

AN EMPIRICAL WAY TO CORRECT SOME DRAWBACKS OF MULLIKEN POPULATION ANALYSIS

JUAN S. GÓMEZ-JERIA

Universidad de Chile, Facultad de Ciencias, Departamento de Química, Casilla 653, Santiago 21, Chile; and Programa de Doctorado en Fisicoquímica Molecular, Universidad Andrés Bello, República 275 Santiago, Chile.

(Received: July 29, 2009 - Accepted: October 30, 2009)

ABSTRACT

The problem of negative electronic populations and of occupation numbers greater than 2 has plagued Mulliken Population Analysis since the very beginning. Through the analysis of three model molecular systems, several basis sets and the relevant literature, we conclude that there is not enough evidence to assign the origin of these errors to the self-consistent scheme, to Mulliken's partition, to the basis set structure or to a combination of these. As Mulliken Population Analysis is still widely used, we have developed an empirical method to eliminate negative electronic populations and occupation numbers greater than 2. This method can be used for any partition of the electron density (not only Mulliken's), for any basis set and for any LCAO-MO methodology (semiempirical or *ab initio*). Finally, the method does not produce any change in the original atomic net charges.

INTRODUCTION

The earliest method of assigning charges to wave functions represented in LCAO bases is due to Mulliken^{1,2}. It is the simplest distribution of the electron density among orbitals. But, as a unitary transformation can be used to construct an equally valid assignment of charges which is different from the original Mulliken assignment, the conclusion is that the Mulliken charges are basis set-dependent and therefore not unique.

The more or less arbitrary partitioning of the total density (an observable) into atom-centered contributions (not observables) raises two problems at the level of the electronic populations of atoms in molecular orbitals (MO). The first problem concerns occupation numbers larger than 2.0, in open violation of the Pauli Exclusion Principle. The second is the occurrence of negative populations which is clearly unphysical. This latter effect arises at the level of the partial gross populations and total gross population of an atom in a MO¹. Several researchers have noted these facts³⁻⁷, and some have claimed that negative populations are a natural effect of the SCF scheme, and not an artifact^{8,9}. Artifacts or not, negative populations may lead to notorious errors, such as the appearance of negative condensed density of states¹⁰ and negative Fukui indices. At the level of quantitative structure-activity relationships, problems of physical interpretation will arise in the case of the so-called orbital superdelocalizabilities (OSDs)^{11,12}, precluding the use of these static reactivity indices at the Extended Hückel, *ab initio* or DFT levels. This is so because OSD are functions of the Fukui indices. Several ways to correct MPA have been proposed, but none of them has been completely successful¹³⁻²¹.

Considering that Mulliken Population analysis (MPA) is still widely used and incorporated into computational packages such as Gaussian, we present here an empirical approach to correcting both the negative gross atomic populations and the MO occupation numbers larger than 2. We present and discuss the results for three model systems and analyze the effects of this approximation on the final net atomic charges.

METHODS, MODELS AND CALCULATIONS

Water, formic acid and n-butane were selected as model systems. The geometry of each system was fully optimized within the Density Functional scheme (DFT) at the B3LYP/6-31G, B3LYP/6-31G*, B3LYP/6-31G**, B3LYP/6-31G**+, B3LYP/6-31G***, B3LYP/6-311G, B3LYP/6-311G*, B3LYP/6-311G**, B3LYP/6-311G*** and B3LYP/6-311G**** basis set levels. The Gaussian 03 package of programs was used²². Appropriate software was written to extract the overlap and density matrices and the eigenvalues and eigenvectors and to calculate all the Mulliken populations, the Fukui indices and the OSDs.

RESULTS AND DISCUSSION

To fully understand the basis of our corrective method it is necessary to divide this section into subsections. We must add that the comments made for one system are also valid for the remaining ones (we have verified case by case and result by result). As we cannot present all the Tables with results, we are

ready to provide, upon request, a pdf file (in Spanish) with all the available information supporting the statements made below.

a. Qualitative analysis of the eigenvectors.

In Table 1 we present, for the occupied MOs of the water molecule and different basis sets, the LCAO coefficients with greater absolute values. Two important facts can be noted in this Table. The first one is that no LCAO coefficient is greater than 1.0. The second is that the addition of polarization and diffuse functions does not modify this situation.

Table 1: The three LCAO coefficients with the greatest values in the occupied MOs of the water molecule^a

Basis set	1	2	3
B3LYP/6-31G	1.00 (1s-O/1)	0.64 (2p _z -O/5)	0.55 (2p _y -O/4)
B3LYP/6-31-G*	1.00 (1s-O/1)	0.64 (2p _z -O/5)	0.55 (2p _y -O/4)
B3LYP/6-31G**	1.00 (1s-O/1)	0.64 (2p _x -O/5)	0.55 (2p _z -O/4)
B3LYP/6-31G***	1.00 (1s-O/1)	0.64 (2p _z -O/5)	0.56 (2p _y -O/4)
B3LYP/6-31G****	0.99 (1s-O/1)	0.64 (2p _z -O/5)	0.55 (2p _y -O/4)
B3LYP/6-311G	0.55 (1s-O/1)	0.54 (2s-O/2)	0.49 (3p _x -O/5)
B3LYP/6-311G*	0.55 (1s-O/1)	0.54 (3s-O/2)	0.50 (4p _y -O/5)
B3LYP/6-311G**	0.55 (1s-O/1)	0.47 (3s-O/2)	0.47 (4p _z -O/5)
B3LYP/6-311G***	0.55 (1s-O/1)	0.54 (3s-O/2)	0.43 (3p _x -O/5)
B3LYP/6-311G****	0.55 (1s-O/1)	0.54 (3s-O/2)	0.47 (2s-O/1)

a. The nomenclature is [LCAO value (Atomic orbital/Molecular orbital)]

In Table 2 we present, for the empty MOs of the water molecule and different basis sets, the LCAO coefficients with greater absolute values. Table 2 shows an entirely different situation. Here, all the absolute values of the LCAO coefficients are greater than 1.0. The addition of diffuse functions worsens the situation in most cases. It is interesting to note that, even in the case in which no polarization or diffuse functions are included, absolute values of the LCAO coefficients greater than 1.0 are obtained (the case of the B3LYP/6-31G basis set).

Table 2.: The three LCAO coefficients with the greatest values in the empty MOs of the water molecule.

Basis set	1	2	3
B3LYP/6-31G	2.71 (3s-O/13)	1.71 (3p _x -O/12)	-1.68 (2s-O/13)
B3LYP/6-31-G*	3.69 (3s-O/19)	3.60 (3s-O/13)	1.62 (3p _x -O/12)
B3LYP/6-31G**	3.54 (3s-O/13)	2.85 (3s-O/23)	2.60 (3s-O/25)
B3LYP/6-31G***+	3.58 (3s-O/17)	3.31 (4s-O/11)	3.06 (3s-O/29)
B3LYP/6-31G****+	5.07 (4s-O/13)	3.84 (3s-O(31)	3.77 (4s-O/19)
B3LYP/6-311G	2.84 (4s-O/13)	2.33 (2s-O/19)	-2.24 (1s-O/19)
B3LYP/6-311G*	2.87 (4s-O/13)	-1.80 (3s-O/13)	1.76 (4p _y -O/12)
B3LYP/6-311G**	3.12 (4s-O/13)	2.33 (2s-O/30)	-2.24 (1s-O/30)
B3LYP/6-311G***+	8.51 (5s-O/11)	5.16 (5s-O/13)	-4.54 (3s-H ₁ /11) (3s-H ₂ /11)
B3LYP/6-311G****+	8.67 (5s-O/13)	4.54 (4s-O/21)	-4.33 (5s-O/15)

Examination of all the results for all the systems allows us to state the following empirical rule: “If in one MO there are one or more LCAO coefficients with a high positive value, in this OM there are also one or more LCAO coefficients with a high negative value”. Table 3 shows three examples for the water molecule. It is worth mentioning that, in all molecules studied, the above empirical rule applies only to empty (or virtual) MOs.

Table 3.: Examples of LCAO coefficients with high negative and positive values in some empty MOs of the water molecule.

B3LYP/6-31G	MO 13	B3LYP/6-31G*	MO 19	B3LYP/6-311G****+	MO 13
1 O 1S	0.05896	1 O 1S	-0.46796	1 O 1S	0.02658
2S	-1.67588	2S	0.28666	2S	0.04181
2PX	0	2PX	0	2PX	0
2PY	0.17862	2PY	0.11661	2PY	-0.01422
2PZ	0	2PZ	0	2PZ	0
3S	2.70866	3S	3.68971	3S	-0.07428
3PX	0.00001	3PX	0	3PX	0
3PY	-0.7902	3PY	-0.33393	3PY	0.02559
3PZ	0	3PZ	0	3PZ	0
2 H 1S	-0.42108	4XX	-1.52617	4S	-1.09482
2S	-0.61572	4YY	-1.54968	4PX	0
3 H 1S	-0.42108	4ZZ	-1.57875	4PY	-0.07371
2S	-0.6157	4XY	0	4PZ	0
		4XZ	0	5S	8.67072
		4YZ	0	5PX	0
		2 H 1S	0.11589	5PY	-2.35473
		2S	-0.55461	5PZ	0
		3 H 1S	0.11589	6D 0	0.02086
		2S	-0.55461	6D+1	0
				6D-1	0
				6D+2	-0.01077
				6D-2	0
				2 H 1S	-0.02147
				2S	0.10048
				3S	-3.94002
				4S	-0.50396
				5PX	-0.02783
				5PY	0.00423
				5PZ	0
				3 H 1S	-0.02147
				2S	0.10048
				3S	-3.94001
				4S	-0.50397
				5PX	0.02783
				5PY	0.00423
				5PZ	0

These results are a direct product of the SCF scheme and the basis set employed.

b. Qualitative analysis of the electronic population of an atomic orbital in a MO.

In MPA, the total number of electrons (or population) in the atomic orbital (AO) r_k of MO ϕ_i is given by:

$$N(i, r_k) = 2.0c_{ir_k}^2 + 2.0 \sum_{l \neq k} c_{ir_k} c_{il} S_{r_k s_l} \quad (1)$$

where c_{ir_k} is the LCAO coefficient of AO r_k in MO i and $S_{r_k s_l}$ is the overlap integral between AOs r_k and s_l . The values of this integral belong to the $[-1, 1]$ closed interval.

The first and second terms on the right side of Eq. 1 will be called, respectively, the diagonal and non-diagonal contributions to $N(i, r_k)$. They are related to Mulliken's net and overlap populations¹. The aim of this separation is to show that the appearance of LCAO coefficients with a high absolute value produces very high values in the diagonal contribution to $N(i, r_k)$. This is illustrated in Table 4 which presents the diagonal and non-diagonal contributions to OA populations for the 13-th MO of the water molecule together with the LCAO-MO coefficients. The most salient result is that negative electronic populations appear at the first stage of the MPA. Also, although the LCAO-MO coefficients for the H atoms are normal (in the sense that their absolute value is between 0.0 and 1.0), the final AO populations are negative. In the case of very high LCAO-MO coefficient values, the diagonal and non-diagonal contributions are also very high. Interestingly the associated AO populations are positive. Therefore, starting only from the LCAO-MO coefficient, we cannot predict the final sign of the AO populations. Also, it is not possible to attribute the origin of these high LCAO-MO coefficients to a particular SCF procedure, to the basis set structure or to both. The analysis of all 30 calculations shows two important facts for the case of the occupied MOs. First, there are no AO populations greater than 2.0. Second, the negative AO populations have a small value.

Table 4.: LCAO coefficients and AO populations of MO 13 of the water molecule (B3LYP/6-31G results).

Atom	AO	LCAO coefficients of MO 13	Diagonal contribution	Non-Diagonal contribution	AO Total Population
1 O	1S	0.05896	0.00695	-0.00579	0.00116
	2S	-1.67588	5.61715	-4.76893	0.84822
	2PX	0	0	0	0
	2PY	0.17862	0.06381	-0.05468	0.00914
	2PZ	0	0	0	0
	3S	2.70866	14.67368	-13.28731	1.38637
	3PX	0.00001	0	0	0
	3PY	-0.7902	1.24883	-1.14563	0.1032
	3PZ	0	0	0	0
2 H	1S	-0.42108	0.35462	-0.35998	-0.00536
	2S	-0.61572	0.75822	-0.9269	-0.16868
3 H	1S	-0.42108	0.35462	-0.35998	-0.00536
	2S	-0.6157	0.75817	-0.92685	-0.16868

c. Qualitative analysis of the electronic population of an atom in a MO.

The next stage in the MPA is to condense all the AO populations in a MO belonging to the same atom in order to get the atomic populations. The set of these atomic populations in each MO constitutes the whole set of Fukui indices, $F_i(\text{MO})$. Those most used in chemical reactivity studies are usually $F_1(\text{HOMO})$ and $F_1(\text{LUMO})$, but in Structure-Affinity relationship studies it has been shown that sometimes Fukui indices of the inner occupied and upper empty MOs are needed²³⁻²⁵. A detailed examination of the condensation process shows that it is the process itself that leads to the arithmetic disappearance of several negative populations. It is important to point this out because several quantum chemical packages provide the MPA results beginning only with the electronic populations of atoms in each MO. In Table 5 we show the atomic populations in each MO of the water molecule (B3LYP/6-31G* results). We can see the appearance of electronic populations greater than 2.0 and of negative populations. As the MPA is dependent on the basis set, different basis sets will produce different values for the Fukui indices²⁶. Fukui indices defined in this way must be either zero or positive²⁷⁻²⁹. The bad condensation results for the LUMO of water, as shown in Table 5, indicate that we certainly face a problem needing a solution.

Table 5.: Atomic populations in each MO of the water molecule (B3LYP/6-31G* results).

	MO 1	MO 2	MO 3	MO 4	MO 5 (HOMO)
1 O	2.0002	1.7065	1.2966	1.7710	2.0000
2 H	-0.0001	0.1467	0.3517	0.1145	0.0000
3 H	-0.0001	0.1467	0.3517	0.1145	0.0000
	MO 6 (LUMO)	MO 7	MO 8	MO 9	MO 10
1 O	-0.1961	0.2091	0.5297	0.8105	2.0000
2 H	1.0980	0.8955	0.7351	0.5948	0.0000
3 H	1.0980	0.8955	0.7352	0.5948	0.0000
	MO 11	MO 12	MO 13	MO 14	MO 15
1 O	1.6814	2.2080	2.4658	2.0007	2.0000
2 H	0.1593	-0.1040	-0.2329	-0.0003	0.0000
3 H	0.1593	-0.1040	-0.2329	-0.0003	0.0000
	MO 16	MO 17	MO 18	MO 19	
1 O	2.0000	1.8030	1.7566	1.9571	
2 H	0.0000	0.0985	0.1217	0.0214	
3 H	0.0000	0.0985	0.1217	0.0214	

d. Set of rules for correcting Population Analysis.

Our approach for correcting the drawbacks of MPA is pragmatic. At this time it has no theoretical foundation, but its advantage lies in that the corrected results can be compared with experimental results coming from the field of chemical reactivity.

The logical bases for building the algorithm and the corresponding computer program are the following.

Rule 1. Carry out the correction separately at the level of each MO.

Rule 2. Classify the atoms of any molecular system in two groups. The first one contains the so called "peripheral atoms", i.e., those that are bonded to only one other atom (the case of H atoms in H₂O and n-butane for example). The second group contains all the remaining atoms.

Rule 3. If in any MO there is a peripheral atom whose atomic population is negative, subtract this value from the atomic population value of the atom to which it is attached. In Table 5 this is the case for the H atoms in MOs 1, 12, 13 and 14. At this level of the algorithm all peripheral atoms must have zero or positive atomic populations in all MOs.

Rule 4a. If any atom of the second group has a negative atomic population and is attached only to peripheral atoms, divide this atomic population by the number of peripheral atoms and subtract the result from all the peripheral atoms. In Table 5 this is the case of MO 6.

Rule 4b. If any atom of the second group has a negative atomic population and it is not attached only to peripheral atoms, distribute the value of the negative atomic populations equally among the non-peripheral atoms attached to it. The existence of equivalent atoms (the H atoms of H₂O for example) must be taken into account in such a way that their final atomic population values are the same.

Rule 4c. If, after all this process, you still have one atom (or more) from the second group having negative atomic populations, divide this value by the number of peripheral atoms having positive atomic populations and subtract the result from each of their populations. In the case that one or more peripheral atoms have a positive atomic population whose value is less than the abovementioned division, subtract from each of them the necessary quantity to guarantee that no peripheral atom will end up with a negative atomic population.

In Table 6 we show the corrected atomic populations for the water molecule. We present here only results for this system and not for formic acid or n-propane because the remaining Tables with original and corrected populations are too large. We finally tested the computer program with a full B3LYP/6-311G**++ calculation of a tetrathiafulvalene-fluorene diad³⁰ which has an exceptionally small HOMO-LUMO gap (molecule 4 of Ref. 30). No errors have been found in the computer program to date.

Table 6.: Corrected atomic populations in each MO of the water molecule (B3LYP/6-31G* results).

	OM 1	OM 2	OM 3	OM 4	OM 5 (HOMO)
1 O	2.0000	1.7065	1.2966	1.7710	2.0000
2 H	0.0000	0.1467	0.3517	0.1145	0.0000
3 H	0.0000	0.1467	0.3517	0.1145	0.0000
	OM 6 (LUMO)	OM 7	OM 8	OM 9	OM 10
1 O	0.0000	0.2091	0.5297	0.8105	2.0000
2 H	1.0000	0.8955	0.7351	0.5948	0.0000
3 H	1.0000	0.8955	0.7352	0.5948	0.0000
	OM 11	OM 12	OM 13	OM 14	OM 15
1 O	1.6814	2.0000	2.0000	2.0000	2.0000
2 H	0.1593	0.0000	0.0000	0.0000	0.0000
3 H	0.1593	0.0000	0.0000	0.0000	0.0000
	OM 16	OM 17	OM 18	OM 19	
1 O	2.0000	1.8030	1.7566	1.9571	
2 H	0.0000	0.0985	0.1217	0.0214	
3 H	0.0000	0.0985	0.1217	0.0214	

It must be noted that Rules 4b and 4c could be modified in the sense that the partitioning of negative atomic populations between atoms of different nature might be ameliorated by including electronegativities or similar indices. A fact deserving attention is that this procedure does not alter in any way the values of the original atomic net charges.

CONCLUSIONS

An algorithm that corrects the problem of negative atomic populations in Mulliken Population Analysis was developed (the associated computer program is available only on the basis of scientific collaboration). The algorithm was tested for different molecules and basis sets, working perfectly in all 31 calculations carried out. The atomic net charges are not modified in this procedure. There is no reason that forbids the use of this algorithm in DFT, Hartree-Fock and/or semiempirical calculations with any basis set. The only requirement is to work within an LCAO-MO scheme. There is also no reason forbidding the extension of this method to any other kind of partition of the charge density.

REFERENCES

1. R. S. Mulliken, *J. Chem. Phys.* **23**, 1833 (1955).
2. R. S. Mulliken, *J. Chem. Phys.* **23**, 2343 (1955).
3. R. F. Fenske, *Pure Appl. Chem.* **60**, 1153 (1988).
4. Q. Wu; T. van Voorhis, *Phys. Rev.* **A72**, 024502 (2005).
5. J. Baker, *Theor. Chim. Acta* **68**, 221 (1985).
6. A. Koizumi; S. Miyaki; Y. Kakutani; H. Koizumi; N. Hiraoka; K. Makoshi; N. Sakai; K. Hirota; Y. Murakami, *Phys. Rev. Lett.* **86**, 5589 (2001).
7. W. Liu; L. Li, *Theor. Chim. Acta* **95**, 81 (1997).
8. M-H. Whangbo; H. Hoffmann, *J. Chem. Phys.* **68**, 5498 (1978).
9. J. H. Ammeter; H-B. Bürgi; J. C. Thibeault; R. Hoffmann, *J. Am. Chem. Soc.* **100**, 3686 (1978).
10. J. C. Santos; R. Contreras; E. Chamorro; P. Fuentealba, *J. Chem. Phys.* **116**, 4311 (2002).
11. J. S. Gómez-Jeria; L. A. Gerli-Candia; M. Hurtado, *J. Chilean Chem. Soc.* **49**, 307 (2004).
12. J. S. Gómez-Jeria; L. Lagos-Arancibia, *Int. J. Quant. Chem.* **71**, 505 (1999).
13. G. Xu, L. Li, D. Wang, *Quantum chemistry: Fundamental principles and ab initio calculations*, Vol 2. Science Press, Beijing, p 798 (1985).
14. E.R. Davidson, *J Chem Phys* **46**, 3320 (1967).
15. K.R. Roby, *Mol Phys* **27**, 81 (1974).
16. R. Heinzmann, R. Ahlrichs, *Theoret Chim Acta* **42**, 33 (1976).
17. C. Ehrhardt, R. Ahlrichs, *Theoret Chim Acta* **68**, 231 (1985).
18. A.E. Reed, B.B. Weinstock, F. Weinhold, *J Chem Phys* **83**, 735 (1985).
19. A.E. Reed, L.A. Curtiss, F. Weinhold, *Chem Rev* **88**, 899 (1988).
20. J. Cioslowski, *J. Am Chem Soc* **111**, 8333 (1989).
21. S. Huzinaga, Y. Sakai, E. Miyoshi, S. Narita, *J Chem Phys* **93**, 3319 (1990).
22. Gaussian 03, Revision B.03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.
23. F. Soto-Morales, J. S. Gómez-Jeria, *J. Chil. Chem. Soc.* **52**, 1214 (2007).
24. J. S. Gómez-Jeria, L. Lagos-Arancibia, E. Sobarzo-Sánchez, *J. Chil. Chem. Soc.* **48**, 61 (2003).
25. J. S. Gómez-Jeria, F. Soto-Morales, J. Rivas, A. Sotomayor, *J. Chil. Chem. Soc.* **53**, 1382 (2008).
26. S. Arulmozhiraja, P. Kollandaivel, *Mol. Phys.* **90**, 55 (1997).
27. R. K. Roy, S. Pal, K. Hirao, *J. Chem. Phys.* **110**, 8236 (1999).
28. R. K. Roy, K. Hirao, S. Pal, *J. Chem. Phys.* **113**, 1372 (2000).
29. R. K. Roy, K. Hirao, S. Krishnamurty, S. Pal, *J. Chem. Phys.* **115**, 2901 (2001).
30. D. F. Perepichka, M. R. Bryce, *Angew. Chem. Int. Ed.* **44**, 5370 (2005).