Theoretical Chemistry Accounts (2026) 145:77
https://doi.org/10.1007/s00214-026-03327-4

RESEARCH

Correcting atomic charges to reproduce molecular dipoles

P. R. Surján[1] · A. Gombás[1,2] · Á. Szabados[1]

Received: 28 May 2026 / Accepted: 3 August 2026

Abstract
Electron densities are often subdivided into smaller units, which can then be allocated to atomic centers. The resulting atomic charges usually do not reproduce the total electronic dipole moment of the molecule. We describe a simple optimization procedure for finding atomic populations closest to some input atomic charges, e.g., the Mulliken or Bader ones, but reproduce the molecular dipole when the atomic coordinates are summed up weighted by the new charges. The proposed procedure can be coded by a few lines and is practically cost-less.

Keywords Atomic charges · Dipole moment · Charge fitting

1 Introduction

Many procedures have been proposed, mainly for interpretational purposes, which assign certain parts of the electron density $\rho(r)$ of a molecule to the constituting atoms identified by the coordinates of their nuclei. While any such divisions of $\rho(r)$ are arbitrary, they facilitate a simple, pictorial representation of a complicated electronic structure. The literature of atomic charges is voluminous, and we do not want to review it here (see e.g., Refs.[1–3]).

Among charges obtained from some partitioning of the molecular density, we mention merely Mulliken charges,[4] Löwdin population analysis,[5] or the atomic charges obtained from Bader's atoms in molecules approach,[6] without completeness. A different type of atomic charges results from the theory of generalized atomic polar tensors (GAPT),[7] in which atomic charges are obtained from dipole moment derivatives, carrying thus some experimental relevance.

Several pro-s and con-s can be listed for various type of atomic charges. One common problem of the above mentioned charges, denoted by q^0, is that they do not reproduce the dipole moment $\vec{D}$ of the molecule, i.e.,

$$\sum_{A}^{\text{atoms}} q_A^0 \, \vec{R}_A \neq \vec{D} := -\langle \Psi | \, \hat{\vec{r}} \, | \Psi \rangle + \sum_{A}^{\text{atoms}} Z_A \, \vec{R}_A \qquad (1)$$

in atomic units, where $\hat{\vec{r}}$ is the coordinate operator of electrons while Z_A, q_A^0 and $\vec{R}_A$-s are atomic numbers, net atomic charges and nuclear coordinates, respectively.

Owing to the fact that reproducing the molecular dipole would enhance the interpretative power of atomic charges, in the following section, we present a simple procedure to obtain a set of corrected atomic charges q that are as close to the original ones as possible, subject to the conditions that

- their sum agrees with the total charge of the system (zero, to have a translationally invariant dipole), and
- their dipole moment agrees with the dipole moment of the given molecule.

Atomic charges reproducing not only monopoles and dipoles, but higher moments also, have been proposed and investigated in Ref.[8] The aim of this paper is to propose a simpler approach reproducing merely the first two multipoles.

✉ P. R. Surján
peter.surjan@ttk.elte.hu

A. Gombás
gombasandras@student.elte.hu

Á. Szabados
agnes.szabados@ttk.elte.hu

[1] Institute of Chemistry, Laboratory of Theoretical Chemistry, Loránd Eötvös University, 112, P.O.B. 32, Budapest 1518, Hungary

[2] Hevesy Gyorgy PhD School of Chemistry, ELTE Eötvös Loránd University, Budapest, Hungary

Published online: 27 August 2026

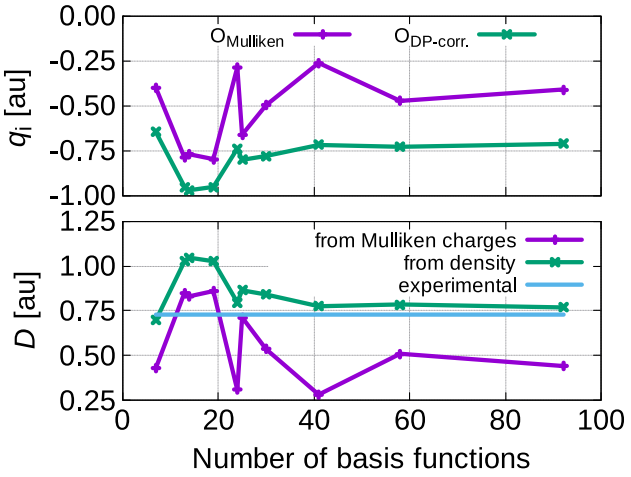

Fig. 1 Basis set dependence of the molecular dipole moment D, Mulliken and dipole-corrected atomic charges q_i of the water molecule. Dipole moments are calculated in each basis set from HF density and Mulliken atomic charges. The results can be compared to the experimental[10] value of 0.727 a.u. The basis sets STO-6 G (7), 6-31 G (13), Dunning-DZ (14), 6-311 G (19), cc-pVDZ (24), Dunning-DZP (25), 6-311 G** (30), aug-cc-pVDZ (41), cc-pVTZ (58) and aug-cc-pVTZ (92) were used, with respective number of basis functions indicated in parentheses

2 Theory

Given some initial atomic charges

$$q_A^0, \quad A = 1, 2, \ldots N_{\text{atoms}}$$

satisfying for a neutral molecule

$$\sum_A^{\text{atoms}} q_A^0 = 0 \tag{2}$$

and producing a dipole moment as

$$\sum_A^{\text{atoms}} q_A^0 \, \vec{R}_A = \vec{D}^0, \tag{3}$$

which is not equal to the known dipole moment $\vec{D}$ of the (neutral) molecule, we search modified charges q_A satisfying the following three conditions:

i) $\sum_A^{\text{atoms}} q_A = 0$, i.e., the total charge of the system remains the same
ii) $\sum_A^{\text{atoms}} q_A \, \vec{R}_A = \vec{D}$, i.e., the new charges reproduce the molecular dipole, and
iii) the new charges q_A are as close to the input charges q_A^0 as possible in a least-square sense.

To find these, consider the Lagrangian

$$\mathcal{L} = \sum_A^{\text{atoms}} \left(q_A^0 - q_A\right)^2 + 2\vec{\lambda}\left(\sum_A^{\text{atoms}} q_A \, \vec{R}_A - \vec{D}\right) + 2\nu \sum_A^{\text{atoms}} q_A. \tag{4}$$

We search for q_A that minimize $\mathcal{L}$. Here, $\vec{\lambda}$ and ν are Lagrangian multipliers, $\vec{\lambda}$ being a 3-component vector. The role of ν is to ensure charge conservation while $\vec{\lambda}$ ensures the reproduction of the molecular dipole $\vec{D}$. To minimize $\mathcal{L}$, we set

$$\frac{\partial \mathcal{L}}{\partial q_B} = 0, \qquad B = 1, 2, \ldots N_{\text{atoms}} \tag{5}$$

to obtain

$$\left(q_B^0 - q_B\right) + \vec{\lambda}\vec{R}_B + \nu = 0. \tag{6}$$

In Eq. (4), the factors of 2 in front of the 2nd and 3rd terms were introduced for convenience. Equation(6) can be solved for q to get

$$q_B = q_B^0 + \vec{\lambda} \cdot \vec{R}_B + \nu. \tag{7}$$

From condition i). one can write:

$$\underbrace{\sum_B q_B}_{0} = \underbrace{\sum_B q_B^0}_{0} + \sum_B \vec{\lambda} \cdot \vec{R}_B + \nu \, N_{\text{atoms}}, \tag{8}$$

from which multiplier ν results as

$$\nu = -\frac{1}{N_{\text{atoms}}} \sum_B \vec{\lambda} \cdot \vec{R}_B. \tag{9}$$

It remains to eliminate the vectorial multipliers $\vec{\lambda}$. This can be done using condition ii).:

$$\underbrace{\sum_B q_B \vec{R}_B}_{\vec{D}} = \underbrace{\sum_B q_B^0 \vec{R}_B}_{\vec{D}^0} + \sum_B \vec{R}_B \left(\vec{\lambda} \cdot \vec{R}_B\right) + \nu \sum_B \vec{R}_B \tag{10}$$

This, upon substituting ν from (9), represents a 3×3 inhomogeneous linear equation for the components of $\vec{\lambda}$ of the form

$$\sum_{\beta=1}^{3} A_{\alpha\beta} \, \lambda_\beta = b_\alpha, \qquad \alpha, \beta = 1, 2, 3 = x, y, z, \tag{11}$$

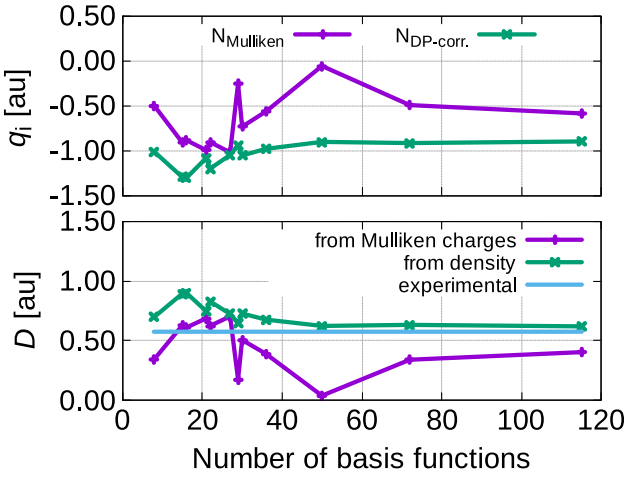

Fig. 2 Same as in Fig. 1., for the ammonia molecule. The experimental dipole moment value of 0.579 a.u. was taken from [11]. The basis sets STO-6 G (8), 6-31 G (15), Dunning-DZ (16), 6-31 G* (21), 6-311 G (22), 6-311 G* (27), cc-pVDZ (29), Dunning-DZP (30), 6-311 G** (36), aug-cc-pVDZ (50), cc-pVTZ (72) and aug-cc-pVTZ (115) were used, with the respective number of basis functions indicated in parentheses

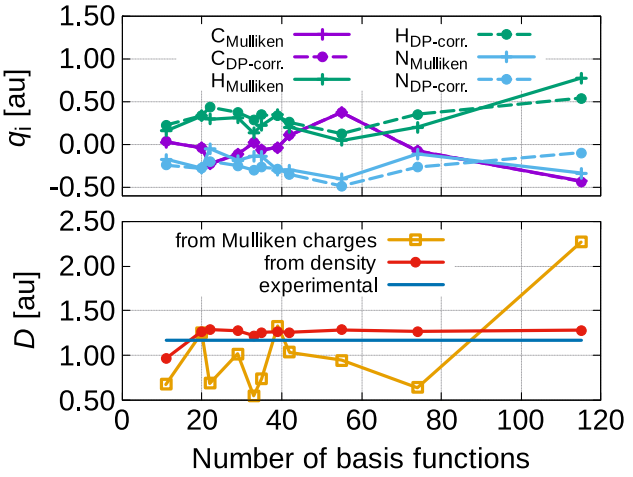

Fig. 3 Basis set dependence of the molecular dipole moment D, Mulliken and dipole-corrected atomic charges q_i of the linear HCN molecule. Dipole moments are calculated in each basis set from HF density and Mulliken atomic charges, results can be compared to the experimental[13] value of 1.174 a.u. The basis sets STO-6 G (11), 6-31 G (20), Dunning-DZ (22), 6-311 G (29), cc-pVDZ (33), Dunning-DZP (35), 6-311 G* (39), 6-311 G** (42), aug-cc-pVDZ (55), cc-pVTZ (74) and aug-cc-pVTZ (115) were used with the respective number of basis functions indicated in parentheses

with $\vec{b} = \vec{D} - \vec{D}^0$ and

$$A_{\alpha\beta} = \sum_B R_{B\alpha} R_{B\beta} - \frac{v_\alpha v_\beta}{N_{\text{atoms}}} \tag{12}$$

with

$$v_\alpha = \sum_B R_{B\alpha}. \tag{13}$$

If the coefficient matrix A in (11) turns out to be non-invertible, which happens for planar or linear molecules, we take its Moore–Penrose pseudoinverse.[9] In situations when the coefficient matrix is *nearly* singular, a careful choice of the threshold for the numerical zero may be necessary when evaluating the Moore–Penrose inverse. Once Eq. (11) is solved for λ-s, ν results from (9), and the final charges are obtained from (7).

It is worth emphasizing that the initial charges q_0 may result from any definition, and if they happen to reproduce the dipole moment, they will not be affected by the above procedure. Also, the molecular dipole moment $\vec{D}$ to be reproduced can be calculated at any level of theory (Hartree–Fock (HF), full configuration interaction (FCI), …), or can originate from measurement.

The above procedure handles all atoms on an equal footing, i.e., we use equal weights for them during minimization. Note that this is only the simplest choice, as appropriate weighting factors could also be introuduced, preferably based on some physical properties of the atoms.

3 Numerical results

Below, we present a few selected examples to illustrate the theory. A case to be nontrivial, the number of unknowns (the atomic charges q) should be larger than the number of relevant constraints, otherwise the constraints determine the final charges independently of the initial ones. In the general case, one has four constraints: the reproduction of the total charge and the three components of $\vec{D}$. This means that at least five nonequivalent atoms are needed for a nontrivial case in a nonplanar molecule. Diatomics, e.g., will not be treated here, as their dipole moments and bond lengths determine the partial atomic charge. Similarly, neither the water nor the ammonia molecule is nontrivial as they contain just two symmetry non-equivalent atoms. Exceptions occur if the molecular symmetry sufficiently reduces the number of relevant constraints. In case of a linear triatomic molecule, for example, two components of the dipole are zero irrespective of the atomic charges, thus the no. of relevant constraints is just two. With its three unknown atomic charges, if the molecule is formed by three different type of atoms, this case constitutes a nontrivial example. We shall illustrate this on the hydrogen cyanide molecule below in subsection 3.3.

3.1 Water

The H_2O molecule, as mentioned above, represents a trivial example in the sense that, taking the C_{2v} symmetry into account, the two constraints ($\sum q_A = 0$ and fixing D_z)

Table 1 Original and corrected Mulliken and Löwdin charges of the HCN molecule computed in cc-pVTZ basis set, each reproducing the HF dipole $\vec{D} = 1.28038$ a.u

Atom	Mulliken		Löwdin	
	Original	Dipole-corrected	Original	Dipole-corrected
H	0.202	0.343	−0.008	0.294
C	−0.076	−0.072	0.015	0.023
N	−0.126	−0.271	−0.007	−0.317

Table 2 Correction of Mulliken type charges obtained from FCI densities of the HCN to reproduce the FCI dipole 0.9695 a.u. (MINI) and 1.1647 a.u. (DZ)

Atom	Huzinaga MINI basis		Dunning DZ basis, frozen core	
	Original	Dipole-corrected	Original	Dipole-corrected
H	0.281	0.286	0.255	0.407
C	−0.105	−0.105	−0.253	−0.249
N	−0.176	−0.181	−0.002	−0.158

uniquely determine the two free parameters: q_O and q_H. Therefore, it is immaterial which charges q^0 initiate the dipole correction procedure. Nevertheless, it is worth checking the variation of atomic charges with the basis set when they are required to reproduce the molecular dipole moment. This is shown in Fig. 1. in comparison with basis set variation of Mulliken charges. The C_{2v} geometry of water was specified as $r_{OH} = 0.940995$ Å and $\alpha_{HOH} = 105.4628°$. It is apparent that the dipole-corrected atomic charges show significantly less sensitivity. It is also clear that the dipoles calculated from Mulliken charges are very much basis set dependent, while the HF dipoles converge fast to their large basis limit, showing a small deviation from the experimental value reflecting a minor correlation effect.

3.2 Ammonia

A second trivial example we show here is the NH_3 molecule. Here again a single component of the dipole moment survives the C_{3v} symmetry, which, together with the charge conservation, determine the two non-equivalent atomic charges. The results are presented in Fig. 2. showing that the dipole moment constraint significantly damps the basis set dependence.

3.3 HCN

For the hydrogen cyanide molecule, a standard linear geometry was used with $r_{HC} = 1.065$ and $r_{CN} = 1.153$ Å. We report below dipole-corrected charges using dipole moments both at the HF and FCI levels.

3.3.1 Hartree–Fock dipole moments

Table 1 shows partial atomic charges in the linear HCN molecule obtained from the HF-SCF wave function in Dunning's cc-pVTZ basis set[12]. Both the Mulliken and Löwdin population analyses are shown. It is apparent that the two analyses result in rather different values for the atomic charge, while after dipole corrections the resulting charges deviate less, agreeing at least in their first decimals. For example, Löwdin population analysis finds the N atom slightly negative, while its Mulliken charge is a much larger negative value. After dipole correction, both became around −0.3. Qualitatively, dipole correction yields the (H,C,N) charge pattern (+,0,–) in both cases. The large positive Mulliken charge on H becomes a small negative number in the Löwdin case, but after dipole correction, both are predicted to be $\sim$ 0.3. In general, reproducing the HF dipole moment of HCN in the cc-pVTZ basis set applied, the dipole correction predicts a much more polarized molecule than either Mulliken or Löwdin population analysis.

Figure 3 depicts the dependence of Mulliken type atomic charges in various basis sets both before and after dipole corrections. The figure also shows the basis set dependence of the HF-SCF dipole moments.

The dipole moments calculated from Mulliken charges illustrate their strong basis set sensitivity due to the instability arising from the classification of large diffuse basis functions to the individual atoms. The dipole moments calculated from HF densities quickly converge with increasing number of basis functions with only the remainder of a minor correlation gap from the experimental value.

Figure 3 also shows that dipole correction of Mulliken charges does not remedy the problem of basis set sensitivity in this example. It practically only corrects the charges on the terminal H and N atoms leaving the charge on the central C atom unchanged.

3.3.2 Full CI dipole moments

First, we generalized the standard, SCF-based Mulliken population analysis to the correlated case simply by replacing the HF density matrix with the FCI one. This is achieved

Table 3 Original and dipole-moment-corrected gross atomic charges [a.u.] in the C_s DMSO molecule computed in 6-311 G** basis set at the HF-SCF level. Only the symmetrically non-equivalent atoms are shown

	Mulliken		GAPT		Bader	
Atom	original	Dipole-corrected	Original	Dipole-corrected	Original	Dipole-corrected
S	0.756	0.689	0.899	0.781	1.447	1.327
O	−0.726	−0.712	−0.846	−0.830	−1.424	−1.201
C	−0.417	−0.420	−0.047	−0.051	0.026	0.000
H	0.147	0.120	0.025	−0.022	−0.023	−0.075
H	0.120	0.070	−0.002	−0.083	0.005	0.279
H	0.135	0.241	−0.002	0.180	−0.019	−0.266

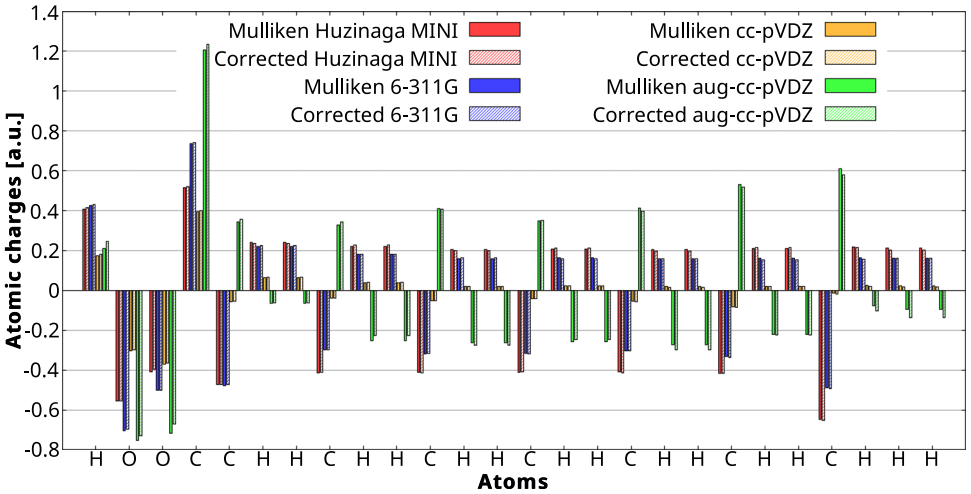

Fig. 4 Mulliken- and dipole-corrected Mulliken atomic charges in the Caprylic acid molecule

by computing matrix P with the FCI wave function in the orthonormal molecular orbital (MO) basis

$$P_{ij}^{\mathrm{MO}} = \langle FCI \mid \sum_{\sigma} a_{i\sigma}^{\dagger} a_{j\sigma} \mid FCI \rangle,$$

where $a^{\dagger}$ (a) create (annihilate) electrons on the MOs, σ is the spin label. This matrix was then transformed to the basis of atomic orbitals:

$$P = C\ P^{\mathrm{MO}}\ C^{\dagger},$$

where C is the matrix of MO coefficients. Though neither P^{MO} nor PS (with S denoting the AO overlap) are idempotent for correlation effects, both of their traces yield the total number of electrons in the system.

The electronic part of the dipole moment components at the FCI level emerge simply by evaluating the (negative) traces of matrix products PX, PY and PZ, with $X, \ldots$ being the AO matrix elements of the respective coordinate operators. Finally, FCI Mulliken atomic charges emerge by computing the partial traces

$$q_A^0 = \sum_{\mu \in A} (PS)_{\mu\mu}$$

for atom A, just like in the SCF case.

Table 2 reports the thus obtained FCI-Mulliken charges before and after dipole corrections in the Huzinaga MINI[14] and Dunning's double zeta[15] bases, in the latter case in the frozen core approximation. Comparing the results to the Mulliken charges in Table 1, we see that the correlation contribution to the atomic charges is moderate. The slight basis set dependence of the total dipole moment is enhanced by Mulliken charges, yielding more polar bonds in Dunning DZ than in MINI. This feature presists at the dipole-corrected level. Watching for the effect of dipole correction, charges of terminal atoms are significantly increased in absolute values in Dunning DZ, while only a modest change is observed in Huzinaga MINI. Original atomic charges are characteristically different for atom N and get similar in the two basis sets upon dipole correction. The converse holds for atom H, showing similar charges originally but getting increased by some 150 % in Dunning DZ.

3.4 DMSO

Table 3. shows dipole-corrected atomic charges of the dimethyl sulfoxide (DMSO) molecule in comparison with the original ones obtained from different methods: (i) standard Mulliken partitioning, (ii) the generalized atomic polar tensors (GAPT) method,[7] where the definition of the charge of atom A is

$$q_A = \frac{1}{3}\,\mathrm{Tr}\,\frac{\partial \vec{D}}{\partial \vec{R}_A} = \frac{1}{3}\sum_{\mu=1}^{3}\frac{\partial D_\mu}{\partial R_{A\mu}}, \tag{14}$$

and (iii) the three dimensional space is partitioned according to Bader's atoms in molecules (AIM) theory.[6] The GAPT and Bader's charges in DMSO are taken from Ref.[16].

According to the table, the S atom is largely positive, while the charge of O is a large negative number according to all the three population analyses. This feature is left unchanged upon dipole correction. Atom C, however, behaves differently: it has a large negative charge according to Mulliken analysis, shows a slight negative charge by GAPT, while Bader analysis predicts a small positive charge. The dipole correction qualitatively preserves these features, too. Mulliken analysis yields rather polarized methyl groups with positive H atoms, while the GAPT and Bader's methods show rather neutral methyls. Dipole correction tries to reflect the original features, though in some cases (two H atoms in GAPT and Bader) it brings an order of magnitude increase in absolute value of q_H.

3.5 Caprylic acid

As a final test, we present results for Caprylic acid, COOH-$(CH_2)_6$-CH_3 in various basis sets in Fig. 4. They indicate that the dipole- correction modifies the original charges only little, and does not induce artificial oscillations on the apolar tail of this molecule.

4 Conclusion

A procedure is described which modifies atomic charges, obtained from any types of partitioning method, to the smallest possible extent, conserving the total (zero) charge but reproducing the dipole moment of the molecule. Based on the pilot calculations presented here, one cannot establish a simple rule on *how* the atomic charges change. In the limiting trivial cases, where the charge- and dipole-constraints determine the values of point charges put on the nuclei (cf. e.g., diatomics, water or ammonia), the resulting charges will be independent of the initial ones. Nontrivial cases emerge in systems having more than four non-equivalent atoms, or for smaller molecules if symmetry reduces the number of constraints. The correction procedure shows certain stabilizing effect, but in the nontrivial cases the dipole correction aims to preserve the features of the initial charges even if these turn out very different by different atomic charge definitions. The algorithm providing the dipole-corrected charges is simple, a short FORTRAN code is available from the authors upon request.

Acknowledgements Part of the computations have been performed with the Budapest version[17] of the MUNGAUSS program system[18]. The full CI computation was done by a code written by Zoltán Rolik (ELTE, Budapest) based on Olsen's algorithm[19]. We thank for partial support from National Research, Development and Innovation Office (NKFIH) under the Hungarian National Excellence Programme (NKKP) ADVANCED grant scheme, No. 150283. A.G. acknowledges the grant from Excellence Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development And Innovation Fund (EKÖP-25).

Author Contributions P.S. initiated the project and wrote the first draft, made the derivations A.G. checked the derivations, performed calculations, revised the paper A.Sz. checked the derivations, supervised the work of A.G., and revised the paper

Funding Open access funding provided by Eötvös Loránd University.

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors do not have conflict of interest in connection with this MS.

References

1. Davidson ER, Clark AE (2022) Int J Quantum Chem 122(8):e26860. https://doi.org/10.1002/qua.26860
2. Mao JX (2014) Journal of Postdoctoral Research 2(2)
3. Edgecombe KE, Boyd RJ (1987) Atomic orbital populations and atomic charges from self-consistent field molecular orbital wavefunctions. J Chem Soc Faraday Trans 2 Mol Chem Phys 83(8):1307
4. Mulliken RS (1955) J Chem Phys 23:1833
5. Löwdin PO (1950) J Chem Phys 18:365

6. Bader R (1990) Atoms in molecules: a quantum theory. Oxford University Press, Oxford
7. Cioslowski J (1989) J Am Chem Soc 111(22):8333. https://doi.org/10.1021/ja00204a001
8. Gruschus JM, Kuki A (1990) J Comput Chem 11(8):978. https://doi.org/10.1002/jcc.540110810
9. Israel AB, Greville T (2003) Generalized Inverses: Theory and Applications, 2nd edn. Springer, New York
10. Clough SA, Beers Y, Klein GP, Rothman LS (1973) J Chem Phys 59(5):2254
11. Weber W (1984) J Mol Spectrosc 107(2):405
12. Dunning TH Jr (1989) J Chem Phys 90(2):1007. https://doi.org/10.1063/1.456153
13. Ebenstein WL, Muenter J (1984) J Chem Phys 80(9):3989
14. Andzelm J, Huzinaga S, Klobukowski M, Radzio-Andzelm E, Sakai Y, Tatewaki H (1984) in *Gaussian Basis Sets for Molecular Calculations*, *Physical Sciences Data*, vol. 16, ed. by S. Huzinaga (Elsevier), pp. 27–426. https://doi.org/10.1016/B978-0-444-42254-5.50008-0
15. Dunning TH (1970) J Chem Phys 53:2823. https://doi.org/10.1063/1.1674408
16. Cioslowski J, Surján PR (1992) J Mol Struct (THEOCHEM) 255:9
17. Surján PR, Program B-M (1995) Laboratory of Theoretical Chemistry. Inst Chem LTE Loránd Eötvös University (Tech. rep.)
18. Poirier R, Peterson M, Program M (1989). Dept. Chemistry, Memorial Univ.St.Johns,(Tech. rep)
19. Olsen J, Roos BO, Jørgensen P, Jensen HJA (1988) J Chem Phys 89:2185