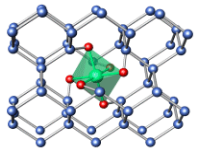
Reinvigorating Norm Conservation: the ONCVPSP Project

D. R. Hamann

Department of Physics and Astronomy, Rutgers University, Piscataway, NJ
and Mat-Sim Research LLC, Murray Hill, NJ

Rutgers
Department of Physics and Astronomy



Mat-Sim Research

Project background

- High-throughput calculations in the Vanderbilt/Rabe Rutgers group were initially carried out using published norm-conserving psp tables
- Testing against LAPW+LO all-electron benchmarks showed inconsistent agreement
- An improved table was initially generated using the norm-conserving OPIUM code
- Better results were obtained using VANDERBILT ULTRASOFT code
- The resulting extensively tested GBRV psp table[1] is available at http://www.physics.rutgers.edu/gbrv/
- So why worry about norm conservation?

1. Garrity, Bennett, Rabe & Vanderbilt, Comput. Mater. Sci. **81**, 446 (2014)

Advantages of norm-conservation and goals

- Ultrasoft and PAW potentials require computations to treat
  - Generalized eigenvalue problems
  - Augmentation of the charge density
  - Self-consistent contributions to each non-local potential
- Norm-conserving computations can use much simpler algorithms
  - Especially important for more complex calculations such as DFPT, GW, BES, QMC
  - Example: DFPT for elastic constants has yet to be implemented for anything but norm-conserving psps
- Accuracy goal – ncpsps should be competitive with ultrasoft and PAW
- Convergence goal – systematic optimization should adequately "soften" ncpsps
- Robustness goal – "tuning" psps to fit certain sets of test data should be unnecessary and disparaged
  - There should be no "black art" in making good psps
  - Any graduate student should be able to do so, **and should!**

Separable psps

- The traditional approach (Kleinman-Bylander, Blöchl):
  - Construct pseudo wave function φ smoothly matching all electron ψ at core radius r_c and its norm inside r_c for each ℓ.*
  - Invert the Schrödinger equation to find the semi-local pseudopotential
  - Choose a local potential matching the all-electron potential outside r_c.
  - Calculate projector from φ, the semi-local psp, and local potential. (KB)
  - This duplicates all-electron scattering and its first energy derivative at the energy ε of the original ψ
  - Additional projectors can duplicate semi-local potential scattering at additional ε (Blöchl)
- The Vanderbilt approach
  - Construct the projectors directly:

$$|\chi\rangle = (\varepsilon - T - V_{\text{loc}})|\varphi\rangle, \quad T = [-d^2/dr^2 + \ell(\ell+1)/r^2]/2$$

  - For one projector, $V_{\text{NL}} = \dfrac{|\chi\rangle\langle\chi|}{\langle\varphi|\chi\rangle}$
  - This is the KB result, but without inverting the Schrödinger eq.

(* *ℓ and m indices will generally be omitted and can be assumed where needed)*

Multi-projector separable psps

- Vanderbilt approach for multiple projectors
  - Calculate ψ_i at several ε_i for each ℓ .
  - Construct φ_i satisfying continuity conditions with the ψ_i at r_c
  - The separable potential can now have the form

$$V_{\mathrm{NL}} = \sum_{i,j} \left|\chi_i\right\rangle (B^{-1})_{ij} \left\langle\chi_j\right| , \quad B_{ij} = \left\langle\varphi_i \middle| \chi_j\right\rangle .$$

- Prove B_{ij} is symmetric and V_{NL} is Hermitian if φ_i also satisfy generalized norm conservation:

$$\left\langle\varphi_i \middle| \varphi_j\right\rangle_{r_c} = \left\langle\psi_i \middle| \psi_j\right\rangle_{r_c}$$

  - Log derivatives and energy derivatives of log derivatives agree with AE results at each ε_i .
- Branch point: "One could stop here, and still have a useful scheme."
  - Violate generalized norm conservation to get *ULTRASOFT* psps
  - Redefine projectors and restore effective Hermiticity with an overlap matrix in a generalized eigenvalue formulation
  - Compute an augmentation operator to add charge to the plane-wave charge density

ONCVPSP – on the other Riemann sheet

- Enforce generalized norm conservation (**O**ptimized "**N**orm-**C**onserving **V**anderbilt **Ps**eudo**p**otentials"[2])
- Find that two projectors give excellent log-derivative agreement over a wide energy range for a wide variety of atoms and reference ε_i choices:
  - Semi-core – valence, valence – scattering, scattering – scattering
- There is a caveat about relativistic all-electron reference ψ_i
- The key matrix element in GNC is

$$B_{ij} = \int_0^{r_c} \varphi_i \left[\varepsilon_j + \frac{1}{2}\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{2r^2} - V_{\text{loc}} \right] \varphi_j$$

- The symmetry of B_{ij} and other good properties follow from integration by parts of this expression and the corresponding ψ_i expression
- For scalar-relativistic and Dirac-equation solutions, this doesn't work
- In practice, B_{ij} asymmetries are ~10^{-4}
- Ad-hoc symmetrization results in acceptable errors ~10^{-5} in eigenvalues, log-derivatives, norms, etc. independent of atomic Z

2. Quoted name from Morrison, Kleinman & Bylander, Phys. Rev. B **47**, 6728 (1993)

V_{NL} format and spin-orbit decomposition

- For easiest use with applications, it is best to find eigenfunctions giving a diagonal expression with orthonormal projectors:

$$V_{\mathrm{NL}}(\mathbf{r},\mathbf{r}') = \sum_{\ell m i} \left|\tilde{\chi}_{\ell m i}\right\rangle e_{\ell i} \left\langle\tilde{\chi}_{\ell m i}\right|$$

- For Dirac-wave-function based psps, the sum is over $j = \ell \pm \frac{1}{2}$
- Most applications require SO psps in the (schematic) form

$$V_{\mathbf{NL}}(\mathbf{r},\mathbf{r}') = \sum_{\ell}\left[V_{\ell}^{\mathrm{SR}}(\mathbf{r},\mathbf{r}') + \mathbf{L}\cdot\mathbf{S}V_{\ell}^{\mathrm{SO}}(\mathbf{r},\mathbf{r}')\right]$$

$$V_{\ell}^{\mathrm{SR}} = \left[(\ell+1)V_{\ell+1/2}^{\mathrm{Rel}} + \ell V_{\ell-1/2}^{\mathrm{Rel}}\right]\Big/(2\ell+1)\,, \quad V_{\ell}^{\mathrm{SO}} = 2\left[V_{\ell+1/2}^{\mathrm{Rel}} - V_{\ell-1/2}^{\mathrm{Rel}}\right]\Big/(2\ell+1)$$

- Direct expression of the SR and SO potentials in terms of Dirac $\tilde{\chi}_{j,i}$ requires 8 projectors per ℓ and subtractions of large, nearly equal terms in the applications
- Instead, we find eigenfunctions $\chi_{\ell,i}^{\mathrm{SR}}$ and $\chi_{\ell,i}^{\mathrm{SO}}$ of V_{ℓ}^{SR} and V_{ℓ}^{SO}, and find that one or two eigenvalues of each are usually negligibly small (< 10^{-5} Ha)

Convergence optimization

- The best method is that of Rappe, Rabe, Kaxiras, and Joannopoulos
  - Adjust the psuedo wave function to minimize the kinetic energy error due to the cutoff of its plane-wave expansion (proxy for total energy)
  - Incorporated in the open-source OPIUM psp code
  - This proved too difficult to adapt to my purposes
- Reformulation of the method introducing a general residual kinetic energy operator:

$$\left\langle \xi_{\ell i} \left| \hat{E}^{\mathrm{R}}(q_{\mathrm{c}}) \right| \xi_{\ell j} \right\rangle = \int_{q_{\mathrm{c}}}^{\infty} \xi_{\ell i}(q)\xi_{\ell j}(q)q^4 dq, \quad \xi_{\ell i}(q) = 4\pi \int_0^{\infty} j_{\ell}(qr)\xi_{\ell i}(r)r^2 dr$$

  where $\xi_{\ell i}$ are some set of basis functions, j_{ℓ} is a spherical Bessel functions, and q_c is the expansion cutoff
- Introduce an initial basis set of N spherical Bessel functions

$$\xi_i^{\mathrm{B}} = j_{\ell}(q_i r), r \le r_c \; ; \; \xi_i^{\mathrm{B}} = 0, r > r_c$$

- Orthogonalize and normalize

$$\xi_i^{\mathrm{O}} = \sum_{j=1}^{N} (S^{-1/2})_{ij} \xi_j^{\mathrm{B}} \; ; \; S_{ij} = \left\langle \xi_i^{\mathrm{B}} \left| \xi_j^{\mathrm{B}} \right. \right\rangle$$

Optimization (single φ) made simple

- M matching conditions $\left.\frac{d^n\varphi}{dr^n}\right|_{r_c} = \left.\frac{d^n\psi}{dr^n}\right|_{r_c}$, $n = 0, M-1$

  give M linear equations for N ξ_i^{O} coefficients solved for matching φ_0 function and N-M orthonormal "null space" functions ξ_i^{N}
- Diagonalize the positive-definite matrix $\left\langle \xi_i^{\mathrm{N}} \middle| \hat{E}^{\mathrm{r}}(q_c) \middle| \xi_j^{\mathrm{N}} \right\rangle$ finding its eigenvalues e_i^{R} and using its eigenfunctions ξ_i^{R} to expand φ

$$\varphi = \varphi_0 + \sum_i x_i \xi_i^{\mathrm{R}}$$

- Residual energy and norm constraint are diagonal quadratic forms

$$\left\langle \varphi \middle| \hat{E}^r \middle| \varphi \right\rangle = \left\langle \varphi_0 \middle| \hat{E}^r \middle| \varphi_0 \right\rangle + \sum_i \left[2 \left\langle \varphi_0 \middle| \hat{E}^r \middle| \xi_i^{\mathrm{R}} \right\rangle x_i + e_i^{\mathrm{R}} x_i^2 \right]$$

$$\sum_i x_i^2 = \left\langle \psi \middle| \psi \right\rangle_{r_c} - \left\langle \varphi_0 \middle| \varphi_0 \right\rangle_{r_c}$$

- Huge e_i^{R} dynamic range (~10^6) demands robust minimization approach
  - Search norm-constrained $\{x_2, \ldots, x_{N-M}\}$ hypersphere on a coarse grid for global minimum and corresponding x_1 sign
  - Finish off with Newton's method
  - Find diminishing returns for N-M>3-4 (2 is often fine)

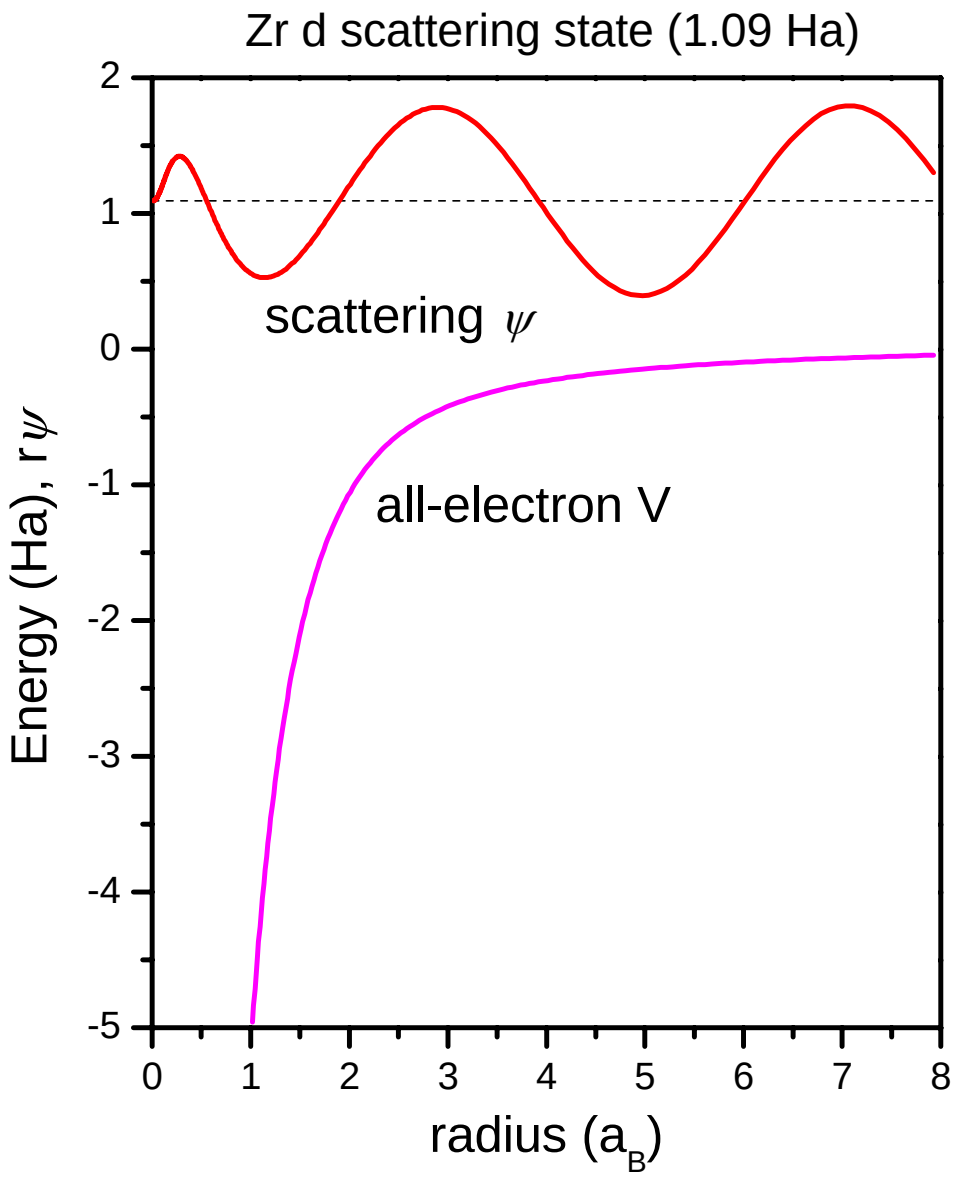
Optimizing two pseudo wave functions

- Choose all-electron ψ_1 and ψ_2, usually with one more node for ψ_2
- Optimize φ_1 first, with only the quadratic $\langle \varphi_1 | \varphi_1 \rangle_{r_c}$ norm constraint.
- Optimize φ_2 combining the linear $\langle \varphi_1 | \varphi_2 \rangle_{r_c}$ overlap constraint with linear $\varphi_2 \leftrightarrow \psi_2$ matching constraints.
- Treat the quadratic $\langle \varphi_2 | \varphi_2 \rangle_{r_c}$ constraint as usual.

Calculating a convergence profile

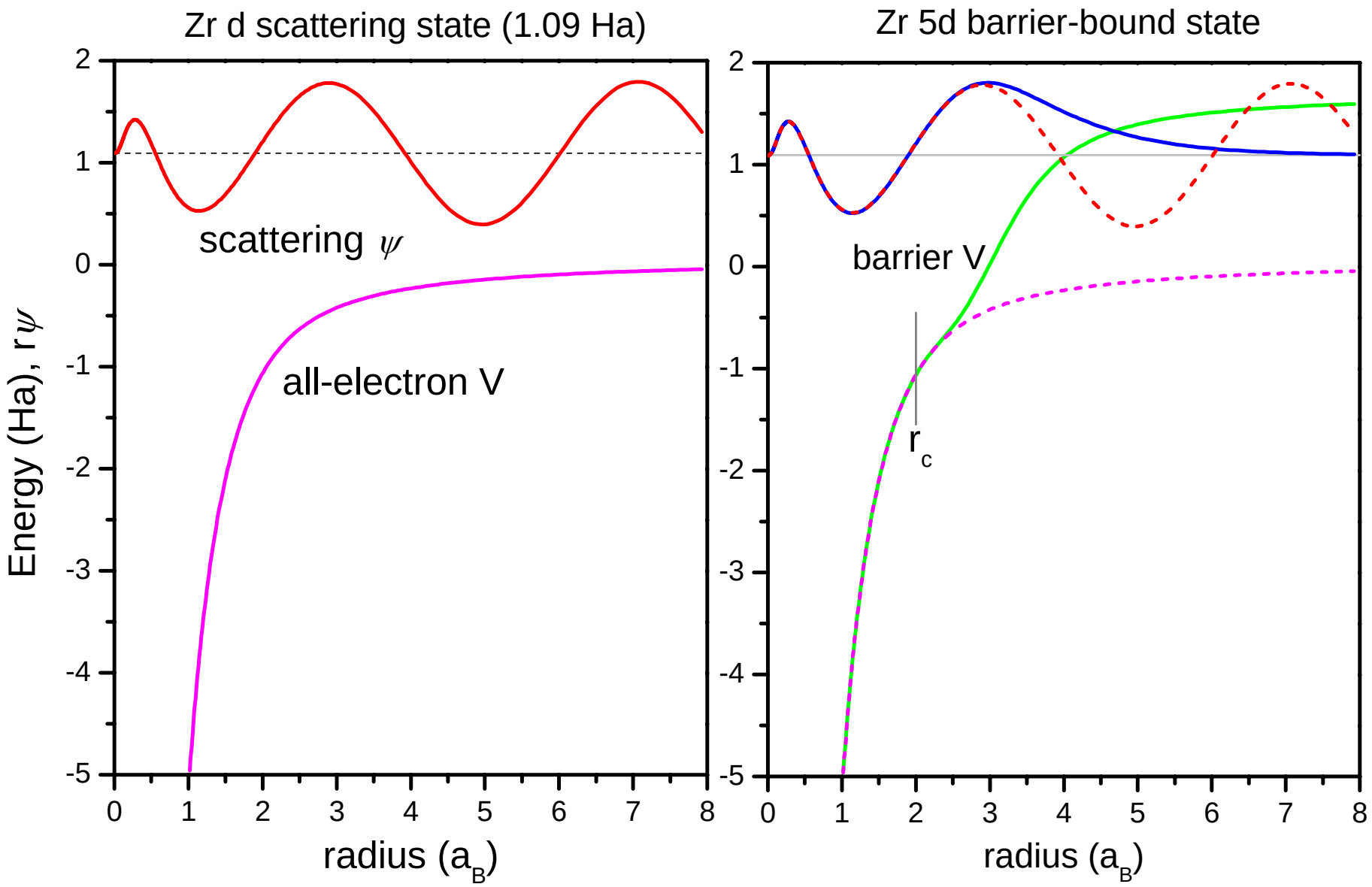
- Optimization is done with a preselected cutoff q_c
- After finding φ , $E^r(q) = \langle \varphi | \hat{E}^r(q) | \varphi \rangle$ is defined for any q
- A set of values can be calculated very efficiently by saving selected data as $\hat{E}^{\mathrm{r}}$ matrix elements are accumulated.

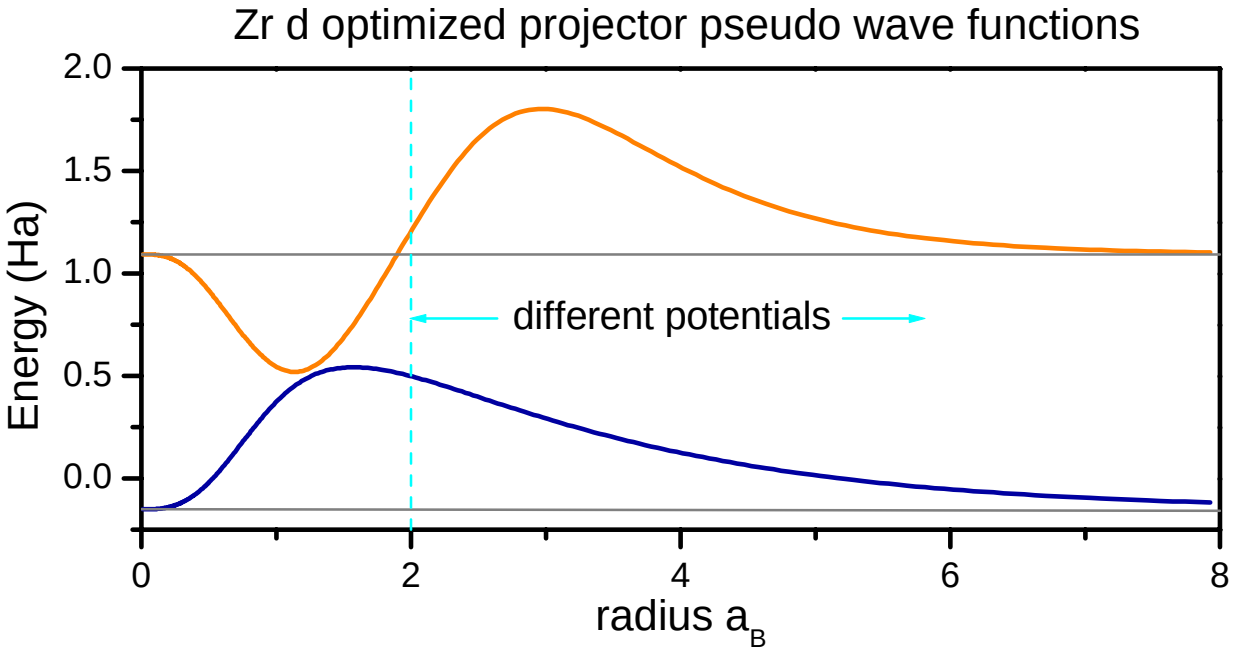
Optimizing positive energy reference states

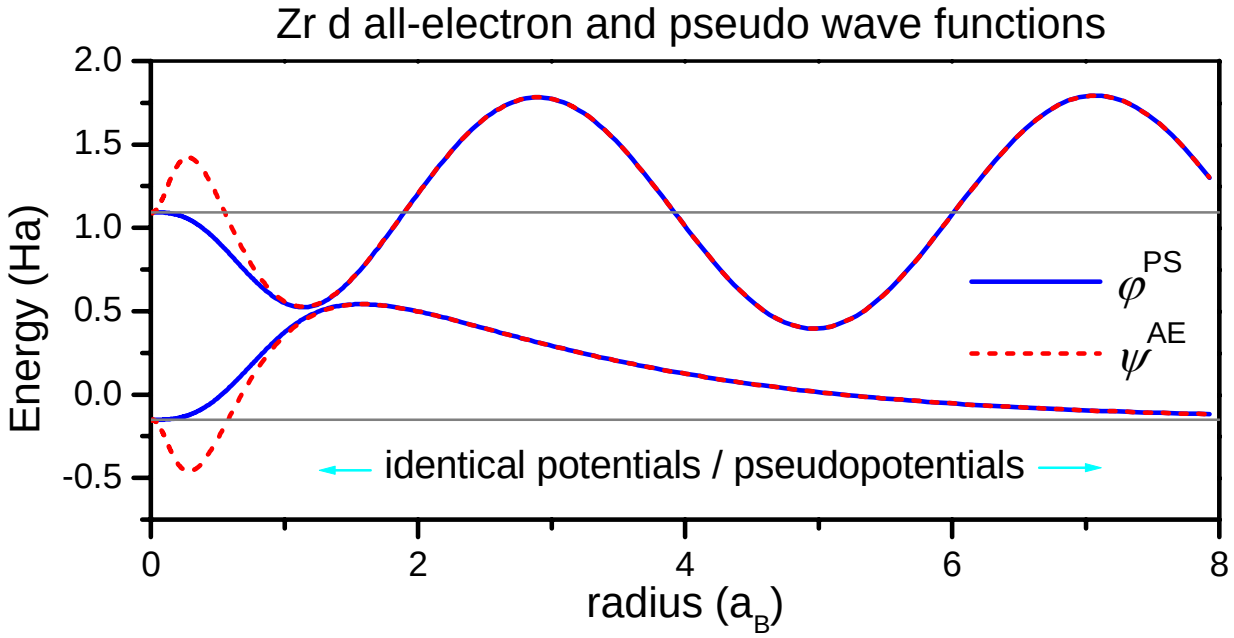


- Fourier transform of this state approaches a delta function and a useful $E^{r}(q)$ for a corresponding scattering pseudo wave function can't be defined
- Solution – create a potential that has a bound state at the desired energy
  - Generally with one more node than lower-energy states
- Barrier to create the bound state should be
  - Additive to all-electron *V*
  - Zero for $r<r_c$
  - Zero with several zero derivatives at r_c
  - Designed so that the bound ψ norm inside r_c is roughly comparable to valence states

Optimizing positive energy reference states

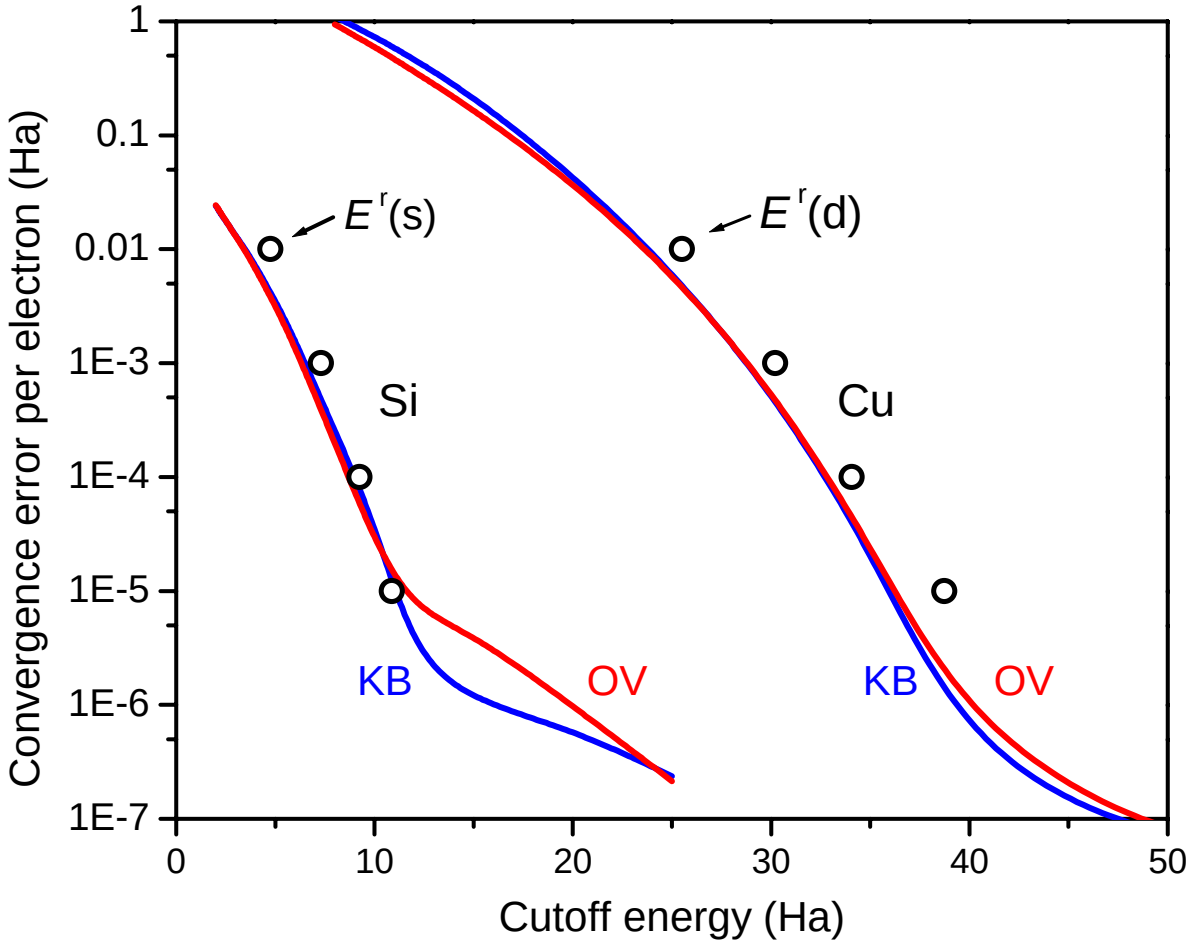






- PSPs and all-electron pots are identical for $r<r_c$ in both with and without barrier
- For barrier-free potentials, bound and scattering states obey the generalized norm conservation condition independent of scattering-state normalization
- All the resulting properties are preserved despite the use of the barrier for optimization
- Can use two positive-energy states with two different barriers

Predicting energy convergence of solids



- Solid lines are smoothed plane-wave convergence results for diamond Si and fcc Cu with one (KB) and two (OV) projectors
- Open circles are $E^r(q)$ for the most slowly convergent first projector
- Second projectors generally have negligible influence on convergence

Continuity

- Continuity in RRKJ paper and OPIUM code was limited to wave function second derivatives and hence psp/projector values
  - Slope discontinuity is a concern, for example for elastic constant DFPT calculations where 1st and 2nd psp derivatives are taken
  - Figure from original paper (below) was a little scary

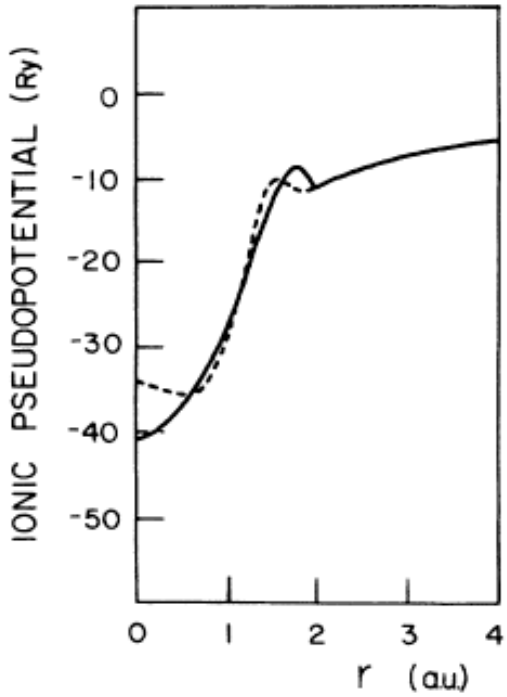
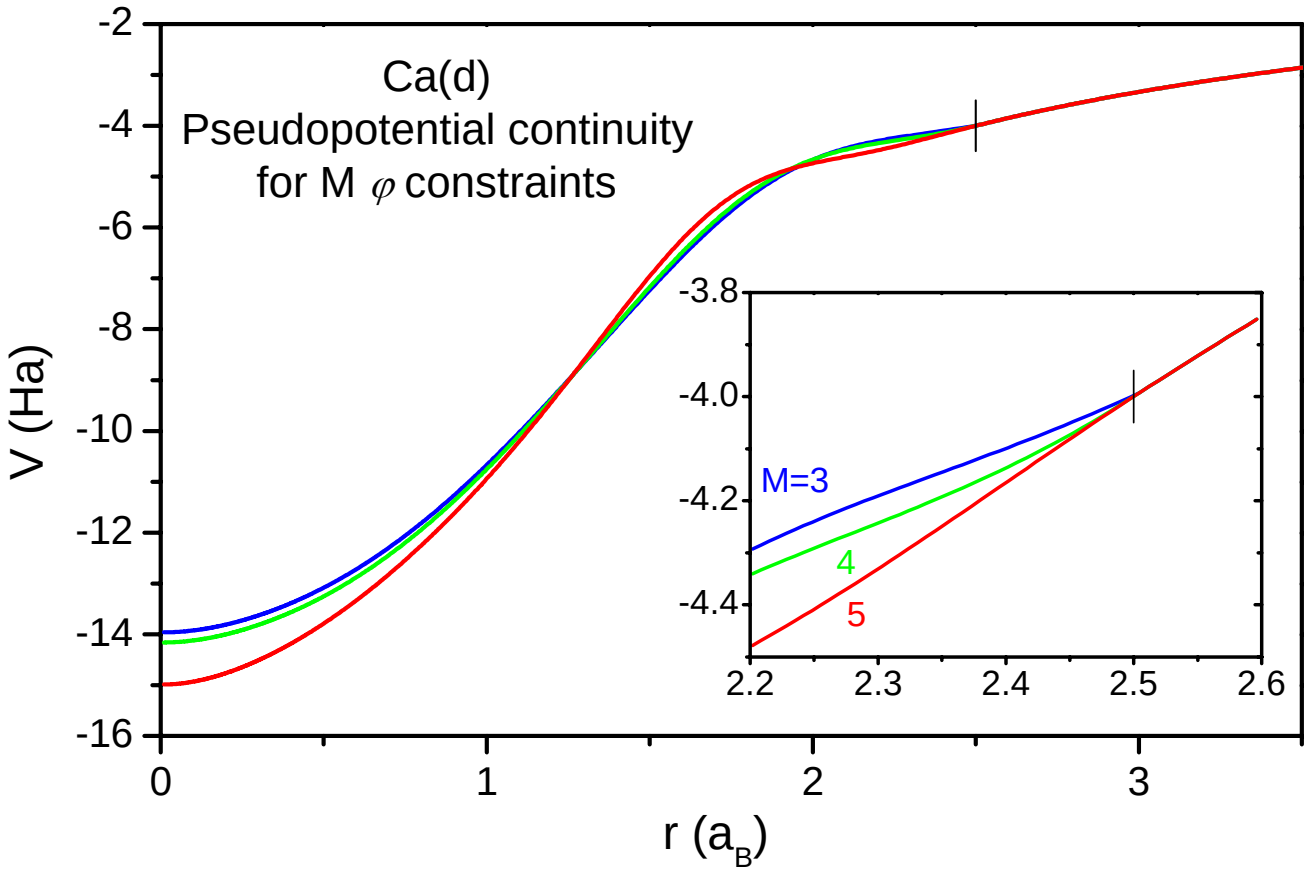


FIG. 2. Pseudopotentials for the copper 3*d* eigenstate using the HSC method (dashed line) and the present approach (solid line).

Continuity

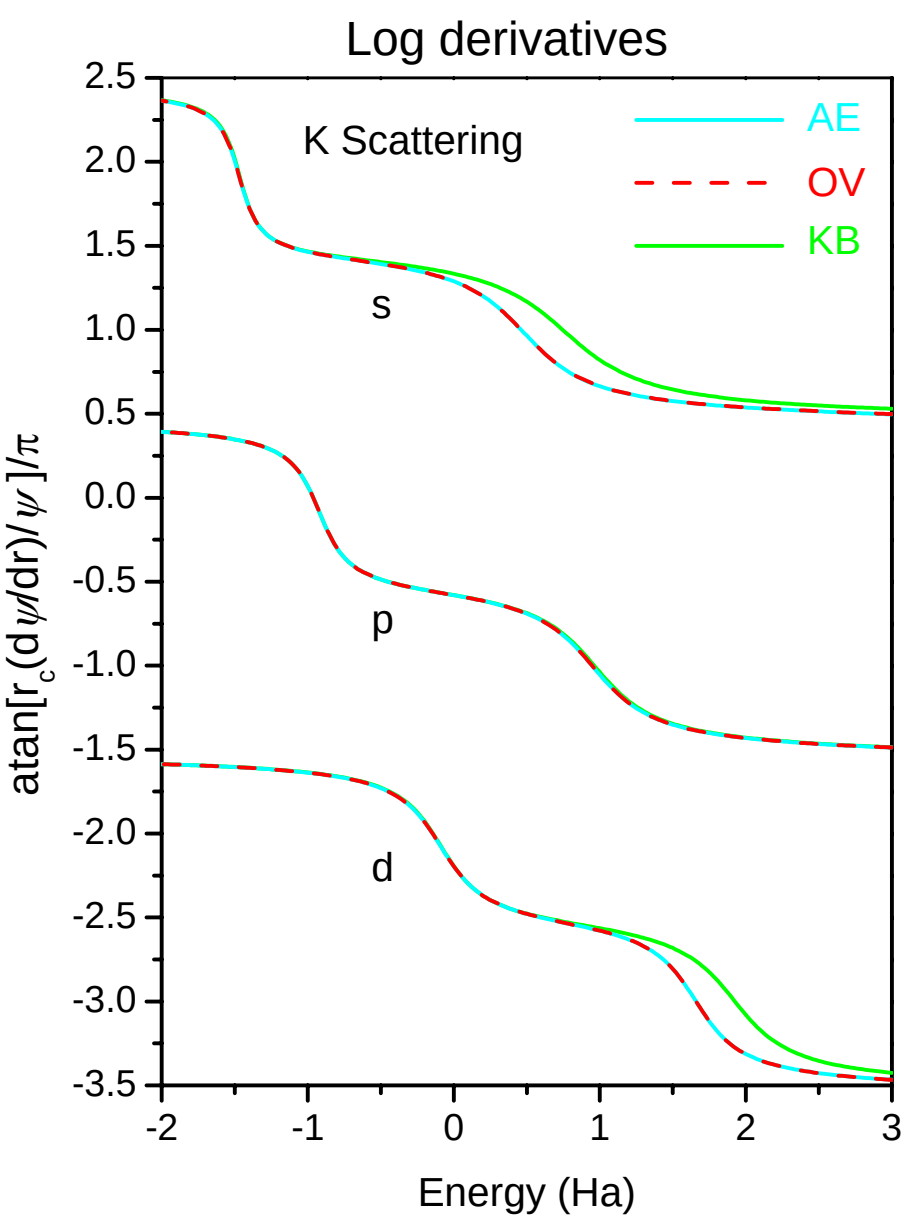
- Additional continuity is nice but in fact has very little effect
  - Present optimization minimizes slope discontinuity, even when it is not enforced (M=3 is equivalent to original)
  - This example was the worst found, and needs a lot of magnification to see the differences



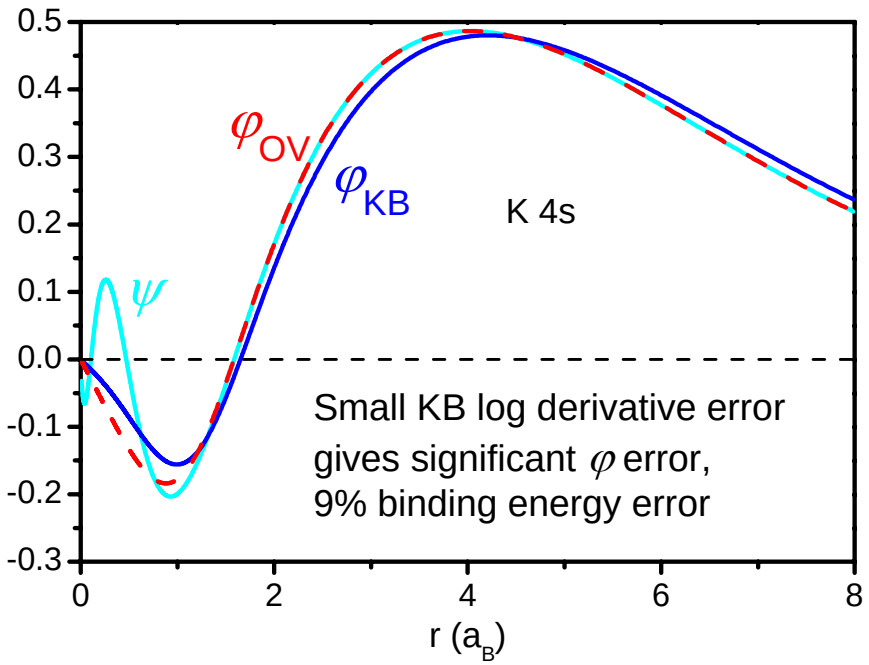
General guidelines for ONCVPSP parameters

- Neutral ground state used as reference for all atoms
  - Formal charge state has very little effect on charge densities in solids (W. Pickett, '14 March Meeting)
  - Scattering states obviate the need for ionic configurations
- Local potential is polynomial extrapolation, not a semi-local V
  - Permits two projectors per ℓ, and avoids conflicts with applications
- Nearest cores treated as valence for groups 1, 2, and transition elements, as are filled d shells for some heavier elements
  - Usually little convergence penalty because of optimization
  - Polynomial model core used otherwise for non-linear correction
- Psp parameters adjusted using built-in graphics
  - Typically, work from some nearby example and adjust r_c, $q_{c,}$ V_{loc} , N, M and projector ε separation
  - Highly "ghost-resistant," but have robust detection by log-derivative scans
- Several excited/ionized configurations are tested
  - Copies OPIUM capability, but doesn't prove that useful
- Post-testing "parameter tuning" should not be necessary, and in fact should not be able to change the results significantly
  - Very short bonds may require somewhat smaller core radii

One vs. Two projectors



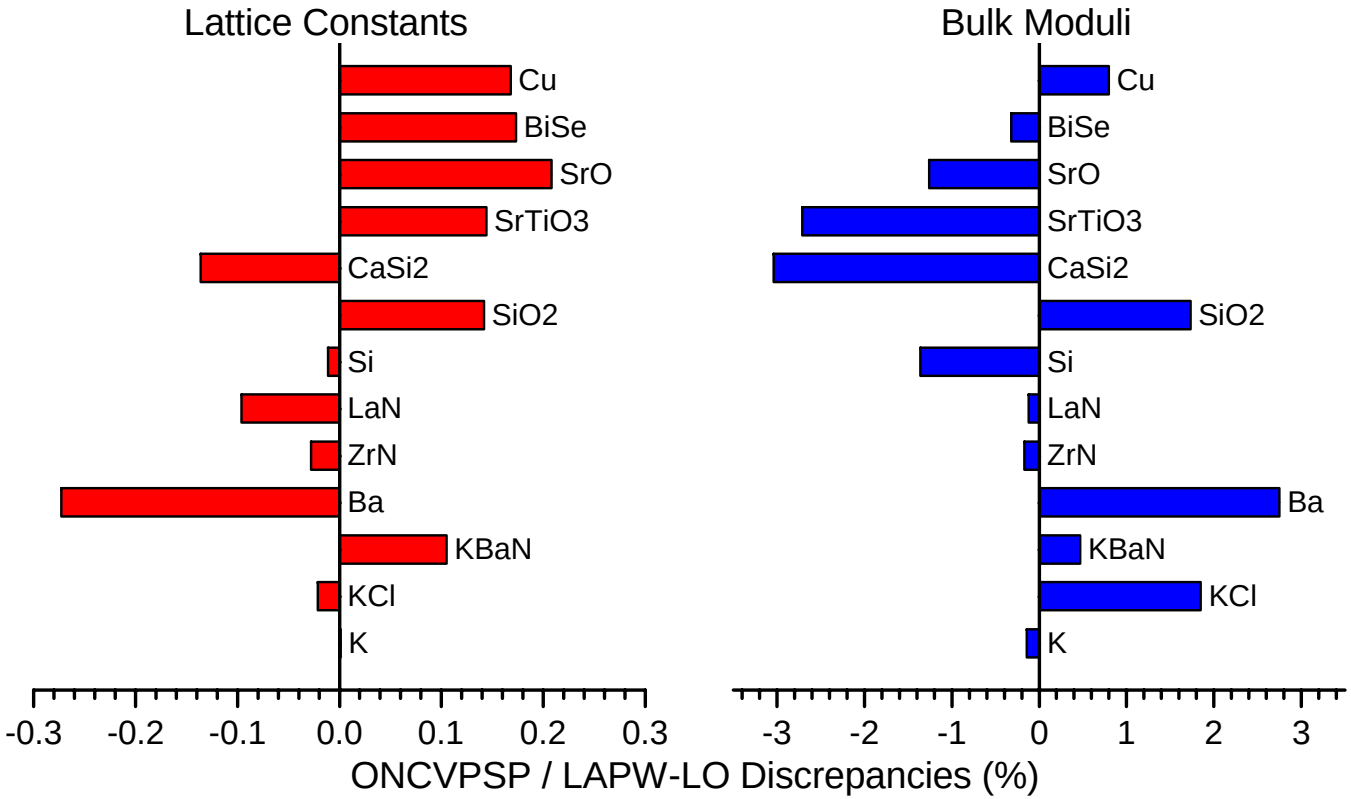
K ground-state all-electron reference,
3s, 3p, 4s treated as valence



Physical properties

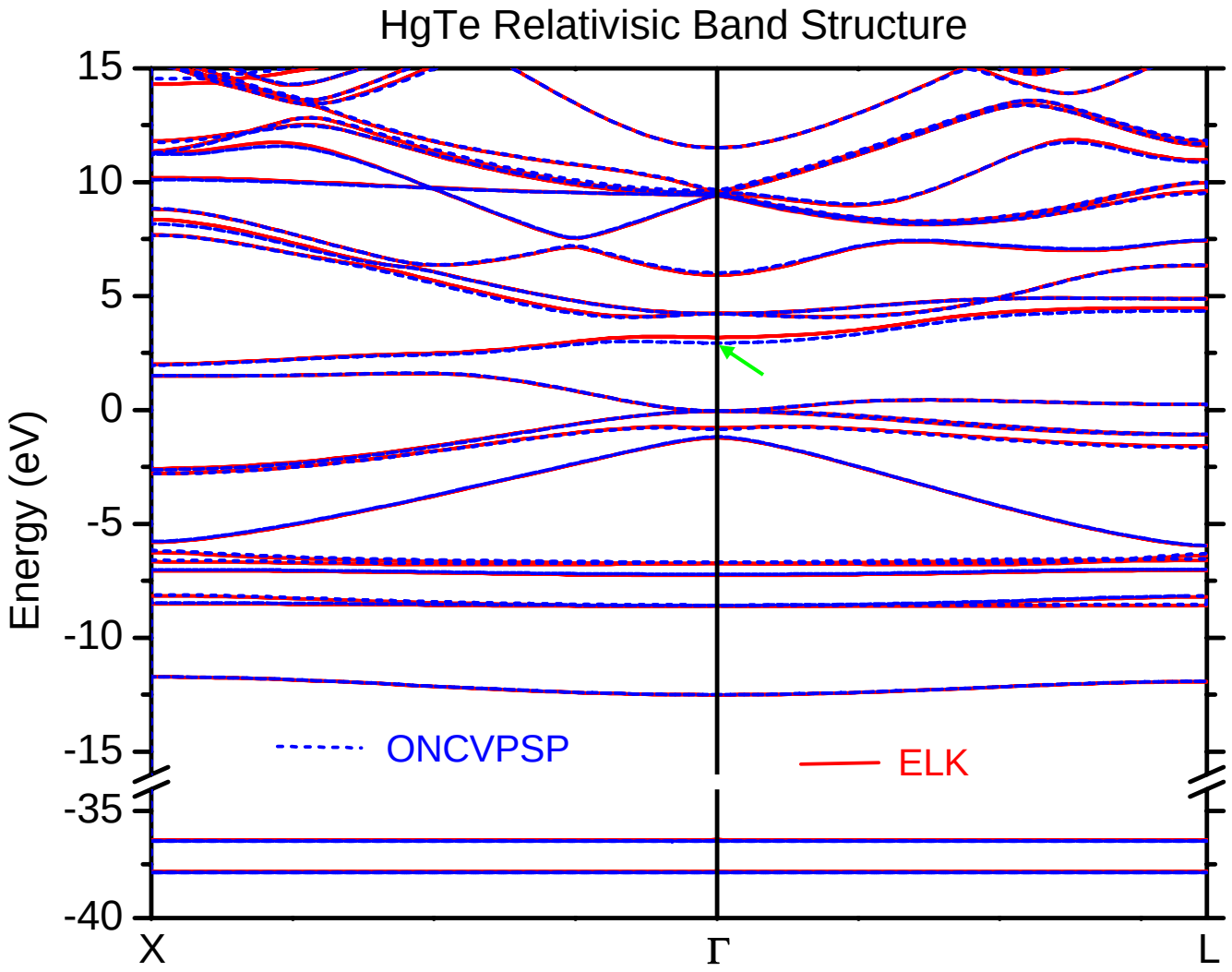
	a (a_B)	B_0 (GPa)
AE	9.58	4.14
KB	10.56	3.00
OV	9.58	4.13

Scalar-relativistic performance tests



- Variety of coordinations and formal valences
  - (Very) coarse sample of periodic table
  - Most atoms tested in several systems
- ELK for lapw+lo, ABINIT for psps with 10 – 30 Ha plane-wave cutoffs
  - Burch-Murnaghan 3rd-order EOS fits for lapw+lo
  - Lattice optimization and DFPT elastic constants fof psps

Spin-orbit comparison



- ONCVPSP using PWSCF with 25Ha cutoff
- ELK using default atomic data and convergence parameters
- 0.07 eV rms agreement over 80 bands at Γ, X, and L
- Maximum discrepancy 0.22eV at Γ **(→)**

Magnetic systems

- Respectable agreement but not as good as results for unpolarized systems
- 3s, 3p, 3d, and 4s treated as valence for metals, with 3d dominating convergence at 30 Ha
- These tests are still a work in progress
  - ELK does not find a minimum in the E(V) plot for NiO, so maybe I don't have things right with it yet for polarized systems
  - ABINIT with these psps gives a very reasonable NiO lattice constant, judging from experiment

System	a (a_B)		B_0 (GPa)		moment (m_B)	
	AE	OV	AE	OV	AE	OV
Fe	5.18	5.21	248	239	1.70	1.92
Co	6.46	6.48	264	271	1.49	1.48
Ni	6.45	6.48	279	253	0.57	0.58
MnO	8.01	8.16	167	181	4.01	4.07

Principles and plans for ONCVPSP

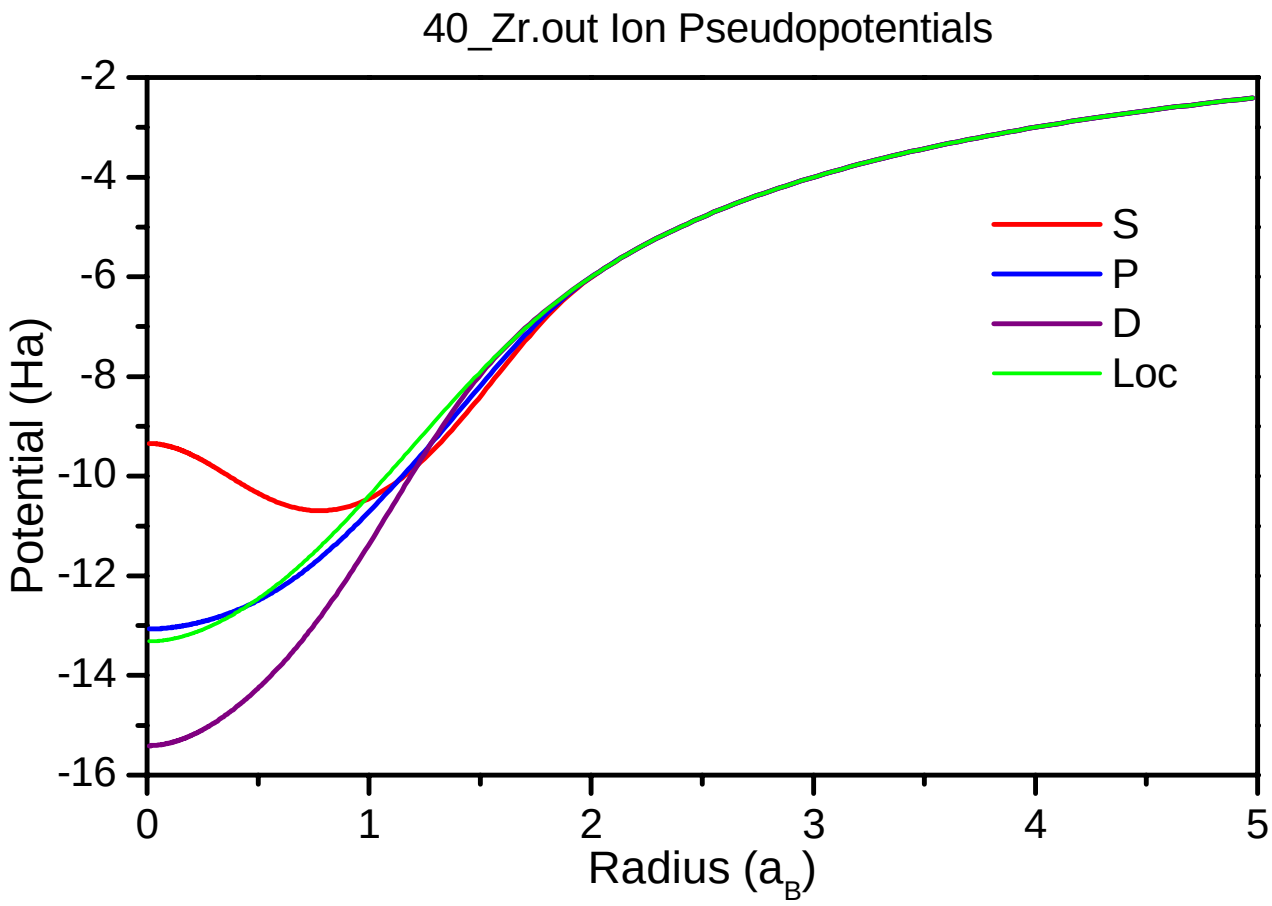
- There are no defaults – all data determining results are input in a simple template, with examples for guidance
  - The same data runs non-, scalar-, or fully-relativistic calculations
- The code is run by simple shell scripts, with a single output file
  - Start of file echoes data and gives diagnostic information
  - Remainder is parsed by script to generate “walk-through” graphics
  - Auxiliary script extracts psp files for ABINIT or QUANTUM ESPRESSO
- Sources are simple Fortran90
  - Lots of documentation and comments, no fancy datatypes
  - Should be easy to add features or psp formats
- Periodic table of these psps?
  - Complete set of input files, eventually yes with volunteered contributions
  - Psp set? **NO!*** Unaccompanied by the code, these become effectively undocumented and can’t be improved (violating open-source policy)
  - Testing of complete set? Please, be my guest! (use GBRV test set)
- Remember, the only tests that ultimately matter are experiments

* *Pseudopotentials That Work: From H to Pu*, Bachelet, Hamann & Schluter, *Phys. Rev. B* **26**, 4199 (1982)

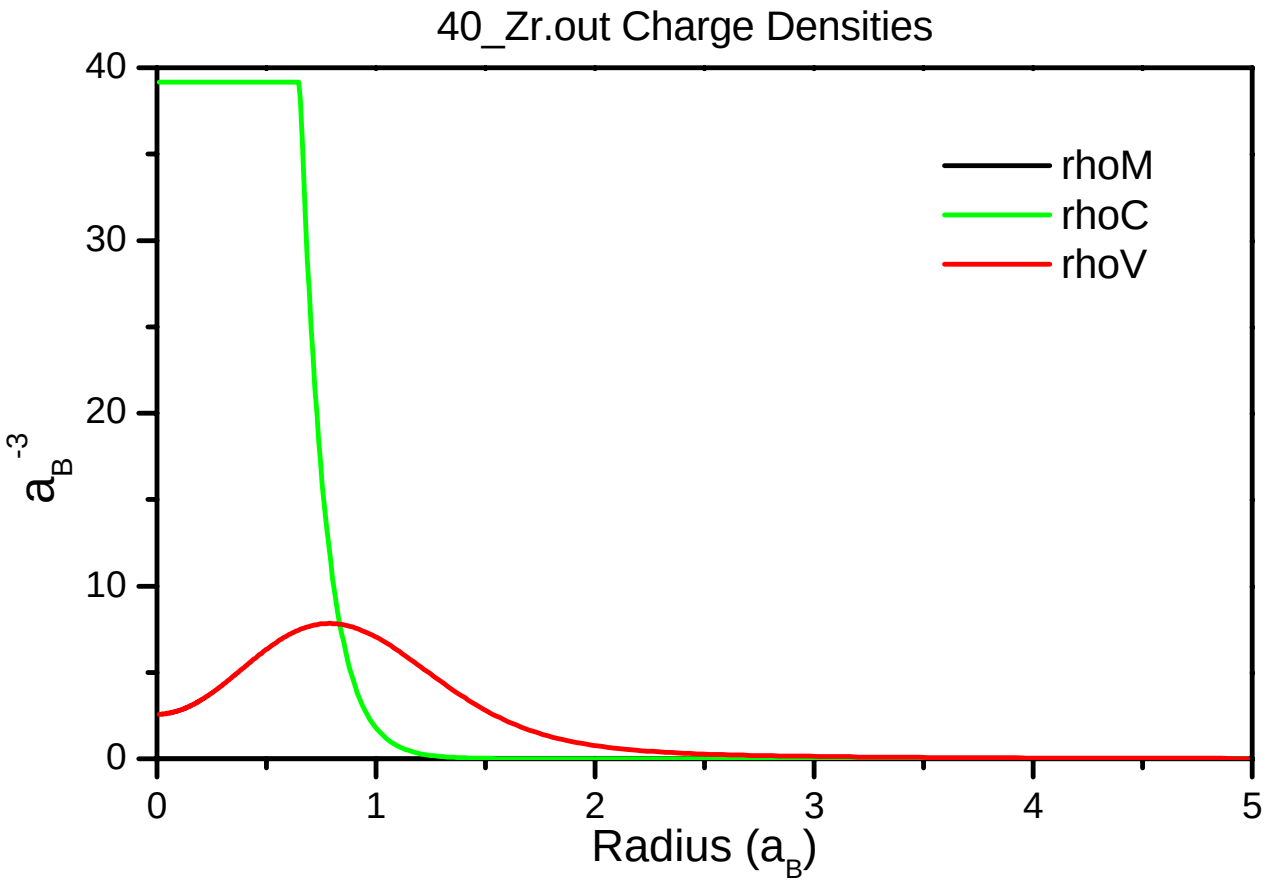
40_Zr input data

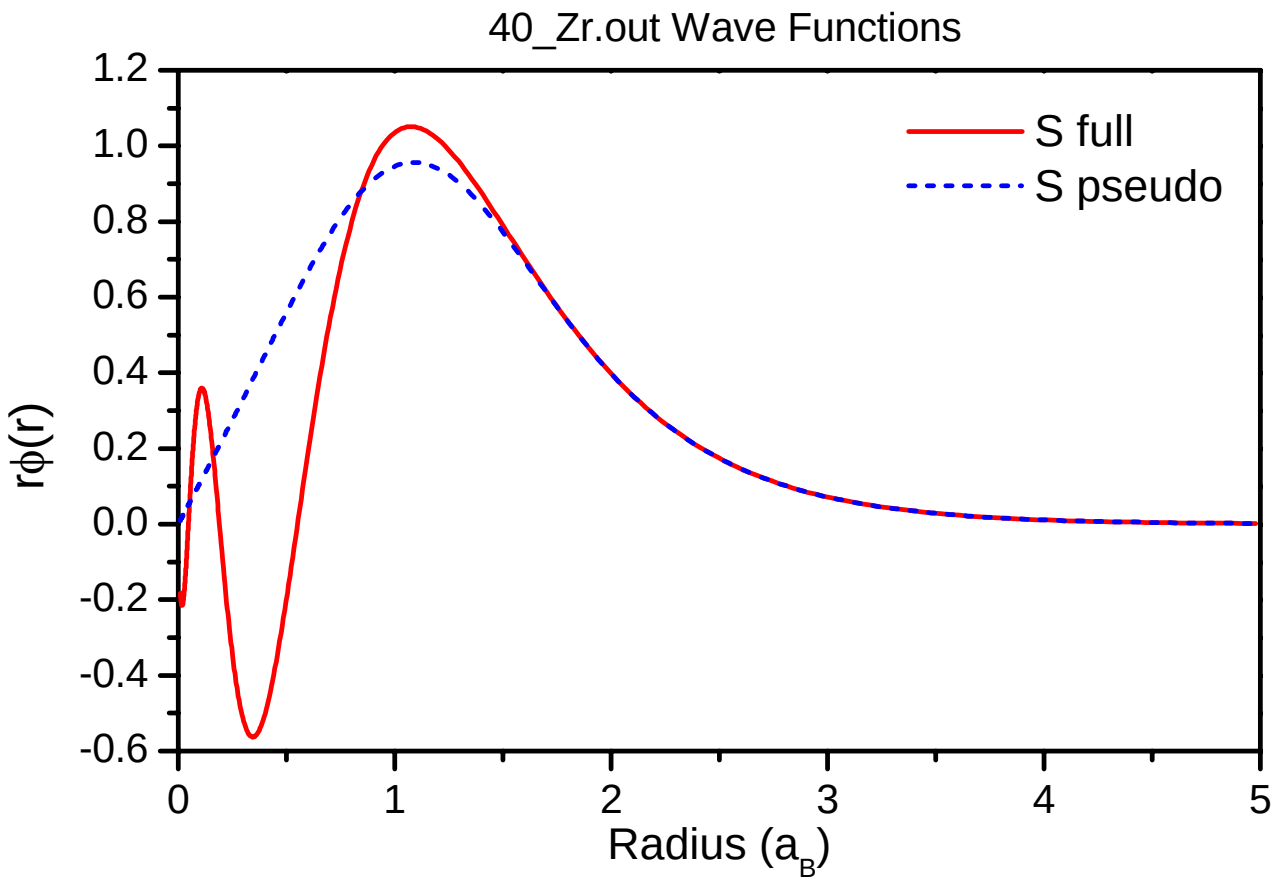
```
# ATOM AND REFERENCE CONFIGURATION
# atsym, z, nc, nv, iexc   psfile
   Zr  40.0   6   4   3   psp8
#
# n, l, f  (nc+nv lines)
   1   0   2.0
   2   0   2.0
   2   1   6.0
   3   0   2.0
   3   1   6.0
   3   2  10.0
   4   0   2.0
   4   1   6.0
   4   2   2.0
   5   0   2.0
#
# PSEUDOPOTENTIAL AND OPTIMIZATION
# lmax
   2
#
# l, rc, ep, ncon, nbas, qcut  (lmax+1 lines, l,s in order)
   0   2.20   0.00   5   8   6.00
   1   2.20   0.00   5   8   6.00
   2   2.00   0.00   5   8   6.50
#
# LOCAL POTENTIAL
# lloc, lpopt, rc(5), dvloc0
   4   5   2.0   0.0
#
# VANDERBILT-KLEINMAN-BYLANDER PROJECTORS
# l, nproj, debl  (lmax+1 lines, l's in order)
   0   2   1.50
   1   2   1.50
   2   2   1.25
#
# MODEL CORE CHARGE
# icmod, fcfact
   0   0.0
#
# LOG DERIVATIVE ANALYSIS
# epsh1, epsh2, depsh
  -3.0  3.0  0.02
#
# OUTPUT GRID
# rlmax, drl
   5.0  0.01
#
# TEST CONFIGURATIONS
# ncnf
   3
#
#  nvcnf (repeated ncnf times)
#  n, l, f  (nvcnf lines, repeated follwing nvcnf's ncnf times)
   4
   4   0   2.0
   4   1   6.0
   4   2   2.0
   5   0   1.0
#
   4
   4   0   2.0
   4   1   6.0
   4   2   1.0
   5   0   2.0
#
   4
   4   0   2.0
   4   1   6.0
   4   2   1.0
   5   0   1.0
```

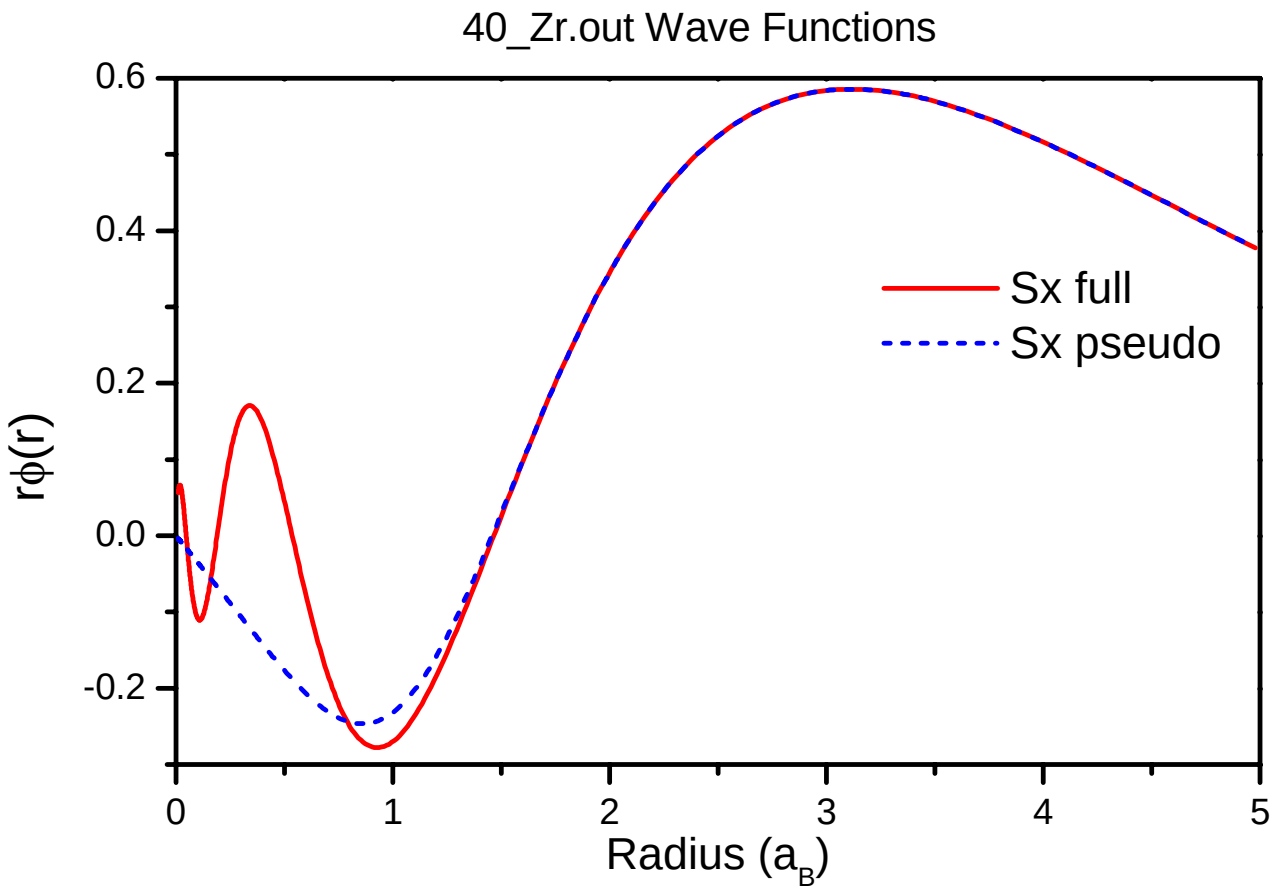
Slide show of automated walk-through graphics

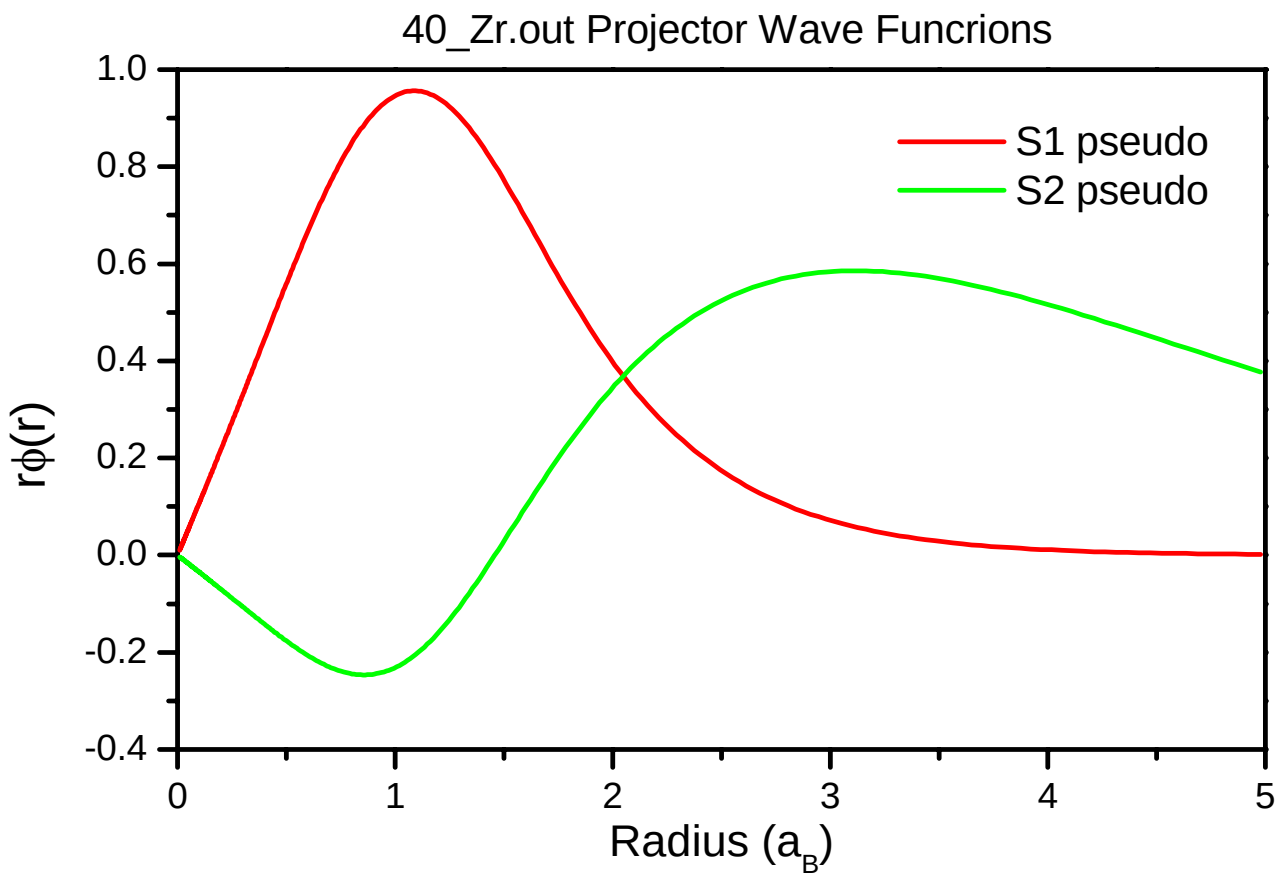


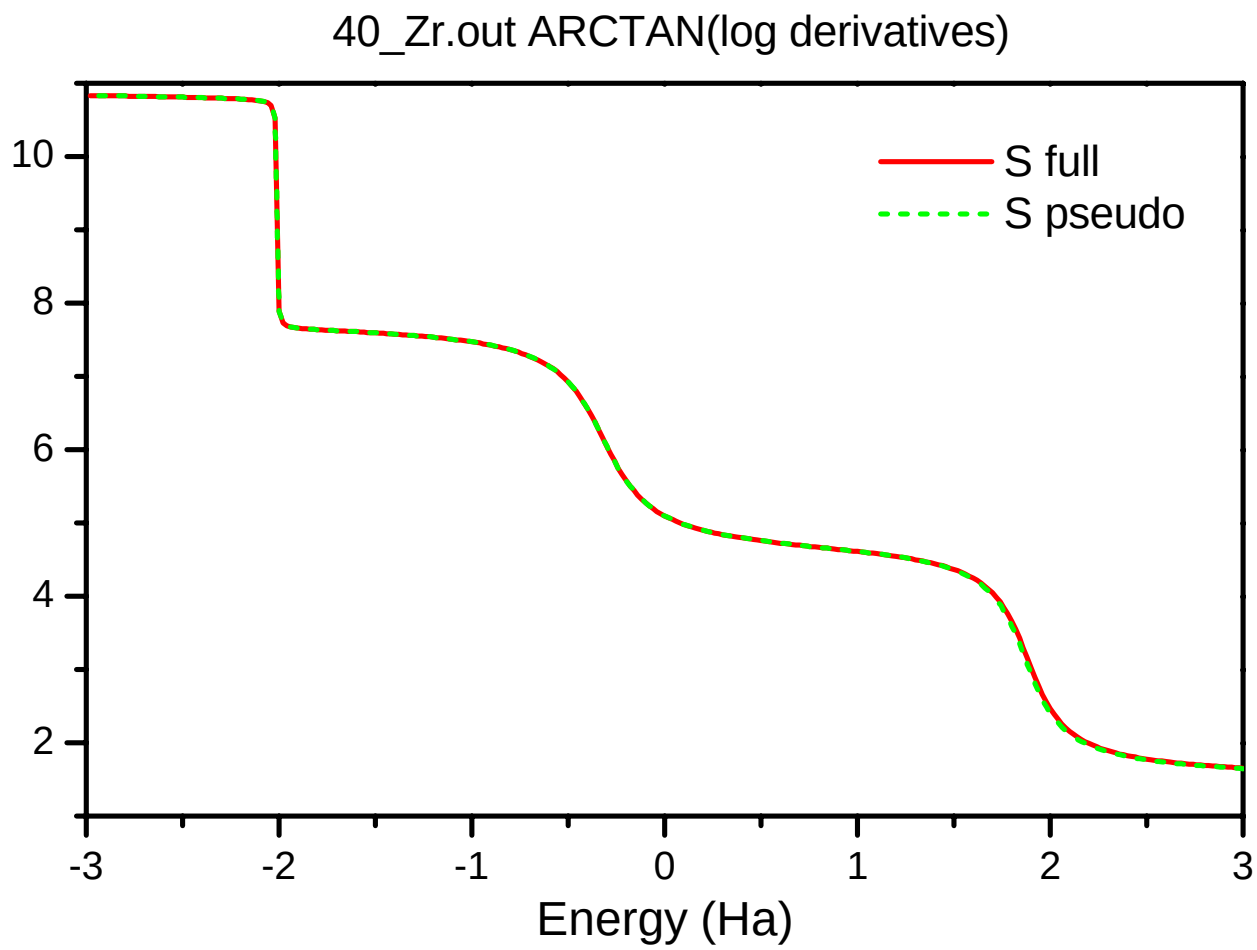
Hit enter to continue

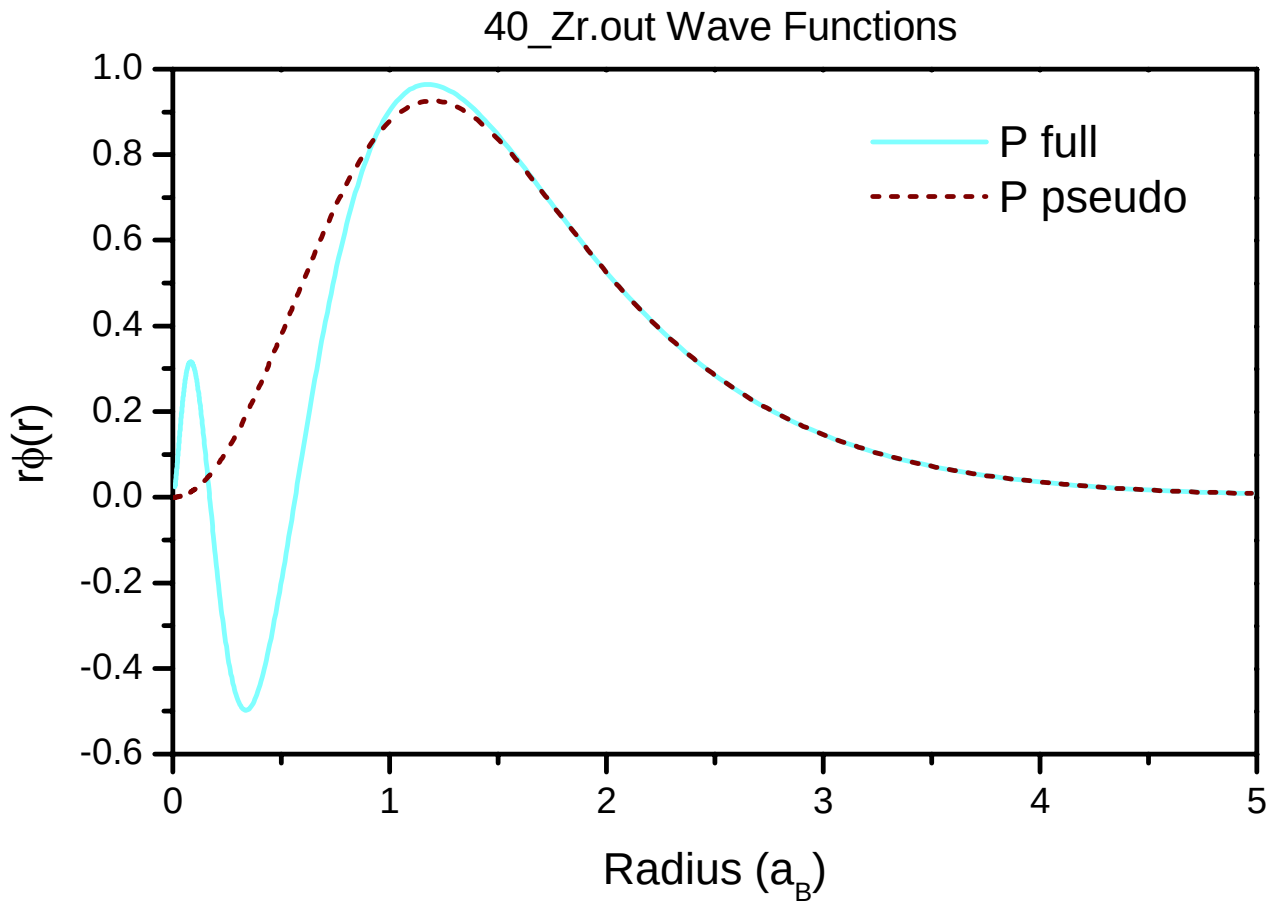

40_Zr.out Charge Densities
rhoM
rhoC
rhoV
40
30
20
10
0
a_B^{-3}
0
1
2
3
4
5
Radius (a_B)

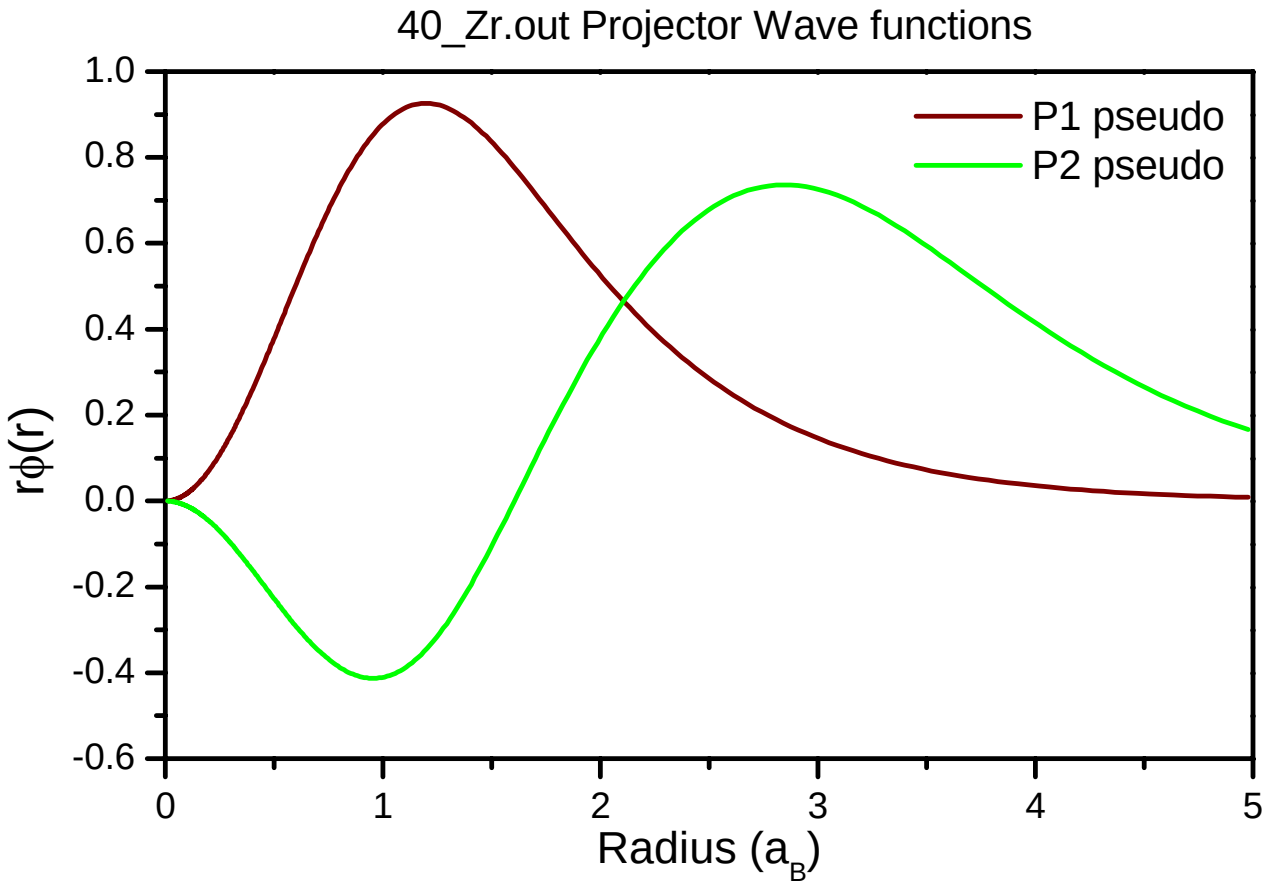

40_Zr.out Wave Functions
S full
S pseudo
rϕ(r)
Radius (a_B)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
1
2
3
4
5

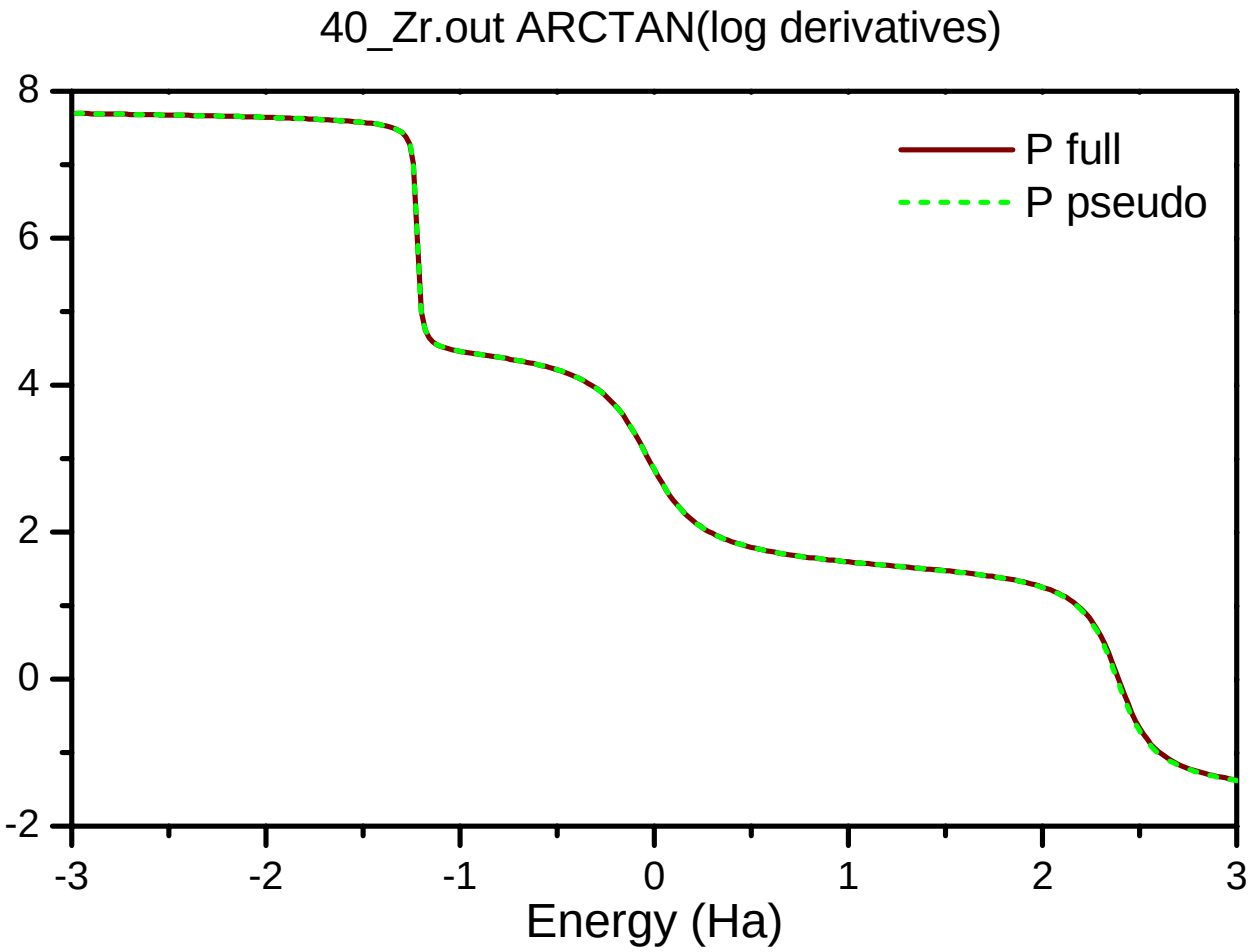

40_Zr.out Wave Functions
0.6
0.4
0.2
0.0
-0.2
rϕ(r)
0
1
2
3
4
5
Radius (a_B)
Sx full
Sx pseudo

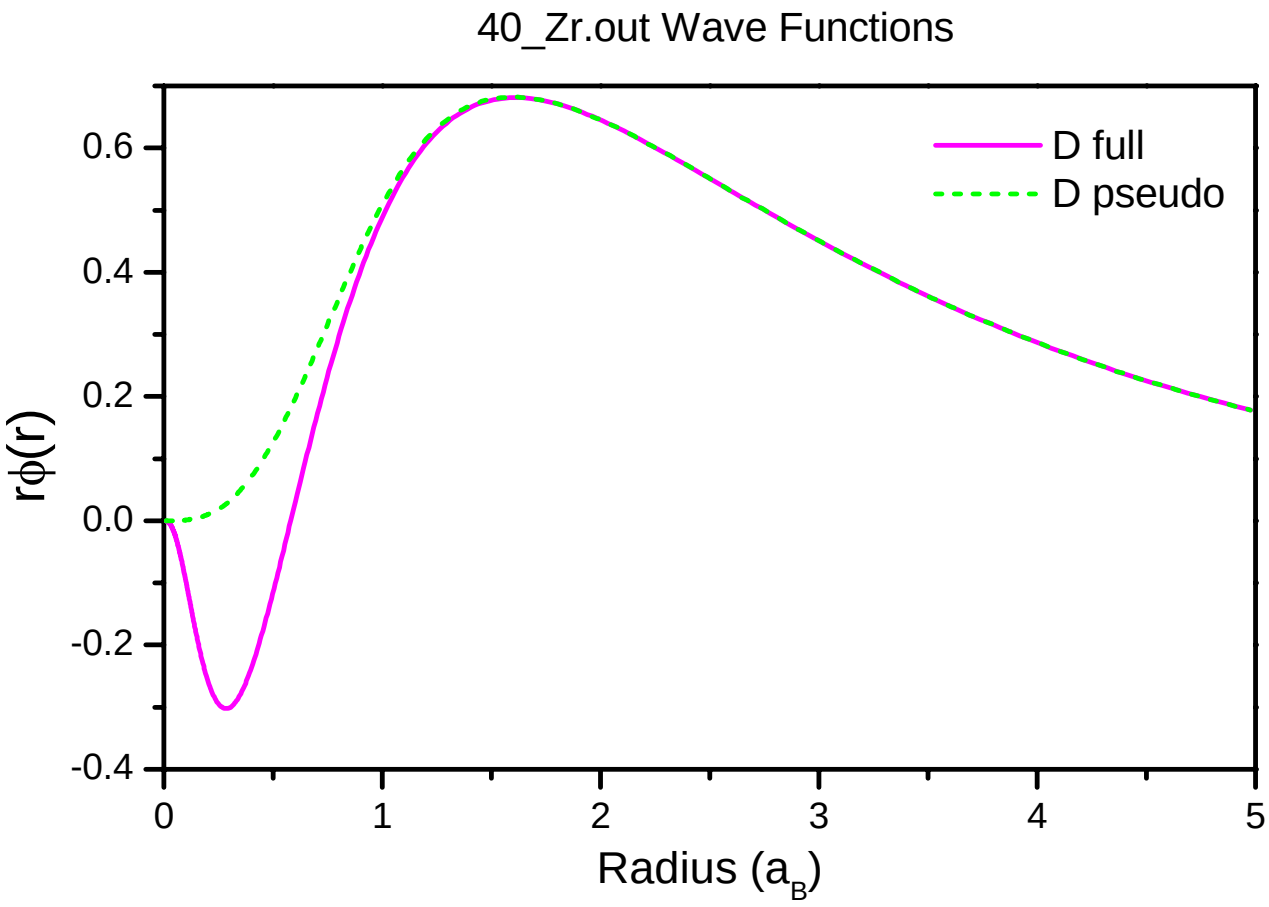

40_Zr.out Projector Wave Funcrions
S1 pseudo
S2 pseudo
rϕ(r)
Radius (a_B)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
1
2
3
4
5

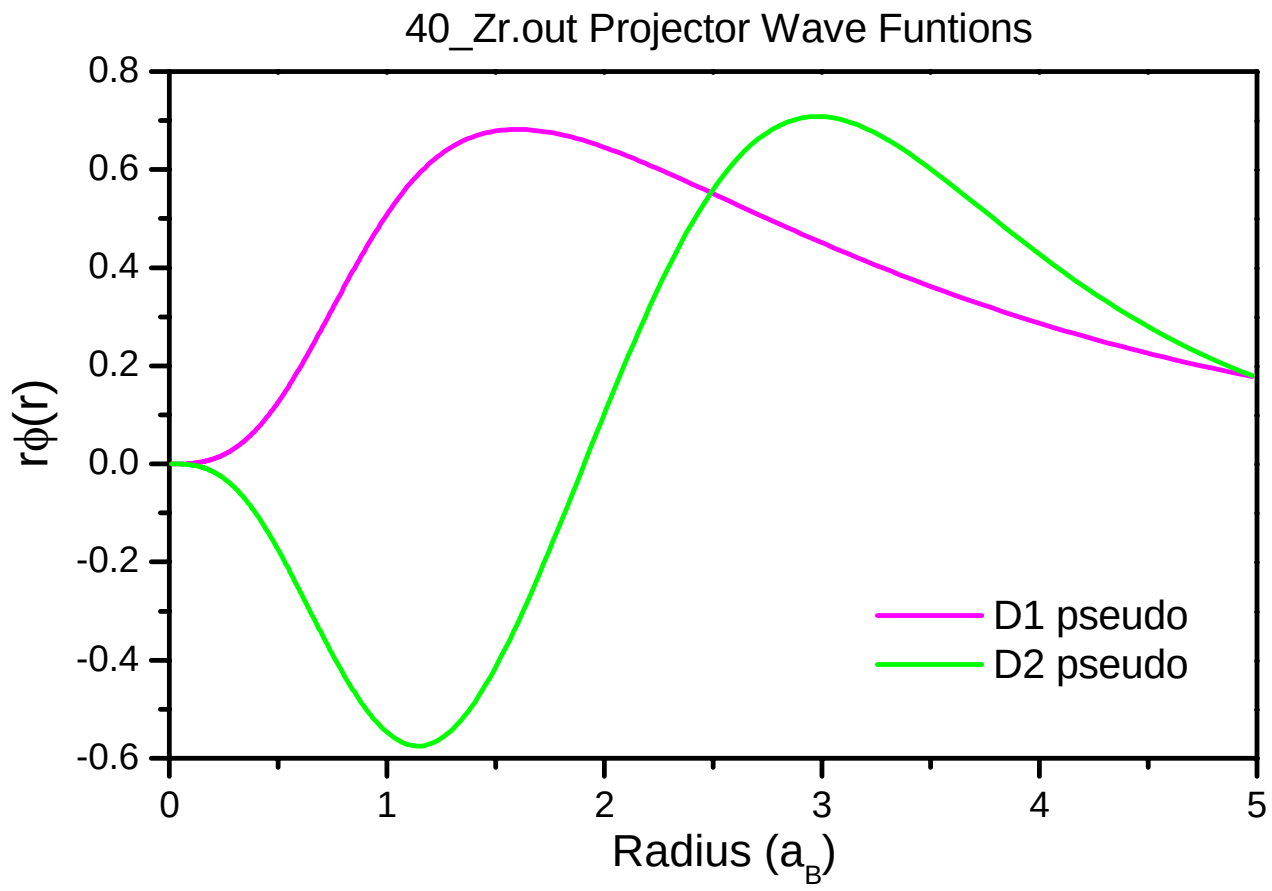

40_Zr.out ARCTAN(log derivatives)
S full
S pseudo
10
8
6
4
2
-3
-2
-1
0
1
2
3
Energy (Ha)

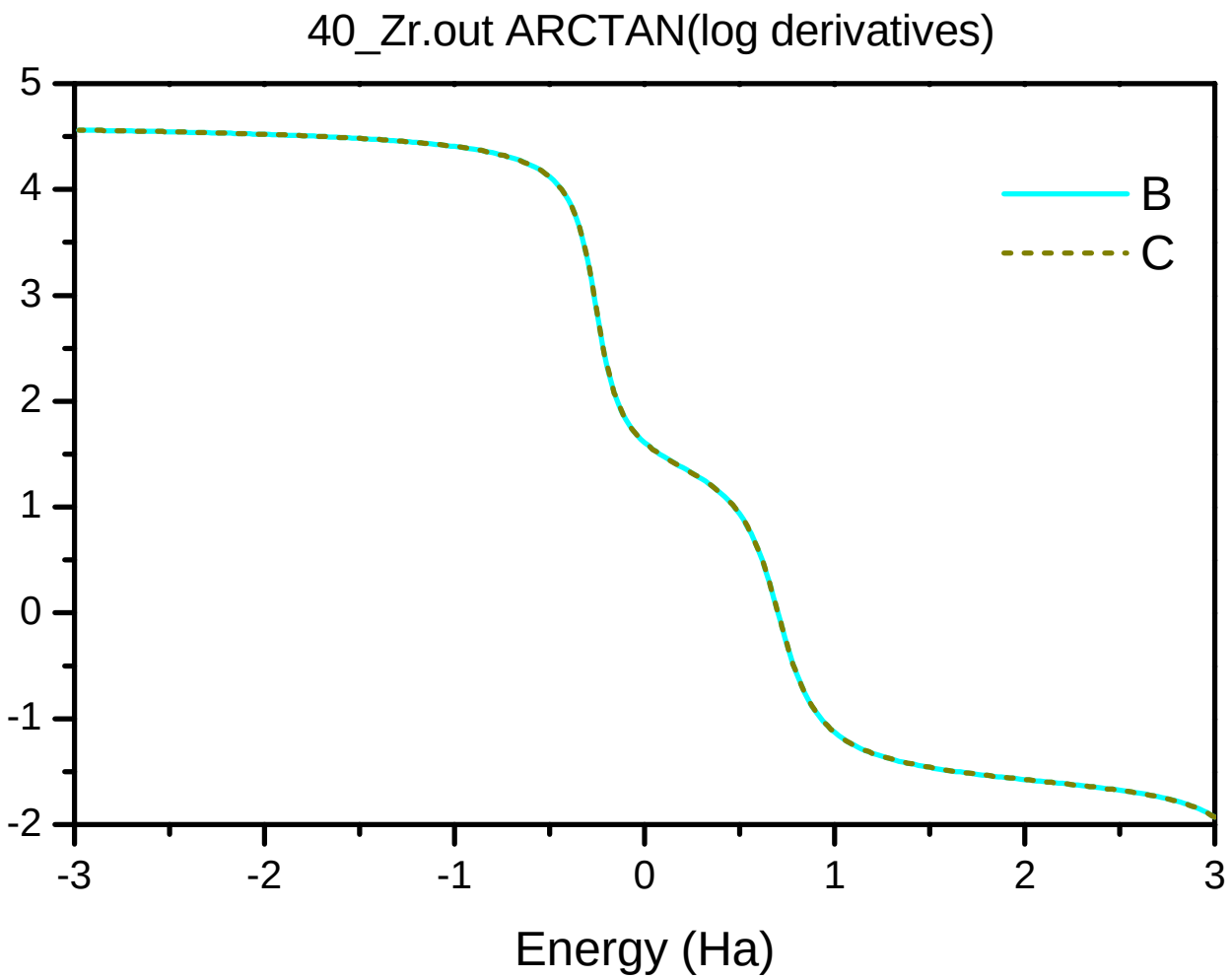

40_Zr.out Wave Functions
P full
P pseudo
rϕ(r)
Radius (a_B)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
1
2
3
4
5

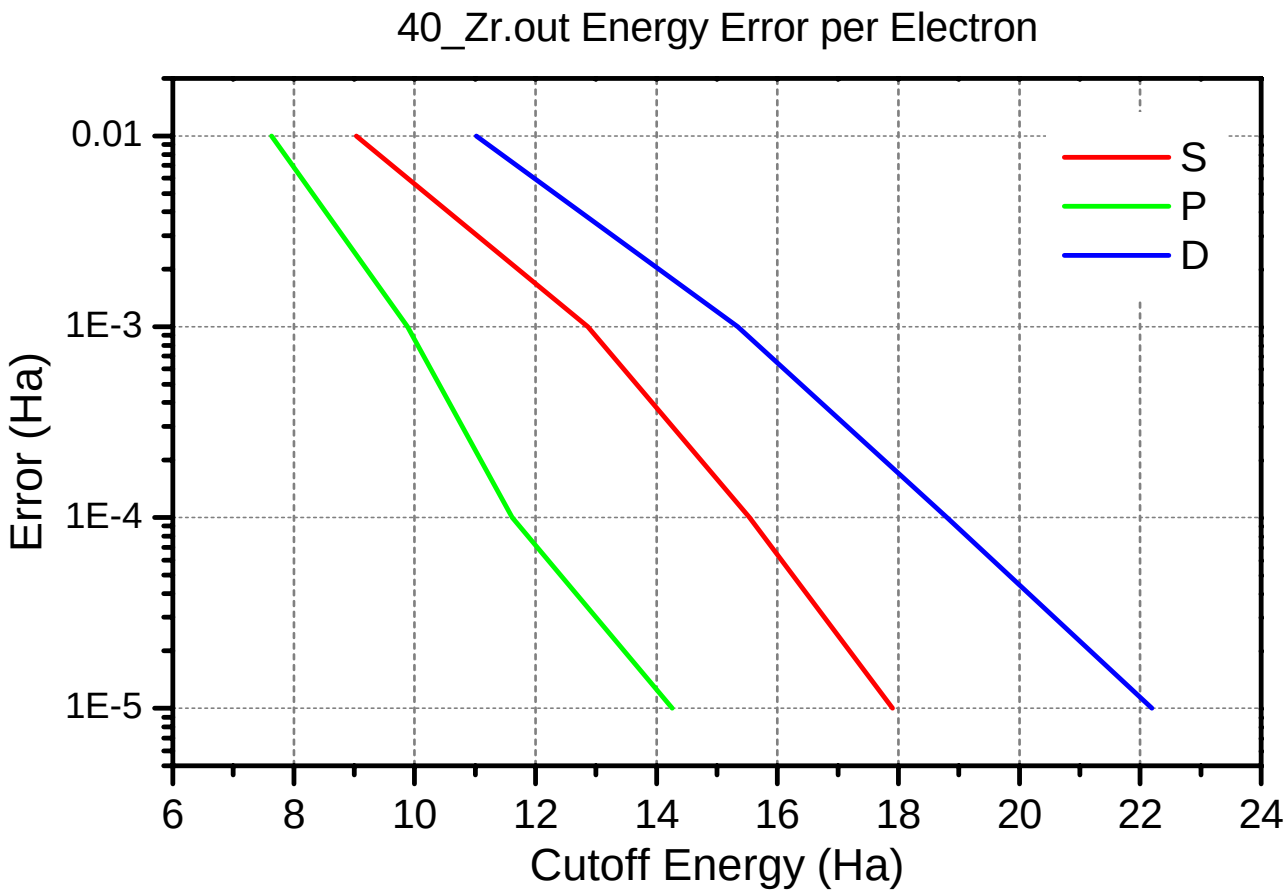

40_Zr.out Projector Wave functions
P1 pseudo
P2 pseudo
rϕ(r)
Radius (a_B)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
1
2
3
4
5


40_Zr.out ARCTAN(log derivatives)
P full
P pseudo
8
6
4
2
0
-2
-3
-2
-1
0
1
2
3
Energy (Ha)


40_Zr.out Wave Functions
D full
D pseudo
rϕ(r)
0.6
0.4
0.2
0.0
-0.2
-0.4
0
1
2
3
4
5
Radius (a_B)


40_Zr.out Projector Wave Funtions
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
rϕ(r)
0
1
2
3
4
5
Radius (a_B)
D1 pseudo
D2 pseudo


40_Zr.out ARCTAN(log derivatives)
5
4
3
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
Energy (Ha)
B
C


40_Zr.out Energy Error per Electron
Error (Ha)
0.01
1E-3
1E-4
1E-5
6
8
10
12
14
16
18
20
22
24
Cutoff Energy (Ha)
S
P
D

Obtaining and using ONCVPSP

You can download the open-source package from
http://www.mat-simresearch.com/
or see me later

The formalism, all the relevant references, and most of these results are in
D. R. Hamann, Phys. Rev. B **88**, 085117 (2013)

The two key papers upon which ONCVPSP is based are
D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990)
A. Rappe *et al*., Phys. Rev. B **41**, 1227 (1990)

For good ultrasoft/paw potentials and good test set see
Garrity, Bennett, Rabe & Vanderbilt, Comput. Mater. Sci. **81**, 446 (2014)