Population Analysis

As one studies Chemistry the significance of the charge hosted by a particular atom in a molecule becomes obvious and in some intuitive way so does the definition of this charge. The question we will address ourselves to is the quantification of this intuitive concept.

Consider the Hartree-Fock wavefunction for the H₂ molecule.

$$\Psi(1,2) = \mathcal{A}\phi\alpha(1)\phi\beta(2) = \phi(1)\phi(2)(\alpha\beta - \beta\alpha)/\sqrt{2}$$

Where ϕ is the spatial molecular orbital. As we have seen the number electron density ρ for this wavefunction is simply the square of the spatial molecular orbital times the number of electrons in this orbital.

$$\rho(\vec{r}) = 2\phi^2(\vec{r})$$

Since ϕ is a linear combination of basis functions on each center we may write it as

$$\phi = (\chi_a + \chi_b) / \sqrt{2(1 + \Delta_{ab})}$$

where $\Delta_{ab} = \langle \chi_a | \chi_b \rangle$ is the overlap integral between basis functions on *a* and *b*. There are no restrictions in taking $\chi_a \& \chi_b$ to be normalized. We may then write

$$\phi = C_a \chi_a + C_b \chi_b$$

where

$$C_a = C_b = 1/\sqrt{2(1+\Delta_{ab})}$$

The electron density is then

$$\rho(\vec{r}) = 2(C_a^2 \chi_a^2 + C_b^2 \chi_b^2 + 2C_a C_b \chi_a \chi_b)$$

The electron density in H₂ is made up of three contributions. Two associated with the individual nuclei, $2C_a^2\chi_a^2$ and $2C_b^2\chi_b^2$ and one which has contributions from both centers, $4C_aC_b\chi_a\chi_b$. Note that if we integrate $\rho(\vec{r})$ over all space the result is the number of electrons in the molecule.

$$\int \rho(\vec{r}) dV = 2(C_a^2 + C_b^2 + 2C_a C_b \Delta_{ab}) = 2$$

This result suggests that we take, $2C_a^2$ as a measure of the number of electrons "on" or "associated" with center a and $4C_aC_b\Delta_{ab}$ as a measure of the number of electrons in the region between and common to both nuclei. This interpretation was put forth by R. Mulliken (J. Chem Phys 1955, <u>23</u>, 1833) and is called a population analysis. Mulliken called $2C_a^2$ the net atomic population, $\eta(a)$ and $4C_aC_b\Delta_{ab}$ the overlap population.

If we assign a radius to an atom in the H₂ molecule and take all charge within this radius as belonging to the atom in question we see that all three components of $\rho(\vec{r})$ will contribute to the "atomic density". This conflict between our intuitively sensible notion that when charge is near an atom it "belongs" to that atom and the mathematically convenient population analysis has spawned numerous attempts at defining a more realistic partitioning. The problem is tied to the inherent ambiguity of partitioning a continuously varying charge distribution among a finite number of centers.

As a measure of the total number of electrons associated with a given center, Mulliken suggested that we take the net atomic population plus one half of the overlap population and call this the gross atomic population N(a) associated with the center in question. Accordingly, we have

$$N(a) = \eta(a) + \eta(a,b)/2 = 2C_a^2 + 2C_a C_b \Delta_{ab}$$

and since in this instance C_a equals C_b we have N(a) = 1, i.e., a covalent bond in which each atom hosts one electron.

Generalization to Arbitrary Wavefunctions

Given a wavefunction for an arbitrary N electron molecule we may form the number electron density as

$$\rho(\vec{r}) = \int \psi^*(1, 2, ..., N) \psi(1, 2, ..., N) d\tau(2, ..., N) d\xi(1)$$

where $d\tau(2,3,...,N)$ indicates an integration over the spatial and spin coordinates of electrons 2 through N and $d\xi(1)$ an integration over the spin of electron 1. $\rho(\vec{r})$ can then be expressed in terms of the natural orbitals of the system as

$$\rho(\vec{r}) = \sum_{i=1}^{\infty} \eta_i \varphi_i^2(\vec{r})$$

where n_i is the occupation number of the ith natural orbital $\varphi_i(\vec{r})$. If the wavefunction is a closed shell Hartree-Fock function the natural orbitals are the molecular orbitals and the occupations numbers are 2 for the occupied orbitals and 0 for the virtual. $\rho(\vec{r})$ then simplifies to

$$\rho(\vec{r}) = 2\sum_{i=1}^{nocc} \varphi_i^2(\vec{r})$$

Where $\varphi_i(\vec{r})$ are the occupied molecular orbitals. We will develop the subsequent details using the general form of $\rho(\vec{r})$.

Let $\chi_{\mu k}$ represent the μ^{th} basis function on the k^{th} nucleus, then the spatial orbitals may be written as

$$\varphi_i = \sum_{k=1}^{nuclei} \sum_{\mu=1}^{n_k} \chi_{\mu k} C_{\mu k i}$$

where n_k is the number of basis functions on center k. Recognizing that $\eta_i \varphi_i^2(\vec{r})$ is the contribution to the electron density due to the electrons in the i^{th} orbital we may write the total density as a sum of orbital contributions.

$$\rho(\vec{r}) = \sum_{i=1}^{orbitals} \rho_i(\vec{r})$$

The total electron density is then

$$\rho(\vec{r}) = \sum_{i=1}^{orbitals} \eta_i \sum_{k=1}^{nuclei} \sum_{\mu=1}^{n_k} \chi_{\mu k} C_{\mu k i} \sum_{l=1}^{nuclei} \sum_{\nu=1}^{n_l} \chi_{\nu l} C_{\nu l i}$$

We observe that the basis functions are independent of the orbital index and sum over i.

$$\rho(\vec{r}) = \sum_{k=1}^{nuclei} \sum_{l=1}^{nuclei} \sum_{\mu=1}^{n_k} \sum_{\nu=1}^{n_l} \chi_{\mu k} \chi_{\nu l} P_{\mu k;\nu l}$$

where

$$P_{\mu k;\nu l} = \sum_{i=1}^{orbitals} \eta_i C_{\mu k i} C_{\nu l i}$$

is called the Coulson density matrix. Note that this matrix is symmetric in that

$$P_{\mu k;\nu l} = P_{\nu l;\mu k}$$

Population Analysis/J. F. Harrison

Integrating $\rho(\vec{r})$ over all space we have

$$\int \rho(\vec{r})dV = \sum_{k=1}^{nuclei} \sum_{l=1}^{nuclei} \sum_{\mu=1}^{n_k} \sum_{\nu=1}^{n_l} P_{\mu k;\nu l} \int \chi_{\mu k} \chi_{\nu l} dV$$

which results in

$$N = \sum_{k=1}^{nuclei} \sum_{l=1}^{nuclei} \sum_{\mu=1}^{n_k} \sum_{\nu=1}^{n_l} P_{\mu k;\nu l} \Delta_{\mu k;\nu l}$$

We may now sum over $\mu \& v$ and define N(k,l) as

$$N(k,l) = \sum_{\mu=1}^{n_k} \sum_{\nu=1}^{n_l} P_{\mu k;\nu l} \Delta_{\mu k;\nu l}$$

so that

$$N = \sum_{k=1}^{nuclei} \sum_{l=1}^{nuclei} N(k, l)$$

This equation partitions the electrons in the molecule among the nuclei so that the number of electrons associated with nucleus *m* is N(m,m), while the number associated with nuclei *m* & *n* is 2N(m,n). Many other partitionings are possible. For example: The number of electrons associated with basis function μ on center *k* is $P_{\mu k;\mu k}$

 $(\Delta_{\mu k;\mu k} = 1)$ and the number of electrons shared between basis function μ on center k and basis function ν on center l is $2P_{\mu k;\nu l}\Delta_{\mu k;\nu l}$ where the factor of two obtains because P and Δ are symmetric.

Example

An SCF calculation (using GAUSSIAN98) on CH₂O using a 6-31g* basis set (at the optimal SCF geometry) results in the following N(m,n) array.

Condensed	to	atoms	(all	e	lectrons)
						_

		1	2	3	4
1	0	8.007728	0.525967	-0.050913	-0.050913
2	С	0.525967	4.596666	0.371821	0.371821
3	Η	-0.050913	0.371821	0.598509	-0.068491
4	Η	-0.050913	0.371821	-0.068491	0.598509

From these data we see that the net electron population on oxygen (atom 1) is 8.007728, while the net on carbon (atom 2) and a H atom is 4.596666 and 0.598509. These numbers correspond to the diagonal elements of the array. The off diagonal elements represent the number of electrons shared between the pairs of atoms and are often called the overlap population. Note that because the matrix is symmetric (N(m,n) = N(n,m)), the total overlap population between oxygen and carbon is 2*(0.525967) and when we form the gross population and divide this equally between the two nuclei, each receives 0.525967. This means that the gross population on an atom is the sum of the row or column entries for that atom. The gross population is then

Atom		gross population	total atomic charge
1	0	8.431869	-0.431869
2	С	5.866275	0.133725
3	Н	0.850926	0.149074
4	Η	0.850926	0.149074

One final observation regarding the N(m,n) array. Those atoms that are directly bonded to one another, (O & C; C & H & H) are characterized by positive overlap populations and the non-bonded pairs of atoms (O & H & H ; H & H) by negative overlap populations. The N(m,n) array seems to reflect the topological connectivity of the molecule.