Projected Augmented Waves (PAW) inspired mixed basis sets for small molecules

Garry Goldstein[1]
garrygoldsteinwinnipeg@gmail.com

The success behind many pseudopotential methods such as Projected Augmented Waves (PAW) and Phillips-Kleinman pseudopotential methods is that these methods are all electron methods in disguise. For the Phillips Kleinman and PAW methods we show that there is an all electron formulation. In the all electron formulation there are regular low wavevector wavefunctions and several, specially chosen, high wavevector wavefunctions that are specialized to the atomic environment. Using this as inspiration here we propose a new, PAW pseudopotential inspired, basis set method for small molecules, where we use the recently introduced LO or lo (Localized orbitals) basis set is paired with a Gaussian basis for a hybrid basis for small molecules.

I. INTRODUCTION

It is of paramount importance to make progress on the basis set problem for Density Functional Theory (DFT) for small molecules. Small molecules larger then diatomic: carbon dioxide, water, ammonia, methane, ethyne, ethene, ethane to name a few have a variety of uses in industry and engineering [1, 2]. However there has been limited progress in optimizing basis sets for small molecule DFT calculations similarly to basis sets that have paved breakthroughs in solid state DFT [3–5]. Here inspired by pseudopotential methods in DFT we propose one such hybrid basis set. Pseudopotentials at their core map a problem into another problem. The exact all electron wavefunction for the system is mapped onto a smooth wavefunction that may well be represented by a small number of plane waves. Two of the pre-eminent examples of pseudopotentials are the Phillips Kleinman and Projected Augmented Waves (PAW) pseudopotentials. Here we show that these pseudopotential methods are all electron methods in disguise where the all electron problem includes smooth plane wave wave functions and specialized wave functions for orbitals near atomic nuclei. Inspired by this observation, and the extensive numerical success of the PAW method [3, 4] we propose new basis sets for small molecules, where localized LO or lo [1, 5–7] basis wave functions are augmented with Gaussian wave functions (where Gaussians have been extensively used for the problem and various configuration integrals for them are well known [2, 3, 8]), in a similar manner to Phillips-Kleinman and PAW all electron basis sets We propose that this new hybrid basis set is efficient for DFT and Hartree-Fock (HF) calculations of small molecules.

II. MANY PSEUDOPOTENTIAL METHODS ARE ALL ELECTRON METHODS IN DISGUISE (MOTIVATION)

In this section we reformulate PAW and the Phillips-Kleinman method as all electron methods where there are plane waves and special wave functions that are chosen to represent states near the core. We show the two methods to be equivalent in their all electron formulation.

A. Phillips-Kleinman as an all electron method in disguise

We now recall the Orthogonalized Plane Wave (OPW) wave functions used for the Phillips-Kleinman method. We know that

$$\begin{aligned}&|\phi_{OPW}\left(\mathbf{k}+\mathbf{K}\right)\rangle\\&=\left|\frac{1}{\sqrt{V}}\exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\rangle\\&-\sum_i\left\langle\phi_i\left(\mathbf{r}\right)\mid\frac{1}{\sqrt{V}}\exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\rangle|\phi_i\left(\mathbf{r}\right)\rangle\end{aligned}\tag{1}$$

Here $|\phi_i\left(\mathbf{r}\right)\rangle$ are localized wave functions relevant to the single site problem. Here V is the volume of the primitive lattice cell and $\mathbf{k}$is a wavevector in the first Brillouin zone while $\mathbf{K}$ is a reciprocal lattice vector. Now the main claim of the Phillips-Kleinman method (greatly numerically supported) is that for a reasonable number of plane waves the exact Khon Sham (KS) wavefunction $|\psi_n\rangle$ ca be written as

$$|\psi_n\rangle\cong\sum_{\mathbf{k},\mathbf{K}}c^{\mathbf{K}}_{\mathbf{k}}|\phi_{OPW}\left(\mathbf{k}+\mathbf{K}\right)\rangle\tag{2}$$

for some reasonable cutoffs $\mathbf{K_{max}}$ that means that

$$\begin{aligned}&|\psi_n\rangle\in Span\left\{|\phi_{OPW}\left(\mathbf{k}+\mathbf{K}\right)\rangle\right\}\\&\subset Span\left\{\left\{\left|\frac{1}{\sqrt{V}}\exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\rangle\right\},\left\{|\phi_i\left(\mathbf{r}\right)\rangle\right\}\right\}\end{aligned}\tag{3}$$

As such in some situations its better and in many ways simpler to do an all electron calculation with the basis being given by:

$$\left\{\left\{\left|\frac{1}{\sqrt{V}}\exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\rangle\right\},\left\{|\phi_i\left(\mathbf{r}\right)\rangle\right\}\right\}\tag{4}$$

As such Phillips-Kleinman is an all electron method in disguise. We shall see below that PAW is very similar.

B. PAW is an all electron method in disguise

1. PAW review

We would like to study highly oscillatory wave functions near the nucleus of an atom using plane waves. In order to obtain a low cutoff of say ~30 Rydberg in energy and as such a tolerable basis set we must have that the wavefunction we represent using plane waves is very smooth. The key idea of PAW is then to introduce:

$$|\psi_n\rangle = \hat{\mathcal{T}} \left|\tilde{\psi}_n\right\rangle \tag{5}$$

where $\hat{\mathcal{T}}$ is some linear transformation (specified below) and $|\psi_n\rangle$ is the exact Khon Sham wave functions (solutions of the KS equations) while $\left|\tilde{\psi}_n\right\rangle$ are the wave functions we approximate efficiently using plane waves. Now suppose there is a set of augmentation spheres and we wish for $\hat{\mathcal{T}}$ to be the identity outside the augmentation spheres, so we write:

$$\hat{\mathcal{T}} = \mathbb{I} + \sum_{\mu} S_{\mu} \tag{6}$$

Here μ are the atomic spheres. Now we want the wavefunction $\left|\tilde{\psi}\right\rangle$ to be smooth inside the sphere and we know that the exact KS wavefunction is close enough to the single atom, atomic wavefunction $|\phi_i\rangle$. As such if we choose S_{μ} to be

$$S_{\mu} |\psi_n\rangle = \sum c_i^n |\phi_i\rangle - \sum c_i^n \left|\tilde{\phi}_i\right\rangle \tag{7}$$

where $\left|\tilde{\phi}_i\right\rangle$ are smooth wave functions. Where we choose

$$c_i^n = \left\langle \tilde{p}_i \mid \tilde{\psi}_n \right\rangle \tag{8}$$

Here $|\tilde{p}_i\rangle$ are the projector wave functions. As such we have that:

$$|\psi_n\rangle = \hat{\mathcal{T}} \left|\tilde{\psi}_n\right\rangle = \left|\tilde{\psi}_n\right\rangle + \sum_i \left(|\phi_i\rangle - \left|\tilde{\phi}_i\right\rangle\right) \left\langle \tilde{p}_i \mid \tilde{\psi}_n \right\rangle \tag{9}$$

2. Reformulation of PAW as an all electron method

The main claim of PAW (which has been extensively numerically verified) is that there is a good expansion in plane waves for the PAW wavefunction given by:

$$|\psi_n\rangle \cong \left|\tilde{\psi}_n\right\rangle + \sum_i \left(|\phi_i\rangle - \left|\tilde{\phi}_i\right\rangle\right) \left\langle \tilde{p}_i \mid \tilde{\psi}_n \right\rangle. \tag{10}$$

That is the wavefunction $\left|\tilde{\psi}_n\right\rangle$ is made of a reasonably small number of plane waves of the form $\frac{1}{\sqrt{V}} \exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)$ in particular we have that:

$$|\psi_n\rangle \in Span\left\{\left\{\frac{1}{\sqrt{V}} \exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\}, \{|\phi_i\rangle\}, \left\{\left|\tilde{\phi}_i\right\rangle\right\}\right\} \tag{11}$$

for small number of plane waves here $|\psi_n\rangle$ are all electron wave functions. Now because $\left|\tilde{\phi}_i\right\rangle$ is very smooth we have that

$$\left|\tilde{\phi}_i\right\rangle \in Span\left\{\frac{1}{\sqrt{V}} \exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\} \tag{12}$$

As such the main claim of PAW can be written as

$$|\psi_n\rangle \in Span\left\{\left\{\frac{1}{\sqrt{V}} \exp\left(i\left(\mathbf{k}+\mathbf{K}\right)\cdot\mathbf{r}\right)\right\}, \{|\phi_i\left(\mathbf{r}\right)\rangle\}\right\} \tag{13}$$

for a reasonable number of plane waves. Since the basis set of PAW and Phillips-Kleinman are similar both methods should yield in their all electron formulation similar accuracy and basis of the same size. They are equivalent within the all electron formulation when they use the same localized orbitals.

III. GLOW AND PLOW BASIS SETS

Here we would like to work backwards we wish to formulate an all electron method as the inverse of a pseudopotential method where we use specialized wave functions near the nuclei and many regular wave functions for the whole space. We focus on small molecules and crystalline solids. Below we will simply describe the relevant basis sets.

A. Glow and GLOW for small molecules

Here G in Glo and GLO stands for Gaussians while lo and LO for localized orbitals. We consider the lo basis set [6]:

$$\Phi_{lo}^{lm} = u_l^{E_l}\left(|\mathbf{r}-\mathbf{r}_{\mu}|\right) \Theta\left(S_l^{\mu} - |\mathbf{r}-\mathbf{r}_{\mu}|\right) Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_{\mu}}\right) \tag{14}$$

where S_l^{μ} is chosen to make the wavefunction continuous everywhere. Here:

$$\left[-\frac{d^2}{dr^2} + \frac{l\left(l+1\right)}{r^2} + \bar{V}_{KS}\left(r\right)\right] r u_{l\mu}^{E_l}\left(r\right) = E_l r u_{l\mu}^{E_l}\left(r\right) \tag{15}$$

and $\bar{V}_{KS}\left(r\right)$ is the spherically average KS potential or HF single particle potential. Alternatively we can use LO basis sets [6]:

$$\begin{aligned}\Phi_{LO}^{lm} = {} & u_l^{E_l}\left(|\mathbf{r}-\mathbf{r}_{\mu}|\right) \Theta\left(S_l^{\mu} - |\mathbf{r}-\mathbf{r}_{\mu}|\right) Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_{\mu}}\right) \\ & + B_l^{\mu} \dot{u}_l^{E_l}\left(|\mathbf{r}-\mathbf{r}_{\mu}|\right) \Theta\left(S_l^{\mu} - |\mathbf{r}-\mathbf{r}_{\mu}|\right) Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_{\mu}}\right)\end{aligned} \tag{16}$$

where B_l^{μ} and S_l^{μ} are chosen to make the wavefunction continuous and continuously differentiable everywhere.

Here $\dot{u}_l^{E_l} = \frac{\partial}{\partial E} u_l^{E_l}$. Higher derivative terms are also possible [6], also regular LO and lo is possible [1, 5, 7]. We also add a large number of Gaussians to our basis set:

$$\phi^{\mu}_{l,m,\sigma} = \frac{1}{\mathcal{N}_l} \exp\left(-\frac{|\mathbf{r}-\mathbf{r}_\mu|^2}{2\sigma}\right) |\mathbf{r}-\mathbf{r}_\mu|^l Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_\mu}\right) \tag{17}$$

to the molecular basis set as well as polynomials times $\phi^{\mu}_{l,m,\sigma}$, these help make configuration integrals cheap. Here $\mathcal{N}_l$ is a normalization constant. This is a basis set for many small molecules. Furthermore Gaussians Coulomb integrals may be efficiently computed [8, 9] so there are "few" coulomb integrals involving complicated functions [10].

B. Plow and PLOW basis sets for crystalline solids

Here P in Plo and PLO stands for plane wave while lo and LO stands for localized orbitals. We start with lo basis wave functions [6]

$$\Phi_{lo} = u_l^{E_l}(|\mathbf{r}-\mathbf{r}_\mu|)\,\Theta(S_l^\mu - |\mathbf{r}-\mathbf{r}_\mu|)\,Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_\mu}\right) \tag{18}$$

where S_l^μ is chosen to make the wavefunction continuous everywhere and LO basis wave functions [6]:

$$\begin{aligned}\Phi_{LO} = {} & u_l^{E_l}(|\mathbf{r}-\mathbf{r}_\mu|)\,\Theta(S_l^\mu - |\mathbf{r}-\mathbf{r}_\mu|)\,Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_\mu}\right) \\ & + B_l \dot{u}_l^{E_l}(|\mathbf{r}-\mathbf{r}_\mu|)\,\Theta(S_l^\mu - |\mathbf{r}-\mathbf{r}_\mu|)\,Y_{lm}\left(\widehat{\mathbf{r}-\mathbf{r}_\mu}\right)\end{aligned} \tag{19}$$

where B_l^μ and S_l^μ are chosen to make the wavefunction continuous and continuously differentiable everywhere. Higher derivative terms are also possible [6, 11, 12]. We can also add plane wave

$$\chi_{\mathbf{k}+\mathbf{K}} = \frac{1}{\sqrt{V}} \exp\left(i\,(\mathbf{k}+\mathbf{K})\cdot\mathbf{r}\right) \tag{20}$$

to the solid state basis set, these help make configuration integrals cheap. This is a basis set for many crystalline solids.

IV. CONCLUSIONS

In this work - inspired by the fact that many pseudopotential methods are all electron methods in disguise with slowly oscillating wave functions and special wave functions to account for the environment of the nucleus - we proposed new basis sets for small molecules and crystalline solids. In this basis set we use the LO or lo wave functions introduced recently in [6] (although regular LO or lo basis wave functions will also do [1, 5, 7]) combined either with Gaussians (for molecules) or plane waves (for solids) to obtain a total basis for the system. This would allow for initial calculations for many molecules as the four center calculation is done for Gaussians and two center density density is done for very generic functions [2, 8, 10]. This should help open small molecules for theoretical DFT or HF like calculations and explorations.

[1] E. Sjostedt, L. Nordstrom, and D. J. Singh, Sol. Sta. Comm. **114**, 15 (2000).

[2] S. M. Blinder and D. E. Huse, Mathematical Physics in Theoretical Chemistry, (Elsevier Inc. Amsterdam, 2019).

[3] R. M. Martin, *Electronic Structures Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2020).

[4] D. Marx and J. Hutter, *Ab Initio Molecular Dynamics Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, 2009).

[5] D. J. Singh, Phys. Rev. B **43**, 6388 (1991).

[6] G. Goldstein, arXiv 2403.1594

[7] D. J. Singh and D. Nordstrom, *Planewaves, pseudopotentials, and the LAPW method* (Springer, New York, 2006).

[8] T. Helgaker, P. Jorgensen, J. Olsen, Molecular Electronic Structure Theory (John Wiley & Sons LTD. New York, 2000).

[9] S. F. Boys, Proc. Roy. Soc. A200, **542** (1950).

[10] G. Goldstein, arXiv 2404.08493

[11] G. Michalicek, *Extending the precision and efficiency of all-electron full-potential linearized augment plane-wave density functional theory* (Aachen University, 2014, thesis).

[12] G. Michalicek, M. Betzinger, C. Freidrich and S. Blugel, Comp. Phys. Comm. **184**, 2670 (2013).