Proc. Natl. Acad. Sci. USA
Vol. 76, No. 6, pp. 2499–2502, June 1979
Chemistry

Integrated spatial electron populations in molecules: The electron projection function

(atomic charges/electron density/Mulliken populations/acetaldehyde)

ANDREW STREITWIESER, JR.*, JOHN B. COLLINS†, JOHN M. MCKELVEY‡, DAVID GRIER*, JOHN SENDER*, AND A. GLENN TOCZKO§

*Department of Chemistry, University of California, Berkeley, California, 94720; †American Cyanamid Medical Research and Development Division, Lederle Laboratories, Pearl River, New York 10965; ‡Color Photography Chemistry Laboratories, Research Division, Eastman Kodak Company, Rochester, New York 14650; and §Department of Chemistry, Cornell University, Ithaca, New York 14853

Contributed by Andrew Streitwieser, Jr., February 21, 1979

ABSTRACT **A "projection function," $P(x,z)$, is defined as the partial integral of the molecular electron density, $\rho(x,y,z)$, over the region $-\infty < y < +\infty$. The projection provides a three-dimensional representation of molecular electron distributions. Chemically useful information can be discerned from graphical displays in either perspective plot or contour format. Numerical integration of the function gives the integrated spatial electron population for any region of interest. The use of the projection function and difference functions is exemplified by application to acetaldehyde.**

Since the earliest calculations of self-consistent field molecular wave functions (1), a need has existed to interpret the electron probability distribution in chemists' familiar terms of atomic charge, polarization, and ionic and covalent bonding; that is, the charge of an atom in a molecule is not a physical observable and, consequently, has no rigorous definition, but the concept has, nevertheless, even in qualitative terms, had important conceptual and heuristic value in understanding molecular chemistry. "Atomic charge" is one of those useful devices that enable us to dissect a complex compound and to comprehend it in terms of its simplified components. Especially important in this regard has been the Mulliken population analysis (2–5), a well-defined and efficient method of partitioning the linear combination of atomic orbitals-molecular orbital (LCAO-MO)-derived electron densities to regions associated with atoms and bonds.

Unfortunately, the Mulliken scheme has important weaknesses; numerous authors (6–23) have pointed out these limitations of Mulliken populations and there is no need to elaborate on them here. There exist several variants of the Mulliken approach (8, 10–12, 17–23) differing from it primarily in the handling of the overlap population. However, all of these methods have a common weakness in partitioning electrons according to contributions from the mathematical functions of the basis set; that is, electrons populating a basis function are generally associated with the atom on which that function is centered even if spatially the function has a large amplitude near another atom.

Two general methods have evolved for dealing with this problem that are more nearly independent of the choice of basis set. The first is the use of both contour and perspective diagrams of orbital density (Eq. 1) or density difference maps (10, 17, 24–36) to explore electron reorganizations that occur during bonding.

$$\rho_k(x,y_o,z) = \phi_k{}^2(x,y_o,z) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \chi_i(x,y_o z)\chi_j(x,y_o,z), \quad [1]$$

in which $c_{i,j}$s are coefficients of the n basis functions $\chi_{i,j}$s in MO ϕ_k, and y_o is an arbitrarily selected value of y. The total electron density ρ is given as the appropriate sum over occupied MOs,

$$\rho(x,y_o,z) = \sum^{\text{occ}} N_k \rho_k, \quad [2]$$

in which N_k is the occupation number of MO ϕ_k. These diagrams are particularly revealing when the molecule of interest has a high-order axis of symmetry such as diatomic and linear polyatomic molecules. In instances in which this is not the case the main disadvantage of this method is obvious. Because $\rho(x,y,z)$ is a four-dimensional function it is necessary to select a fixed value of one of the coordinates (e.g., $y = y_o$) in order to reduce the problem to three dimensions. Because this coordinate often corresponds to the selection of the molecular plane in which to display the electron density, it is not possible to show any contribution from orbitals having a node in the molecular plane; e.g., σ and π MOs cannot be depicted on the same scale. Thus, unless one examines a spectrum of parallel planes, misleading interpretations could result because entire orbital representations (in the group theoretical sense) might be omitted.

The second alternative method of dealing with charge distribution is direct integration (21, 22, 36–47) of $\rho(x,y,z)$ over some region of interest. In some applications this region is a spherical volume element in spherical coordinate space surrounding an atom (Eq. 3) (38–41, 43–45).

$$N_e = \int_0^R \rho(r,\theta,\phi)R^2 dR \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} d\phi, \quad [3]$$

in which N_e is the integrated electron population in a sphere of radius R. In the "planar density function" of Brown and Shull (37) the region of interest is a linear element in Cartesian space; in effect, the electron density is projected onto a line (Eq. 4) (21, 42, 46, 47).

$$N_e = \int_{x_1}^{x_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y,z)dx\, dy\, dz, \quad [4]$$

in which N_e is the electron population in a volume element spanning $-\infty < y, z < \infty$ and $x_1 < x < x_2$. Both methods are an improvement over Mulliken-like basis set populations because they make use of integrated spatial electron populations, but they still have limitations of their own. The spherical region, Eq. 3, is applicable primarily to obtaining charges on atoms but provides little information about bonding regions. The planar density function, Eq. 4, has greater utility in that both core and bonding regions can be examined so long as a molecule can be

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: LCAO, linear combination of atomic orbitals; MO, molecular orbital; e, electron; a.u., atomic unit.

reasonably represented as having a cylindrical distribution of charge.

We introduce here an electron projection function (Eq. 5) that enjoys the features of direct integration as in Eqs. 3 and 4 but includes an additional Cartesian dimension for added versatility and applicability.

$$P(x,z) = \int_{-\infty}^{\infty} \rho(x,y,z)dy. \quad [5]$$

Properties of the electron projection function

Projecting the four-dimensional function, $P(x,y,z)$, into the three-dimensional function, $P(x,z)$, has some important advantages. It corresponds to summing the aforementioned electron density function of a plane (Eq. 1) over all parallel planes. Hence, there is no arbitrariness about the choice of one of these planes in which to display the distribution. σ and π MOs can be represented in the same coordinate frame. Because the projection function (Eq. 5) has units of electrons per unit area, the magnitude of $P(x,z)$ is a direct measure of the number of electrons in a region, not just their density. Numerical integration of $P(x,z)$ (Eq. 5) over x reduces to the linear projection (21, 37, 42, 46, 47) of Eq. 4.

Application of the projection function parallels that of density diagrams and density difference diagrams with one important distinction; namely, charge is conserved whereas density is not; that is, the total volume of $P(x,z)$ is a constant equal to the number of electrons in the system, unlike $\rho(x,y,z)$, in which density can "leak" into parallel planes. The consequence of this distinction is that difference maps of the projection function have zero net volume; charge can be gained in one region only at the expense of another region. Thus, difference maps represent a direct measure of the amount of charge redistribution rather than the amount of density redistribution.

Some of these features are exemplified below with applications to acetaldehyde.

Technical considerations: Efficiency and accuracy of numerical integration

In general, the analytical integration in Eq. 5 can be accomplished readily for Gaussian basis sets by application of Huzinaga's formula (48). We found it advantageous to utilize the separability of coordinates of Gaussian functions to rewrite Eq. 5 as Eq. 6.

$$P(x,z) = \sum_{i=1}^{n}\sum_{j=1}^{n} F_{ij}(x)\cdot F_{ij}(z) \cdot \int_{-\infty}^{\infty} F_{ij}(y)dy \sum_{k=1}^{\text{occ}} N_k \cdot C_{ki} \cdot C_{ki}, \quad [6]$$

in which $F_{ij}(x)$ is the x component of Huzinaga's formula for the ith and jth Gaussian primitives, and C_{ki} is the ith coefficient in the kth LCAO-MO having occupation number N_k. This treatment has several useful consequences. First, because the integral in y is independent of x and z it need be evaluated only once for each pair of basis functions i,j. Second, it permits a refinement procedure that renders subsequent numerical integrations quantitative. We superimpose the molecule on a grid in a convenient plane. Because our concern is to measure distribution of charge, we wished to have each of our grid points represent the average value of the projection function $P(x,z)$ in the neighborhood $x \pm \Delta x/2$, $z \pm \Delta z/2$ rather than the value of the function at the arbitrarily selected grid points. By utilizing the separability of Eq. 6 it is possible to compute the averages of $F_{ij}(x)$ and $F_{ij}(z)$ in their respective neighborhoods $x \pm \Delta x/2$ and $z \pm \Delta z/2$. This allows one to rewrite Eq. 6 as Eq. 7.

$$\bar{P}(x,z) = \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{F}_{ij}(x) \cdot \bar{F}_{ij}(z) \cdot \int F_{ij}(y)dy \sum_{k=1}^{\text{occ}} N_k C_{ki} C_{kj}. \quad [7]$$

The averages of $F_{ij}(x)$ and $F_{ij}(z)$ were computed by dividing the respective grid intervals into a large number of subintervals (40–100), computing $F_{ij}[x + (b\cdot\Delta x/m)]$ (in which m is the number of subintervals and b ranges from $-m/2$ to $+m/2$) and storing the average $\bar{F}_{ij}(x)$.

Valence orbitals have slowly varying derivatives and little refinement is needed to achieve quantitative (error less than 10^{-4} electrons per orbital) numerical integration. However, even core regions of second-row (Na–Ar) atoms can be integrated to within 0.01 electron if the number of subintervals m = 100. Given these considerations, it was possible to examine a grid of sufficient size so as to encompass all of the electron distribution for molecules of chemical interest while retaining the precision necessary for valid interpretation. We found that a grid spacing of 0.2 atomic unit (a.u.; 1 a.u. = 0.529 Å) and extension of the grid 3.0 a.u. beyond the nearest atom gives satisfactory results in all cases. Further details of the computational strategy are being published separately (49).

Example: Acetaldehyde

The projection function for acetaldehyde in the C–C–O plane is shown in Fig. 1 in perspective plot form as calculated with the 4-31G basis set (50) for the "standard" geometry (51). The function shows a distinct minimum between carbon and oxygen that could be used as a defined boundary for evaluating the electronic charge "on oxygen." This boundary differs from that given by covalent bond radii and would lead to an effectively varying size for the oxygen in different carbonyl compounds. This problem emphasizes the limitations of any method that attempts to define the atomic population or charge on an atom in a molecule—a physically nonobservable quantity. However, far less ambiguity is associated with the use of any consistent definition in determining the *change* in the oxygen charge in a structural change such as between formaldehyde and acetaldehyde. Fig. 2 shows the projection function difference plot

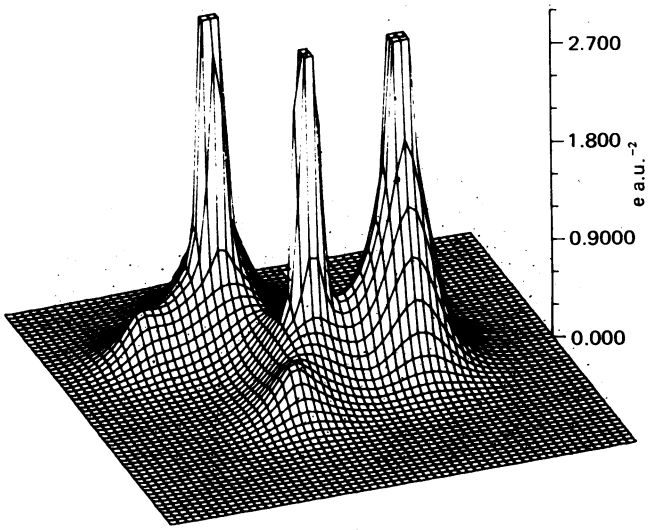

FIG. 1. Electron projection plot in perspective format for acetaldehyde. The carbonyl oxygen is at the far right, and the methyl group is in the back left. e, Electrons; a.u., atomic units.

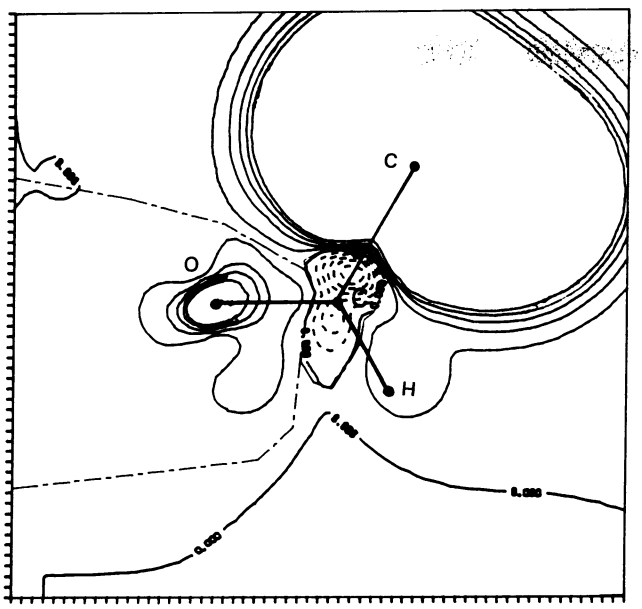

FIG. 2. Difference projection function plot for acetaldehyde minus formaldehyde. - - - contours (units are e a.u.$^{-2}$) are negative; — · — line shows limits for numerical integration about oxygen. Contour levels are from −0.012 to 0.012 at intervals of 0.002.

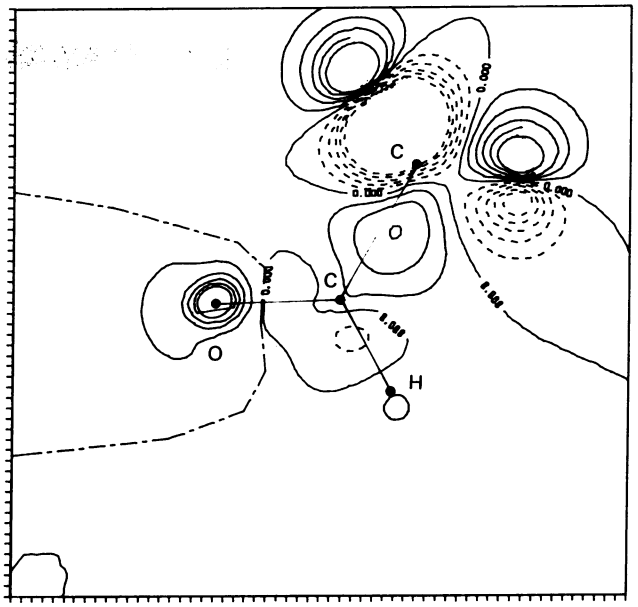

FIG. 3. Difference projection function plot for acetaldehyde enolate ion minus acetaldehyde. - - - contours (units are e a.u.$^{-2}$) are negative; — · — line shows limits for numerical integration about oxygen. Contour levels are from −0.012 to 0.012 at intervals of 0.002.

in contour form for P(acetaldehyde) − P(formaldehyde) using the same basis set and same aldehyde geometry. Only the contours in the carbonyl region are important because of the difference in numbers of atoms and electrons in the methyl region. The positive difference contours around oxygen show clearly that the oxygen of acetaldehyde has more electrons than that of formaldehyde. Moreover, the contours between oxygen and carbon in the carbonyl group are now rather shallow and the difference population is not as sensitive to the precise definition of the atomic boundaries. Integration gives a difference of 0.044 e for oxygen. The negative contours at carbon show that the carbonyl carbon of acetaldehyde is more positive than that of formaldehyde. Integration with reasonable boundaries gives a difference population of about −0.009 e. These values for oxygen and carbon may be compared to the 4-31G Mulliken population differences of 0.037 e and −0.146 e, respectively. The shallow contour at the aldehyde hydrogen shows that this hydrogen is essentially similar in acetaldehyde and formaldehyde, with the acetaldehyde hydrogen being slightly more negative. The corresponding Mulliken population difference is −0.011 e; that is, with the formaldehyde hydrogen more negative.

Another interesting comparison is that between acetaldehyde and its enolate anion. Fig. 3 shows the difference projection plot P(enolate) − P(acetaldehyde) for a planar enolate structure having the same CCHO skeleton as acetaldehyde. The true structure would be a perturbation of this, but use of the same skeleton facilitates interpretation of the difference function. We note that, as expected, the enolate oxygen is more negative than that of acetaldehyde. Integration gives a difference of 0.24 e; moreover, the shallow nature of the difference contours between oxygen and carbon shows that this number is not sensitive to the precise definition of the boundaries of the oxygen atom. Again, a direct comparison of population differences between comparison systems is less sensitive to the precise definition of an atom in a molecule than is the absolute value in any one compound. The shallow contours at the carbonyl carbon show that this carbon and its attached hydrogen are much the same in enolate and aldehyde. The positive contours between the carbons are consistent with the increased double bond character of the C—C bond in the enolate ion.

The foregoing examples suffice to demonstrate the value of the projection function. Other applications to showing the effects of basis set variations, to carbonium ions and hyperconjugation, and to carbanions and anionic hyperconjugation will be presented elsewhere.

This work was supported in part by U.S. Public Health Service National Institutes of Health Grant GM-12855 and National Science Foundation Grant CHE 76-10997.

1. Roothann, C. C. J. (1951) *Rev. Mod. Phys.* **23**, 69–89.
2. Mulliken, R. S. (1955) *J. Chem. Phys.* **23**, 1833–1840.
3. Mulliken, R. S. (1955) *J. Chem. Phys.* **23**, 1841–1846.
4. Mulliken, R. S. (1955) *J. Chem. Phys.* **23**, 2338–2342.
5. Mulliken, R. S. (1955) *J. Chem. Phys.* **23**, 2343–2346.
6. Streitwieser, A., Jr., Williams, J. E., Jr., Alexandratos, S. & McKelvey, J. M. (1976) *J. Am. Chem. Soc.* **98**, 4778–4784.
7. Lowden, P.-O. (1955) *J. Chem. Phys.* **21**, 374–375.
8. Mulliken, R. S. (1962) *J. Chem. Phys.* **36**, 3428–3439.
9. Peters, D. (1963) *J. Chem. Soc.* 2015–2023.
10. Kern, C. W. & Karplus, M. (1964) *J. Chem. Phys.* **40**, 1374–1389.
11. Ros, P. & Schuit, G. C. A. (1966) *Theor. Chim. Acta* **4**, 1–12.
12. Davidson, E. R. (1967) *J. Chem. Phys.* **46**, 3320–3324.
13. Pollack, M. & Rein, R. (1967) *J. Chem. Phys.* **47**, 2045–2052.
14. Smith, P. R. & Richardson, J. W. (1967) *J. Phys. Chem.* **71**, 924–930.
15. Veillard, A. (1968) *J. Chem. Phys.* **48**, 2012–2016.
16. Cusachs, L. C. & Politzer, P. (1968) *Chem. Phys. Lett.* **1**, 529–531.
17. Politzer, P. & Cusachs, L. C. (1968) *Chem. Phys. Lett.* **2**, 1–4.
18. Stout, E. W., Jr. & Politzer, P. (1968) *Theor. Chim. Acta* **12**, 379–386.
19. Coulson, C. A. & Doggett, G. (1968) *Int. J. Quantum Chem.* **2**, 825–843.
20. Doggett, G. (1969) *J. Chem. Soc. A*, 229–233.
21. Politzer, P. & Harris, R. R. (1970) *J. Am. Chem. Soc.* **92**, 6451–6454.
22. Politzer, P. & Mulliken, R. S. (1971) *J. Chem. Phys.* **55**, 5135–5136.

23. Roby, K. R. (1974) *Mol. Phys.* **27,** 81–104.
24. Smith, P. R. & Richardson, J. W. (1967) *J. Phys. Chem.* **69,** 3346–3357.
25. Bader, R. F. W. & Henneker, W. H. (1965) *J. Am. Chem. Soc.* **87,** 3063–3068.
26. Ransil, B. J. & Sinai, J. J. (1967) *J. Chem. Phys.* **46,** 4050–4074.
27. Bader, R. F. W. & Keaveny, I. (1967) *J. Chem. Phys.* **47,** 3381–3402.
28. Politzer, P. (1970) *Theor. Chim. Acta* **16,** 120–125.
29. Boyd, D. B. (1970) *J. Chem. Phys.* **52,** 4846–4857.
30. Politzer, P. & Harris, R. R. (1971) *Tetrahedron* **27,** 1567–1572.
31. Absar, I. & Van Wazer, J. R. (1971) *J. Phys. Chem.* **75,** 1360–1365.
32. Maismann, H., Robert, J.-B. & Van Wazer, J. R. (1971) *Tetrahedron* **27,** 4377–4386.
33. Dunning, T. H., Jr. & Winter, N. W. (1971) *J. Chem. Phys.* **55,** 3360–3371.
34. Bader, R. F. W. & Beddall, P. M. (1973) *J. Am. Chem. Soc.* **95,** 305–315.
35. Streitwieser, A., Jr. & Owens, P. H. (1973) *Orbital and Electron Density Diagrams* (Macmillan, New York).
36. Bader, R. F. W. (1975) *Acc. Chem. Res.* **8,** 34–40.
37. Brown, R. E. & Shull, H. (1968) *Int. J. Quantum Chem.* **2,** 663–685.
38. Politzer, P. (1971) *Theor. Chim. Acta* **23,** 203–207.
39. Politzer, P. & Reggis, P. H. (1972) *J. Am. Chem. Soc.* **94,** 8308–8311.
40. Politzer, P. & Politzer, A. (1973) *J. Am. Chem. Soc.* **95,** 5450–5455.
41. Politzer, P., Elliot, J. D. & Maroney, B. F. (1973) *Chem. Phys. Lett.* **23,** 331–334.
42. Russeger, P. & Schuster, P. (1973) *Chem. Phys. Lett.* **19,** 245–259.
43. Dean, S. M. & Richards, W. G. (1975) *Nature* **256,** 473–475.
44. Boyd, R. J. (1977) *J. Chem. Phys.* **66,** 356–358.
45. Politzer, P., Reuther, J. & Kasten, G. T. (1977) *J. Chem. Phys.* **67,** 2385–2387.
46. Lischka, H. (1977) *J. Am. Chem. Soc.* **99,** 353–360.
47. Polak, R. (1978) *Theoret. Chim. Acta* **50,** 21–30.
48. Taketa, H., Huzinaga, S. & Ohata, K. (1966) *J. Phys. Soc. Jpn,* **21,** 2313–2324.
49. Collins, J. B., Streitwieser, A., Jr. & McKelvey, J. (1979) *Comput. Chem.*, in press.
50. Ditchfield, R., Hehre, W. J. & Pople, J. A. (1970) *J. Chem. Phys.* **52,** 5001–5007.
51. Pople, J. A. & Gordon, M. (1967) *J. Am. Chem. Soc.* **89,** 4253–4261.