Chapter 1

Atomic basis sets

about this chapter, lcao, mention that we use pseudo wfs for lcao, etc

1.1 Localized functions

Basis functions, projectors, compensation charges and several other entities the nature of which was never meant to be understood by mortals, are compactly supported functions defined on real space, each being centered on an atom. Generally, each such function is represented as an arbitrary radial part times an angular part, which is a spherical harmonic function.[1]

Implementation-wise, the radial part is stored as a spline on a one-dimensional grid, whereas the spherical harmonics can be accounted for in ways that do not require tabulations of function values.

For any such function X on any atom a we may write

$$\mathrm{X}(\mathbf{r}) = \mathrm{X}(\mathbf{r}^a - \mathbf{R}^a) = \chi(r^a) Y_{lm}(\hat{\mathbf{r}}^a), \tag{1.1}$$

where $\mathbf{r}^a = \mathbf{r} - \mathbf{R}^a$ are nucleus-centered coordinates.

1.2 Two-center integrals

Inner products between localized functions appear in several places, most notably the total energy expression. It is for this reason necessary to calculate the product integral of two localized functions $\Phi(\mathbf{r})$ and $\mathrm{X}(\mathbf{r})$. In all relevant cases, one of the functions is a basis function whereas the other is either a basis function, the laplacian of a basis function, or a projector function. The functions are centered on atoms a and b with coordinates $\mathbf{R}^a$ and $\mathbf{R}^b$, where a and b are probably distinct. Define

$$\Phi^a(\mathbf{r}^a) = \phi(r^a) Y_{L_1}(\hat{\mathbf{r}}^a), \tag{1.2}$$

$$\mathrm{X}^b(\mathbf{r}^b) = \chi(r^b) Y_{L_2}(\hat{\mathbf{r}}^b). \tag{1.3}$$

---

[1] *More* generally, the spherical harmonics are *complete*, i.e. *every* angular function can be expanded as a linear combination of several or infinitely many spherical harmonics, but for our purposes it suffices to consider functions with only a single spherical harmonic. Redeemingly, the more complicated case can be reduced to multiple single-spherical harmonic cases.[**?**]

The overlap integral is nominally a function, say Θ, of both nuclear coordinates $\mathbf{R}^a$ and $\mathbf{R}^b$, though in fact it only depends on the atomic separation vector $\mathbf{R} = \mathbf{R}^b - \mathbf{R}^a$:

$$\begin{aligned}
\langle \Phi | \mathrm{X} \rangle &= \int \Phi^*(\mathbf{r} - \mathbf{R}^a)\mathrm{X}(\mathbf{r} - \mathbf{R}^b)\,\mathrm{d}^3\mathbf{r} && (1.4)\\
&= \int \Phi^*(\mathbf{r})\mathrm{X}(\mathbf{r} - \mathbf{R}^b + \mathbf{R}^a)\,\mathrm{d}^3\mathbf{r}, && (1.5)
\end{aligned}$$

which is then

$$\Theta(\mathbf{R}) = \int \Phi^*(\mathbf{r})\mathrm{X}(\mathbf{r} - \mathbf{R})\,\mathrm{d}^3\mathbf{r} \tag{1.6}$$

The inner product is calculated through a Bessel transform blahblah explain the whole thing...

$$\begin{aligned}
\Theta(\mathbf{R}) &= 4\pi \int \phi^*(k)\chi(k) Y^*_{L_1}(\hat{\mathbf{k}}) Y_{L_2}(\hat{\mathbf{k}}) \sum_L \mathrm{i}^l J_l(kr) Y_l^{m*}(\hat{\mathbf{k}}) Y_l^m(\hat{\mathbf{R}})\,\mathrm{d}^3\mathbf{k}\\
&= 4\pi \sum_L \int_0^\infty \phi^*(k)\chi(k) J_l(kr) k^2\,\mathrm{d}k\\
&\times \quad Y_l^{m*}(\hat{\mathbf{R}}) \int Y^*_{L_1}(\hat{\mathbf{k}}) Y_{L_2}(\hat{\mathbf{k}}) Y_l^{m*}(\hat{\mathbf{k}})\,\mathrm{d}^2\hat{\mathbf{k}} && (1.7)
\end{aligned}$$

FiXme: *Use language*

thus if... blahblah Gaunt coefficients...

$$G^L_{L_1 L_2} = \int Y^*_{L_1}(\hat{\mathbf{r}}) Y_{L_2}(\hat{\mathbf{r}}) Y^*_L(\hat{\mathbf{r}})\,\mathrm{d}^2\hat{\mathbf{r}} \tag{1.8}$$

and blahblah radial part

$$\begin{aligned}
\Theta(R) &= 4\pi \mathrm{i}^l \int_0^\infty J_l(kr)\phi^*(k)\chi(k) k^2\,\mathrm{d}k && (1.9)\\
& && (1.10)
\end{aligned}$$

and finally

$$\Theta(\mathbf{R}) = \langle \Phi | \mathrm{X} \rangle = \sum_L \Theta_L(R) Y_L(\hat{\mathbf{R}}) \tag{1.11}$$

The summation over L includes all $m = -l \ldots l$ for each $l = 0 \ldots \infty$, but for some reason all contributions above a certain value, $l_{\max} = l_1 + l_2$, are zero. In fact the only relevant values of l are $l_{\max}, l_{\max} - 2, \ldots, l_{\min}$, where $l_{\min}$ may be 0 or 1 depending on whether $l_{\max}$ is even or odd.

FiXme: *which reason, to be exact? Reference?*

FiXme: *So this is actually correct. Or what?*

Interestingly, if two distinct atoms possess identical localized functions, then those functions will share the radial function for the purposes of all their overlaps. Thus it is necessary only to store splines for every *unique* pair of localized functions, even though the final arrangement of atoms would involve superficially different overlap integrals.

FiXme: *does this make sense?*

Specifically we define the *kinetic energy overlap matrix* $T_{\mu\nu}$, the *projector* *overlap matrix*[2] $P_{i\mu}$ and the *basis function overlap matrix* $S_{\mu\nu}$

$$\begin{aligned}
T_{\mu\nu} &= \langle \Phi_\mu | -\tfrac{1}{2}\nabla^2 | \Phi_\nu \rangle, && (1.12)\\
P^a_{i\mu} &= \langle \tilde{p}^a_i | \Phi_\mu \rangle, && (1.13)\\
S_{\mu\nu} &= \langle \Phi_\mu | \hat{S} | \Phi_\nu \rangle = \langle \Phi_\mu | \Phi_\nu \rangle + \sum_{aij} \langle \Phi_\mu | \tilde{p}^a_i \rangle \Delta O_{ij} \langle \tilde{p}^a_j | \Phi_\nu \rangle. && (1.14)
\end{aligned}$$

The use of these matrices will become evident in Chapter Many

FiXme: *which chapter, or maybe explain their use here*

1.3 Derivatives of two-center integrals

For the purposes of atomic force calculations, it is necessary to differentiate overlaps with respect to nuclear coordinates $\mathbf{R}^a$. This also involves a derivative of the two-center integral expansions (1.11). Since the real solid spherical harmonics $\bar{Y}_{lm}$ are simply polynomials of the cartesian coordinates, it proves convenient to transfer a factor of R^l from the radial part to the formerly angular part:

FiXme: *reference to sth abt RSSH*

$$\begin{aligned}
\bar{\Theta}_{lm}(R) &= \frac{\Theta_{lm}(R)}{R^l}, && (1.15)\\
\bar{Y}_{lm}(\mathbf{R}) &= R^l Y_{lm}(\hat{\mathbf{R}}). && (1.16)
\end{aligned}$$

Then the overlap integral takes the form

$$\Theta(\mathbf{R}) = \sum_L \Theta_L(R) Y_L(\hat{\mathbf{R}}) = \sum_L \bar{\Theta}_L(R) \bar{Y}_L(\hat{\mathbf{R}}). \tag{1.17}$$

Differentiation with respect to $\mathbf{R}$ yields

$$\begin{aligned}
\frac{\mathrm{d}\Theta(\mathbf{R})}{\mathrm{d}\mathbf{R}} &= \sum_L \left\{ \frac{\mathrm{d}\bar{\Theta}_L(R)}{\mathrm{d}\mathbf{R}} \bar{Y}_L(\mathbf{R}) + \bar{\Theta}_L(R) \frac{\mathrm{d}\bar{Y}_L(\mathbf{R})}{\mathrm{d}\mathbf{R}} \right\}\\
&= \sum_L \left\{ \frac{\mathrm{d}\bar{\Theta}_L(R)}{\mathrm{d}R} \bar{Y}_L(\mathbf{R})\hat{\mathbf{R}} + \bar{\Theta}_L(R) \frac{\mathrm{d}\bar{Y}_L(\mathbf{R})}{\mathrm{d}\mathbf{R}} \right\}, && (1.18)
\end{aligned}$$

where we have used $\frac{\mathrm{d}r}{\mathrm{d}x} = \frac{x}{r}$ and so on. In this form all variables are trivially evaluated: spline derivatives are blahblah...

NB! The following should be moved to some later chapter, since it contains as of yet undefined variables.

Differentiation of an overlap with respect to a nuclear coordinate $\mathbf{R}^a$ works out somewhat differently depending on matters of localized function ownership: if e.g. both or none of the localized functions reside on the atom in question, the overlap is translation invariant. This combined with (1.5) trivially yields

$$\frac{\partial \Theta_{\mu\nu}(\mathbf{R}^\nu - \mathbf{R}^\mu)}{\partial \mathbf{R}^a} = \begin{cases} \mathbf{0} & \mu \in a, \quad \nu \in a \\ -\frac{\mathrm{d}\Theta_{\mu\nu}}{\mathrm{d}\mathbf{R}_{\mu\nu}} & \mu \in a, \quad \nu \notin a \\ \frac{\mathrm{d}\Theta_{\mu\nu}}{\mathrm{d}\mathbf{R}_{\mu\nu}} & \mu \notin a, \quad \nu \in a \\ \mathbf{0} & \mu \notin a, \quad \nu \notin a \end{cases} \tag{1.19}$$

where blahblah then

FiXme: *wait, where do these sums suddenly come from? Move section to force calculations or something*

---

[2] blahblah PAW article defines this differently, but that definition is bad because we like our matrices to have antilinear/linear behaviour wrt. rows and columns, thank you very much

$$\sum_{\mu\nu} \frac{\partial\Theta_{\mu\nu}}{\partial\mathbf{R}^a}\rho_{\nu\mu} = \sum_{\substack{\mu\in a\\ \nu\notin a}} \frac{\partial\Theta_{\mu\nu}}{\partial\mathbf{R}^a}\rho_{\nu\mu} + \sum_{\substack{\mu\notin a\\ \nu\in a}} \frac{\partial\Theta_{\mu\nu}}{\partial\mathbf{R}^a}\rho_{\nu\mu} \tag{1.20}$$

$$= -\sum_{\substack{\mu\in a\\ \nu\notin a}} \frac{\mathrm{d}\Theta_{\mu\nu}}{\mathrm{d}\mathbf{R}^{\mu\nu}}\rho_{\nu\mu} + \sum_{\substack{\mu\notin a\\ \nu\in a}} \frac{\mathrm{d}\Theta_{\mu\nu}}{\mathrm{d}\mathbf{R}^{\mu\nu}}\rho_{\nu\mu} \tag{1.21}$$

The former term can be rewritten by interchanging symbols, then reversing the atomic separation vector:

$$-\sum_{\substack{\mu\in a\\ \nu\notin a}} \frac{\partial\Theta_{\mu\nu}}{\partial\mathbf{R}^{\mu\nu}}\rho_{\nu\mu} = -\sum_{\substack{\mu\notin a\\ \nu\in a}} \frac{\partial\Theta_{\nu\mu}}{\partial\mathbf{R}^{\nu\mu}}\rho_{\mu\nu} = \sum_{\substack{\mu\notin a\\ \nu\in a}} \frac{\partial\Theta^*_{\mu\nu}}{\partial\mathbf{R}^{\mu\nu}}\rho^*_{\nu\mu}, \tag{1.22}$$

which yields the final expression

$$\sum_{\mu\nu} \frac{\partial\Theta_{\mu\nu}}{\partial\mathbf{R}^a}\rho_{\nu\mu} = 2\Re\sum_{\substack{\mu\notin a\\ \nu\in a}} \frac{\mathrm{d}\Theta_{\mu\nu}}{\mathrm{d}\mathbf{R}^{\mu\nu}}\rho_{\nu\mu}, \tag{1.23}$$

FiXme: *why does the real part not appear in siesta?*

where $\Re$ denotes the real part.

1.4 Matrix formulation

Fundamentally, each of the N single-particle pseudo wave functions $|\tilde{\Psi}_n\rangle$ is expressed as a linear combination of the M basis functions $|\Phi_\mu\rangle$:

$$|\tilde{\Psi}_n\rangle = \sum_\mu c_{\mu n}|\Phi_\mu\rangle, \quad c_{\mu n} \in \mathbb{L}, \tag{1.24}$$

where $\mathbb{L}$ is either the reals $\mathbb{R}$ or complex numbers $\mathbb{C}$. The matrix of expansion coefficients $\mathbf{C} = [c_{\mu n}] \in \mathbb{L}^{M\times N}$ corresponds to the transformation matrix from the pseudo wavefunctions $\{|\tilde{\Psi}_n\rangle\}$ to the LCAO basis $\{|\Phi_\mu\rangle\}$, though it does not technically represent a change of basis since it is not generally a square matrix.[3]

Now, for any operator $\hat{A}$, we may write

$$\langle\tilde{\Psi}_n|\hat{A}|\tilde{\Psi}_m\rangle = \sum_{\mu\nu} c^*_{\mu n}\langle\Phi_\mu|\hat{A}|\Phi_\nu\rangle c_{\nu m}, \tag{1.25}$$

which in matrix notation becomes the familiar-looking

$$\mathbf{A}' = \mathbf{C}^\dagger\mathbf{A}\mathbf{C}, \quad (\mathbf{A})_{\mu\nu} = \langle\Phi_\mu|\hat{A}|\Phi_\nu\rangle, \quad (\mathbf{A}')_{mn} = \langle\tilde{\Psi}_m|\hat{A}|\tilde{\Psi}_n\rangle. \tag{1.26}$$

Specifically, the pseudo wavefunction orthogonality condition becomes

$$\langle\tilde{\Psi}_n|\hat{S}|\tilde{\Psi}_m\rangle = \sum_{\mu\nu} c^*_{\mu n}\langle\Phi_\mu|\hat{S}|\Phi_\nu\rangle c_{\mu m} = \delta_{nm} \tag{1.27}$$

or

$$\mathbf{C}^\dagger\mathbf{S}\mathbf{C} = \mathbf{I}^{N\times N}, \quad S_{\mu\nu} = \langle\Phi_\mu|\hat{S}|\Phi_\nu\rangle. \tag{1.28}$$

---

[3] In practice some circumstances, such as matrix diagonalization, prompt the adoption of square matrices. The band count N is then increased to M for the purposes of these operations, while the extraneous columns are discarded afterwards.

Let $\mathbf{F} = \mathrm{diag}(\{f_n\})$ be the diagonal matrix of occupation numbers f_n and define

$$\boldsymbol{\rho} = \mathbf{C}\mathbf{F}\mathbf{C}^\dagger, \quad \rho_{\mu\nu} = \sum_n c_{\mu n} f_n c^*_{\nu n}. \tag{1.29}$$

Observe that for any operator $\hat{A}$,

$$\sum_n f_n \langle \tilde{\Psi}_n | \hat{A} | \tilde{\Psi}_n \rangle = \mathrm{Tr}[\mathbf{F}\mathbf{C}^\dagger \mathbf{A}\mathbf{C}] = \mathrm{Tr}[\mathbf{C}\mathbf{F}\mathbf{C}^\dagger \mathbf{A}] = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{A}], \tag{1.30}$$

i.e. $\boldsymbol{\rho}$ represents the state operator.

FiXme: *well, except that it doesn't include the core states. Also it is probably not normalized meaningfully.*

1.5 Variational problem

The energy blahblah by varying the wave functions which, in our case, means the coefficients. The energy is, however, determined by the state operator[4] $\boldsymbol{\rho} = \mathbf{C}\mathbf{F}\mathbf{C}^\dagger$. We may write

$$\begin{aligned} \Omega &= E(\boldsymbol{\rho}) - \sum_{mn} \lambda_{nm} \left(\langle \tilde{\Psi}_m | \hat{S} | \tilde{\Psi}_n \rangle - \delta_{mn} \right) && (1.31) \\ &= E(\boldsymbol{\rho}) - \sum_{mn\mu\nu} \lambda_{nm} \left(c^*_{\mu m} \langle \Phi_\mu | \hat{S} | \Phi_\nu \rangle c_{\nu n} - \delta_{mn} \right). && (1.32) \end{aligned}$$

The minimal energy occurs at a point where the energy is stationary with respect to the coefficients $\{c_{\mu m}\}$. Obviously we want to differentiate by $c_{\mu m}$, but it turns out that we might as well differentiate by its complex conjugate [5] $c^*_{\mu m}$, which happens to look slightly less ugly on paper (the other is left as an exercise to the reader).

$$\begin{aligned} \frac{\partial \Omega}{\partial c^*_{\xi k}} &= \sum_{\mu\nu} \frac{\partial E}{\partial \rho_{\mu\nu}} \frac{\partial \rho_{\mu\nu}}{\partial c^*_{\xi k}} \\ & \quad - \sum_{mn\mu\nu} \lambda_{nm} \left(\frac{\partial c^*_{\mu m}}{\partial c^*_{\xi k}} c_{\nu n} + c^*_{\mu m} \frac{\partial c_{\nu n}}{\partial c^*_{\xi k}} \right) \langle \Phi_\mu | \hat{S} | \Phi_\nu \rangle = 0 && (1.33) \end{aligned}$$

Observe that $\frac{\partial c^*_{\nu n}}{\partial c^*_{\xi k}} = \delta_{\nu\xi}\delta_{nk}$, whereas $\frac{\partial c_{\mu n}}{\partial c^*_{\xi k}} = 0$. Then

$$\frac{\partial \rho_{\mu\nu}}{\partial c^*_{\xi k}} = \sum_n \left(\frac{\partial c_{\mu n}}{\partial c^*_{\xi k}} c^*_{\nu n} + c_{\mu n} \frac{\partial c^*_{\nu n}}{\partial c^*_{\xi k}} \right) f_n = \delta_{\nu\xi} c_{\mu k} f_k \tag{1.34}$$

---

[4] Hmmm. We know that this is not the *true* state operator since it does not include core state contributions. However, it appears to be the state operator for the isolated Kohn-Sham Hamiltonian, i.e. it can be called the *Kohn-Sham state operator*. While this only applies to the Kohn-Sham energy functional, I seem to remember something about the KS energy having the same minimum as the actual total energy functional, for which reason the KS state operator is sufficient to characterize the true minimum. Does this make sense at all?

[5] Ω, being a function of multiple complex variables, is here perceived as an explicit function not of MN whole complex numbers $\{c_{\mu n}\}$, but rather all the $2MN$ conjugate pairs $\{c_{\mu n}, c^*_{\mu n}\}$. This ensures that the differently-conjugated partial derivatives $\frac{\partial c^*_{\mu n}}{\partial c_{\nu m}}$ are well-defined (and indeed zero) in spite of complex conjugation not being an analytical operation in the sense that it violates the Cauchy-Riemann equations everywhere.

If we define $H_{\nu\mu} = \frac{\partial E}{\partial \rho_{\nu\mu}}$,

$$0 = \frac{\partial \Omega}{\partial c^*_{\xi k}} = \sum_{\mu} H_{\xi\mu} c_{\mu k} f_k - \sum_{nm\nu} \langle \Phi_\xi | \hat{S} | \Phi_\nu \rangle c_{\nu n} \lambda_{nk} \tag{1.35}$$

which is seen to be the (ξ, k)'th element of a matrix product. Since this is supposed to be the case for all (ξ, k), we can reformulate it as a matrix equation

$$\mathbf{HCF} = \mathbf{SC\Lambda}. \tag{1.36}$$

Differentiation by $c_{\mu k}$ rather than $c^*_{\mu k}$ simply results in the "adjoint" equation $\mathbf{FC}^\dagger\mathbf{H} = \mathbf{\Lambda C}^\dagger\mathbf{S}$, which is completely equivalent to (1.36) for real $\mathbf{H}$, $\mathbf{S}$ and $\mathbf{\Lambda}$. If all variables are assumed real, the result (1.36) is also obtained, a factor of 2 appearing on both sides due to the now quadratic dependence on coordinates.

Chapter 2

Force calculations in LCAO

2.1 Differentiation of big expression

$\mathbf{F}^a$ is calculated as the derivative of the total energy with respect to the nuclear coordinates $\mathbf{R}^a$ of atom a. The plan is to use the chain rule after identifying a suitable set of parameters through which to differentiate. Recall that the total energy E is

$$E = \tilde{E} + \sum_a \left(E^a - \tilde{E}^a \right), \tag{2.1}$$

where

$$\begin{aligned} \tilde{E} = {} & \sum_{\mu\nu} \rho_{\nu\mu} \left\langle \Phi_\mu \middle| -\tfrac{1}{2}\nabla^2 \middle| \Phi_\nu \right\rangle + \frac{1}{2} \int \tilde{v}^{\mathrm{Ha}}(\mathbf{r}) \tilde{\rho}(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} \\ & + \sum_a \int \tilde{n}(\mathbf{r}) \bar{v}^a(r^a) \, \mathrm{d}^3\mathbf{r} + E_{\mathrm{xc}}[\tilde{n}]. \end{aligned} \tag{2.2}$$

The atomic contributions E^a and $\tilde{E}^a$ can be lobbed together in an expression $\Delta E^a = E^a - \tilde{E}^a$, which is a function of the atomic density matrices D^a_{ij} and some atomic variables. It is important that apart from D^a_{ij}, *all other variables in* ΔE^a *are defined in a purely atomic context*, i.e. the only dependence on environment, and thus nuclear position, is through D^a_{ij}. Thus ΔE^a is a function *only* of D^a_{ij} for the purposes of differentiation with respect to nuclear coordinates.

With this and (2.2) in mind, the total energy will be regarded as a function of the following variables, where each variable is itself dependent on the nuclear coordinates in some manner yet to be described:

- The density matrix elements $\rho_{\mu\nu}$
- The kinetic energy overlap matrix $T_{\mu\nu} = \langle \Phi_\mu | -\frac{1}{2}\nabla^2 | \Phi_\nu \rangle$
- The pseudo charge distribution $\tilde{\rho}(\mathbf{r})$
- The electron density $\tilde{n}(\mathbf{r})$
- The zero potentials $\bar{v}^a$ (which move rigidly with the atoms)

- The atomic density matrices D^a_{ij}

The force is

$$\mathbf{F}^a = -\frac{\partial E}{\partial \mathbf{R}^a}, \tag{2.3}$$

$$\begin{aligned}\frac{\partial E}{\partial \mathbf{R}^a} = {} & \sum_{\mu\nu} \frac{\partial E}{\partial \rho_{\mu\nu}} \frac{\partial \rho_{\mu\nu}}{\partial \mathbf{R}^a} + \sum_{\mu\nu} \frac{\partial E}{\partial T_{\mu\nu}} \frac{\partial T_{\mu\nu}}{\partial \mathbf{R}^a} + \int \frac{\partial E}{\partial \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial \mathbf{R}^a} \,\mathrm{d}^3\mathbf{r} \\ & + \int \frac{\partial E}{\partial \tilde{n}(\mathbf{r})} \frac{\partial \tilde{n}(\mathbf{r})}{\partial \mathbf{R}^a} \,\mathrm{d}^3\mathbf{r} + \int \frac{\partial E}{\partial \bar{v}^a(\mathbf{r})} \frac{\partial \bar{v}^a(\mathbf{r})}{\partial \mathbf{R}^a} \,\mathrm{d}^3\mathbf{r} + \sum_{ij} \frac{\partial E}{\partial D^a_{ij}} \frac{\partial D^a_{ij}}{\partial \mathbf{R}^a},\end{aligned} \tag{2.4}$$

where we have used that the only non-zero atomic contributions are those that correspond to atom a in particular. The force contributions from each of these partial derivatives will now be evaluated in turn.

FiXme: *So what about th Hartree potential? For reason it can be regarde constant when chang This must be because symmetry of the Co term which, combine the product different rule, takes care of t*

FiXme: *siesta rambles about partial/total deriv and free energy, fir what this is about. have to find out wheth use hard or*

State operator contribution

FiXme: *insert eqrefs*

Using the $\frac{\partial E}{\partial \rho_{\mu\nu}} = H_{\nu\mu}$ as well as $\mathbf{HCF} = \mathbf{SC\Lambda}$ and its adjoint,

$$\begin{aligned}\sum_{\mu\nu} \frac{\partial E}{\partial \rho_{\mu\nu}} \frac{\mathrm{d}\rho_{\mu\nu}}{\mathrm{d}\mathbf{R}^a} &= \sum_{\mu\nu} H_{\nu\mu} \frac{\mathrm{d}}{\mathrm{d}\mathbf{R}^a} \sum_n c_{\mu n} f_n c^*_{\nu n} \\ &= \sum_{\mu\nu n} \left\{ \frac{\mathrm{d}c_{\mu n}}{\mathrm{d}\mathbf{R}^a} f_n c^*_{\nu n} H_{\nu\mu} + H_{\nu\mu} c_{\mu n} f_n \frac{\mathrm{d}c^*_{\nu n}}{\mathrm{d}\mathbf{R}^a} \right\} \\ &= \mathrm{Tr}\left[\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\mathbf{R}^a} \mathbf{F}\mathbf{C}^\dagger \mathbf{H} \right] + \mathrm{Tr}\left[\mathbf{HCF} \frac{\mathrm{d}\mathbf{C}^\dagger}{\mathrm{d}\mathbf{R}^a} \right] \\ &= \mathrm{Tr}\left[\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\mathbf{R}^a} \mathbf{\Lambda}\mathbf{C}^\dagger \mathbf{S} \right] + \mathrm{Tr}\left[\mathbf{SC\Lambda} \frac{\mathrm{d}\mathbf{C}^\dagger}{\mathrm{d}\mathbf{R}^a} \right] \\ &= \mathrm{Tr}\left[\mathbf{\Lambda} \left(\mathbf{C}^\dagger \mathbf{S} \frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\mathbf{R}^a} + \frac{\mathrm{d}\mathbf{C}^\dagger}{\mathrm{d}\mathbf{R}^a} \mathbf{SC} \right) \right].\end{aligned} \tag{2.5}$$

This horrible mess, however, can be cleared up by using the orthogonality condition $\mathbf{C}^\dagger\mathbf{SC} = \mathbf{I}$. By differentiation with respect to $\mathbf{R}^a$,

$$\frac{\mathrm{d}\mathbf{C}^\dagger}{\mathrm{d}\mathbf{R}^a} \mathbf{SC} + \mathbf{C}^\dagger \frac{\mathrm{d}\mathbf{S}}{\mathrm{d}\mathbf{R}^a} \mathbf{C} + \mathbf{C}^\dagger \mathbf{S} \frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\mathbf{R}^a} = \frac{\mathrm{d}\mathbf{I}}{\mathrm{d}\mathbf{R}^a} = \mathbf{0}, \tag{2.6}$$

thus allowing the simplification

$$\sum_{\mu\nu} \frac{\partial E}{\partial \rho_{\mu\nu}} \frac{\mathrm{d}\rho_{\mu\nu}}{\mathrm{d}\mathbf{R}^a} = -\,\mathrm{Tr}\left[\mathbf{C}^\dagger \frac{\mathrm{d}\mathbf{S}}{\mathrm{d}\mathbf{R}^a} \mathbf{C\Lambda} \right] \tag{2.7}$$

$$= -\sum_{\mu\nu n} \epsilon_n f_n c^*_{\mu n} \frac{\partial S_{\mu\nu}}{\partial \mathbf{R}^a} c_{\nu n}. \tag{2.8}$$

FiXme: *Maybe we should drop some of the matrix equations and stick to index-heavy single-element notation.*

The expression for the PAW-corrected overlaps might complicate matters somewhat... also, the $\mathbf{\Lambda}$ elements are $\Lambda_{nn} = \epsilon_n f_n$. Regarding the PAW-corrected overlaps:

$$\frac{\partial S_{\mu\nu}}{\partial \mathbf{R}^a} = \frac{\partial \Theta_{\mu\nu}}{\partial \mathbf{R}^a} + \sum_{ij} O^a_{ij} \left\{ \frac{\partial P^{a*}_{i\mu}}{\partial \mathbf{R}^a} P_{j\nu} + P^{a*}_{i\mu} \frac{\partial P_{j\nu}}{\partial \mathbf{R}^a} \right\}, \tag{2.9}$$

where $\Theta_{\mu\nu} = \langle \Phi_\mu | \Phi_\nu \rangle$. (luckily this entire thing is calculated in the code somewhere, though it'll be difficult to see how the summation set is reduced.)

Kinetic energy contribution

The kinetic energy contribution is simply

$$\sum_{\mu\nu} \frac{\partial E}{\partial T_{\mu\nu}} \frac{\partial T_{\mu\nu}}{\partial \mathbf{R}^a} = \sum_{\mu\nu} \frac{\partial T_{\mu\nu}}{\partial \mathbf{R}^a} \rho_{\nu\mu} = 2\Re \sum_{\substack{\mu \notin a \\ \nu \in a}} \frac{\mathrm{d} T_{\mu\nu}}{\mathrm{d} \mathbf{R}^{\mu\nu}} \rho_{\nu\mu}. \tag{2.10}$$

Charge contribution

In this case we use that

$$\frac{\partial E}{\partial \tilde{\rho}} = v^{\mathrm{Ha}}. \tag{2.11}$$

Recall that $\tilde{n}$ is a separate parameter in chain rule context, so $\tilde{n}$ is constant in this expression. Then $\tilde{\rho} = \tilde{n} + \sum_a \tilde{Z}^a(\mathbf{r}^a)$ is a function only of the compensation charges $\tilde{Z}^a$ which are completely independent from the LCAO scheme. Therefore the expression is completely equivalent to the ordinary non-LCAO case

$$\int \frac{\partial E}{\partial \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} = \int v^{\mathrm{Ha}}(\mathbf{r}) \sum_L Q_L^a \frac{\mathrm{d}\hat{g}_L^a(\mathbf{r})}{\mathrm{d}\mathbf{R}^a} \, \mathrm{d}^3\mathbf{r}. \tag{2.12}$$

Pseudo density contribution

In this case our magical expression is

$$\frac{\partial E}{\partial \tilde{n}} = \tilde{v}_{\mathrm{eff}}. \tag{2.13}$$

We know

$$\begin{aligned}
\tilde{n}(\mathbf{r}) &= \sum_n f_n \langle \tilde{\Psi}_n | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\Psi}_n \rangle + \sum_a \tilde{n}_c^a && (2.14)\\
&= \sum_{n\mu\nu} f_n c_{\mu n}^* \langle \Phi_\mu | \mathbf{r} \rangle \langle \mathbf{r} | \Phi_\nu \rangle c_{\nu n} + \sum_a \tilde{n}_c^a && (2.15)\\
&= \sum_{\mu\nu} \rho_{\nu\mu} \Phi_\mu^*(\mathbf{r}) \Phi_\nu(\mathbf{r}) + \sum_a \tilde{n}_c^a && (2.16)
\end{aligned}$$

Then

$$\int \frac{\partial E}{\partial \tilde{n}(\mathbf{r})} \frac{\partial \tilde{n}(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} = \int \tilde{v}_{\mathrm{eff}}(\mathbf{r}) \left[\frac{\partial \tilde{n}_c^a}{\partial \mathbf{R}^a} + \sum_{\mu\nu} \rho_{\nu\mu} \frac{\partial}{\partial \mathbf{R}^a} \left(\Phi_\mu^*(\mathbf{r}) \Phi_\nu(\mathbf{r}) \right) \right] \mathrm{d}^3\mathbf{r} \tag{2.17}$$

$$\begin{aligned}
&= \int \tilde{v}_{\mathrm{eff}}(\mathbf{r}) \frac{\partial \tilde{n}_c^a(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} && (2.18)\\
&\quad + \sum_{\mu\nu} \rho_{\nu\mu} \left[\left\langle \frac{\partial \Phi_\mu}{\partial \mathbf{R}^a} \middle| \tilde{v}_{\mathrm{eff}} \middle| \Phi_\nu \right\rangle + \left\langle \Phi_\mu \middle| \tilde{v}_{\mathrm{eff}} \middle| \frac{\partial \Phi_\nu}{\partial \mathbf{R}^a} \right\rangle \right] && (2.19)
\end{aligned}$$

The code appears to contain an LCAO effective potential matrix. I wonder if this matrix can be used for the integral. Unfortunately this part is grid-based, so it'll probably be inefficient to evaluate this. Until now we can just write

$$\int \frac{\partial E}{\partial \tilde{n}(\mathbf{r})} \frac{\partial \tilde{n}(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} = \int \tilde{v}_{\mathrm{eff}}(\mathbf{r}) \frac{\partial \tilde{n}_c^a(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} + \sum_{\mu\nu} \frac{\partial V_{\mu\nu}}{\partial \mathbf{R}^a} \rho_{\nu\mu}, \tag{2.20}$$

where $V_{\mu\nu} = \langle \Phi_\mu | \tilde{v}_{\mathrm{eff}} | \Phi_\nu \rangle$.

Zero potential contribution

The $\bar{v}^a$ contribution is straightforwardly

$$\int \frac{\partial E}{\partial \bar{v}^a(\mathbf{r})} \frac{\partial \bar{v}^a(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} = \int \tilde{n}(\mathbf{r}) \frac{\partial \bar{v}^a(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r}, \tag{2.21}$$

and is therefore not dependent on LCAO specifics.

Atomic density contribution

We know that

$$\frac{\partial E}{\partial D_{ij}^a} = \frac{\partial \Delta E^a}{\partial D_{ij}^a} = H_{ji}^a, \tag{2.22}$$

where H_{ij}^a are expressions containing D_{ij}^a, plus a multitude of variables that pertain to the isolated atom, being independent from LCAO specifics. Specifically, H_{ij}^a is a second order polynomial in $\{D_{ij}^a\}$ with the coefficients precalculated for the isolated atom. Next,

FiXme: *if I remember correctly*

$$\begin{aligned} D_{ij}^a &= \sum_n f_n \langle \tilde{p}_i^a | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \tilde{p}_j^a \rangle = \sum_{n\mu\nu} \langle \tilde{p}_i^a | \Phi_\mu \rangle c_{\mu n} f_n c_{\nu n}^* \langle \Phi_\nu | \tilde{p}_j^a \rangle \\ &= \sum_{\mu\nu} \langle \tilde{p}_i^a | \Phi_\mu \rangle \rho_{\mu\nu} \langle \Phi_\nu | \tilde{p}_j^a \rangle \end{aligned} \tag{2.23}$$

which is equivalent to the matrix expression

$$\mathbf{D}^a = \mathbf{P}^a \boldsymbol{\rho} \mathbf{P}^{a\dagger} \tag{2.24}$$

where $P_{i\mu} = \langle \tilde{p}_i^a | \Phi_\mu \rangle$.[1] Therefore

$$\frac{\partial \mathbf{D}^a}{\partial \mathbf{R}^a} = \frac{\partial \mathbf{P}^a}{\partial \mathbf{R}^a} \boldsymbol{\rho} \mathbf{P}^{a\dagger} + \mathbf{P} \boldsymbol{\rho} \frac{\partial \mathbf{P}^{a\dagger}}{\partial \mathbf{R}^a} \tag{2.25}$$

and finally

$$\sum_{ij} \frac{\partial E}{\partial D_{ij}^a} \frac{\partial D_{ij}^a}{\partial \mathbf{R}^a} = \sum_{ij\mu\nu} H_{ji}^a \rho_{\mu\nu} \left(\frac{\partial P_{i\mu}^a}{\partial \mathbf{R}^a} P_{j\nu}^{a*} + P_{i\mu}^a \frac{\partial P_{j\nu}^{a*}}{\partial \mathbf{R}^a} \right) \tag{2.26}$$

$$= \mathrm{Tr} \left[\mathbf{H}^a \left(\frac{\partial \mathbf{P}^a}{\partial \mathbf{R}^a} \boldsymbol{\rho} \mathbf{P}^{a\dagger} + \mathbf{P}^a \boldsymbol{\rho} \mathbf{P}^{a\dagger} \frac{\partial \mathbf{P}^{a\dagger}}{\partial \mathbf{R}^a} \right) \right] \tag{2.27}$$

depending on whether we like matrices or not.

---

[1] Note that the variable $P_{ni}^a = \langle \tilde{p}_i^a | \tilde{\Psi}_n \rangle$ used in [?], confusingly, corresponds to the transpose of this.

Big formula

This gives a total force expression (NB! there's some fuss with the complex conjugation of the $P_{i\mu}$, so this might not be entirely correct)

$$
\begin{aligned}
\mathbf{F}^a = & \sum_{ij} \left(\epsilon_n O^a_{ij} - H^a_{ij}\right) \sum_{\mu\nu} \rho_{\nu\mu} \left(\frac{\partial P^{a*}_{i\mu}}{\partial \mathbf{R}^a} P^a_{j\nu} + P^{a*}_{i\mu} \frac{\partial P^a_{j\nu}}{\partial \mathbf{R}^a} \right) \\
& - \sum_{\mu\nu} \frac{\partial T_{\mu\nu}}{\partial \mathbf{R}^a} \rho_{\nu\mu} - \sum_{\mu\nu n} \epsilon_n \rho_{\nu\mu} \frac{\partial \Theta_{\mu\nu}}{\partial \mathbf{R}^a} \\
& - \sum_{\mu\nu} \frac{\partial V_{\mu\nu}}{\partial \mathbf{R}^a} \rho_{\nu\mu} - \int \tilde{v}_{\mathrm{eff}}(\mathbf{r}) \frac{\partial \tilde{n}^a_c(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} \\
& - \int v^{\mathrm{Ha}}(\mathbf{r}) \sum_L Q^a_L \frac{\partial \hat{g}^a_L(\mathbf{r})}{\partial \mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} - \int \tilde{n}(\mathbf{r}) \frac{\mathrm{d}\bar{v}^a(\mathbf{r})}{\mathrm{d}\mathbf{R}^a(\mathbf{r})} \, \mathrm{d}^3\mathbf{r} \\
& - \int \tilde{n}(\mathbf{r}) \sum_L Q_L \frac{\mathrm{d}\hat{v}^a_L(\mathbf{r})}{\mathrm{d}\mathbf{R}^a} \, \mathrm{d}^3\mathbf{r} - \sum_{a'LL'} Q^a_L \frac{\mathrm{d}V^{aa'}_{LL'}}{\mathrm{d}\mathbf{R}^a} Q^{a'}_{L'} \qquad (2.28)
\end{aligned}
$$

where we have magically included the last two terms because they are in JJ's article. Specifically Equation (34) from JJ's article. This is a correction to $\tilde{E}$ which I have overlooked so far because it's not on the developer documentation page! Anyway, this should be correct I think, since they are LCAO-independent.

Miscellaneous bad things

$\frac{\partial V_{\mu\nu}}{\partial \mathbf{R}^a}$ probably has to be evaluated on the grid or something, depending on what it actually contains.

The partial derivatives are written in terms of $\mathbf{R}^a$ here. In reality we want use $\mathbf{R}^{\mu\nu}$ instead, thus changing variables to those used for the overlap splines $\Theta_{\mu\nu}(\mathbf{R})$. But this is quite tricky for the $S_{\mu\nu}$ case with all the projector overlap products and summations over intra-atomic entities.