between ion and neutral molecule is quite different. In the ion molecule the carbon atom is about neutral (0.03 negative), but there is a two-way charge transfer of considerable intensity. The C<sub>1</sub> (or C<sub>2</sub>) carbon accepts 0.25  $\sigma$  electron and donates 0.22  $\pi$  electron. The C<sub>3</sub> carbon atom (or C<sub>4</sub>) again exhibits this two-way charge transfer. The C<sub>3</sub> carbon atom is over-all 0.12 electron negative, but this excess charge is obtained by accepting 0.37  $\sigma$  electron and donating 0.24  $\pi$  electron. The carbon atom opposite to the nitrogen atom, C<sub>5</sub>, is negatively charged. It accepts 0.21  $\sigma$  electron and donates 0.03  $\pi$  electron. The nitrogen and the carbon atoms are, therefore, all negatively charged in the pyridine ion. Comparison with the neutral molecule indicates that this is not sufficient to account for one electron ionized. The total

excess electronic charge on the ring atoms is 0.68 electron in the ion; this should be compared with an excess of 1.10 electrons in the atoms of the ring for the neutral molecule. Therefore, the ionization of one  $\pi$  electron has decreased the excess charge of the ring atoms from 1.10 electrons to 0.68 electron. In other words, the ring atoms contribute only 0.42 of the electron which has been ionized. The remainder is offered by the hydrogen atoms. From Table XXVI the hydrogens are much more positively charged in the ion than in the molecule.

In the neutral molecule, there was no two-way charge transfer: to make the ring atoms negative, the hydrogens supply  $\sigma$  charge that is accepted by the ring atoms. However, in the ion one  $\pi$  electron is removed from the ring atoms. This loss is equilibrated by additional  $\sigma$ -charge donation from the hydrogens, which offset partly the  $\pi$ -electron loss following ionization. This mechanism takes place mainly on C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, and C<sub>4</sub>. The nitrogen atom and C<sub>5</sub> are rather insensitive to the ionization. This is not surprising since N and C<sub>5</sub> are not included in the  $a_2$  MO which is the one ionized.

We shall first comment on the value of the ionization potential. The computed total energy for the positive ion is  $-245.19708$  or  $0.42486$  au (or  $11.46$  eV) above the computed ground-state energy. These quantities correspond to the computed one-electron particle model contribution to the vertical ionization potential. Therefore, it ignores (a) the correlation effects and (b) the relativistic effects. The latter can be ignored as indicated in a previous section of this paper. The correlation contribution to the ionization potential is, however, substantial as previously discussed. From atomic data on the first-row atoms it is about 2 eV per pair for p electrons.<sup>68</sup> However, we have at present no sufficient data on the pair-pair correction in aromatic systems. We have, therefore, to conclude that the vertical ionization potential should be in the neighborhood of  $11.46 + 2 = 13.46$  eV. With more computations of ionized states, we shall probably be in a position to give a more accurate estimate of the correlation correction. The experimental ionization potential is reported to be 9.266 eV.

It is quite simple to reconcile the two seemingly contradictory values of the computed and experimental ionization potential. First, we have to realize that the 0.66-electron rearrangement in the A<sub>1</sub> symmetry MO's (and equivalent in the B<sub>2</sub> symmetry MO's) is mainly in the sense of shifting charges from the hydrogens to the carbons and from s-type orbitals to p-type orbitals. This will bring about an increase of correlation energy in the ion, which partially offsets the pair correlation

TABLE XXVII  
MOLECULAR GEOMETRY FOR THE PYRAZINE MOLECULES<sup>a</sup>

|                | $z$       | $y$       |                | $z$       | $y$       |
|----------------|-----------|-----------|----------------|-----------|-----------|
| N <sub>1</sub> | 0.0       | 2.588950  | H <sub>1</sub> | -3.993240 | 2.343280  |
| N <sub>2</sub> | 0.0       | -2.588950 | H <sub>2</sub> | 3.993240  | 2.343280  |
| C <sub>1</sub> | -2.209380 | 1.313370  | H <sub>3</sub> | -3.993240 | -2.343280 |
| C <sub>2</sub> | 2.209380  | 1.313370  | H <sub>4</sub> | 3.993240  | -2.343280 |
| C <sub>3</sub> | -2.209380 | -1.313370 |                |           |           |
| C <sub>4</sub> | 2.209380  | -1.313370 |                |           |           |

<sup>a</sup> Distances are given in atomic units; the value of the  $z$  coordinate is 0.

energy in the  $\pi$  systems (let us indicate this quantity as  $e_1$ ). In addition, this SCF computation for the ion uses the same basis set we have selected for the neutral molecule. A more adequate set for the ion would have lowered the ion SCF total energy by a nonnegligible amount,  $e_2$ . Finally, the charge shift from the hydrogens to the carbon atoms will bring about some variation in the equilibrium distance of the ion (let us indicate with  $e_3$  the energy lowering due to this effect). Therefore, the vertical ionization potential is  $-13.46$  eV +  $e_1$  +  $e_2$  and the 0-0 ionization potential is  $-13.46$  +  $e_1$  +  $e_2$  +  $e_3$ . We estimate  $e_3 \geq e_2 > e_1$  with  $e_1$  in the neighborhood of 1 eV. Therefore, the experimental value of  $\sim 9.3$  eV for the ionization potential is in reasonable agreement with our computation and preliminary analysis.

From our discussion on the ionization potential and from our analysis of the differences in the electronic distribution between the ion and the neutral molecule, it is clear that the quantity  $\theta$  given in eq 64 is different from zero. (It is noted that  $H_{a_2}$  and  $H_{a_2}^+$  are nearly equal in the ion and in the neutral molecule:  $H_{a_2} = -9.799327$  au and  $H_{a_2}^+$  is larger by 0.0003 au). Therefore, the quantity  $\theta$  of eq 66 is certainly different from zero. It is noted that the computed  $\epsilon_{a_2}$  for the neutral molecule is  $-0.44725$  au (or  $-12.17$  eV). Assuming again a 2-eV contribution to the ionization potential from correlation effects, there is, therefore, about 0.71 eV due to the rearrangement effects. *The rearrangement involves more than one full electron and is accomplished therefore rather inexpensively.* This should caution us against equating near agreement of  $\epsilon$  and the experimental ionization potentials with the assumption of small electronic rearrangement in the molecule. It is finally noted that the energy of about 1 eV (0.71 eV) needed to accomplish the rearrangement of about one electron is quite compatible with spectroscopic excitations (below the Rydberg excitation) where again we transfer (and therefore rearrange) one electron at the expense of a few electron volts, in general.

## VIII. PYRAZINE GROUND STATE

In Table XXVII we report the molecular geometry we have used in this computation. The molecule has been computed in the C<sub>2v</sub> symmetry, and we have chosen the principal axis of symmetry as the  $y$  axis.

(68) The value of 2 eV is likely on the upper limit. It is noted that the same basis set was used for the ion and the neutral molecule. An accurately selected set would bring the computed energy difference of 11.46 eV to a somewhat smaller value.

TABLE XXVIII

TOTAL ENERGY AND ORBITAL ENERGIES FOR PYRAZINE  
(IN ATOMIC UNITS)

| A <sub>1</sub> (σ)         | B <sub>2</sub> (σ) | B <sub>1</sub> (π) | A <sub>2</sub> (π) |
|----------------------------|--------------------|--------------------|--------------------|
| -15.69544                  | -11.47649          | -0.64940           | -0.46176           |
| -15.69527                  | -11.47540          | -0.49355           |                    |
| -11.47642                  | -1.13257           |                    |                    |
| -11.47541                  | -0.93097           |                    |                    |
| -1.37619                   | -0.75519           |                    |                    |
| -1.27295                   | -0.68371           |                    |                    |
| -0.95711                   | -0.61071           |                    |                    |
| -0.77967                   |                    |                    |                    |
| -0.72282                   |                    |                    |                    |
| -0.53275                   |                    |                    |                    |
| -0.44125                   |                    |                    |                    |
| Total energy -261.55432 au |                    |                    |                    |

As previously, the wave function is analyzed by considering the orbital energies and the gross population charges on the individual atoms. The gross population for the a<sub>1</sub> MO's is given in Table XXIX: for the b<sub>2</sub> MO's, in Table XXX: and for the π-electrons MO's, in Table XXXI.

The 1a<sub>1</sub>, 2a<sub>1</sub>, 3a<sub>1</sub>, 4a<sub>1</sub>, 1b<sub>2</sub>, and 2b<sub>2</sub> are core MO's, which represent the 1s electrons for the nitrogen and for the carbon atoms. The remaining orbitals resemble those of the pyridine molecule. (The two molecules are isoelectronic: one nitrogen of pyrazine is replaced by a C-H group in pyridine.) Let us consider the MO's of symmetry A<sub>1</sub>. In pyrazine, 5a<sub>1</sub> is more stable than 5a<sub>1</sub>

TABLE XXIX

PYRAZINE. GROSS POPULATION ANALYSIS (A<sub>1</sub>)

|   | 1a <sub>1</sub> | 2a <sub>1</sub> | 3a <sub>1</sub> | 4a <sub>1</sub> | 5a <sub>1</sub> | 6a <sub>1</sub> | 7a <sub>1</sub> | 8a <sub>1</sub> | 9a <sub>1</sub> | 10a <sub>1</sub> | 11a <sub>1</sub> | Total   |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|---------|
| N <sub>1</sub> (1s)   | 1.9890          | 0.0046          | 0.0             | 0.0             | 0.0012          | 0.0017          | 0.0006          | 0.0001          | 0.0001          | 0.0001           | 0.0001           | 1.9975  |
| N <sub>1</sub> (2s)   | 0.0072          | 0.0             | 0.0             | 0.0             | 0.4118          | 0.5685          | 0.1752          | 0.0551          | 0.0155          | 0.1534           | 0.1060           | 1.4927  |
| C <sub>1</sub> (1s) + C <sub>3</sub> (1s)                             | 0.0             | 0.0             | 1.0792          | 0.9182          | 0.0007          | 0.0004          | 0.0004          | 0.0             | 0.0002          | 0.0001           | 0.0              | 1.9992  |
| C <sub>1</sub> (2s) + C <sub>3</sub> (2s)                             | -0.0003         | 0.0             | 0.0015          | 0.0011          | 0.4002          | 0.1703          | 0.2494          | 0.0311          | 0.1590          | 0.0381           | -0.0006          | 1.0498  |
| C <sub>1</sub> (1s) + C <sub>3</sub> (1s)                             | 0.0             | 0.0             | 0.9180          | 1.0794          | 0.0007          | 0.0004          | 0.0004          | 0.0             | 0.0002          | 0.0001           | 0.0              | 1.9992  |
| C <sub>3</sub> (2s) + C <sub>4</sub> (2s)                             | 0.0             | -0.0003         | 0.0013          | 0.0013          | 0.3998          | 0.1709          | 0.2494          | 0.0311          | 0.1592          | 0.0381           | -0.0006          | 1.0502  |
| N <sub>2</sub> (1s)   | 0.0046          | 1.9890          | 0.0             | 0.0             | 0.0012          | 0.0018          | 0.0006          | 0.0001          | 0.0001          | 0.0001           | 0.0001           | 1.9976  |
| N <sub>2</sub> (2s)   | 0.0             | 0.0072          | 0.0             | 0.0             | 0.4109          | 0.5694          | 0.1752          | 0.0551          | 0.0154          | 0.1536           | 0.1059           | 1.4927  |
| H <sub>1</sub> (1s) + H <sub>3</sub> (1s)                             | 0.0             | 0.0             | 0.0             | 0.0             | 0.0148          | 0.0123          | 0.0652          | 0.1769          | 0.3697          | 0.0149           | 0.1127           | 0.7665  |
| H <sub>3</sub> (1s) + H <sub>4</sub> (1s)                             | 0.0             | 0.0             | 0.0             | 0.0             | 0.0148          | 0.0124          | 0.0652          | 0.1765          | 0.3705          | 0.0149           | 0.1128           | 0.7671  |
| C <sub>1</sub> (2p <sub>x</sub> ) - C <sub>3</sub> (2p <sub>x</sub> ) | -0.0005         | 0.0             | 0.0             | 0.0             | 0.0937          | 0.1119          | 0.1953          | 0.2636          | 0.3290          | 0.0021           | -0.0018          | 0.9933  |
| C <sub>3</sub> (2p <sub>x</sub> ) - C <sub>4</sub> (2p <sub>x</sub> ) | 0.0             | -0.0005         | 0.0             | 0.0             | 0.0937          | 0.1122          | 0.1952          | 0.2626          | 0.3298          | 0.0022           | -0.0019          | 0.9933  |
| N <sub>1</sub> (2p <sub>y</sub> )                                     | 0.0             | 0.0             | 0.0             | 0.0             | 0.0734          | 0.0427          | -0.0031         | 0.1705          | 0.0300          | 0.7866           | 0.5815           | 1.6816  |
| C <sub>1</sub> (2p <sub>y</sub> ) + C <sub>3</sub> (2p <sub>y</sub> ) | 0.0001          | 0.0             | 0.0             | 0.0             | 0.0048          | 0.0913          | 0.3169          | 0.3041          | 0.0955          | 0.0041           | 0.2022           | 1.0190  |
| C <sub>3</sub> (2p <sub>y</sub> ) + C <sub>4</sub> (2p <sub>y</sub> ) | 0.0             | -0.0001         | 0.0             | 0.0             | 0.0050          | 0.0911          | 0.3174          | 0.3031          | 0.0957          | 0.0041           | 0.2026           | 1.0189  |
| N <sub>2</sub> (2p <sub>y</sub> )                                     | 0.0             | 0.0             | 0.0             | 0.0             | 0.0732          | 0.0428          | -0.0031         | 0.1700          | 0.0302          | 0.7875           | 0.5811           | 1.6817  |
| Total   | 2.0001          | 1.9999          | 2.0000          | 2.0000          | 1.9999          | 2.0001          | 2.0002          | 1.9999          | 2.0001          | 2.0000           | 2.0001           | 22.0003 |

TABLE XXX

PYRAZINE. GROSS POPULATION ANALYSIS (B<sub>2</sub>)

|   | 1b <sub>2</sub> | 2b <sub>2</sub> | 3b <sub>2</sub> | 4b <sub>2</sub> | 5b <sub>2</sub> | 6b <sub>2</sub> | 7b <sub>2</sub> | Total   |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------|
| C <sub>1</sub> (1s) - C <sub>3</sub> (1s)                             | 1.0736          | 0.9237          | 0.0013          | 0.0007          | 0.0             | 0.0             | 0.0             | 1.9993  |
| C <sub>1</sub> (2s) - C <sub>3</sub> (2s)                             | 0.0015          | 0.0011          | 0.7061          | 0.3386          | 0.0             | -0.0026         | 0.0167          | 1.0614  |
| C <sub>3</sub> (1s) - C <sub>4</sub> (1s)                             | 0.9235          | 1.0738          | 0.0013          | 0.0007          | 0.0             | 0.0             | 0.0             | 1.9993  |
| C <sub>3</sub> (2s) - C <sub>4</sub> (2s)                             | 0.0013          | 0.0013          | 0.7061          | 0.3388          | 0.0             | -0.0026         | 0.0167          | 1.0616  |
| H <sub>1</sub> (1s) - H <sub>3</sub> (1s)                             | 0.0             | 0.0             | 0.0715          | 0.1118          | 0.0078          | 0.2728          | 0.3151          | 0.7790  |
| H <sub>3</sub> (1s) - H <sub>4</sub> (1s)                             | 0.0             | 0.0             | 0.0715          | 0.1118          | 0.0076          | 0.2722          | 0.3159          | 0.7790  |
| N <sub>1</sub> (2p <sub>x</sub> )                                     | 0.0             | 0.0             | 0.1446          | 0.3624          | 0.3658          | -0.0016         | 0.1854          | 1.0566  |
| C <sub>1</sub> (2p <sub>x</sub> ) + C <sub>3</sub> (2p <sub>x</sub> ) | 0.0             | 0.0             | -0.0073         | -0.0009         | 0.2619          | 0.3636          | 0.4781          | 1.0954  |
| C <sub>3</sub> (2p <sub>x</sub> ) - C <sub>4</sub> (2p <sub>x</sub> ) | 0.0             | 0.0             | -0.0073         | -0.0009         | 0.2623          | 0.3626          | 0.4786          | 1.0953  |
| N <sub>2</sub> (2p <sub>x</sub> )                                     | 0.0             | 0.0             | 0.1445          | 0.3624          | 0.3661          | -0.0016         | 0.1854          | 1.0568  |
| C <sub>1</sub> (2p <sub>y</sub> ) - C <sub>3</sub> (2p <sub>y</sub> ) | 0.0             | 0.0             | 0.0837          | 0.1874          | 0.3640          | 0.3686          | 0.0040          | 1.0077  |
| C <sub>3</sub> (2p <sub>y</sub> ) - C <sub>4</sub> (2p <sub>y</sub> ) | 0.0             | 0.0             | 0.0839          | 0.1871          | 0.3643          | 0.3685          | 0.0041          | 1.0079  |
| Total   | 1.9999          | 1.9999          | 1.9999          | 1.9999          | 1.9998          | 1.9999          | 2.0000          | 13.9993 |

The second axis in the molecular plane is the *x* axis, and the axis perpendicular to the molecular plane is the *z* axis. This choice of axes and symmetry has been made in order to allow easy comparison with the pyrrole and pyridine computations. The basis set used in this computation is given in Tables VIII and IX.

The total energy and the orbital energies are given in Table XXVIII. The electronic configuration of pyrazine is given below.

|             |  |
|-------------|--|
| σ electrons | 1a <sub>1</sub> <sup>2</sup> 2a <sub>1</sub> <sup>2</sup> 3a <sub>1</sub> <sup>2</sup> 4a <sub>1</sub> <sup>2</sup> 5a <sub>1</sub> <sup>2</sup> 6a <sub>1</sub> <sup>2</sup> 7a <sub>1</sub> <sup>2</sup> 8a <sub>1</sub> <sup>2</sup> 9a <sub>1</sub> <sup>2</sup> 10a <sub>1</sub> <sup>2</sup> 11a <sub>1</sub> <sup>2</sup> |
| σ electrons | 1b <sub>2</sub> <sup>2</sup> 2b <sub>2</sub> <sup>2</sup> 3b <sub>2</sub> <sup>2</sup> 4b <sub>2</sub> <sup>2</sup> 5b <sub>2</sub> <sup>2</sup> 6b <sub>2</sub> <sup>2</sup> 7b <sub>2</sub> <sup>2</sup>   |
| π electrons | 1b <sub>1</sub> <sup>2</sup> 2b <sub>1</sub> <sup>2</sup> 1a <sub>2</sub> <sup>2</sup>   |

TABLE XXXI

PYRAZINE. GROSS POPULATION ANALYSIS (π) A<sub>2</sub> AND B<sub>1</sub>

|   | 1b <sub>1</sub> | 2b <sub>1</sub> | 1a <sub>2</sub> |
|---|-----------------|-----------------|-----------------|
| N <sub>1</sub> (2p <sub>x</sub> )                                     | 0.3402          | 0.6550          |                 |
| C <sub>1</sub> (2p <sub>x</sub> ) ± C <sub>3</sub> (2p <sub>x</sub> ) | 0.6602          | 0.3446          | 1.0001          |
| C <sub>3</sub> (2p <sub>x</sub> ) ± C <sub>4</sub> (2p <sub>x</sub> ) | 0.6599          | 0.3452          | 0.9999          |
| N <sub>2</sub> (2p <sub>x</sub> )                                     | 0.3398          | 0.6552          |                 |
| Total   | 2.0001          | 2.0000          | 2.0000          |

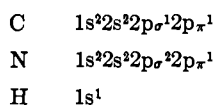
in pyridine by about 1.2 eV. This added stability is due to the increased delocalization of the 5a<sub>1</sub> MO, which was concentrated around the nitrogen in pyridine and now is about equally distributed in the four carbon and the two nitrogen atoms. The 6a<sub>1</sub> has the same



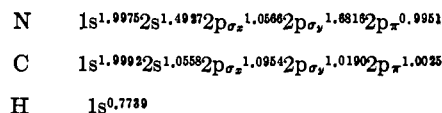
characteristics in being more delocalized and more stable than the  $6a_1$  pyridine. In pyridine the  $11a_1$  MO is the lone-pair MO; in pyrazine, the  $10a_1$  and the  $11a_1$  MO's are the two lone-pair MO's. In pyridine the lone pair is delocalized on the neighboring carbon atom (1.41 electrons are on the N atom in pyridine and the remaining 0.59 is delocalized on the molecule). In pyrazine the lower lone pair ( $10a_1$ ) has 0.94 electron on one nitrogen and 0.94 electron on the second nitrogen atom. The s character in the sp hybrid is relatively low (0.15 s character and 0.79 p character) if compared with diagonal hybridization. The second lone pair, the  $11a_1$ , has 0.687 electron on each nitrogen. Therefore 1.374 electrons are on the nitrogens and 0.626 electron is delocalized on the rest of the molecule (0.22 electron on the hydrogens and 0.40 on the carbons). These results are of interest because they indicate that the two lone pairs in the diazine molecules are quite different in character, so that one cannot expect to describe the two lone pairs by simple assumptions [such as assuming the two to be orthogonal and of the forms (1) hybrid( $N_1$ ) + hybrid( $N_2$ ) and (2) hybrid( $N_1$ ) - hybrid( $N_2$ )].

Here we have identified the  $10a_1$  and the  $11a_1$  as the lone-pair MO's, because of their large 2p population on the nitrogen atom and of their high energy. It should be noted that this identification is open to question. The MO calculation, by its nature, should refer to "spectroscopic" lone pairs, rather than "chemical" lone pairs. The two chemical lone pairs should be of  $A_1$  and  $B_2$  symmetry. Therefore, we could possibly attempt to identify the lone pairs with the  $11a$  and  $7b_2$ . But then other questions remain unsolved; for example, we should analyze how the lone pair is deformed in the excited states and study its transformation properties on a localized basis. We shall give no answer to this point but call the reader's attention to the author's ambiguous use of the term "lone" pair.

We can summarize the population analysis results of the pyrazine molecule: the electronic configuration for the separated atoms is



and for the atoms in pyrazine is



## IX. CONCLUSIONS

The first conclusion is that we have reached the time for *ab initio* computations in aromatic molecules, with explicit consideration of all the electrons. This is presently done in the SCF formalism.

As a second conclusion, we have described a technique whereby *ab initio* computations in molecules can include a large fraction of the correlation energy. Since the amount of labor needed to obtain such functions is not much greater than the labor required to obtain SCF functions, and since our present program has been extended so as to include the multiconfiguration SCF technique (the program has already been tested for few small molecules), then we conclude that the all-electron SCF-LCAO-MO era for aromatic molecules will be rather a short one and we can realistically look forward to much more accurate work than that here presented.<sup>68</sup>

The effect of having reached the present all-electron SCF stage for aromatic molecules has already proven most important. It has clearly demonstrated that the  $\pi$ -electron approximation is untenable and that the  $\sigma$ - $\pi$  electrons interact so strongly that any approximation which will not give as much attention to the  $\sigma$  electrons as is now given to the  $\pi$  electrons is on unsound grounds.

The price for this reversal in our thinking is, however, very large. Much of the previous and present empirical work on large molecules and, in particular, on the aromatic molecules of biological significance is at best open to question, not only on its quantitative validity but even on its qualitative validity. Finally, the economical price for the computations here reported is by far higher than the one required for more conventional work. This opens the question of readiness both financially and technically for many of the institutions of higher learning.