Atomic charges in molecules: Hirshfeld-style partitioning

Sergei Vyboishchikov

- The *total charge* of a system is a fundamental quantity in physics;
- A charge ("partial charge") of an individual atom in a polyatomic molecule is not well-defined in quantum mechanics.
  - ▶ *Question*: Do we need **atomic charges** at all and why?
    - – Interpretative purposes (*a posteriori* analysis of the wavefunction);
    - – Very important for descriptive chemistry ($\delta+$, $\delta-$);
    - – Description of electrostatic interaction in molecular mechanics and dynamics and for solvent effects.
- The total molecular charge must be the sum of the atomic charges.
  - ▶ *Question*: how to attribute charges to various atoms within a molecule?
    - – A nucleus (and its charge) is attributed to the respective atom.
    - – But what to do with electrons? – this is a question.
      - · *Answer*: we usually (not always) have to *partition* the electron density *or* the density matrix, though some other approaches also exist.
- Example of a charge scheme *not* based on partitioning:
  Electrostatic (ESP) charges (Cox–Williams, Merz–Kollman): designed to *fit* the atomic charges such that the *electrostatic potential* of the real molecule is reproduced as well as possible by the atonic charges.

- **Bader** charges:

$$Q_A^{\text{Bader}} = Z_A - \int\limits_{\Omega_A} \rho(\mathbf{r})d^3\mathbf{r}$$
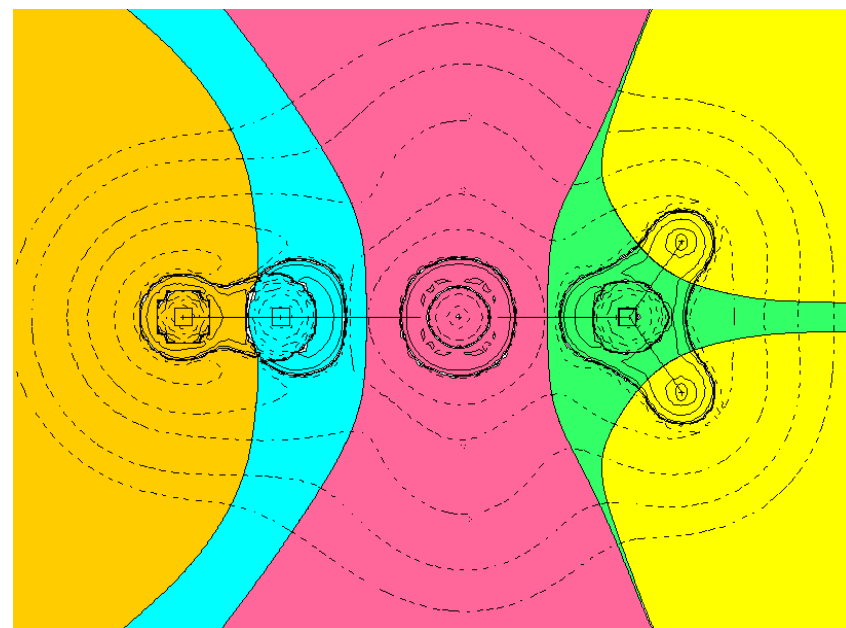
- "Stockholder" (**Hirshfeld**-style) methods:

$$Q_A = Z_A - \int \rho_A(\mathbf{r})d^3\mathbf{r}$$

$$\rho_A(\mathbf{r}) = \rho(\mathbf{r})\omega_A(\mathbf{r}), \quad \text{where } \omega_A(\mathbf{r}) \text{ are } \textit{weighting functions}$$

$$\sum_A \omega_A(\mathbf{r}) = 1$$

- Significant freedom in choosing *weighting functions* $\omega_A(\mathbf{r})$

- **Hirshfeld** partitioning – available in Gaussian™: `pop=Hirshfeld`

$$Q_A = Z_A - \int \rho_A(\mathbf{r}) d^3\mathbf{r}$$

$$\rho_A(\mathbf{r}) = \rho(\mathbf{r})\, \omega_A(\mathbf{r}) = \rho(\mathbf{r}) \frac{w_A(\mathbf{r})}{\sum_I w_I(\mathbf{r})}$$

Weighting functions $w_A(r)$: spherically symmetric free-atom densities

(must be pre-calculated $\Rightarrow$ a reference is needed)

$\sum_I w_I(\mathbf{r})$ is a *promolecule* density
*Problem*: whether to form $w_I(\mathbf{r})$ from neutral atoms or ions?

- The electron density in atomic ground states *always* has a single sharp maximum at the nuclear position.

- The electron density in molecular ground states has sharp maxima at the nuclear positions. The density values in the bond area are relatively small.

N_2 density in atomic units
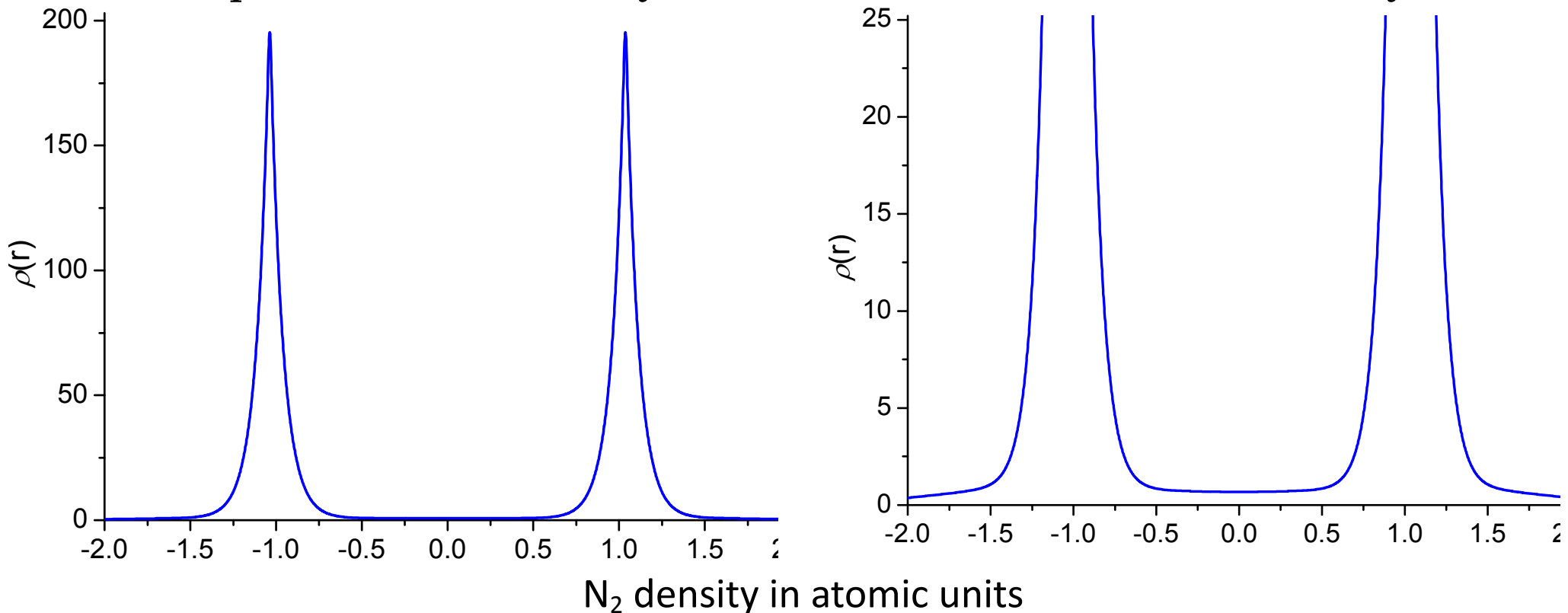
- In the case of the hydrogen atom, the maximum is very weakly pronounced due to a lack of the core density. But it still exists always:
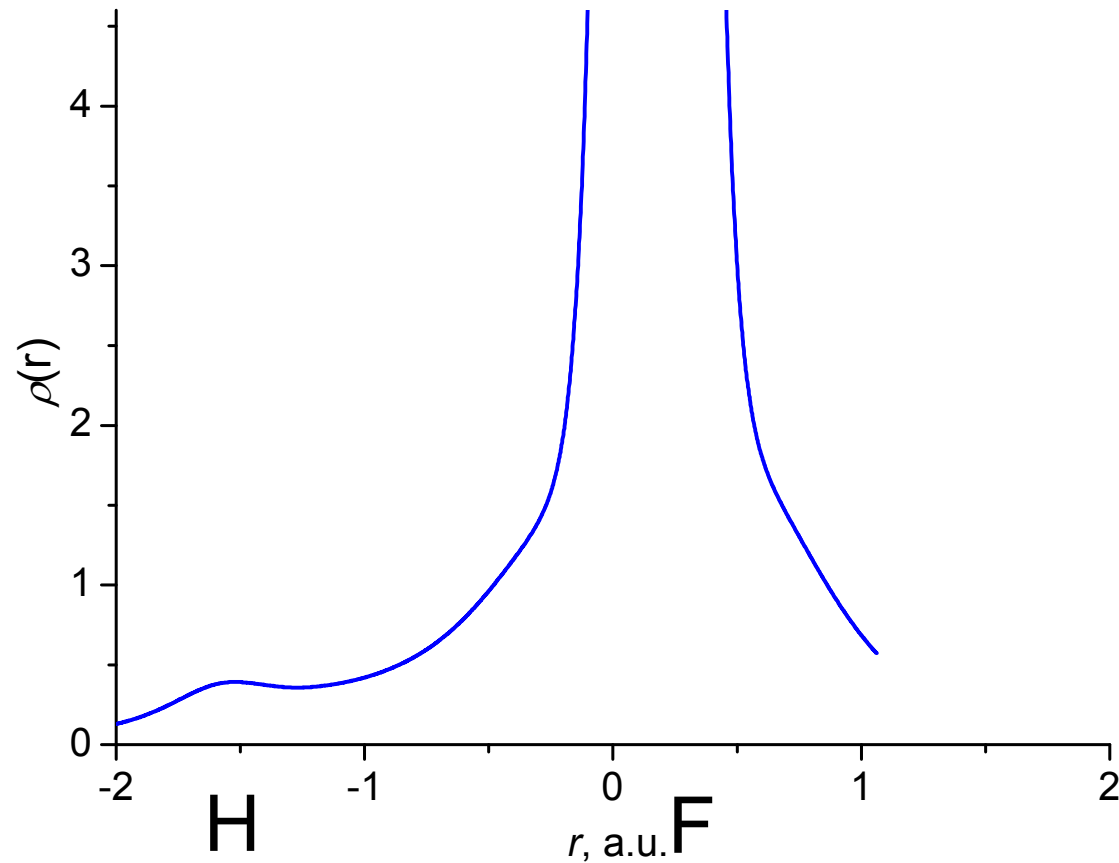
- Further developments based on the Hirshfeld partitioning:

  ▶ Hirshfeld Iterative method (Bultinck *et al*.) – available in Gaussian™

  The iterative procedure solves the issue of atomic/ionic pro-atom.

  ▶ CM5 (Truhlar and co-w.) – available in Gaussian™

  ▶ ISA (*Iterative stockholder analysis*, Lillestolen and Wheatley)

  available from me

  ▶ MBIS (*Minimal basis iterative stockholder*, Ayers, Bultinck *et al*.)

  ▶ ACP (*Adjusted charge partitioning*).

  available from https://github.com/vyboishchikov/ACP

  ▶ Iterative ACP. Available from https://github.com/vyboishchikov/ACP

  · ACP, I-ACP, and ISA require a `wfn` file (not `xfn`!) as input.

- ISA – *Iterative stockholder analysis* (Lillestolen and Wheatley 2009):
  - ► Like Hirshfeld, but without predefined reference atomic densities!
  - ► ISA algorithm:
    - All initial weights are set to unity: $w_A(r) = 1$.
    - Initial atomic densities are obtained:

$$\rho_A(\mathbf{r}) = \rho(\mathbf{r}) \frac{w_A(\mathbf{r})}{\sum_I w_I(\mathbf{r})}$$

    - Atomic densities are spherically averaged to produce new weights:

$$w_A(r) = \langle \rho_A(\mathbf{r} - \mathbf{R}_A) \rangle_{\text{sphe}}$$

    The center of the sphere is put in the position $\mathbf{R}_A$ of nucleus A.

    - Back to the previous step; to repeat until converged.

+ No reference densities; self-adjusted weights instead.

+ $\sum_A w_A(\mathbf{r})$ gives the best-possible approximation of $\rho(\mathbf{r})$ by a sum of spherically symmetric atomic densities (see next page for more details).

+ Quite good dipole moments for CHNO-containing molecules.

– Not so good dipole moments for other molecules.

– Often very (!) slow convergence.

- Lillestolen and Wheatley's $\sum_A w_A(\mathbf{r})$ is the best-possible promolecule: approximation of $\rho(\mathbf{r})$ by a sum of spherically symmetric atomic densities.

  ▶ why is it so important?

  – Because for spherical densities $\rho^{\mathrm{sphe}}(\mathbf{r}) = \sum_A \rho_A(\mathbf{r})$

$$\mu^{\mathrm{ch}} = \sum_A Q_A \mathbf{R}_A \qquad Q_A = Z_A - \int \rho_A(\mathbf{r}) d^3\mathbf{r}$$

  is identical to

$$\mu(\rho^{\mathrm{sphe}}) = \sum_A Z_A \mathbf{R}_A - \int \mathbf{r} \rho^{\mathrm{sphe}}(\mathbf{r}) d^3\mathbf{r}$$

  Exercise: recall the general definition of the *dipole moment* for a molecule with nuclear charges $\{Z_A\}$ at positions $\{\mathbf{R}_A\}$ and electron density $\rho(\mathbf{r})$.

- Hirshfeld, Hirshfeld Iterative, CM5, MBIS, ACP, and Iterative ACP are all various *approximations* to Lillestolen and Wheatley's ISA method.

- Electron density: valence density ρ^{val} – no core electrons!

- Partitioning: the same as Hirshfeld, but with a Slater function as an analytical weighting function:

$$w_A(r) = c_A N_A^{\mathrm{val}} \mathcal{N}\, r^{2n_A - 2} e^{-\alpha_A r} \qquad \rho_A^{\mathrm{val}}(\mathbf{r}) = \rho^{\mathrm{val}}(\mathbf{r})\frac{w_A(r)}{\sum\limits_I w_I(r)}$$

$$Q_A = N_A^{\mathrm{val}} - \int \rho_A^{\mathrm{val}}(\mathbf{r})d^3\mathbf{r}$$

N_A^{val} : the number of valence electrons;

n_A : is the principal quantum number of the valence shell;

c_A: related to electronegativity (fixed parameter)

α_A: related to atomic radius (fixed parameter)

- Adjustment: non-linear fitting α_A and c_A are empirically adjusted for 17 elements: H, Li–F, Na–Cl, Br, and I, minimizing the root mean-square error of dipole moments $\sum_i (\mu_i^{\mathrm{ch}} - \mu_i^{\mathrm{ref}})^2$.

- Partitioning: the same equation as before, but c_A is dynamically adjustable, not fixed!

$$w_A(\mathbf{r}) = c_A N_A^{\text{val}} \mathcal{N} \, r^{2n_A - 2} e^{-a_A r} \qquad \rho_A^{\text{val}}(\mathbf{r}) = \rho^{\text{val}}(\mathbf{r}) \frac{w_A(\mathbf{r})}{\sum\limits_I w_I(\mathbf{r})}$$

$$Q_A = N_A^{\text{val}} - \int \rho_A^{\text{val}}(\mathbf{r}) d^3\mathbf{r}$$

(1) All c_A are initially set to 1;

(2) The weighting functions are calculated:

$$w_A(\mathbf{r}) = c_A N_A^{\text{val}} \mathcal{N} \, r^{2n_A - 2} e^{-a_A r}$$

(3) The atomic densities $\rho_A^{\text{val}}(\mathbf{r})$ are calculated :

$$\rho_A^{\text{val}}(\mathbf{r}) = \rho^{\text{val}}(\mathbf{r}) \frac{w_A(\mathbf{r})}{\sum\limits_I w_I(\mathbf{r})}$$

(4) The new c_A values are computed:

$$c_A = \frac{\int \rho_A^{\text{val}}(\mathbf{r}) d^3\mathbf{r}}{N_A^{\text{val}}}$$

(5) Step (2) and the subsequent steps are repeated with the new c_A values until the procedure converges.

- Statistics (μ^{ch} *vs* μ^{ref} root mean-square error)