

$$\begin{aligned}
E_{tot} &= E_{TB} + E_{disp} + E_{charge} + E_{rep} \\
E_{TB} &= \\
E_{disp} &= \\
E_{charge} &= \\
E_{GFN2-xTB} &= E_{rep}^{(0)} + E_{disp}^{(0,1,2)} + E_{EHT}^{(1)} + E_{IES+IXC}^{(2)} + E_{AES+AXC}^{(2)} + E_{IES+IXC}^{(3)} \\
&= E_{rep} + E_{disp}^{D4'} + E_{EHT} + E_{\gamma} + E_{AES} + E_{AXC} + E_{\Gamma}^{GFN2} \quad (1)
\end{aligned}$$

$$\begin{aligned}
E_{rep} &= \frac{1}{2} \sum_{A,B} \frac{Z_A^{eff} Z_B^{eff}}{R_{AB}} e^{-\sqrt{a_A a_B} (R_{AB})^{(k_f)}} \\
k_f &= \begin{cases} 1 & \text{if } A, B \in \{\text{H, He}\} \\ \frac{3}{2} & \text{otherwise} \end{cases} \quad (2)
\end{aligned}$$

$Z^{eff}$  and  $a$  are variables fitted for each element. A,B are the labels of atoms. Since we only have C and H in our systems we can simplify this quite a bit in code.  $R_{AB}$  is the distance between the A and B atoms

$$E_{EHT} = \sum_{\mu\nu} P_{\mu\nu} + H_{\mu\nu}^{EHT} \quad (3)$$

where  $\mu$  and  $\nu$  are AO indecies,  $l$  and  $l'$  index shells. Both AO's are associated with an atom labled A and B.

$$\begin{aligned}
P_{\mu\nu} &= P_{\mu\nu}^{(0)} + \delta P_{\mu\nu} \\
H_{\mu\nu} &= \frac{1}{2} K_{AB}^{ll'} S_{\mu\nu} (H_{\mu\mu} + H_{\nu\nu}) \\
&\cdot X(EN_A, EN_B) \\
&\cdot \Pi(R_{AB}, l, l') \\
&\cdot Y(\zeta_l^A, \zeta_{l'}^B), \forall \mu \in l(A), \nu \in l'(B)
\end{aligned} \quad (4)$$

$K_{AB}^{ll'}$  is a element and shell specific fitted constant however, in GFN2 it only depends on the shells.  $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$  is just the overlap of the orbitals. In GFN2  $H_{\kappa\kappa} = h_A^l - \delta h_{CN'_A}^l CN'_A$  where  $CN'_A$  is the modified GFN2-type Coordinate Number for the element of atom A.  $h_A^l$  and  $\delta h_{CN'_A}^l$  are both fitted constants.  $EN_A$  is the electronegativity of the element of atom A.

$$\begin{aligned}
X(EN_A, EN_B) &= 1 + k_{EN} \Delta EN_{AB}^2 \\
k_{EN} &= 0.02 \text{ in GFN2} \\
\Delta EN_{AB}^2 &= (EN_A - EN_B)^2
\end{aligned} \quad (5)$$

The electronegativity for C and H are 2.55 and 2.20 according to wikipedia. Thus here is a table for the combinations we will be working with:

A	B	$X(EN_A, EN_B)$
C	C	1
C	H	$1 + 0.02 \cdot (0.35^2)$
H	C	$1 + 0.02 \cdot (0.35^2)$
H	H	1

$$\Pi(R_{AB}, l, l') = \left( 1 + k_{A,l}^{\text{poly}} \left( \frac{R_{AB}}{R_{\text{cov},AB}} \right)^{\frac{1}{2}} \right) \left( 1 + k_{B,l'}^{\text{poly}} \left( \frac{R_{AB}}{R_{\text{cov},AB}} \right)^{\frac{1}{2}} \right) \quad (6)$$

$R_{\text{cov},AB}$  are the summed covalent radii and taken from Reference 61<sup>1</sup>, the covalent radii is the second number in the table for each element,  $H_{\text{cov}} = 0.32$ ,  $C_{\text{cov}} = 0.75$  (can be found on wiki as 0.75).  $k_{A,l}^{\text{poly}}$  and  $k_{B,l'}^{\text{poly}}$  are element and shell specific constants.

$$Y(\zeta_l^A, \zeta_{l'}^B) = \left( \frac{2\sqrt{\zeta_l^A \zeta_{l'}^B}}{\zeta_l^A + \zeta_{l'}^B} \right)^{\frac{1}{2}} \quad (7)$$

Here,  $\zeta_l^A$  are the STO exponents of the GFN2-xTB AO basis. Slater Type Orbitals are defined as such:

$$\chi_{\zeta,n,l,m}(r, \theta, \varphi) = N Y_{l,m}(\theta, \varphi) r^{n-1} e^{-\zeta r}$$

N is a normalisation constant, Y are spherical harmonic functions, n, l, m are the quantum numbers for the AO.  $r, \theta, \varphi$  are polar 3D coordinates.  $\zeta$  determines the radial extent of the STO, a large value gives rise to a function that is "tight" around the nucleus and a small value gives a more "diffuse" function. This  $\zeta$  is the one mentioned in the Y term of  $E_{\text{EHT}}$  and is a value fitted when constructing the basis set.

---

<sup>1</sup>p.2109, <https://www.taylorfrancis.com/books/mono/10.1201/b12286/crc-handbook-chemistry-physics-william-haynes>