1 Total Energy for GFN2-xTB

Most equations etc. are from the xTB review paper or the GFN2-xTB paper.

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

1.1 Repulsion Energy

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

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.

1.2 Extended Hückel Theory Energy

$$E_{EHT} = \sum_{\mu\nu} P_{\mu\nu} + H_{\mu\nu}^{EHT} \tag{3}$$

where μ and ν are AO indecies, l and l' index shells. Both AO's are associated with an atom labled A and B.

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

QUESTION: How do we calculate the density matrix or the zeroth order and delta terms?

1.3 Isotropic electrostatic and XC energy

1.3.1 Second order

$$
\begin{aligned}
E_\gamma &= \frac{1}{2}\sum_{A,B}^{N_{atoms}}\sum_{l\in A}\sum_{l'\in B} q_l q_{l'}\gamma_{AB,ll'} \\
\gamma_{AB,ll'} &= \frac{1}{\sqrt{R_{AB}^2+\eta_{AB,ll'}^{-2}}} \\
\eta_{AB,ll'} &= \frac{1}{2}\left[\eta_A(1+k_A^l)+\eta_B(1+k_B^{l'})\right]
\end{aligned}
\tag{5}
$$

q_l is a partial muliken charge. η_A and η_B are element-specific fit parameters, while k_A^l and $k_B^{l'}$ are element-specific scaling factors for the individual shells ($k_A^l = 0$ when $l = 0$).

1.3.2 Third order

$$
E_\Gamma^{GFN2} = \frac{1}{3}\sum_{A}^{N_{atoms}}\sum_{l\in A}(q_l)^3 K_l^\Gamma \Gamma_A \tag{6}
$$

K_l^Γ is a shell specific constant common for all elements and Γ_A is an element specific constant.
QUESTION: We can do all of this, except get the partial charges. We assume we have to do Self Consistent Charge, but how we do that is not clear from the GFN2-xTB paper, the xTB review paper or Frank Jensen CH 7.6. Any hints?

1.4 Anisotropic electrostatic energy

$$
\begin{aligned}
E_{AES} &= E_{q\mu} + E_{q\Theta} + E_{\mu\mu} \\
&= \frac{1}{2}\sum_{A,B}\{f_3(R_{AB})[q_A(\boldsymbol{\mu}_B^T\boldsymbol{R}_{BA}) + q_B(\boldsymbol{\mu}_A^T\boldsymbol{R}_{AB})] \\
&\quad + f_5(R_{AB})[q_A\boldsymbol{R}_{AB}^T\boldsymbol{\Theta}_B\boldsymbol{R}_{AB} + q_B\boldsymbol{R}_{AB}^T\boldsymbol{\Theta}_A\boldsymbol{R}_{AB} \\
&\quad - 3(\boldsymbol{\mu}_A^T\boldsymbol{R}_{AB})(\boldsymbol{\mu}_B^T\boldsymbol{R}_{AB}) + (\boldsymbol{\mu}_A^T\boldsymbol{\mu}_B)R_{AB}^2]\}
\end{aligned}
\tag{7}
$$

QUESTION: Is $\boldsymbol{R}_{AB} = (x_i - x_j, y_i - y_j, z_i - z_j)^T$ or maybe $\boldsymbol{R}_{AB} = (x_i + x_j, y_i + y_j, z_i + z_j)^T$, if atoms A and B are centered in $(x_i, y_i, z_i)^T$ and $(x_j, y_j, z_j)^T$? The latter would be weird as then $\boldsymbol{R}_{AB} = \boldsymbol{R}_{BA}$ and why both then?
$\boldsymbol{\mu}_A$ is the cumulative atomic dipole moment of atom A and $\boldsymbol{\Theta}_A$ is the corresponding traceless quadrupole moment. The curly braces and brackets are used in the same way as normal parenthesis for showing order of operations. q_A is

the atomic charge of atom A.

$$\Theta_A^{\alpha\beta} = \frac{3}{2}\theta_A^{\alpha\beta} - \frac{\delta_{\alpha\beta}}{2}\left(\theta_A^{xx} + \theta_A^{yy} + \theta_A^{zz}\right) \tag{8}$$

$$\theta_A^{\alpha\beta} = \sum_{\kappa\in A}\sum_{\lambda} P_\lambda \left(\alpha_A D_{\lambda\kappa}^{\beta} + \beta_A D_{\lambda\kappa}^{\alpha} - \alpha_A\beta_A S_{\lambda\kappa} - Q_{\lambda\kappa}^{\alpha\beta}\right) \tag{9}$$

$$q_A = Z_A - \sum_{\kappa\in A}\sum_{\lambda} P_{\kappa\lambda} S_{\lambda\kappa} \tag{10}$$

$$\mu_A^{\alpha} = \sum_{\kappa\in A}\sum_{\lambda} P_{\kappa\lambda}\left(\alpha_A S_{\kappa\lambda} - D_{\kappa\lambda}^{\alpha}\right) \tag{11}$$

$$D_{\lambda\kappa}^{\alpha} = \langle\phi_\lambda|\alpha_i|\phi_\kappa\rangle \tag{12}$$

$$Q_{\lambda\kappa}^{\alpha\beta} = \langle\phi_\lambda|\alpha_i\beta_i|\phi_\kappa\rangle \tag{13}$$

α and β are Cartesian components.
QUESTION: Are we correct in the following, assuming that atom A is centered in $(x_i, y_i, z_i)^T$?

$$\boldsymbol{\Theta}_A = \begin{pmatrix} \Theta_A^{xx} & \Theta_A^{xy} & \Theta_A^{xz} \\ \Theta_A^{yx} & \Theta_A^{yy} & \Theta_A^{yz} \\ \Theta_A^{zx} & \Theta_A^{zy} & \Theta_A^{zz} \end{pmatrix}$$

$$\boldsymbol{\mu}_A = \begin{pmatrix} \mu_A^x \\ \mu_A^y \\ \mu_A^z \end{pmatrix}$$

QUESTION: Is δ 1 if the labels match and 0 otherwise, i.e. a delta function?

1.5 Anisotropic XC energy

$$E_{AXC} = \sum_A (f_{XC}^{\mu_A}|\boldsymbol{\mu}_A|^2 + f_{XC}^{\Theta_A}||\boldsymbol{\Theta}_A||^2) \tag{14}$$

Where $f_{XC}^{\mu_A}$ and $f_{XC}^{\Theta_A}$ are fitted values.
QUESTION: What norms are these, what are the formulas for calculating them?

1.6 Dispersion Energy

$$
\begin{aligned}
E_{disp}^{D4'} = &- \sum_{A>B} \sum_{n=6,8} s_n \frac{C_n^{AB}(q_A, CN_{cov}^A, q_B, CN_{cov}^B)}{R_{AB}^n} f_{damp,BJ}^{(n)}(R_{AB}) \\
&- s_9 \sum_{A>B>C} \frac{(3cos(\theta_{ABC})cos(\theta_{BCA})cos(\theta_{CAB}) + 1) C_9^{ABC}(CN_{cov}^A, CN_{cov}^B, CN_{cov}^C)}{(R_{AB}R_{AC}R_{BC})^3} \\
&\times f_{damp,zero}^{(9)}(R_{AB}, R_{AC}, R_{BC}).
\end{aligned} \tag{15}
$$

The term in the second line is the three-body Axilrod– Teller–Muto (ATM) (What is this??????) term and the last line is the corresponding zero-damping function for this term.

C_6^{AB} is the pairwise dipole-dipole dispersion coefficients calculated by numerical integration via the Casimir-Polder relation.

$$
C_6^{AB} = \frac{3}{\pi} \sum_j w_j \overline{\alpha}_A(i\omega_j, q_A, CN_{cov}^A) \overline{\alpha}_B(i\omega_j, q_B, CN_{cov}^B) \tag{16}
$$

w_j are the integration weights, which are derived from a trapeziodal partitioning between the grid points $j (j \in [1, 23])$.
The isotropically averaged, dynamic dipole-dipole polarizabilites $\overline{\alpha}$ at the jth imaginary frequency $i\omega_j$ are obtained from the self-consistent D4 model; i.e., they are depending on the covalent coordination number and are also charge dependent.

$$
\overline{\alpha}_A(i\omega_j, q_A, CN_{cov}^A) = \sum_r^{N_{A,ref}} \xi_A^r(q_A, q_{A,r}) \overline{\alpha}_{A,r}(i\omega_j, q_{A,r}, CN_{cov}^{A,r}) W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) \tag{17}
$$

The Gaussian weighting for each reference system is given by:

$$
W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) = \sum_{j=1}^{N_{gauss}} \frac{1}{\mathcal{N}} \exp\left[-6j \cdot (CN_{cov}^A - CN_{cov}^{A,r})^2\right] \tag{18}
$$

with

$$
\sum_r^{N_{A,ref}} W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) = 1 \tag{19}
$$

$\mathcal{N}$ is a normalization constant.
The number of Gaussian function per reference system N_{gauss} is mostly one, but equal to three for $CN_{cov}^{A,r} = 0$ and reference systems with similar coordination number.

The carge-dependency is included via the empirical scaling function ξ_A^r.

$$\xi_A^r(q_A, q_{A,r}) = \exp\left[3\left\{1 - \exp\left[4\eta_A\left(1 - \frac{Z_A^{eff} + q_{A,r}}{Z_A^{eff} + q_A}\right)\right]\right\}\right] \tag{20}$$

where η_A is the chemical hardness taken from ref 98.
Z_A^{eff} is the effective nuclear charge of atom A, which has been determined by subtracting the number of core electrons represented by the def2-ECPs in the time-dependent DFT reference calculations.

C_8^{AB} is calculated recursively from the lowest order C_6^{AB} coefficients.

$$C_8^{AB} = 3C_6^{AB}\sqrt{\mathcal{Q}^A\mathcal{Q}^B} \tag{21}$$

$$\mathcal{Q}^A = s_{42}\sqrt{Z^A}\frac{\langle r^4\rangle^A}{\langle r^2\rangle^A} \tag{22}$$

$\sqrt{Z^A}$ is the ad hoc nuclear charge dependent factor. (How do we get Z?????)

$\langle r4\rangle$ and $\langle r2\rangle$ are simple multipole-type expectation values derived from atomic densities which are averaged geometrically to get the pair coefficients. (What is 'r', how we get??????) (what is s_{42}???)

CN_{cov}^A is the covalent coordination number for atom A.

q is the atomic charge, so q_A is the atomic charge for atom A.

The damping and scaling parameters in the dispersion model are:

$$a1 = 0.52 \quad | \quad a2 = 5.0 \quad | \quad s6 = 1.0 \quad | \quad s8 = 2.7 \quad | \quad s9 = 5.0$$

C_9^{ABC} is the triple-dipole constant[1]:

$$C_9^{ABC} = \frac{3}{\pi}\int_0^\infty \alpha^A(i\omega)\alpha^B(i\omega)\alpha^C(i\omega)d\omega \tag{23}$$

The three-body contribution is typically $< 5-10\%$ of E_{disp}, so it is small enough that we can reasonably approximate the coefficients by a geometric mean as[1]:

$$C_9^{ABC} \approx -\sqrt{C_6^{AB}C_6^{AC}C_6^{BC}} \tag{24}$$

θ_{ABC} is the angle between the two edges going from B to the other two atoms.
θ_{BCA} is the angle between the edges going from C to the other two and so on.

[1] https://www.researchgate.net/publication/43347348_A_Consistent_and_Accurate_Ab_Initio_Parametrization_of_Density_F
D_for_the_94_Elements_H-Pu

BJ = Becke-Johnson

$$f_n^{damp,BJ}(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + (a_1 \cdot R_{AB}^{crit} + a_2)^6} \tag{25}$$

$$R_{AB}^{crit} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} \tag{26}$$

$$f_9^{damp,zero}(R_{AB}, R_{AC}, R_{BC}) = \left(1 + 6\left(\sqrt{\frac{R_{AB}^{crit} R_{BC}^{crit} R_{CA}^{crit}}{R_{AB} R_{BC} R_{CA}}}\right)^{16}\right)^{-1} \tag{27}$$